

**Risk**  
books

# Energy Markets

BY VINCENT KAMINSKI



# *Energy Markets*



# *Energy Markets*

---

**Vincent Kaminski**



Published by Risk Books, a Division of Incisive Media Investments Ltd

Incisive Media  
32–34 Broadwick Street  
London W1A 2HG  
Tel: +44(0) 20 7316 9000  
E-mail: [books@incisivemedia.com](mailto:books@incisivemedia.com)  
Sites: [www.riskbooks.com](http://www.riskbooks.com)  
[www.incisivemedia.com](http://www.incisivemedia.com)

© 2012 Incisive Media

ISBN 978 1 906348 79 3

British Library Cataloguing in Publication Data  
A catalogue record for this book is available from the British Library

Publisher: Nick Carver  
Commissioning Editor: Sarah Hastings  
Editorial Development: Alice Levick  
Managing Editor: Lewis O'Sullivan

Designer: Lisa Ling  
Copy-edited by Laurie Donaldson

Typeset by Mark Heslington Ltd, Scarborough, North Yorkshire  
Printed and bound in the UK by Berforts Group Ltd

#### Conditions of sale

*All rights reserved. No part of this publication may be reproduced in any material form whether by photocopying or storing in any medium by electronic means whether or not transiently or incidentally to some other use for this publication without the prior written consent of the copyright owner except in accordance with the provisions of the Copyright, Designs and Patents Act 1988 or under the terms of a licence issued by the Copyright Licensing Agency Limited of Saffron House, 6–10 Kirby Street, London EC1N 8TS, UK.*

*Warning: the doing of any unauthorised act in relation to this work may result in both civil and criminal liability.*

*Every effort has been made to ensure the accuracy of the text at the time of publication, this includes efforts to contact each author to ensure the accuracy of their details at publication is correct. However, no responsibility for loss occasioned to any person acting or refraining from acting as a result of the material contained in this publication will be accepted by the copyright owner, the editor, the authors or Incisive Media.*

*Many of the product names contained in this publication are registered trade marks, and Risk Books has made every effort to print them with the capitalisation and punctuation used by the trademark owner. For reasons of textual clarity, it is not our house style to use symbols such as TM, ®, etc. However, the absence of such symbols should not be taken to indicate absence of trademark protection; anyone wishing to use product names in the public domain should first clear such use with the product owner.*

*While best efforts have been intended for the preparation of this book, neither the publisher, the editor nor any of the potentially implicitly affiliated organisations accept responsibility for any errors, mistakes and or omissions it may provide or for any losses howsoever arising from or in reliance upon its information, meanings and interpretations by any parties.*

# Contents

<b>About the Author</b>	vii
<b>Abbreviations</b>	xix
<b>Introduction</b>	xxi

## SECTION 1: ENERGY TRADING: THE PAST AND THE FUTURE

<b>1 Energy Trading and Marketing : The Macro View</b>	3
<b>2 Energy Trading: The Organization</b>	51
<b>3 Weather Information in Energy Trading</b>	85

## SECTION 2: PARTICIPANTS AND INSTRUMENTS

<b>4 Energy Markets: The Instruments</b>	109
<b>5 Energy Markets: Structured Transactions</b>	157
<b>6 Energy Markets: Exchanges</b>	199
<b>7 Energy Markets: Market Participants and Regulatory Developments</b>	253

## SECTION 3: NATURAL GAS

<b>8 Natural Gas: Upstream</b>	283
<b>9 Non-Conventional Natural Gas</b>	325
<b>10 Natural Gas Transportation and Storage</b>	365
<b>11 US Natural Gas Markets</b>	413
<b>12 International Natural Gas Markets</b>	463

**SECTION 4: OIL MARKETS**

<b>13 Oil Markets: Properties, Production and Reserves</b>	<b>503</b>
<b>14 Non-Conventional Oil</b>	<b>533</b>
<b>15 Oil Processing</b>	<b>565</b>
<b>16 Oil Transportation and Storage</b>	<b>601</b>
<b>17 Oil Pricing</b>	<b>625</b>
<b>18 Transactions in the Oil Markets</b>	<b>663</b>

**SECTION 5: ELECTRICITY, EMISSIONS AND COAL**

<b>19 Electricity: The Basics</b>	<b>695</b>
<b>20 Power Generation</b>	<b>723</b>
<b>21 Transmission, Loads and Power Pools</b>	<b>755</b>
<b>22 Analytical Tools</b>	<b>801</b>
<b>23 Electricity Markets Transactions</b>	<b>837</b>
<b>24 Manipulation and Gaming of Energy Markets</b>	<b>865</b>
<b>25 Emission Markets</b>	<b>895</b>
<b>26 Coal</b>	<b>941</b>
<b>Conclusion</b>	<b>963</b>
<b>Index</b>	<b>967</b>

## About the Author

Mr Vincent Kaminski has spent 14 years working in different positions related to quantitative analysis and risk management in the merchant energy industry. The companies he worked for include Citigroup, Sempra Energy Trading, Reliant Energy, Citadel Investment Group, and Enron (from 1992 to 2002) where he was the head of the quantitative modeling group. Prior to starting a career in the energy industry, Mr. Kaminski was a vice president in the research department, bond portfolio analysis group, of Salomon Brothers in New York (from 1986 to 1992). As of September 15, 2006, Mr. Kaminski has accepted an academic position with Rice University in Houston (Jesse H. Jones Graduate School of Business) where he is teaching MBA level classes on energy markets, energy risk management and valuation of energy derivatives.

Mr Kaminski holds an M.S. degree in international economics, a Ph.D. degree in theoretical economics from the Main School of Planning and Statistics in Warsaw, Poland, and an MBA from Fordham University in New York. He is a recipient of the 1999 James H. McGraw award for Energy Risk Management (Energy Risk Manager of the Year). Mr. Kaminski published a number of papers, and contributed to several books, on energy markets.



# Abbreviations

AAPG	American Association of Petroleum Geologists
ACA	Annual charge adjustment
ACER	Agency for the Cooperation of Energy Regulators
ACRP	Average contract reference price
ADP	Alternative delivery procedure
AFRA	Average freight rate assessment
AGO	Atmospheric gas oil
AKI	Anti-knock index
AMSEC	Annual Monthly System Entry Capacity
ANC	Adjusted net capital
ANE	Available net exports
ANGA	American National Gas Association
AO	Authorised overrun
AOC	Attestation of Compliance
AP	Authorised participant
APERC	Asia Pacific Energy Research Centre
API	American Petroleum Institute
API	Application programming interface
ARA	Antwerp–Rotterdam–Amsterdam
ASCI	Argus Sour Crude Index
ASPO	Association for Study of Peak Oil & Gas
ASTM	American Society for Testing and Materials
ATN	Alternative trading network
ATRS	American Tanker Rate Schedule
ATS	Alternative trading system
BAFA	Bundesamt für Wirtschaft und Ausfuhrkontrolle
Bcf	Billion cubic feet
BCTI	Baltic Clean Tanker Index
BDTI	Baltic Dirty Tanker Index
BFA	Baltic Forward Assessments
BFO	Baltic Freight Index
BG	British Gas

Biffex	Baltic International Freight Futures Exchange
BIS	Bank for International Settlements
BNO	Brent Oil Fund
BOEM	Bureau of Ocean Energy Management
BOEMRE	Bureau of Ocean Energy Management, Regulation and Enforcement
BOM	Balance of month
BPI	Baltic Panamax Index
BSEE	Bureau of Safety and Environmental Enforcement
BTL	Biomass-to-liquids
Btu	British Thermal Unit
CAA	Clean Air Act
CAPM	Capital asset pricing model
CAT	Cumulative average temperature
CCP	Clearing counterparty
CCRO	Committee of Chief Risk Officers
CCS	Cyclic steam simulation
CDD	Cooling degree day
CDG	Centro de Gravedad
CEA	Commodity Exchange Act
CEC	Commodity Exchange Commission
CEER	Council of European Energy Regulators
CEGH	Central European Gas Hub
CFD	Contract-for-difference
CFMA	Commodity Futures Modernization Act
CFTC	Commodity Futures Trading Commission
CHI	Carvill Hurricane Index
CHOPS	Cold heavy oil production on site
CIF	Cost, insurance, freight
CIM	Commodity index multiplier
CIP	Commodity index percentage
CL	Crude oil
CLOB	Central limit order book
CLP	Commodity liquidity percentage
CMA	Calendar Merc Average
CME	Chicago Mercantile Exchange
CNG	Compressed natural gas
CO	Carbon monoxide
CoA	Contract of affreightment

COP	ConocoPhillips
COT	Commitment of traders
CPI	Consumer Prices Index
CPP	Commodity production percentage
CPUC	California Public Utilities Commission
CPW	Contract production weight
CRB	Commodity Research Bureau
CTA	Commodity trading advisor
CTL	Coal-to-liquids
CVA	Credit valuation adjustment
CWA	Clean Water Act
DADSEC	Day Ahead Daily System Entry Capacity
DCM	Designated contract market
DCM	Direct clearing members
DCS	Distributed control system
DDG	Dried distillers grain
DFL	Dated-to-frontline
DIPE	Diisopropyl ether
DME	Dubai Mercantile Exchange
DNO	United States Short Oil Fund
DOE	Department of Energy
DPA	Department of Petroleum Affairs
DSRO	Designated self-regulatory organization
Dwt	Deadweight tonnage
E&P	Exploration and production
EBOT	Exempt board of trade
ECE	Eligible commercial entity
ECM	Exempt commercial market
ECMWF	European Center for Medium-Range Weather Forecasts
ECN	Electronic communications network
EEI	Edison Electric Institute
EFP	Exchange for physical
EFS	Exchange for swaps
EIA	Energy Information Administration
EMIR	European Market Infrastructure Regulation
ENSO	El Niño Southern Oscillation
EOL	EnronOnLine
EPA	Energy Policy Act

EPA	Environmental Protection Agency
EPD	Enterprise Product Partners
EPSQ	Elapsed prorated scheduled quantity
ERCOT	Electric Reliability Council of Texas
EROI	Energy return on investment
ETBE	Ethyl tertiary butyl ether
ETF	Electronically traded fund
ETF	Exchange-traded fund
ETN	Exchange-traded note
ETP	Energy Transfer Partners
EU	European Union
EUR	Estimated ultimate recovery
Eurostat	Statistical Office of the European Commission
FASB	Financial Accounting Standards Board
FBOT	Foreign board of trade
FCC	Fluid catalytic cracker
FCM	Futures commission merchant
FDIC	Federal Deposit Insurance Corporation
FEMA	Federal Agency Management Service
FERC	Federal Energy Regulatory Commission
FFA	Forward freight agreement
FLNG	Floating LNG
FOB	Free on board
FSA	Financial Services Authority
FSRU	Flexible storage and regasification unit
FTS	Fischer-Tropsch synthesis
G&A	General and administrative
G20	Group of 20
GCM	General clearing member
GCSI	Goldman Sachs Commodity Index
GFS	Global Forecast System
GGE	Gasoline gallon equivalent
GIIP	Gas initially in place
GISB	Gas Industry Standards Board
GJ	Giga joules
GLD	Gold
GM	General Motors
GNE	Global net exports
GPW	Gross product worth

GTL	Gas-to-liquids
GTS	Gas Transport Services
GUI	Graphical user interface
HCDP	Hydrocarbon dew point
HDD	Heating degree day
HFAU	Hydrofluoric alkylation unit
HHI	Herfindahl–Hirschman Index
HIQ	Hourly injection quantity
HO	Heating oil
HSC	Houston Ship Channel
HTQ	Hourly transportation quantity
HWQ	Hourly withdrawal quantity
ICE	Intercontinental Exchange
IDB	Interdealer broker
IEA	International Energy Agency
IEF	International Energy Forum
Imarex	International Maritime Exchange
IMC	Internal model control
IOR	Improved oil recovery
IOSCO	International Organization of Securities Commissions
IPE	International Petroleum Exchange
IPRO	Independent price reporting organisation
ISDA	International Swap and Derivatives Association
ITS	Interruptible transportation service
JAS	Joint Association Survey
JCC	Japanese Crude Cocktail
JODI	Joint Oil Data Initiative
kWh	Kilowatt-hour
L&U	Lost and unaccounted for
LADE	Latin American Energy Organization
LCH	London Clearing House
LDC	Local distribution company
LHV	Lower heating value
Libor	London Interbank Offered Rate
Liffe	London International Financial Futures Exchange
LLS	Light Louisiana Sweet
LNG	Liquefied natural gas
LOOP	Louisiana Offshore Oil Port
LP	Linear programming

LPG	Liquefied petroleum gases
LSFO	Low sulphur fuel oil
LTBP	London Tanker Brokers' Panel
MDID	Maximum daily injection quantity
MDQT	Maximum daily transportation quantity
MDWQ	Maximum daily withdrawal quantity
MEND	Movement for Emancipation of the Niger Delta
MFT	Mature fine tailings
MFV	Modified fixed-variable
MiFID	Markets in Financial Instruments Directive
MJO	Madden–Julian Oscillation
MMS	Materials Management Service
MMS	Minerals Management Service
MOC	Market-on-close
MON	Motor Octane Number
MOT	Ministry of Transport
MPC	Model predictive control
MSLP	Mean sea level pressure
MSP	Major swap participant
MSQ	Maximum storage quantity
MTBE	Methyl tertiary butyl ether
MTM	Mark-to-market
MVC	Model variable control
MW	Megawatt
NAAQS	National Ambient Air Quality Standard
NAESB	North American Energy Standards Board
NAM	Nederlandse Aardolie Maatschappij
NAO	North Atlantic Oscillation
NAV	Net asset value
NBP	National Balancing Point
NBS	National Bureau of Statistics
NCEP	National Center for Environmental Prediction
NCG	NetConnect Germany
NCM	Non-clearing member
NEPA	National Environmental Policy Act
NEV	Net energy value
NG	Natural gas
NGA	Natural Gas Act of 1938
NGI	Natural Gas Intelligence

NGL	Natural gas liquid
NGPA	Natural Gas Policy Act
NGPL	Natural gas plant liquids
NGPSA	Natural Gas Pipeline Safety Act of 1968
NGX	Natural Gas Exchange
NH	Northern Hemisphere
NOAA	National Oceanic and Atmospheric Administration
NOC	National oil company
NOPR	Notice of proposed rulemaking
NORM	Naturally occurring radioactive material
NPV	Net present value
NTS	National Transmission System
Nymex	New York Mercantile Exchange
OCM	On-the-day commodity market
ODAC	Oil Depletion Analysis Centre
OGEM	Office of Gas and Electricity Markets
OFO	Operational flow order
OIIP	Oil initially in place
OOG	Office of Oil and Gas
OPEC	Organization of the Petroleum Exporting Countries
OSP	Official selling price
OSPW	Oil sands process wastewater
OTC	Over-the-counter
P&A	Plug and abandon
P&L	Profit and loss
PADD	Petroleum Administration Defense District
PCG	Potential Gas Committee
PDO	Pacific Decadal Oscillation
PDP	Proved developed producing
PDW	Percentage dollar weight
PG&E	Pacific Gas and Electric
POI	Percentage-of-index
POP	Percentage-of-proceeds
PPEG	Points d'Echange Gaz
PRA	Price reporting agency
PRMS	Petroleum Resources Management System
PSA	Petroleum Supply Annual
PSM	Petroleum Supply Monthly
PSRS	Petroleum Supply Reporting System

PSV	Punto di Scambio Virtuale
PUD	Proved undeveloped
QSEC	Quarterly System Entry Capacity
RBOB	Reformulated gasoline blendstock for oxygen blending
RFG	Reformulated gasoline
RFP	Request for proposal
RMTNTSEC	Rolling Monthly Trades & Transfer System Entry Capacity
RON	Research Octane Number
RPCA	Rotated principal component analysis
RPDW	Reference percentage dollar weight
RPMS	Refinery and Petrochemical Modeling System
RVP	Reid vapor pressure
SAAU	Sulphuric acid alkylation unit
SAGD	Steam-assisted gravity drainage
SAP	System average price
SBP	Specific boiling point
SCADA	Supervisory control and data acquisition
Scf	standard cubic foot
SD	Standard deviation
SDR	Swap data repository
SDWA	Safe Drinking Water Act
SEC	Securities and Exchange Commission
SEF	Swap execution facility
SFV	Straight fixed-variable
SIFMA	Securities Industry & Financial Management Association
SLP	Sea level pressure
SLV	Silver
SMBP	System marginal buy price
SMSP	System marginal sell price
SO	System operator
SOI	Southern Oscillation Index
SPA	Sale and purchase agreement
SPAN	Standard Portfolio Analysis of Risk
SPDC	Significant price discovery function contract
SPE	Society of Petroleum Engineers
SPE	Special-purpose entity

SPEE	Society of Petroleum Evaluation Engineers
SPM	Single point mooring
SPR	Strategic Petroleum Reserve
SR	Simple ratio
SRG	Straight-run gasoline
SRN	Straight-run naphtha
SST	Sea surface temperature
SSU	Saybolt Seconds Universal
STS	Ship-to-ship
TAME	Tertiary amyl methyl ether
TAN	Total acid number
TAP	Trans-Alaska Pipeline
TAS	Trading at settlement
TBA	Tertiary-butyl alcohol
TCE	Time charter equivalent
TDVT	Total dollar value traded
TEL	Tetra ethyl lead
THAI	Toe-to-heel air injection
TPA	Third-party access
TRRC	Texas Railroad Commission
TSO	Transmission system operator
TTF	Title Transfer Facility
UGA	United States Gasoline Fund
UHN	United States Heating Oil Fund
ULCC	Ultra-large crude oil carriers
ULSD	Ultra-low sulphur diesel
UNC	Uniform Network Code
UNG	United States Natural gas
UNL	United States 12 Month Natural Gas Fund
UNSD	United Nations Statistics Division
URR	Ultimately recoverable resources
USGC	US Gulf Coast
USGS	US Geological Survey
USL	United States 12 Month Oil Fund
USMC	United States Maritime Commission
USO	United States Oil Fund
VaR	Value-at-risk
VGO	Vacuum distillation
VLCC	Very-large crude oil carriers

VOC	Volatile organic compound
VPP	Volumetric production payments
VWAP	Volume-weighted average price
WDG	Wet distillers grain
WPC	World Petroleum Council
WPSR	Weekly Petroleum Status Report
WTI	West Texas Intermediate

# Acknowledgements

We would like to thank Michael Giesinger of Barclays Capital for contributing the section on Basel III. We are also grateful to the following executive practitioners for the knowledge gained through direct discussion and our observation of their work, both of which stimulated our thought processes:

Tom Flynn, executive vice-president and chief financial officer of BMO Financial Group, for integrating the finance and risk disciplines and always asking the right questions. Terry Bulger, executive vice-president, US risk management and chief risk officer at BMO Financial Corp and Mike Stramaglia, former executive vice-president and chief risk officer of Sun Life Financial Group, for their appetite for difficult conceptual issues and dedication to the improvement of risk management practices.

We would also like to thank Sarah Hastings, Alice Levick, and Lewis O'Sullivan of Risk Books.



# *Introduction*

*Energy Markets* is the first of two books aimed at providing a comprehensive, systematic and extensive review of these most complex and constantly evolving markets, markets which are critical to maintaining the standard of living achieved by highly developed societies, as well as the future fortunes of emerging economies. The objective of this introduction is to explain the scope, organisation and, most importantly, the philosophy behind the book – the product of over 25 years we have spent working in, or with, merchant energy companies and large financial institutions.

The topics covered in *Energy Markets* include:

- ❑ the production, transportation, distribution and storage of different energy commodities, such as natural gas, crude oil and refined products, coal, electricity and certain types of environmental products (emission credits);
- ❑ certain aspects of the laws and regulations affecting energy markets;
- ❑ price formation and the price discovery process in the energy markets;
- ❑ a review of the entities which form the infrastructure of the energy markets – such as price reporting agencies (PRAs), exchanges, voice brokers, etc;
- ❑ the most important types of transactions used in the energy markets; and
- ❑ a general introduction to energy-related derivatives.

*Energy Markets* is organised into five sections, covering the following component parts of the energy complex:

- ❑ the basics of energy trading:
  - economic and social forces shaping energy markets;
  - organisation of an energy trading company;
  - the “plumbing” of the energy markets (the basic components of market infrastructure); and
  - recent legal and regulatory developments.
- ❑ natural gas markets;
- ❑ crude oil and refined products markets;
- ❑ electricity markets; and
- ❑ emission markets.

A detailed discussion of the quantitative aspects of energy markets (derivative valuation, modelling of price processes, risk management models and principles) is deferred until the second book. The structure of the entire book is not an accident: it is rather a reflection of the fundamental beliefs concerning how the energy markets function and what are the critical skills required to navigate these turbulent waters.

## THE SCOPE

One obvious question is why this book – representing as it does an effort to identify, describe and synthesise the main aspects of different parts of the energy complex – should be written at all. Why should one not settle for a number of specialised books covering separate energy commodity markets, or even segments of different markets? The answer is that energy traders need both types of books. The main reason why detailed, narrow studies of different elements of the energy complex are not only highly useful but also necessary is that no single, general book can explore in depth all the details a practitioner has to master to function in a specific sector of the industry. There are, however, many good reasons to write a book that represents an attempt to synthesise the related developments unfolding in front of us at an ever-increasing speed. Reliance on narrow, highly specialised texts presents a risk of failing to see the forest for the trees: being unable to discern a general pattern behind a large volume of detailed information. One should recognise that there are important similarities and interactions between different energy markets that require a holistic treatment of this field. This introduction therefore offers a review of the reasons

why a comprehensive book on the modern energy markets is needed.

Many excellent contributions to this field are available,<sup>1</sup> and should be on the bookshelf of any energy trader or energy industry analyst. We have personally benefitted immensely from the collective wisdom of the industry and learned a lot from a number of books and papers on commodity markets. The only complaint we might have is that some of them have become somewhat obsolete due to the rapid pace of change in the industry. The old adage that nobody steps twice into the same river is particularly true of the energy business. Also, writing books is a time-consuming task, one that offers mediocre financial rewards compared to working in the trenches of energy trading and risk management.

A personal reflection may explain how we came up with the outline for the book. When we started in the energy industry in 1992, our job description was very short and straightforward. We were to be responsible for the development of option-pricing models for natural gas. We did not think that we needed to develop any insights into crude or electricity markets to do our job. We also believed, somewhat naively, that the tools from the financial markets could be easily transferred to the commodity markets. Over time, the firm we worked for, Enron Corporation, expanded its tiny trading operation into several giant trading floors, with the printout of trading positions reaching a height of several feet. Each time the addition of a new set of commodities was contemplated, we had to struggle to prepare the group we were managing for new challenges, for the development of a new suite of valuation, risk management and decision-support models for a new business. Initially, each project was a snowflake, with the only common denominator being the set of skills in financial engineering, statistics and computer programming required to solve the problems before us. In time, many commonalities and similarities between different parts of the energy complex became obvious and we, along with our colleagues, realised that we were dealing with one integrated industry that can be understood as a single complex system, with multiple interactions between its separate parts. The days are long gone when one person can spend their career working for one sector of the energy industry – for example, in natural gas – without closely following other sectors, across the full spectrum of commodities and geographical

locations. This makes the life of professionals in this business more interesting, but also very stressful. A few observations explain why this is the case.

Energy commodities are used to satisfy the same fundamental needs of heat, light and motion. Electricity is also used, in addition to meeting these needs, in supporting the information infrastructure of the modern society through the creation and propagation of electromagnetic fields. The proportions in which different energy commodities are used vary over time and from country to country, depending on their availability and costs of production, and on the emergence of new technologies. The final energy mix is the result of interactions between market forces and political processes, and one has to understand the entire dynamic system. Understanding the history of the industry is helpful, but is not a substitute for critical and unconventional thinking. After all, as Samuel Taylor Coleridge said, “History [...] is but a lantern on the stern, which shines on the waves behind us.”<sup>2</sup> Our industry is far too volatile to rely on insights from history alone. We are always one technological development, one price shock, or one major industrial accident away from drastically revising our expectations of the way in which the future energy needs of our society will be satisfied.

Energy commodities compete against each other as substitutes and, at the same time, are critical inputs to the production of other energy commodities. For example, some refined products compete against natural gas as sources of heat for residential and commercial dwellings, and as fuels used in transportation (although the share of natural gas is still very small compared to gasoline and diesel). At the same time, natural gas is used as an input in the production of electricity and represents an alternative to coal and certain distillates. Natural gas is also used as a generation fuel and the source of heat in many refineries. Substitution between competing sources of energy happens at the margin, under the influence of relative prices, and subject to the technological constraints of the physical infrastructure. This explains why one cannot simply focus on price developments to explain the evolution of demand and supply: one has to understand the operations side of the industry to make useful predictions about market developments. It is equally important to understand the dynamics of other sectors of the world economy that are important consumers of energy commodities. This is why energy traders spend

so much time thinking, for example, about trends in the plastic industry and economic growth in China.

These general statements will be exemplified many times in the subsequent chapters of this book. Examples will show why disruptions in the production and flows of one energy commodity can shift demand to competing energy sources, trigger flows from one location to another and produce a realignment of prices across both space and time. An understanding of different possible market interactions and the ability to predict them produces the potential for profitable trading opportunities and/or the means to avoid significant losses. Linkages between different markets are not static but evolve constantly under the impact of technological, economic and social forces.<sup>3</sup> The art of energy trading is the ability to anticipate such changes ahead of the rest of the crowd. A person with the talent to see around the corner tends to be the most valuable asset on any energy-trading floor.

Energy commodities also compete for the inputs to production processes (materials such as steel and cement), the pool of experienced labour and specialised infrastructure (railway lines, ports and water channels). Different parts of the energy complex compete for the finite pool of investment funds and the capacity of the financial markets to provide risk capital in the form of insurance underwriting lines and funding, in addition to the trading volumes critical to the ability to control price and credit exposures.

The following observations help to explain the philosophy underlying this book. First, energy markets have evolved over the last 15 years into a highly integrated, global system with shocks propagating across different geographical locations and specific commodity markets through very complicated and constantly evolving channels of transmission. One cannot understand one energy commodity market in isolation if other related parts of the entire system are ignored. One of the consequences of these developments is growing price volatility that calls for new creative ways of managing both market and credit risks. This new market reality creates enormous profit opportunities for financial institutions engaged in energy trading and provides risk management tools to the producers and end users of energy commodities. One of the objectives of this book is the identification and description of the links among these different markets.

---

Second, the production and delivery of energy takes place within a complex system with three interacting layers: (i) a physical layer consisting of the hard assets used for production, transportation and storage of primary energy sources, and for the transformation of one form of energy into another; (ii) markets for energy that consist of interacting spot, forward, option and long-term structured transactions; and (iii) the system of national laws, regulations and international treaties. The market transactions can be purely financial or physical. The financial transactions are cash-settled, with the cashflows calculated through formulas referencing energy prices established in different markets (and possibly other prices and variables). The physical transactions require delivery of physical commodities (spot transactions are, by definition, physical transactions). These transactions are often structured using the templates developed for the financial markets, including swaps and options. Both physical and financial transactions happen in a system governed by multiple laws, regulations and conventions, which require the constant involvement of legal and compliance professionals. The laws that apply are often the same rules that are written for the financial markets, often to the chagrin of energy professionals – who believe, quite correctly, that their business is different and cannot be painted with the same brush.

Energy commodities are critical to our standard of living, directly or because they are used as critical inputs in the production and distribution of such essential goods as water and food. Most citizens of the developed countries see cheap energy as their birthright, and increases in prices trigger an immediate and strong social and political reaction. Emerging nations hope to raise their citizens out of poverty and reach a standard of living comparable to those of the North Atlantic region within a generation or two. The examples of Japan and Singapore demonstrate that this is feasible, but may not be easy to replicate on a much larger scale. Barriers to growth may be reached in the short-run because of the difficulty of increasing certain energy supplies (to be discussed later). An understanding of the social and geopolitical ramifications of the energy industry is very important to any trader and risk manager. We must be cognisant of the fact that we are not making our decisions in a social and political vacuum. There is far more to trading and risk management than mathematical formulas.

One critical difference between the financial and commodity markets explains why an understanding of the physical layer of the industry is imperative. Financial instruments exist as entries in computer databases and can be transferred from one counterparty to another with a keystroke. At the end of the day, transactions in the commodity markets have to obey the constraints imposed by the laws of physics and chemistry. Even the cleverest financial engineering cannot suspend the laws of preservation of energy or preservation of matter. From time to time, the energy markets are invaded by men and women from the fixed income or equity markets, with exceptional mathematical skills, but with a limited or non-existent understanding of the messy realities of the energy supply chain. In the interest of full disclosure, we must include ourselves in this group. The consequences can sometimes be catastrophic, both for them and their institutions and clients. This book can facilitate interactions between different groups of professionals collaborating, more by necessity than choice, in the energy markets, contributing different skills and perspectives.

There are further reasons to cover all the energy markets within one book. A newcomer to the industry is bewildered by the complexities of each specialised market, the variety of transaction types, highly specialised procedures for discovery and dissemination of prices, details of logistical arrangements and the cryptic language used in the contracts. At some point, if an energy trader has an opportunity to rotate through multiple desks and acquire familiarity with different sectors of the industry, commonalities between separate energy markets emerge; many diverse instruments and deal types become special cases of certain general transaction types. Once this is recognised, many complex tasks can be simplified and similarities between different markets will facilitate the more efficient design of trading and risk management platforms. This book provides assistance in the task of managing an energy trading business as one integrated operation.

The philosophy of this book has been greatly influenced by almost 10 years spent working for Enron Corporation. It is not our objective to write a panegyric for Enron, given the many grave failings of this company and the arrogance of some of its employees. One has to recognise, however, that the energy trading landscape would have been very different without the "Big E." Their model of energy

---

trading and risk management was based on an integrated trading floor, a combination of physical and financial transactions and heavy reliance on the analysis of fundamental factors. This model has propagated across the industry through replication (after all, imitation is the subtlest form of flattery) and through Enron DNA being scattered across the industry when former employees migrated to other firms after 2001. Of course, were it not for Enron, other firms would have gravitated towards the same business model over time, but somebody has to be first at the finish line. Such praise for Enron's contribution to energy trading does not mean that one should be blind to other aspects of their legacy. The trading organisations managed by the company's alumni often tend to be too big for their own good. Enron was not a very efficient company, trading in too many markets and present in too many market niches (at the end of its existence, Enron was trading everything in the commodity space, except probably the Ebola virus), and a cost-cutting ability was not a path to quick promotion. Enron veterans also display too much appetite for risk. The outcome, given the zero-sum-game nature of many competing strategies, translates often into a great success accompanied by its mirror image: a great failure somewhere else. In hindsight, the randomness of the markets creates sometimes an illusion of great skill. One should never forget, that "under the sun, that the race is not to the swift, nor the battle to the strong, neither yet bread to the wise, nor yet riches to men of understanding, nor yet favour to men of skill; but time and chance happeneth to them all."<sup>4</sup>

## THE TARGET AUDIENCE

This is a book written by a practitioner for practitioners. The target audience includes energy traders and widely defined support personnel, ranging from quantitative and fundamental analysts, risk managers, accountants, lawyers and compliance officers to back-office staff. All these professionals are highly specialised in their respective fields but need common foundations and understanding of some basic facts and principles of the industry in order to communicate and co-operate in an effective way. They can operate in two different types of organisations. The first includes trading desks in large financial institutions (commercial and investment banks, hedge funds, mutual funds) engaging in speculative transactions and/or

providing risk management instruments and investment vehicles to their clients. The second type includes producers and consumers of energy commodities, who rely on specialised units to market their production, buy fuel and feedstocks, and control price and credit risks. Both types of organisations need employees with similar skills and experience, and hopefully this book will be a useful source of information for them.<sup>5</sup>

The author worked in many energy-trading operations and often witnessed young associates with MBA degrees starting their careers and being petrified during discussions at the morning traders' meetings. Such meetings typically last between 15 and 20 minutes and start with a weather update, followed by reports from different markets. Natural gas discussion is typically split between the Nymex overview and an update on the basis prices (locational price differentials) by different regions of the country. Power market reports follow, organised usually by different regions and power pools. Oil markets are discussed in the global context, with traders from London and/or Singapore joining in over the telephone in the case of large global organisations. Inexperienced employees are often overwhelmed by technical jargon, the abundance of details related to financial and physical operations and the fast-moving pace of the meeting. If the book meets its objectives, they will be able to participate in such meetings from day one.

Quantitative experts supporting trading desks and risk management teams face similar challenges. Their education makes them often very dangerous. They are skilled in building complex systems based on a few axioms that are accepted at the starting point of analysis and, in some cases, never subsequently questioned.<sup>6</sup> They study a few basic introductory textbooks before joining a trading organisation, and accept the simplified, stylised description of the commodity markets as a perfect representation of reality. Some of them have critical minds and learn, some of them will continue to dwell in the fictional world of commodity markets 101 forever. On many occasions, we have heard groups of quants leaving business meetings on energy trading floors and mumbling about understanding roughly 10% of what had been said. This would not be a serious problem were it not for the fact that the commercial people leave the room with the same conclusions about the statements made by the quants. Both groups, however, feel unqualified to challenge

each other. Bridging the gap between both sides is another key objective of this book.

*Energy Markets* also targets employees responsible for hedging energy risk at companies producing or using energy, developing business strategies and making investment decisions. The growing level of volatility in energy prices will make energy risk management an imperative in the coming years, and understanding how the energy markets operate will be of critical importance. Such employees are not traders in the strict sense of this word, but are often responsible for the execution of hedging transactions. Understanding the option valuation models for energy markets and risk management fundamentals is a critical precondition for such jobs. Finally, the book targets the policymakers, lawyers and economic and financial analysts who have to develop a tangential awareness of the energy markets in order to operate in their professions or make public policy choices.

One special audience that may find this book useful is the community of academics specialising in finance and economics, and interested in the energy markets. One of the frequent complaints heard from practitioners is that the models offered in academia are excessively removed from the realities of the energy markets, and do not capture the texture and richness of the underlying systems. Such complaints are, unfortunately, justified, with a few notable exceptions (almost all the exceptions being our good friends). This can be explained by the way the theory of energy markets and energy derivatives has evolved over the last 20 years. The industry was growing through many cycles of expansions and contractions, with academics being attracted to the industry only to be pushed away. Relatively few had the perseverance and wisdom to stay the course.

## SELECTION OF TOPICS

A book of this scope cannot explain in detail all the exceptionally complicated issues related to the organisation of the energy markets. One of the most difficult decisions we had to make was the selection of topics and the amount of space to devote to each of them. We had no choice but to rely on our judgement and the experience accumulated over many years spent in the industry. We are sure that our choices were often suboptimal, and we have spent too much time

elaborating on points of minor importance (which appeared to be the most important thing going on at the time) or have missed a major mountain by concentrating our attention on a molehill. The very rapid pace of change in the energy business will make some statements contained in this book obsolete by the time the ink has dried. This is inevitable, and there is no easy solution to this dilemma. There will be some mistakes because the world evolves faster than modern printing presses. Some mistakes will, of course, reflect our own personal failures.

Another challenge we faced was the highly controversial nature of many issues facing the industry – and the heated, and on some occasion quite vitriolic, debate around them. This is not surprising given both the stakes and the amount of money riding on the answers. In most cases, we can just express our personal view without offering the final word on the subject, given the range and complexity of many topics covered in this book. Our objective is to be an honest broker of information and to alert the readers to the debates regarding certain critical problems, ranging from global warming to the safety of some technologies used in the extraction of hydrocarbons. Failure to do this would amount to professional malpractice and intellectual dishonesty. Whether we like it or not, these issues are not going away and any decision-maker must be aware of them and pay attention to them.

One complaint we received from friends whilst they were reviewing the drafts of this book was that the flow was often curtailed as the narrative began to get interesting. This may have been true; however this seemed inevitable given the scope of the book. Exhaustive treatment of every interesting issue would have been impossible without doubling or tripling the size of the book (and even then we would not have satisfied everyone). Our objective is to signal to the reader an important issue and provide enough information to allow them to investigate the topic further.

We do not intend the discussion of regulatory developments to be construed as legal advice (we have no legal training and it would be a futile exercise anyway, given that the rules may change significantly at very short notice). The numerical examples and numerical information (such as unit conversions) are provided to illustrate the concepts and the orders of magnitude, and not to offer final, exhaustive solutions. The measurement units used in the industry are

usually defined with respect to specific ambient conditions and should be carefully defined in any contract.

Many important developments took place following the deadline for completion of the final draft of this book (for example, finalisation of the swap definition by the CFTC). We shall occasionally comment on developments in this author's blog which we invite the reader to visit from time to time, as well as the other excellent blogs mentioned in this book, in order to stay current.

On a personal note, we would like to express our gratitude and deep appreciation to the many friends and associates who helped us by reading the successive drafts of the book, offering generous comments and linguistic advice: Tony Hamilton, Anders Johnson, Stinson Gibner, Barbara Ostdiek, Richard Dziurzynski, Jozef Lieskovsky, Sheetal Nasta, Jesus Melendrez, Robert Sharp, Jose Marquez and Ehud Ronn. Suffice it to say that any remaining errors, omissions, foolish mistakes and abuse of the English language are entirely our fault.

Special gratitude is reserved for the Risk Books editors Alice Levick and Lewis O'Sullivan without whom the book would never have been completed.

- 1 Just to mention a few: Fletcher J. Sturm, 1997, *Trading Natural Gas: Cash, Futures, Options and Swaps* (Tulsa, OK: PennWell Books); Alexander Eydeland and Krzysztof Wolyniec, 2003, *Energy and Power Risk Management: New Developments in Modeling, Pricing and Hedging* (Hoboken, NJ: Wiley); Héllyette Geman, 2005, *Commodities and Commodity Derivatives: Modelling and Pricing for Agriculturals, Metals and Energy* (Chichester, England: John Wiley); Neil C. Schofield, 2007, *Commodity Derivatives: Markets and Applications* (Chichester, England: John Wiley); Morgan Downey, 2009, *Oil 101* (Wooden Table Press).
- 2 Samuel Taylor Coleridge, 1835, *Specimens of the Table Talk* (<http://www.gutenberg.org/cache/epub/8489/pg8489.txt>).
- 3 The first part of the book contains graphical representations of interdependencies between different parts of the energy complex.
- 4 Ecclesiastes 9:11.
- 5 There is another class of energy trading organisation (some large utilities and oil majors) that do market and hedge their production and also provide risk management services to the rest of the industry.
- 6 To a mathematician, the practical validity of the underlying assumptions often does not matter. One can build a consistent system of geometry making either an assumption that only parallel line can be drawn through a point (Euclides) or a number of parallel lines (Nikolai Lobachevsky). Both theories will be equally satisfying as axiomatic systems, but in practice we have to choose one of them.

## Section 1

# **Energy Trading: The Past and the Future**



# *Energy Trading and Marketing: The Macro View*

In this chapter we shall elaborate further on the central message of this book, signalled in the Introduction: markets for different energy commodities represent an integrated, intertwined system, with connections between each component of that constantly changing system. The differences between physical and chemical properties of different commodities and differences in production and transportation technology do not change the basic fact that energies traded in the markets share many common features – and that these markets can be analysed using identical or similar tools. Interdependencies between different parts of the energy complex have profound implications for price formation and price dynamics. One cannot successfully trade one type of energy without paying attention to the developments related to other parts of the system. This translates, in turn, into the principles underlying the procedures of energy trading organisations, the design of trading floors, required skills and data flows.

We will start by discussing the common features of different energy markets. These commonalities explain why it is possible to uproot a natural gas trader and move them to an electricity trading desk and get them up to speed within a few months. This is not equally easy in the case of fixed income or equity traders migrating to energy trading, as we have witnessed on many occasions. We will continue by outlining major economic and social forces shaping the world energy markets. From the 1990s onwards, we have witnessed a number of revolutions that changed the ways energy trading is organised and how energy traders carry out their business.<sup>1</sup>

- A worldwide transition towards systems of economic organisation based on market solutions, often associated with the

establishment of property rights, the rule of law and democratic institutions.<sup>1</sup>

- ❑ Technological revolution related to computing and communications, with supercomputers, data storage technology and the Internet being the most obvious examples.
- ❑ The development of derivatives and modern techniques of risk management.
- ❑ A revolution in the production of certain hydrocarbons (hydraulic fracturing and horizontal drilling).

Without these new technologies and geopolitical changes, we would have been operating very differently in our trading jobs. We shall discuss certain aspects, specific manifestations and consequences of these developments below.

The final message to be conveyed is that energy markets are important not just as a way of earning a living and having fun at the same time. For those of us who live in economically advanced countries, modern energy markets are critical to maintaining our standards of living. For those of us who live in the emerging economies, energy markets are crucial in making a transition to modernity. This is why we feel strongly about preserving and protecting the integrity of these markets as a fundamental tenet of public policy.

### **ENERGY MARKETS: COMMON FEATURES**

As signalled in the Introduction<sup>2</sup> the energy markets share many common features. What are the similarities between different parts of the energy complex?

#### **Contracts, prices and markets**

In many cases, the needs of energy users are satisfied, and the production volumes delivered, under long-term contracts that are based on formulaic prices (defined in terms of current or lagged spot and forward prices). These prices are established in short-term forward and spot markets which may be dominated by a small group of participants who control the price discovery process (more about this below). The spot markets may be very thin but they determine the value of a very large volume of outstanding long-term transactions based on floating prices. They also often serve as

balancing markets. The end-users may face temporary shortages due to an unexpected spike in demand (for example, due to weather conditions) or because they have lost the source of supply. Sometimes, they have a surplus to dispose of when demand drops but it is necessary, because of inflexibility in supply arrangements, to accept deliveries. The temporary surplus can be liquidated by transacting in the spot markets.

The producers may face the same predicament. The loss of a buyer, interruptible contracts and accumulation of excess inventories require frequent reliance on the short-term markets.<sup>3</sup> The incremental transactions may have a limited impact on the bottom line (because of the small size relative to the overall throughput) and, even if the impact is significant, there is no choice as they are dictated by logistical necessities. The combination of supply chain rigidities and the need to make marginal adjustments to the contracted volumes reduce in the short run the price sensitivities of both producers and end users of energy – and this translates into increased price volatility. This “tail-wags-the-dog” quality of energy markets (ie, amplified impact of a relatively small subset of transactions on the overall market) is important to understanding the properties of demand and supply curves for energy commodities.

### **Price formation process**

Price formation processes in different energy markets share not only many similarities but are also interdependent. A simple stylised graph, which we have found over time to be a very powerful way of organising thinking about energy prices,<sup>4</sup> illustrates how prices are determined in the short and medium run, given constant production capacity (see Figure 1.1). In the figure, a supply curve typically has a horizontal branch, with prices being relatively insensitive to the fluctuations in demand. Changes in demand can be accommodated through adjustments to the inventory levels or adjustments to output levels in production units with similar average and marginal cost curves. As demand increases, the price elasticity of supply<sup>5</sup> decreases: it takes increasingly greater changes in price to induce a given change in output. This can be explained by the need to provide additional supply by expanding production in the firms with higher marginal costs. At some points, the supply curve becomes vertical as

---

the rigidities and limitations of physical transportation and production infrastructure become binding.

A demand curve for a generic energy commodity is typically characterised by low price elasticity. This can be explained by the critical role energy plays in our lives and by the design of many energy markets. In many cases, energy consumers are not receiving price signals in real time (electricity and natural gas bills usually arrive usually once a month, or even less frequently). Depending on the configuration of demand and supply curves, the same increment of demand can produce dramatic changes in prices or can be a non-event. There are several important implications of this simple model for energy trading and for the organisation of an energy trading floor.

### **Supply and demand**

It is important to recognise that the market participants pay close attention to the balance of supply and demand and, when they see a shrinking spare production capacity margin, they react by hoarding and buying forward. This behaviour may be related to speculation and/or the need to procure additional reserves in anticipation of potential supply disruptions.<sup>6</sup> It may result in rapidly increasing prices combined with no concurrent and evident shortages. The “speculators” are singled out as the culprits for price increases that result from the combination of prudent behaviour (building an inventory buffer as a protection against potential supply shocks) and the desire to make profits (a perfectly legal behaviour) from the likely future market developments. The speculators are, in most cases, the bearers of bad news (tight supplies) and are not themselves responsible for creating shortages. Having said this, our culture has strong built-in inhibitions against profiting from the misfortune of others under unusual circumstances.<sup>7</sup> Any energy trader should be aware of public perceptions about their activities.

Supply and demand curves for energy commodities are dependent on the prices of other energy commodities. One example is the change in the shape and level of electricity supply curve, called supply stack. As will be explained in the chapters on electricity markets, a drop in prices of natural gas in the US at the time of writing changed the order of dispatch of power generation units, with important consequences for electricity, coal and natural gas

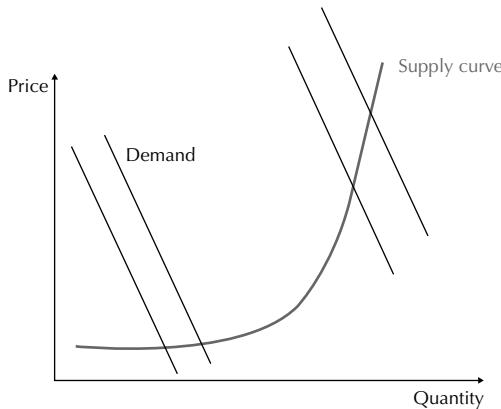
prices. This is an illustration of one of the central themes of this book: no market for any energy commodity is an island. All the energy markets are closely connected and cannot be analysed in isolation. An analyst has to understand the entire energy complex to make price predictions and offer meaningful advice. More examples of interdependencies between different energy markets will follow in subsequent chapters.

The shape of a supply curve depends critically on the conditions of the physical infrastructure. Production plant outages, transportation disruptions and additional commodity flows from other markets can change from minute to minute the slope of a supply curve and its level. This underscores the importance of the access to fundamental and competitive intelligence information (as discussed in Chapter 2).

### **Transactions for spreads**

Many transactions in the energy markets are transactions for spreads, which trade as a single underlying. The spreads may be related to locational and calendar differences in prices and in some cases may represent more active markets than the markets where transactions happen at the flat (also called absolute or outright) prices of a given commodity. The absolute prices may be sometimes constructed indirectly from the prices of actively traded spreads. Transactions as diverse as the natural gas basis contract (see chapters on natural gas) and the Brent contract for difference (see chapters on

**Figure 1.1** Demand and supply curves, energy markets



the oil markets) are examples of the same underlying basic concept, with common valuation and risk management principles applying to each. Once the importance of spread-related transactions is recognised, the design of risk management systems and valuation procedures can be implemented in a more coherent and logical way.

### **Price discovery**

The price discovery process happens often in the forward and futures markets. Historically, the development of vibrant spot markets preceded the development of forward and futures markets. This creates a temptation to assume that information flows from the spot markets to the forward markets, with these perceptions being reinforced through the way the relationships between spot and forward prices are explained in the standard textbooks. This topic, the arbitrage relationships between spot and forward prices, will be discussed in Chapter 4. What must be recognised is that futures (ie, the forward contracts traded on the organised exchanges) evolved in many cases from conduits providing access to physical supplies into platforms performing functions of price discovery and information processing. Spot prices are often influenced by the futures markets, with causality being reversed compared to what is implied by conventional wisdom. A trader buying a physical cargo of oil sometimes does not realise that they become an unwilling participant in the derivative markets through the reverse link between forward and spot prices. This means that the traditional distinction between the physical and derivative traders becomes fuzzy. Both groups of traders watch hawk-like the monitors displaying futures prices (the “screens” in the industry jargon). Understanding the process of price formation in each specific market is a critical skill any trader has to master, and we shall document the details of this process for each market.

### **Market manipulation**

Low price sensitivity of supply and demand in the short-term and frequent instances of small trading volumes make some energy markets vulnerable to manipulation. In practice, manipulation can happen in a number of different ways.

- ❑ Aggressive trading strategies, such as rapid, high-volume trading or the injection of false information into the marketplace.
-

- ❑ Exploitation of imperfections in market design (especially in the electricity markets, with convoluted, highly complicated rules, written in a very difficult technical jargon by committees of experts, and constantly evolving).
- ❑ The ability to use dominating position in certain market pockets created through the rigidities and limitations of the underlying physical infrastructure. For example, a US intrastate pipeline company dominating access to, and storage around, a natural gas market hub has the ability to influence prices at specific locations.

This dark side of the commodity markets is often vigorously denied for reasons ranging from excessive reliance on an idealised notion of market efficiency<sup>8</sup> to a natural tendency to defend one's industry against accusations of price gouging and the extraction of market rents by critically positioned institutions. This book is designed for traders and risk managers who have to survive in the rough markets and cannot afford to be blinded by ideology (although those who are tend to be sometimes exceptionally useful as expert witnesses in defending alleged perpetrators. Nothing is as effective as dogma masquerading as science).<sup>9</sup> We personally feel strongly about manipulation seen as a threat to the integrity and efficiency of markets that are critical to our economic future.

We subscribe to the opinion that a strong case exists for vigorous government action against market manipulation. No efficient market can last for long without strong ethical foundations. This obvious truth was stated many times by Friedrich Hayek, who said "to create conditions in which competition will be as effective as possible, to prevent fraud and deception, to break up monopolies – these tasks provide a wide and unquestioned field for state activity."<sup>10</sup> He also felt strongly about the threat posed by manipulation: "Deception, like coercion, is a form of manipulating the data on which a person counts, in order to make him do what the deceiver wants him to do. Where it is successful, the deceived becomes in the same manner the unwilling tool, serving another man's ends without advancing his own."<sup>11</sup>

These comments should not be construed as a statement that market manipulation is currently a systemic problem. This is certainly possible and is known to have happened, with the

---

California electricity markets being the most obvious example. Manipulation in the energy markets represents a risk of episodic events, which may be bad enough to threaten the career and reputation of an honest trader who is on the wrong side of a transaction. It also threatens the good name and bottom line of the employer of a trader who engages in market manipulation. It is certainly serious enough to warrant attention.

### **Price reporting agencies**

Price discovery in many energy markets is facilitated by price reporting agencies (PRAs), which specialise in collection of transaction data, and the calculation and dissemination of market prices (often called indexes or index prices) representing current market levels. The role of the PRAs in the energy markets is a topic of growing interest to governments and regulatory bodies across the developed world. Starting with the Pittsburgh meeting in 2009, the Group of 20 (G20) countries at successive summit meetings requested the systematic review of the role played by these agencies. In the Final Declaration issued after the 2011 G20 Cannes Summit, the G20 leaders stated:<sup>12</sup>

Recognising the role of Price Reporting Agencies for the proper functioning of oil markets, we ask IOSCO, in collaboration with the IEF, the IEA and OPEC, to prepare recommendations to improve their functioning and oversight to our Finance Ministers by mid-2012.

Also, the International Organization of Securities Commissions (IOSCO) will consider “the governance of PRAs including consideration of the ownership of oil PRAs, their board and executive management structures, how they manage conflicts of interest, their complaints handling procedures, and their systems and controls. Secondly, IOSCO will consider the impact of the current functions of the PRAs on price transparency in the physical and derivative oil markets.”<sup>13</sup>

PRAs, given their importance to the energy markets, will be extensively covered in subsequent chapters. The critical question any trader has to answer is to what extent energy markets can be described as an inverted pyramid, with a relatively few transactions in the short-term forward and spot markets (with only a subset of these transactions reported to the PRAs) dominating price formation

and discovery processes. In other words, the question is whether energy markets can in fact be characterised as a case of “the-tail-wagging-the-dog.” This observation should not be construed as criticism of PRAs, which play the role of an important lubricant in the commodity markets. We all have to operate within the constraints of the existing market framework.

### **Complexity**

As mentioned in the Introduction, energy markets represent an integrated system that facilitates the rapid propagation of shocks. One of the objectives of this book is to explain and document what is meant by the complexity of the energy markets. There are many competing definitions of complexity, but there is a general consensus that the following features describe best this concept:<sup>14</sup>

- ❑ complicated and evolving feedback loops between different components of the system;
- ❑ non-stationarity – the statistical properties of the system evolve over time;
- ❑ multiplicity of interacting agents;
- ❑ adaptation – the market participants learn and adjust their behaviour in order to survive and prosper in a changing world;
- ❑ evolution – the system evolves and remains far removed from a stable equilibrium at any point in time; and
- ❑ openness – the boundary between a given system and its environment is fuzzy at best, and one has to understand many external influences to explain the behaviour of the system (a clever trading manager will actively seek to gather a diverse team representing many backgrounds and cultures).

One of our objectives is to provide many examples of these features and explain their consequences.

Energy markets represent a very complex system that requires many years of tedious data collection and thinking about it in order to develop the understanding of how it works. It is also an example of a “learning by doing” concept, with knowledge passed, like folk wisdom, in unstructured form from one generation of traders and analysts to another. There is no university that can teach how these markets really work. It takes many years of immersion and an ability

to generalise to develop insights into the mechanics of these markets. This book is not a substitute for such years of “toil and trouble” on trading floors, or intended as a magic pill substituting for many long days and evenings at the office. However, it will hopefully help to jump-start the process of initiation for aspiring energy market specialists.

## ENERGY FLOWS AND CONSUMPTION

There are two very useful graphs used extensively in the industry that illustrate dependencies between different parts of the energy complex. Before we discuss these, we have to explain the units in which energy is measured and which are routinely used in statistical publications.

### British Thermal Unit

The British Thermal Unit (Btu) is a universal unit used in some countries for some individual commodities (for example, for natural gas in the US) and for aggregation of energy flows. The term Btu is sometimes used as a generic term for energy, as, for example, in description of taxes on energy use.<sup>15</sup> Many domestic appliances, such as air conditioners and stoves, are rated in Btu units.

One Btu is the amount of energy required to heat one pound of water from 39°F to 40°F at the pressure of one atmosphere. The definition may vary from source to source, as the initial temperature of water may be chosen differently. For example, the Canadians use a Btu definition based on the starting point of 60°F. This explains why every energy contract using this unit should contain a definition of Btu, in order to avoid potential misunderstandings and costly legal disputes. One Btu is equal to 1.05505585257348 kilojoules.<sup>16</sup> In the natural gas industry, one uses a unit called MMBtu (one million Btus<sup>17</sup>), equal by convention to about 1.055 giga joules (GJ).

One match produces an amount of energy equal to about one Btu.<sup>18</sup> The US Energy Information Administration (EIA) website contains a summary document providing a good intuitive illustration of the energy content of different energy sources. More information on detailed conversions will be provided in subsequent chapters.

**Table 1.1** Conversion table of common energy sources to Btu

<b>Energy source</b>	<b>Physical units and Btu equivalents</b>
Gasoline	1 gallon = 124,000 Btu
Diesel fuel	1 gallon = 139,000 Btu
Heating oil	1 gallon = 139,000 Btu
Electricity	1 kilowatt-hour (kWh) = 3,412 Btu (but 7,000 to 10,000 Btu of primary energy to generate the electricity)
Natural gas	1 cubic foot ( $\text{ft}^3$ ) = 1,028 Btu; 1 cubic foot = 0.01 therms

Source: U.S. Energy Information Administration

Total energy used by one average US car per year:

- ❑ Driven 12,300 miles per year at 22.5 miles per gallon equals 547 gallons of gasoline at 124,000 Btu per gallon = 67.8 million Btu

Total electricity used by one average US household in 2008:

- ❑ Total US residential electricity used is 1,379 billion kWh divided by 115 million households at 3,412 Btu per kWh = 40.9 million Btu

Total primary energy used to provide the electricity used by one average US household per year:

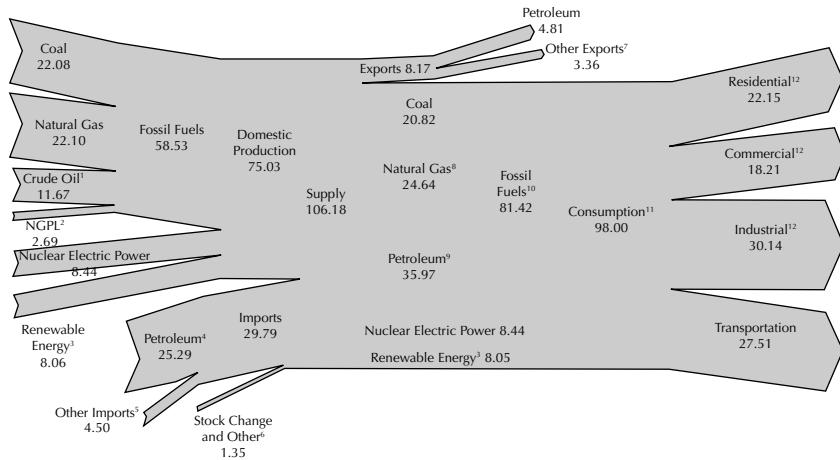
- ❑ Total energy input to electricity production is 40.7 quadrillion Btu times the residential share of electricity use of 37% divided by 115 million households = 131.0 million Btu

## EIA US energy flows

Figure 1.2 emphasises the commonalities between different energy commodities: the ability to satisfy the same basic human needs. Energy entering the US economy can be compared to a vast river with many large and small tributaries contributing to the main flow, and with the flow undergoing many transformations before reaching the final users. This figure allows the tracking of the sources of energy used by the US economy and distribution of energy among competing uses, defined as residential, commercial, industrial and transportation. Along the way, conversion of coal and natural gas to electricity takes place (although this process is not captured effectively in the graph). Energy injections and consumption are measured in common units, which are quadrillion Btu.

Figure 1.2 serves two objectives:

- ❑ to illustrate the magnitude of annual energy flows through the US economy; and
- ❑ to visualise how different parts of the energy complex interact,<sup>19</sup>

**Figure 1.2** US energy flows 2010 (quadrillions Btu)<sup>1</sup> Includes lease condensate.<sup>2</sup> Natural gas plant liquids.<sup>3</sup> Conventional hydroelectric power, biomass, geothermal, solar/photovoltaic, and wind.<sup>4</sup> Crude oil and petroleum products. Includes imports into the Strategic Petroleum Reserve.<sup>5</sup> Natural gas, coal, coal coke, biofuels, and electricity.<sup>6</sup> Adjustments, losses, and unaccounted for.<sup>7</sup> Coal, natural gas, coal coke, electricity, and biofuels.<sup>8</sup> Natural gas only; excludes supplemental gaseous fuels.<sup>9</sup> Petroleum products, including natural gas plant liquids, and crude oil burned as fuel.<sup>10</sup> Includes 0.01 quadrillion Btu of coal coke net exports.<sup>11</sup> Includes 0.09 quadrillion Btu of electricity net imports.<sup>12</sup> Total energy consumption, which is the sum of primary energy consumption, electricity retail sales, and electrical system energy losses. Losses are allocated to the end-use sectors in proportion to each sector's share of total electricity retail sales. See Note, "Electrical Systems Energy Losses," at end of Section 2.

Notes: • Data are preliminary. • Values are derived from source data prior to rounding for publication. • Totals may not equal sum of components due to independent rounding.

Sources: Tables 1.1, 1.2, 1.3, 1.4 and 2.1a

**Source:** U.S. Energy Information Administration (2010)

Notes: NGPL – natural gas plant liquids; Renewable energy = conventional hydroelectric power, biomass, geothermal, solar/photovoltaic and wind; Crude oil flows include lease condensate; Petroleum categories entering the US energy stream include imported crude oil and petroleum products (including imports into the Strategic Petroleum Reserve, SPR).

These objectives are a roadmap to what we want to accomplish with this book, to explain how different parts of the energy complex interact and how one part of the industry impacts the rest of the system. The central message, and one well worth repeating, is that the days of insular energy markets are over, for good or bad.

Energy flows entering the US economy include coal, natural gas, crude oil, natural gas liquids, electricity from nuclear power plants and different renewable energy sources, petroleum and other imports. These will be discussed in detail across different chapters of this book. Other imports include natural gas, coal, coal coke, biofuels and electricity. These flows form total supply, equal to 106.18

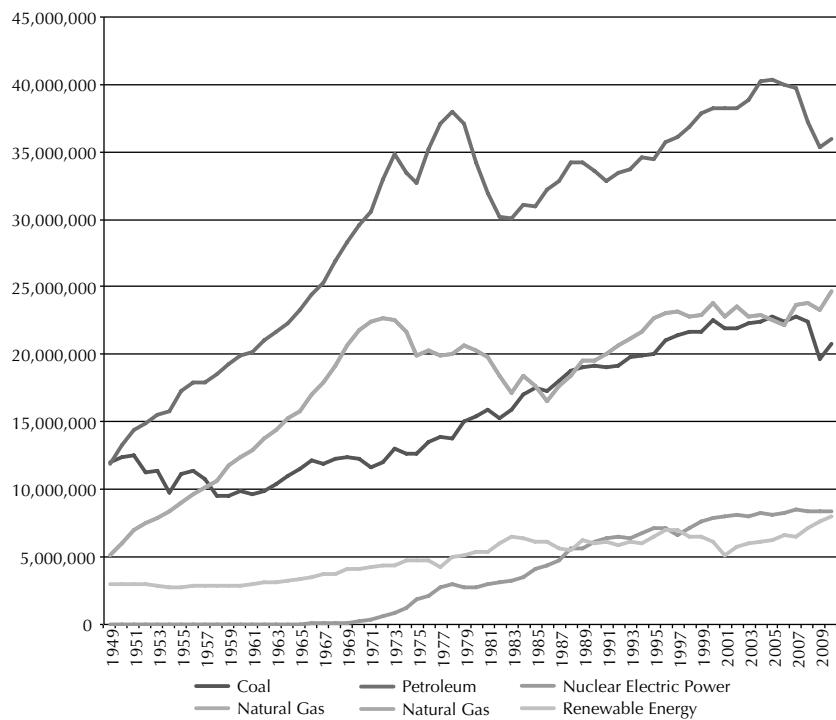
quadrillion Btu in 2010. This category breaks down into domestic production and imports plus stock change and other<sup>20</sup> ( $75.03 + 29.79 + 1.35 = 106.18$ ). The total supply is broken into several categories as different streams are utilised through the US economy and diverted to different uses, such as exports, domestic fossil fuel, nuclear electric power and renewable energy.

Exports include petroleum and other exports (coal, natural gas, coal coke, electricity and biofuels). Fossil fuel flow (81.42 quadrillion Btu) is equal to the sum of coal, petroleum (which at this stage includes mostly refined products) and natural gas (yes, natural gas is a fossil fuel, although some politicians believe otherwise). The breakdown of fossil fuels is given by the sum of these three components ( $20.82 + 24.64 + 35.97 = 81.43$ ).<sup>21</sup> The petroleum category in this midstream flow contains petroleum (refined) products, including natural gas plant liquids, and crude oil burned as fuel. Consumption equal to 98.00 quadrillion Btu is broken into the residential, commercial, industrial and transportation categories. Consumption includes 0.09 quadrillion Btu of electricity net imports. Total energy consumption is a sum of primary energy consumption, electricity retail sales and electrical system energy losses. As explained in the EIA documentation, “losses are allocated to the end-use sectors in proportion to each sector’s share of total electricity retail sales.”

Another useful piece of information from the EIA contains the summary of historical evolution of the use of different primary energy sources in the US (see Figure 1.3). This graph demonstrates vividly the dependence of the US economy on fossil fuels in general, and specifically on petroleum (both produced domestically and imported). Petroleum is critical to the US transportation system and our standard of living, to an extent unparalleled in human history. This book is focused primarily on fossil fuels and nuclear energy. Renewable energy sources are likely to remain marginal for the lifetime of the author and most of the readers (unless this book attracts a large number of primary school students).

Figure 1.4 also presents a graph showing the evolution of the US energy production by primary source.

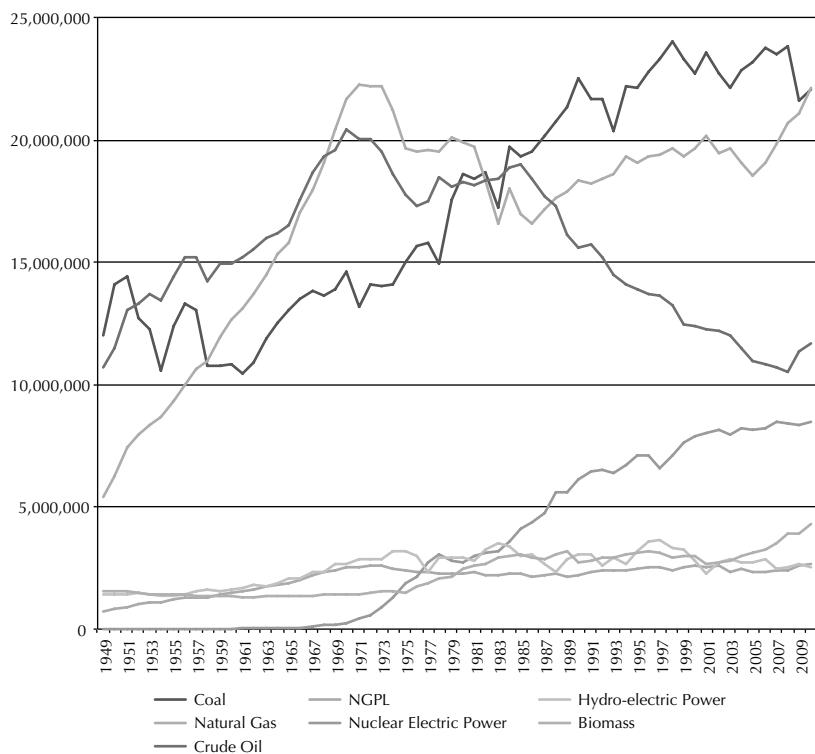
As for the US, the global dependence on fossil fuel is striking, although the world at large is more dependent than the US on renewable energy and, to a smaller degree than the US, on nuclear power. These differences can be attributed to a number of factors.

**Figure 1.3** US primary energy consumption by source (billion Btu)

Source: U.S. Energy Information Administration 2010

Nuclear energy is an option available only to the most technologically advanced nations. In many cases, the pressure of public opinion slows down or completely frustrates the development of nuclear reactors, following the Three Mile Island, Chernobyl and Fukushima disasters. The US is blessed with a vast hydrocarbon resource base, reducing the pressure to invest in alternative energy sources.

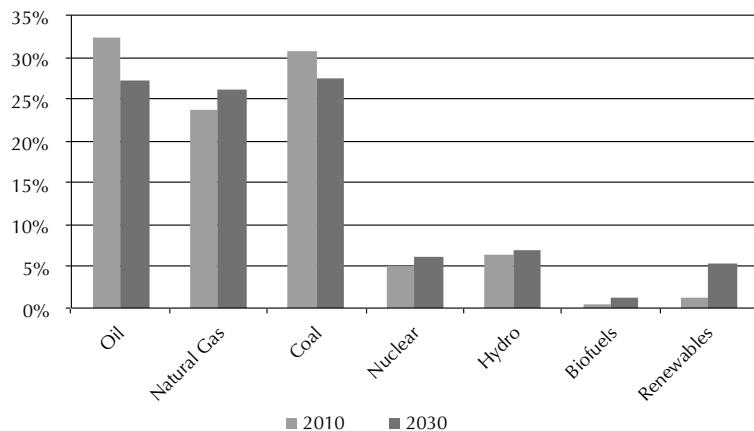
The dependence of the US and world economy on fossil fuels will continue for a long time, the efforts to promote development of alternative sources of energy notwithstanding. Projections made available by BP (January 2012) illustrate this inevitable dependence of the entire world on hydrocarbons (see Figure 1.5). Figure 1.6 shows the world reliance on different energy sources between 2010 and 2030.

**Figure 1.4** US energy production by primary source (billion Btu)

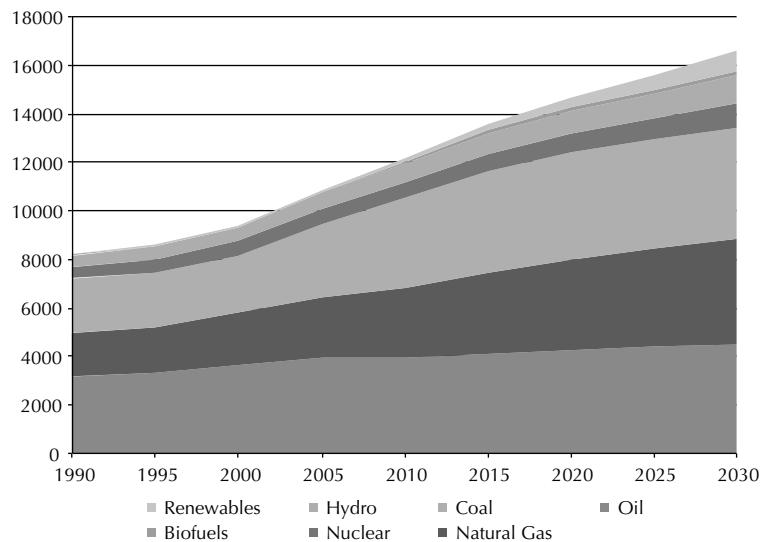
Source: U.S. Energy Information Administration 2010

### Lawrence Livermore National Laboratory: US energy sources and uses

Figure 1.7 illustrates both complementarity and substitutability of different energy commodities in the US economy (although it can be easily generalised to the rest of the world). This graph, produced annually by the Lawrence Livermore National Laboratory,<sup>22</sup> is one of the best visualisations of a national energy system that we have seen. Electric power generation relies on coal, uranium and natural gas as the sources of heat that, through production of steam or very hot gases, produce electricity, converting mechanical energy into electric energy. Another source of electricity is the renewables sector – such as wind, solar, geothermal, ocean waves and tides and hydro. These sources are characterised by a high level of intermittency and it is

**Figure 1.5** Total world energy production by source, 1990–2030 (%)

Source: BP, "Energy Outlook 2030," January 2012

**Figure 1.6** Total world energy production by source, 1990–2030 (million tonnes oil equivalent)

Source: BP, "Energy Outlook 2030," January 2012

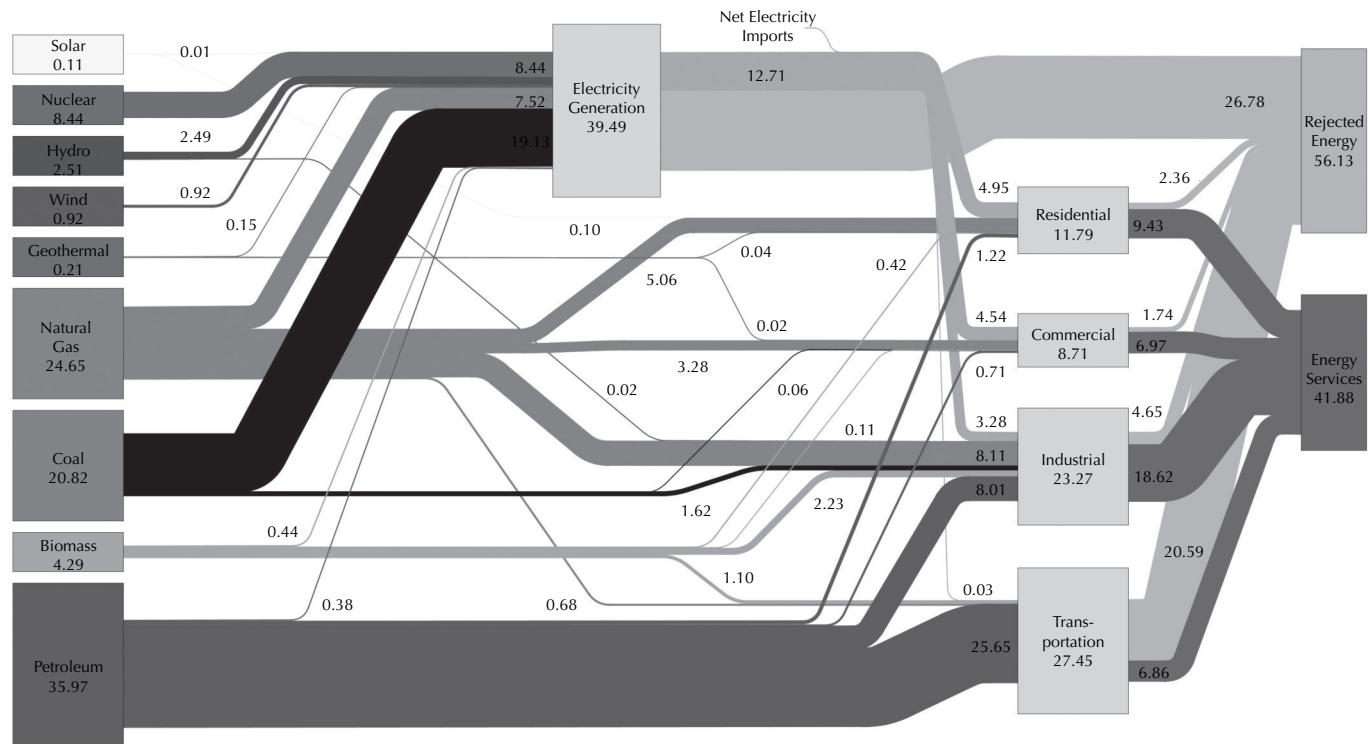
difficult to predict their output in the short run. This creates the need for back-up that comes mostly in the form of other renewable or natural gas plants which can be dispatched at short notice. Recognition of this mutual dependency and competition is critical in modelling the energy systems and in investment decisions. Electricity uses are classified by convention as residential, industrial and commercial. There is an alternative way to look at the disposal of electricity using the classification based on how electricity is used: as a source of heat, light, motion and information transfer. This classification can be further combined with a classification based on the end-user type, providing a useful breakdown for forecasting and rulemaking in the regulatory process.

Natural gas, in addition to electricity production, is used by the residential sector and industry directly as a fuel for space heating. The industrial sector also uses natural gas as a feedstock in chemical plants (primarily in plastics and fertiliser production). The transportation sector currently consumes rather small volumes of natural gas but this is likely to change with the growing penetration of natural gas vehicles. The usage of natural gas liquids (not shown here, but discussed later in chapters on natural gas), is divided primarily between residential and industrial consumption. The complication in the case of natural gas liquids is that some of the supply is a byproduct of crude oil processing, creating interesting price fluctuations.

Petroleum products are used primarily in the transportation sector (as gasoline and diesel) and in the industrial sector as a feedstock to a variety of chemical processes. The residential and commercial sectors use petroleum products (primarily heating oil) as a source of energy for space heating. A small amount of distillates is used in the production of electricity (residual fuel oil). This use of petroleum products has been to a large extent eliminated in most countries following the oil price shocks.

Renewables, in addition to electricity generation, are used in the transportation sector (biodiesel and ethanol), where they compete directly with petroleum. Some consumption of renewable energy occurs in the residential sector (for example, geothermal energy used for home heating or solar energy used to heat water supply or swimming pools).

The category *rejected energy* denotes energy that is wasted and not used (for example, energy losses in the thermal power plants). This

**Figure 1.7** US energy flows 2010 (quadrillions Btu)

Source: The Lawrence Livermore National Laboratory and the US Department of Energy  
[https://flowcharts.llnl.gov/content/energy/energy\\_archive/energy\\_flow\\_2010/LLNLUSEnergy2010.png](https://flowcharts.llnl.gov/content/energy/energy_archive/energy_flow_2010/LLNLUSEnergy2010.png)

category points out a very significant potential for improving energy efficiency in the US.

## THE BRIEF HISTORY OF ENERGY TRADING

The objective of this section is to explain the importance and evolution of energy trading since the 1980s. In principle, this section could be very short. Energy is critical to our standard of living and the efficient production and use of energy is impossible without a well-oiled (no pun intended) market, sending signals about the costs and scarcity of different commodities to end users and producers. The standard of living in the Western world would be impossible to achieve and maintain without ample supplies of cheap and reliable energy. The following calculation helps to illustrate this fact: an average American uses about 25 barrels of oil every year, translating into energy content of 145 MMBtus (using a conversion factor of 5.8 MMBtu/barrel of oil). An average person working for one hour generates 240 Btus. This means that one barrel of oil translates into 25,000 hours of labour (12.5 years at 40 hours/week). In a manner of speaking, an average American has more than 300 "energy slaves" just through consumption of oil, and close to 700+ "slaves" if natural gas and coal is counted.<sup>23</sup> This makes an average citizen of the US a person of fabulous riches by the standard of the Roman Empire.<sup>24</sup> This is a back-of-the-envelope calculation illustrating the order of magnitude of the contribution of energy to our collective well-being. Even if we allow for inefficient use of energy and cut the number of "slaves" by 75%, we are still a very rich society by the standards of ancient Rome.

Another example, borrowed from a book by Joseph A. Tainter and Tadeusz W. Patzek, illustrates the importance of energy for the US agriculture, a critical sector of the economy and one of the foundation stones of our standard of living:

On average, a US resident uses 100 time more energy than he needs to live,  $100 \times 100 = 10,000$  watts continuously, or 0.01 megawatts of power. As an industrial worker, a single person uses and outputs many times more power. For example, it can be calculated that an average agricultural worker in the Midwest uses 0.8 MW of power as fuel, machinery, electricity and field chemicals, and outputs 3 MW of power as crops. [...] Therefore, an agricultural worker in the United States has at her disposal the power of 8,000 ordinary people [...] This external, or exosomatic, use of mostly fossil fuel power has no

parallel in human history, and will be a short-lived phenomenon by evolutionary and historical standards.<sup>25</sup>

Energy trading is as old as the energy business itself. After all, energy commodities have to travel through the supply chain from the producers, through the processors, to the end users, and in most societies this happens through market transactions. A comprehensive study of the history of this business would have to cover at least a few centuries. This section has rather modest objectives: a review of the major historical forces shaping the energy markets since the 1980s and of the major milestones in the growth of this business. The topics covered here include:

- ❑ deregulation of commodity markets and financialisation of major developed economies;
- ❑ emergence of a number of rapidly growing developing countries (primarily India and China);
- ❑ emergence of commodities as an asset class;
- ❑ financial innovation;
- ❑ growing environmental concerns; and
- ❑ technological revolution in production of certain energy commodities (primarily natural gas).

### **Deregulation**

At the end of the decade of the 1970s most energy markets in the developed countries were either regulated (natural gas and electricity) or effectively administered by a group of the most powerful producers (coal and crude oil). Deregulation, often combined with, or following, privatisation created a powerful incentive for the development of energy markets. Deregulation created a risk management vacuum: in a regulated industry a standard procedure for managing exposures is to maintain good relationships with the regulators and the ability to shift costs to the ratepayers. As long as the supplies are reliable, the quality of service is satisfactory and the rates do not grow at an excessive speed, the ability to manage the regulatory oversight process was more than sufficient to be a successful executive in a regulated company.<sup>26</sup> Under deregulation, reliance on markets to procure supplies of energy commodities and to find outlets for production flows required investment in risk management capabilities with growing demand for financial instruments

linked to energy. The financial industry and a number of other entrants (to be covered later) responded to the growing need for risk management tools by expanding their energy trading desks.

Deregulation in the energy industry started in the English speaking countries, spreading to Western Europe and other parts of the world. The critical developments in different markets will be covered briefly in subsequent chapters. At this point, it is important to mention that the experience of deregulation of the energy industry worldwide is generally positive, with the exception of some fiascos – the most obvious example being the Western US electricity markets crisis of 2000–01. The deregulation of the energy industry, which cannot be separated from the deregulation of the financial industry, started in the 1980s and gained steam in the 1990s. Both trends converged in the US in 2000 in the Commodity Futures Modernization Act (CFMA), which removed a significant part of the US energy markets from the effective regulatory oversight. CFMA is seen today as an important underlying factor in the financial crisis of 2007–08. Many provisions of the CFMA were reversed by the Dodd–Frank Act of 2010. Some of the most important provisions of this act will be covered in later chapters.<sup>27</sup>

### **Financialisation**

The second critical trend affecting energy trading was the process of financialisation of the advanced economies, which can be defined as follows:

The increasing dominance of the finance industry in the sum total of economic activity, of financial controllers in the management of corporations, of financial assets among total assets, of marketised securities and particularly equities among financial assets, of the stock market as a market for corporate control in determining corporate strategies, and of fluctuations in the stock market as a determinant of business cycles.<sup>28</sup>

Financialisation is reflected in the growing contribution of the financial firms to the national income creation and their growing share in overall corporate profits.<sup>29</sup> The available statistics are likely to under-represent this trend for the reasons discussed below.

One of the manifestations of the trend towards financialisation in the developed economies was a process that can be described as commoditisation of finance and financialisation of the commodity

business. These processes manifest themselves in growing importance of commodity oriented business in the financial firms and expansion of commodity producers and processors into financial markets. The changes occurring in the energy industry related to this process can be summarised as follows.

- ❑ The growing use of risk management tools among the producers and end-users of energy. This trend reflects growing awareness of high price volatility of energy commodities and increasing supply of industry professionals with good understanding of derivatives, hedge accounting and related disciplines. The users of energy-related derivatives are supported by a cottage industry of consulting firms, freelancing professionals and media firms offering advice, training courses and outsourcing services.
- ❑ Growing importance of financial strategies structured around the use of energy derivatives. The examples of such strategies (covered in more detail in subsequent chapters) include:
  - creative funding strategies such as commodity-linked bonds and loans or volumetric production payments;
  - fuel purchases by utilities and municipalities combining the use of fixed income instruments and energy derivatives in one complex transaction; and
  - asset acquisitions financed with loans and combined with implementation of simultaneous hedging programmes, which reduce credit risk and allow the lenders to reduce interest rates.
- ❑ Proliferation of energy trading units embedded inside energy companies (big utilities, integrated oil companies, independent producers). Such units perform two related functions: on one hand, they function as service providers to other parts of the company, marketing and hedging their production or energy purchases; on the other, they may engage in activities related to provision of risk management tools to energy producers and consumers or in outright speculation.

The growing importance of activities that were historically associated with business carried out on Wall Street and in the City of London was one of the most important developments reshaping energy industry since the 1980s. However, this is a trend that can be

easily overlooked. As mentioned above, some big utilities, energy producers and large integrated oil firms have units that engage in directional energy trading and which offer risk management tools to their clients. These units are important profit centres but their activities are sometimes toned down in the financial statements and in conference calls with stock analysts. Shareholders sometimes get nervous about the risks of trading operations; the regulators worry about the cross-subsidisation of unregulated business by the traditional utilities still operating behind the tariff walls.

The proliferation of energy trading units outside the confines of the narrowly-defined financial industry was fuelled by a number of factors, including the realisation that traditional energy companies could leverage their experience in the physical markets and gain access to information associated with the control and/or regular use of the industry infrastructure. The dash to create “Wall Street firms away from Wall Street” was also driven by the differences in the compensation structure in the financial industry and the traditional energy companies. In many merchant energy companies, the employees started to trickle into new unregulated units, and at some point this trend changed into a stampede. Smith and Louisiana streets in Houston became known as the energy corridor, the home to a new breed of merchant energy firm, transforming the landscape of the US energy industry. The term merchant energy company applies to the producers, processors, transporters, etc, of energy commodities, who, in addition to participating in the physical supply chain of the industry, are actively involved in trading, marketing and financial operations, and treat such activities as a separate profit centre. These activities go beyond supporting their core physical business. Many merchant companies are sophisticated participants in the financial and derivative markets. Similar operations emerged in other locations in the US and in the City of London, trends supported by the deregulation of the UK energy markets in the 1990s.

One of the consequences of the trend towards financialisation of traditional energy companies was the growing importance of units reporting to chief financial officers/treasurers, which in some cases grew to dominate the entities they were supposed to serve (often to their regret).<sup>30</sup> These developments cannot be unambiguously described as positive or negative. Risk-reducing activities produce

---

obvious private and social benefits; strategies leading to accumulation of risks inside one company beyond reasonable levels can produce outcomes very costly both to individuals and organisations. As John Kay observed in one of his columns,<sup>31</sup> there is nothing wrong if a small casino is attached to a public utility. A small public utility attached to a big casino is usually a prescription for a disaster. Unfortunately, some energy companies chose to transform themselves into casinos and, for a period of time in the late 1990s, were very proud of this transformation.<sup>32</sup>

The merchant energy companies were initially very successful and in the 1990s succeeded in displacing many commercial and investment banks, early entrants into the business of energy trading and energy risk management. However, the most successful merchant players happened to be giants with the feet of play dough. Their initial growth was fuelled by the ability to mark-to-market the contracts they entered into in the first years in the new business and the blatant disregard of the levels of risk capital required to support energy trading. *Mark-to-market* accounting rules had been first granted by the Securities and Exchange Commission (SEC) to Enron in 1992, and were immediately embraced by the rest of the industry. Mark-to-market accounting allowed the energy merchants to recognise in profits the future cashflows from a given deal at its inception, following the industry insistence that long-term, sufficiently liquid and deep energy markets were in place, allowing for reliable valuation and hedging of the long-term contracts. Unfortunately, in practice this quickly evolved into “mark-to-model” and “mark-to-myth.” This approach would give a new business an initial boost, inflating quickly reported profits and creating perceptions of unrealistic future potential.

As the new merchant businesses grew, maintaining the initial rates of growth was difficult and this led to a “kill-to-eat” mentality: the tendency to enter into new transactions while sweeping potential problems in the old deals under the rug, in the hope that any future challenge would be somehow addressed through a mixture of daring and good luck. Another troubling aspect of this practice was the tendency to harvest the existing book of business by selling the best contracts (to raise cash) and adjusting the valuation of the remaining contracts through convenient modifications of the initial modelling assumptions. When this was difficult, there was always a

decision of last resort: moving an unprofitable transaction to the accrual book.<sup>33</sup> Another questionable practice was raising cash through loans masquerading as proceeds from commodity transactions (this is covered in the chapters on the natural gas markets).

The rise of the merchant energy companies ended in a series of disasters, starting with the bankruptcy of Enron in December 2001, followed by the demise of many other companies, with most remaining players limping for many years like destitute zombies through the graveyard of a once vibrant market, weakened by the short-lived experiment with energy trading. Enron's bankruptcy was followed by a dramatic drop in trading volumes, a widening of bid-offer spreads and the shortening of maturity of available contracts. The slump lasted for about two years, with trading volumes recovering very quickly, supported by a number of new entrants to the industry and with other market participants expanding their trading desks. The recovery of energy trading is not only testimony to the resilience of this industry but also proof of the importance of the energy markets to the US economy. The new or expanding entrants included hedge funds, commercial and investment banks, integrated oil companies and some big utilities. The energy trading crisis of 2001–02 extended to Europe, as many US-based companies discontinued or curtailed operations. Since then, energy trading in Europe has recovered, supported by the push of the European Union towards deregulation and the integration of the European energy markets.

### **Globalisation**

The third important trend shaping energy trading was globalisation. This process can be defined as:

The acceleration and intensification of interaction and integration among the people, companies, and governments of different nations. This process has effects on human well-being (including health and personal safety), on the environment, on culture (including ideas, religion, and political systems), and on economic development and prosperity of societies across the world.<sup>34</sup>

This trend has profound impact on energy consumption and is another reason why the world needs robust energy markets. Globalisation expands the trade flows between different countries and regions and increases the dependence on transportation and

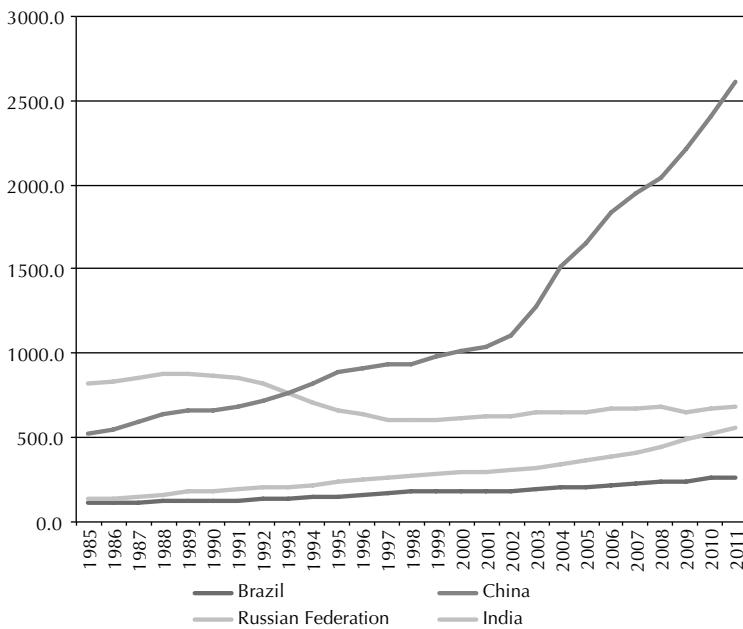
---

communications networks. The shift of production activities to certain locations (China being the most obvious example), with the more mature economies relying increasingly on services, makes the global economy more dependent on moving large volumes of goods and, by extension, on reliable energy supplies. Globalisation, in spite of some negative aspects, has lifted hundreds of millions of people from poverty. This in turn has resulted in the rapid growth of energy consumption in the emerging economies. Growing energy consumption is driven by a number of related factors. General increase in the levels of consumption per capita is associated with growing energy consumption, both in relative (per unit of GDP, per capita, etc) and absolute terms. This happens not only as a result of increased overall levels of consumption and GDP, but also due to structural changes.

The process of globalisation is associated with the shifting of industrial production to the emerging economies, resulting in an economic system that is much more energy intensive than the traditional agricultural economies prior to rapid industrialisation. Participation in the world economy requires development of infrastructure (ports, highways, power plants), which require not only significant energy input during the construction stage, but also increase the demand for energy-intensive commodities (steel, aluminium, copper, etc). Even if a developing country has significant deposits of coal, natural gas and oil, and starts as a net exporter, the domestic supplies are eventually diverted to satisfy the local needs, and this leads in turn to greater reliance on the world markets to support economic growth.

Another important factor underlying rapid increase in demand for energy commodities is the growing level of affluence and the emergence of a middle class. This, in turn, changes the structure of consumption. The percentage of proteins in the daily diet grows, effectively increasing the consumption of energy, given that the production of milk, eggs and meat is energy intensive. Table 1.1 provides the data regarding energy input required to produce a given amount of energy in the form of animal protein. Growing affluence is also associated with the demonstration effect, the copying of the consumption model of the more developed countries, with bigger and better quality dwellings, higher quality diet and, last but not least, growing reliance on personal modes of transportation. Private cars, motorcycles and scooters are not only a necessity given

**Figure 1.8** Primary energy consumption, BRIC countries (million tonnes oil equivalent)



Source: BP, "Statistical Review of World Energy 2011"

the low quality of public transportation networks but are seen as a proof of transition to a better life, a way to achieve personal freedom and symbol of high status.

The most obvious example of ever-increasing demand for commodities is China (see Figure 1.9 for information about historical and projected production and consumption of energy in China). According to statistics compiled by Michael Pettis, China – representing only 9.4% of the global economy (in terms of GDP) and with a population equal to 19% of the world population – represents the following percentages of global demand for raw materials and food products:<sup>35</sup>

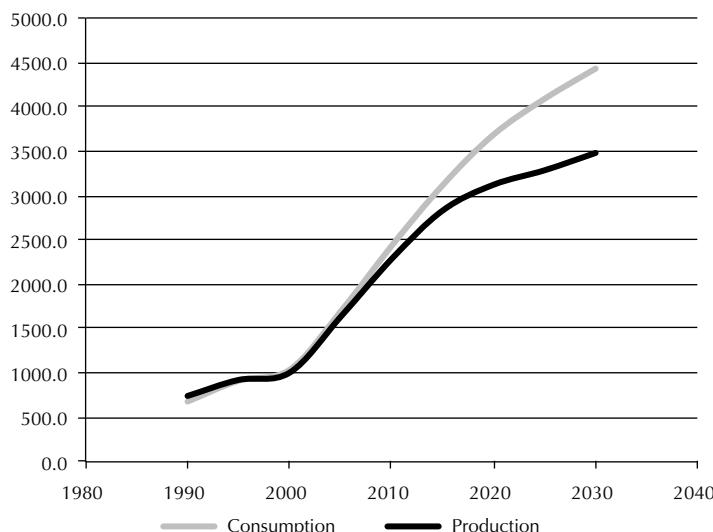
- cement = 53.2%;
- iron ore = 47.7%;
- coal = 46.9%;
- pigs = 46.4%;

- steel = 45.4%;
- lead = 44.6%;
- zinc = 41.3%;
- aluminium = 40.6%;
- copper = 38.9%;
- eggs = 37.2%; and
- nickel = 36.3%.

It is not an accident that many of our friends who trade oil or liquefied natural gas (LNG) start their day by looking at the news from China and poring through the most recent Chinese statistics. Soon they will have to monitor to the same extent developments in India, Russia and other countries with large populations and the will to bring their people out of poverty.

The demand pressure from emerging economies and particularly from China will continue for the next few decades and will constitute one of the defining factors of the energy markets for years to come. Nothing illustrates this better than a forecast contained in the BP "Energy Outlook 2030: January 2012." (see Figure 1.9)

**Figure 1.9** Projections of Chinese consumption and production of energy (million tonnes oil equivalent)



Source: BP, "Energy Outlook 2030," January 2012

The gap between production and demand will continue to widen. Even if the initial boost to energy consumption arising from development of modern infrastructure and the industrial base tapers off, the changes in consumption levels and structure will continue to increase directly and indirectly the demand for energy. A shift to a more protein-rich diet, private car ownership, larger and better-equipped homes and a lifestyle that emphasises free personal choices cannot be stopped. Once on a cold December day I was trying to cross the Ponte Vecchio in Florence, walking slowly, by necessity and not by choice, in a large crowd of tourists. This place, and many others, will be even more crowded in a few years from now, when tens of millions of Chinese tourists will join those from other countries. A better and more open world will result, but the pressure on the finite planetary resources will be enormous.

The very high rates of growth of some emerging economies and the resulting rapid increase in energy demand test the ability of producers of energy to adjust levels of output. The investments in the energy industry have long gestation periods, are capital intensive and require extensive planning and environmental studies. High capital costs are addressed through joint ventures and partnerships that require difficult, time-consuming negotiations.<sup>36</sup> This results in markets operating at the limits of their capacity, with increasing price levels and price spikes in reaction to even the most remote potential reduction in supply. The challenge of meeting the rapidly growing direct and indirect demand for energy is compounded by

**Table 1.2** Annual production in the US and the fossil energy required to produce one kcal of animal protein

Livestock and animal products	Production volume ( $\times 10^6$ )	Ratio of energy input to protein output (kcal)
Lamb	7	57:1
Beef cattle	74	40:1
Eggs	77,000	39:1
Swine	60	14:1
Dairy (milk)	13	14:1
Turkeys	273	10:1
Broilers	8,000	4:1

Source: David Pimentel and Marcia Pimentel, 2003, "Sustainability of meat-based and plant-based diets and the environment," *American Journal of Clinical Nutrition*, 78(3), 660S–663S, September.

market inefficiencies. Many emerging countries rely, for historical and social reasons, on systems of subsidies targeting certain groups of customers. The subsidies send misleading price signals to consumers and industrial users of energy, contributing to the perpetuation of economic structures characterised by high energy intensity. Removing the subsidies, as the experience of Bolivia in 2010 demonstrates,<sup>37</sup> is often not a politically viable solution.

Most governments of emerging economies realise that the rationalisation of energy markets is a critical precondition of continuation of high rates of economic growth. Reduction of energy intensity is possible in the long run only through reliance on an efficient market, with the price system sending correct signals to the producers and consumers of energy. The developing countries often have no institutions that could support the developments of such markets and no experienced traders and risk managers who could support the organic growth of the new markets. This creates an opportunity for many financial institutions willing to make long-term investments and take the risk of political and economic instability to invest in the development of new energy markets.

The expansion of energy markets will not happen without impediments and setbacks. One of the troubling trends is growing resource nationalism, which will not only slow down the progress towards market-oriented solutions, but will also stress available supplies. Resource nationalism can be defined as a doctrine that natural resources should be used for the benefit of the citizens of the country that owns the deposits.<sup>38</sup> At this level of abstraction, it is difficult to argue against this statement. However, the policies inspired by this doctrine often produce negative consequences for countries aggressively embracing this position. First, most policies are oriented towards reduction of income distribution inequality and, judged from this point of view, they are unquestionably very successful (in most cases). The downside is that in most countries they are pursued at the expense of the development of basic infrastructure and human capital. In the long run, they are likely to be unsustainable. Second, the tendency to rely on domestic resources and state-owned companies to develop available mineral deposits leads to inefficiencies, waste and production levels that fall short of the true potential of a country. In case companies from other countries are invited, the specific deals are often seen as extensions of foreign policy, with

commodity flows based on barter or strictly bilateral arrangements, often with side deals (related sometimes to weapon procurement) which slow down the growth of the global markets for energy commodities.<sup>39</sup> It can be expected that resource nationalism, to the extent it will curtail future supplies of energy commodities, will contribute to upward price pressures.

Resource nationalism should not be seen exclusively as an emerging economic policy trend in the developing countries (as some energy companies hit by higher taxes in countries seen as friendly to private business could attest). Many advanced and industrialising countries follow actively policies designed to secure sources of supply of raw materials and energy through a web of long-term contracts, political alliances, direct purchases of deposits and land in other countries. Efforts of private companies are coordinated and supported by government agencies<sup>40</sup> and nurtured through proactive foreign policy (including foreign aid) programmes. Deployment of military resources to protect the safety of sea-lanes and counteract threats to friendly countries is part of the same strategy.<sup>41</sup>

### **Commodities as an Asset Class**

The emergence of commodities as an asset class during the 2000s was one of critical developments in the financial markets, contributing to both the expansion of energy trading volumes and price volatility. Commodities are being included on an increasing scale, along with more traditional assets such as equities, bonds and real estate, in the portfolios of institutional and retail investors. Exposure to commodities may be acquired through a number of different instruments, which will be covered in detail in the next chapters. At this point, however, let us analyse the factors behind this trend.

There are several major developments one can point to which could explain the growing interest in commodities, ranging from disappointments with the performance of other assets to a recognition of major fundamental changes in the world economy that create unique investment opportunities.<sup>42</sup> These are described below.

#### *The Quest for High Returns*

The beginning of the first decade of the 21st century was marked by an implosion of technology and Internet-related stocks, which provided a dramatic illustration of the dangers of investment

---

strategies characterised by an excessive concentration of positions in one specific asset class. The Federal Reserve responded to the recessions following the tech stock bubble and the financial crisis of 2007–08 with continuation of its policy of low interest rates, seen as a solution not only to anaemic investment levels and persistent high rates of unemployment but also as a solution to the undercapitalisation of major banks. Low short-term interest rates, combined with a steep yield curve, allow the banks to capture the spread between the rates paid on retail deposits and rates earned through investment in long-term bonds. One can argue about the wisdom of this policy, which amounts to redistribution of wealth from the savers (primarily the more affluent households) to financial institutions, and reduces the incentives for banks to lend more aggressively.

One cannot, however, deny that the policy of low interest rates creates incentives and opportunities to engage in more aggressive investment strategies. Very low (and possibly negative) real interest rates reduce the opportunity cost of investing in assets producing low or zero cashflows, such as gold and silver. In many cases, as will be explained in detail in subsequent chapters, one can take advantage of the shape of the forward price curves for different commodities, buying physical supply in the spot markets, using storage and selling the commodities kept in storage in the forward markets. The perceptions of pending shortages of many commodities, arising from the recognition of major shifts in the world economy (as described above), contributed to widely held views of secular upward price trend in commodities and prompted pre-emptive buying, producing what some critics believe to be a self-fulfilling prophecy.

### *Hedging Opportunities*

Investment managers have always worried about neutralising exposures to certain risks for which they could not identify effective hedging instruments. A typical portfolio manager in a pension fund is concerned with hedging inflation risk and with increasing nominal value of long-term liabilities, which often depend on the pre-retirement compensation of the plan participants – which will become known after many years into the future. Traditional asset classes such as bonds and equities may not provide effective immunisation against generalised price increases. Another factor increasing the interest in commodities was persistence of imbalances

in the global economy: increasing twin fiscal and external deficits in the US and certain West European countries, associated with a corresponding accumulation of financial claims on the debtor countries by emerging economies in Asia. Many investors expect inevitable adjustments, which may result in realignment of exchange rates and spikes in inflation. Commodities are seen as an ultimate hedge against such risks that the investors can neither fully assess nor effectively neutralise with traditional financial instruments.

The most powerful argument for treating commodities as an asset class, however, was provided by a number of studies demonstrating their very low or negative correlations with other investment opportunities. This regularity was pointed out over time by many investment advisors and analysts, but gained wide acceptance with the publication of two very influential and exhaustive studies.<sup>43</sup>

The findings of the paper by Gary Gorton and K. Geert Rouwenhorst (2005), can be summarised as follows.

- The average annualised returns on a collateralised investment in commodity futures<sup>44</sup> for a 45-year period was comparable to returns on the S&P500 index and much superior to returns on bonds. During the period 1959/7–2004/12, the average annual returns on the commodity futures portfolio were equal to 5.23% versus 5.65% for stocks and 2.22% for bonds. The standard deviations of these returns were equal to 12.10%, 14.85% and 8.45%, respectively. Higher volatility for stocks, somewhat surprisingly to the author, results most likely from the inclusion of the October 1987 crash and the tech stock bubble.
- The most important finding was about correlations between commodities and other asset classes. One-year return correlations between the commodity portfolio and stocks, bonds and inflation<sup>45</sup> were, respectively, equal to -0.06, -0.27 and 0.29. Five-year correlations were equal to -0.42, -0.25 and 0.45. The findings about negative correlations provided a powerful argument in favour of commodities, supporting the diversification thesis. The authors also established that “the negative sensitivities of stocks and bonds to inflation stem mainly from sensitivities to unexpected inflation. The correlations with unexpected inflation exceed the raw inflation correlations.”<sup>46</sup> The same is true of commodities, except that the sign of the correla-

tion coefficient is positive. The diversification hypothesis was further supported by the findings about the downside risk, which can be summarised as follows: "During the 5% of the months of worst performance of equity markets, when stocks fell on average by 8.98% per month, commodity futures experienced a positive return of 1.03%, which is slightly above the full sample average return of 0.89% per month. [...] During the 1% of months of lowest performance of equity markets, when equities fell on average by 13.87% per month, commodity futures returned an average of 2.36%."<sup>47</sup>

The study by Ibbotson Associates contains similar findings:

In addition to impressive historical returns, commodities had the lowest average correlation to the other asset classes; yet, the positive correlation to inflation supports the idea that commodities result in real inflation adjusted returns. For three typical types of portfolios (an all US equity portfolio, a 50/50 US stock/US bond mix, and all US bond portfolio), commodities provided positive returns when they were needed most. We also demonstrated that the correlation of commodities with other asset classes evolves over time. Including commodities in the opportunity set resulted in superior historical efficient frontier, which included large allocation to commodities. Over the common standard deviation range, the average improvement in historical returns at each of the risk level was approximately 133 basis points.<sup>48</sup>

A critical question any portfolio manager has to answer at this point is whether these historical regularities reflected historical neglect of commodities as an asset class and whether growing interest from portfolio managers and hedge funds is likely to result in convergence of return profiles and correlations for commodities and other assets over time. Should this happen, the diversification investment benefits of commodities would have been reduced and possibly completely eliminated. The 2007–09 financial crisis resulted in a dramatic increase of correlations between certain commodities and equities, leading some researchers to conclude that the arguments about diversification benefits of commodities were based on a flawed theory. According to Parantap Basu and William T. Gavin (2011), there are no immutable statistical relationships between returns of different asset classes:

[T]he correlation between returns to equity and commodity futures can change sign over time. In a general equilibrium model in which

there are no unexploited hedging opportunities, it is straightforward to show that the equilibrium correlation can be either negative or positive, depending on the nature of shocks to the world economy.<sup>49</sup>

At this point, we cannot have a conclusive answer to this question. The increase in correlations observed in the last few years could simply be a manifestation of developments typical for any financial crisis: all assets march to the same tune. This may be, however, a symptom of a more fundamental shift reflecting the growing influence of professional portfolio managers trading across multiple asset classes. The fluctuations in liquidity available to them and the waves of pessimism and optimism affecting them translate into synchronised returns. It is important not only what is traded but also to who trades and why.<sup>50</sup>

## FINANCIAL INNOVATION

During the 1980s, an investor could acquire a significant upside exposure to commodities in one of two ways: investing in a portfolio of energy-related stocks or buying physical commodities. Both investment strategies had certain undesirable aspects. It may be difficult to identify companies engaged exclusively in the production of a single commodity or a basket of targeted commodities. A typical stock represents multiple risks, such as potential for environmental impact, labour disputes, nationalisation or expropriation, inability to replace reserves or to control production costs. These risks can, in principle, be addressed by holding a diversified portfolio of many different stocks – but this solution creates a new set of problems. An investor interested in maintaining exposure to commodities may attempt to neutralise exposure to other factors, such as the overall stock market level, interest rate or exchange rate fluctuations. Such strategies require implementation of sophisticated hedging strategies and active trading, potentially associated with significant transaction costs. Investment in physical commodities requires an investment in technical know-how related to handling complicated logistical issues, management of storage and transportations assets, investment in teams of qualified professionals with necessary skills and compliance with multiple rules and regulations. Taking possession of physical commodities is not without risk and creates exposure to potential environmental damage, spoilage and theft, not to mention the potential for endless

disputes related to the quality, quantity and location of the specific commodity volumes.

Given all these problems, commodities would never reach the current status of an attractive asset class without the creativity and drive of the financial industry. The financial institutions, including exchanges and commercial and investment banks, responded to increased interest of the fund managers with the development of multiple instruments, allowing the investors to acquire exposure to commodities at reduced transaction costs without incurring undesirable risks. The detailed analysis of certain instruments, and their advantages, is provided in the subsequent chapters. The most important financial instruments, which offer exposure to energy markets and are actively used by retail and institutional investors, include the following.

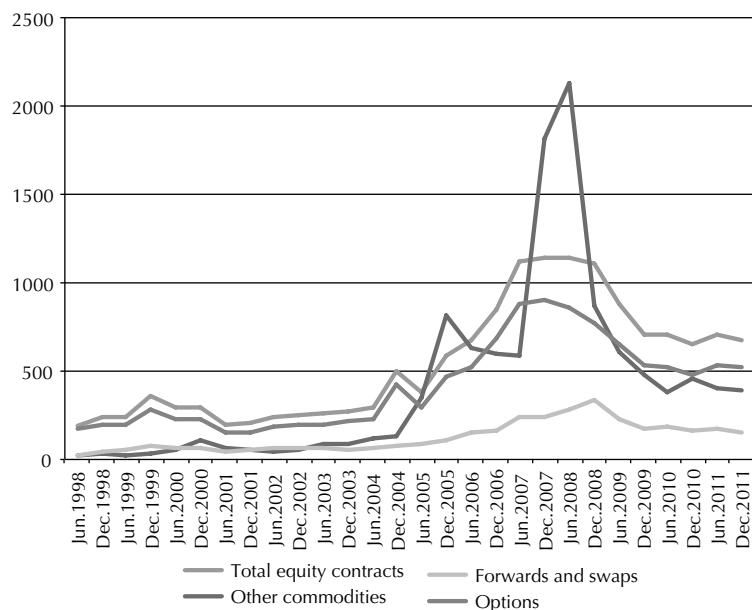
- ❑ *Commodity futures* have been known for centuries and agricultural futures have been actively traded in the US since the 19th century. The last 30 years have witnessed the introduction of a number of exceptionally successful futures contracts on a number of exchanges, including:
  - heating oil contract (New York Mercantile Exchange, Nymex, 1978);
  - WTI contract (Nymex, 1983);
  - natural gas (Nymex, 1990); and
  - Brent contract (International Petroleum Exchange, 1988).The futures contracts with associated options have not only expanded the spectrum of hedging alternatives but also allowed the speculators to acquire leveraged exposure to certain commodities.
- ❑ *Commodity-linked bonds and notes*. Many institutional investors are barred under their charters from making direct investment in physical commodities. The financial industry addressed this limitation by redesigning conventional instruments such as bonds, linking their returns to commodity prices. Commodity-linked bonds became popular with pension funds and mutual funds.
- ❑ *Commodity index funds*. A growing number of commodity-related futures and improving liquidity facilitated development of commodity indexes, such as the Goldman Sachs Commodity

Index and the AIG–Dow Jones Commodity Index (to be covered in detail in the subsequent chapters). The acceptance of commodity indexes facilitated the creation of commodity-related funds and derivative instruments using different indexes as benchmarks or as settlement prices. Index funds are designed for passive investors who have a long investment horizon and desire to match the performance of a benchmark.

- Exchange traded funds (ETFs) and electronically exchange traded notes (ETNs) – these are covered in Chapter 5.

Proliferation of multiple platforms, investment vehicles and instruments facilitated rapid growth in energy related derivatives. Available statistics often commingle energy-related positions with other commodities. Given the importance of energy in the modern economy and the bias of most commodity indexes towards energy,

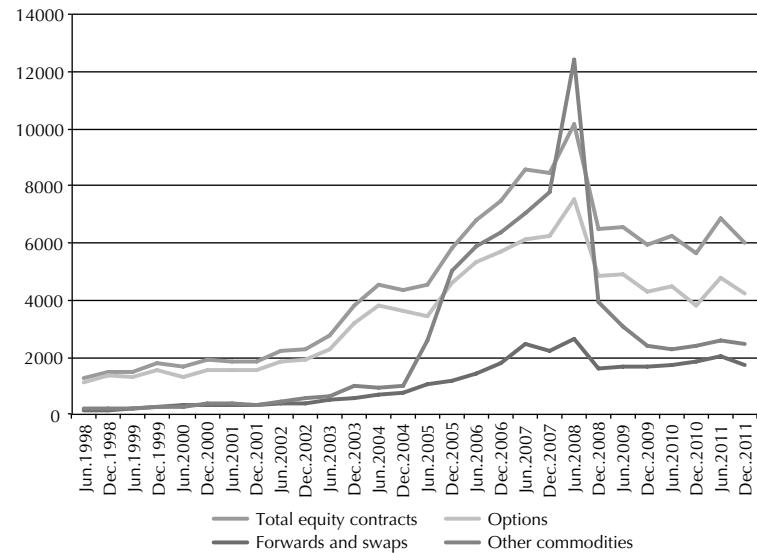
**Figure 1.10** Gross market value, equity and commodity derivatives (US\$ bn)



Source: BIS Quarterly Review: June 2012; Table 22A: Amounts outstanding of OTC equity-linked and commodity derivatives.

Note: The category "Other commodities" excludes precious metals. Forwards, swaps and options add up to equity derivatives.

**Figure 1.11** Notional amount outstanding, equity and commodity derivatives (US\$ bn)



Source: BIS Quarterly Review: June 2012; Table 22A: Amounts outstanding of OTC equity-linked and commodity derivatives

there are good reasons to believe that the expansion in this area is led by energy-related activity. The Bank for International Settlements (BIS) publishes information on the size of derivatives markets. Figure 1.10 shows gross market value for outstanding equity-linked and commodity derivatives, while Figure 1.11 shows the same information for notional amounts outstanding.<sup>51</sup> The financial crisis produced a drop in the gross market value of outstanding commodity contracts after June 2008, which at some point exceeded the corresponding number for equities.

The US Commodity Futures Trading Commission (CFTC) started collecting information about index-related futures positions for 25 (initially 24) commodities at the end of 2007. Figure 1.12 shows aggregate notional index-related positions for four energy commodities: WTI, heating oil, natural gas and RBOB gasoline, with long, short and net positions plotted separately. Figure 1.13 shows the corresponding numbers disaggregated down to the level of the four individual commodities mentioned above.

The proliferation and rapid growth of index funds is a source of

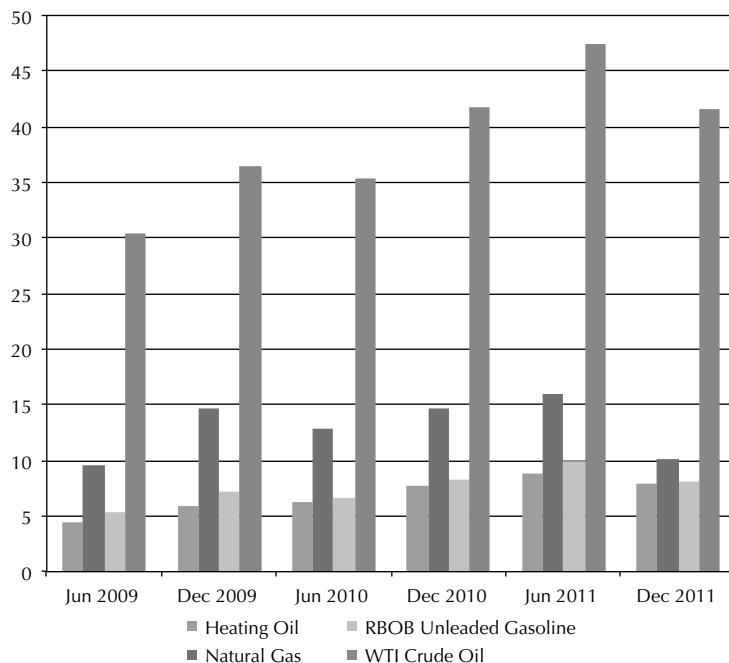
continuing controversy. A spike in many commodity prices, including WTI – which peaked at US\$147 per barrel in 2008 – and the 2012 run-up in agricultural commodity prices led many observers of the commodity markets to conclude that the financialisation of commodity markets was the factor behind recent price trends. This position can be best summarised using the words of one of the most vocal critics of index-related investment strategies:

Institutional investors, with nearly US\$30 trillion in assets under management, have decided en masse to embrace commodities futures as an investable asset class. While individually these investors are trying to do the right thing for their portfolios (and stakeholders), they are unaware that collectively they are having a massive impact on the futures markets. [...] In the last 4 1/2 years assets allocated to commodity index replication trading strategies have grown from US\$13 billion in 2003 to US\$317 billion in July 2008. At the same time, the prices for the 25 commodities that make up these indices have risen by an average of over 200%.<sup>52</sup>

**Figure 1.12** Index investment data: natural gas, WTI, RBOB gasoline, heating oil (US\$ bn)



Source: CFTC

**Figure 1.13** Index investment data (US\$ bn)

Source: CFTC

This hypothesis is a subject of heated debate among the commodity markets professionals, journalists, academics and politicians, details of which we will examine in Chapter 5 and in our next book published by Risk Books. Some of the studies of this topic rely on certain statistical techniques (such as Granger causality), which will also be described in the next book.

### **Environmental issues**

Growing concerns regarding the environmental impact of energy production and consumption activities have had a profound impact on the direction of evolution of the energy markets. There are two aspects to this problem:

- ❑ the production and consumption of energy commodities is associated with externalities, ranging from different forms of air and water pollution to degradation of the environment through irre-

- versible destruction of landscapes, invasive production activities and accumulation of huge volumes of industrial waste; and
- ❑ the burning of fossil fuels leads to accumulation of CO<sub>2</sub> in the atmosphere and potentially permanent climate change (see Chapter 25).

The emergence of a well-organised and very effective environmental movement resulted in the adoption of many policy measures designed to address the problems of pollution and climate change. This trend has manifested itself through regulatory measures ranging from constraints on the use of certain technologies and the scope of certain activities, to the creation of new markets designed to internalise social costs ignored in the past by private agents in their individual cost assessment. Development of in-depth understanding of environmental challenges facing the energy industry is a critical task facing any trader or analyst: future levels of supply and demand for different energy commodities and their relative prices, the patterns of international trade flows and the choice of technologies and location of fixed assets, will depend on the policy choices related to protection of the environment. Understanding the combined impact of different policy measures over time is not an easy task. Effective environmental policies require coordination across national boundaries and across different industries. Any solution will produce losers and winners and this will, in turn, lead to very acrimonious confrontations, a heavy use of propaganda machines and an injection of false information into public debate. In later chapters we will identify and discuss major concerns related to the production and consumption of different types of energy commodities. These different threads will converge in Chapter 25, which covers the most important policy tools developed to address the environmental degradation of our planet, including emission markets and carbon taxes.

No environmental issue produces a greater controversy than the theory of global warming,<sup>53</sup> which postulates the anthropogenic impact of human activity on climate on a planetary scale, primarily due to emission of greenhouse gases. From the point of view of this book, it does not matter if this theory is correct or if it is a figment of imagination of a small but vocal group (and we are not competent to opine on this topic). What matters is that a critical mass of sufficiently

influential public opinion movers has been mobilised behind this theory, and policies designed to reduce emissions of CO<sub>2</sub> and other greenhouse gases will be applied on a sufficiently large scale to produce a profound impact on energy markets.

There are two principal channels of transmission of the global warming phenomenon on the energy markets. The most important is through the mix of administrative and market-based measures designed to reduce CO<sub>2</sub> emissions and other forms of pollution, ranging from national and local government policies to voluntary actions undertaken by companies and individuals due to public pressures or guided by moral principles and the desire to be a good citizen. The second channel, which may be locally very important, is the change in the weather patterns affecting the production and consumption of energy commodities, precipitation and hurricane intensity. Exceptionally cold weather in some European countries (for example, in February 2012), a predictable consequence of global warming, had a huge impact on the natural gas markets as the volumetric flows from Russia diminished and consumption spiked.<sup>54</sup>

The global warming controversy affects not only government policies at a global level but may also have an unexpected impact on relatively local decisions, with significant market consequences. The debate in the US regarding the Keystone XL Pipeline (see chapters on the oil markets) was largely fuelled by concerns about its contribution to the production levels of oil from the Canadian tar sands, seen as a very environmentally harmful source of hydrocarbons.<sup>55</sup> The discussion regarding global issues had suddenly very specific and localised consequences.

## TECHNICAL PROGRESS

Last but not least we cover developments in the production of natural gas and oil, a technological breakthrough related to hydraulic fracturing and horizontal drilling in shale rock formations, which reached critical mass after 2007 and is still in progress. The new technological processes will change relative prices of hydrocarbons and trade flows of commodities, and will have profound geopolitical consequences, enhancing the energy security of the US and many other nations. Every day we detect in the markets tremors caused by the shale revolution, creating great trading opportunities

and potential for big losses for those on the wrong side of this historical change. One example of the speed at which the market can react to these fundamental developments was the drop in the US price of natural gas in 2008 when the scale of the shale revolution became obvious. The decline in prices was exacerbated by delays in reporting production data by the government agencies. The price of the prompt futures contract dropped from US\$13.91/MMBtu on July 1, 2008, to US\$6.59/MMBtu on December 1, 2008. This underscores the importance of following the technological developments and having access to accurate and timely data. We will cover this topic in more detail in the next chapter.

We also further explore these topics in more detail in the chapters on non-conventional oil and natural gas. These developments are very controversial for a number of reasons (again covered later in the book). As always in the case of fundamental breakthroughs, there will be some long-term consequences we cannot even guess at. However, we can venture some predictions. Experience has taught us never to underestimate the creative energies of scientists and engineers, or to bet against them. We think that not only productivity in drilling and well completion will be further enhanced, but also safety concerns will be addressed through improved processes and monitoring devices. Our experience has also taught us never to underappreciate the human ability to overhype any significant invention. It is difficult to avoid comparisons with the dot com boom of the late 1990s. This was undoubtedly one of the most important developments in human history, a wave of innovations that transformed the way we live and work, created great wealth and will continue to contribute to economic growth.

## CONCLUSIONS

At this point it makes sense to step back and weave together some of the different threads of this chapter. The discussion of the varied forces shaping the energy industry leads us to the following conclusions.

- ❑ We are rapidly entering an era of economic scarcity (although not yet the physical scarcity of different energy commodities). Growing demand for energy, seen as a passport to a better life, will require moving to more expensive sources of supply charac-

terised by higher risks, bigger potential impact on environment and using more intrusive technology.

- The competition for sources of energy may lead to an escalation of tensions as different nations scramble to secure long-term supplies of different commodities.

These trends underlie the importance of efficient, unobstructed and vibrant energy markets, sending correct price signals to the consumers and producers. There is no other feasible way to ensure the most efficient use of the energy sources we still have available. Alternative solutions sometimes have a magnetic attraction for politicians, but in the long run they always lead to waste and scarcities. The development and support of efficient energy markets should be a top priority for each nation. Ensuring the integrity of the energy markets should be seen as a primary objective for the regulators.

We shall end this chapter with words from more than 50 years ago, which convey the message of this chapter in a much more effective way than we ever could (and this justifies the length of the citation):

I think no further elaboration is needed to demonstrate the significance of energy resources for our own future. Our civilisation rests upon a technological base which requires enormous quantities of fossil fuels. What assurance do we then have that our energy needs will continue to be supplied by fossil fuels: The answer is – in the long run – none. The earth is finite. Fossil fuels are not renewable. [...] Fuel that has been burned is gone forever. Fuel is even more evanescent than metals. [...] In the face of the basic fact that fossil fuel reserves are finite, the exact length of time these reserves will last is important in only one respect: the longer they last, the more time do we have, to invent ways of living off renewable or substitute energy sources and to adjust our economy to the vast changes which we can expect from such a shift. Fossil fuels resemble capital in the bank. A prudent and responsible parent will use his capital sparingly in order to pass on to his children as much as possible of his inheritance. A selfish and irresponsible parent will squander it in riotous living and care not one whit how his offspring will fare.<sup>56</sup>

The next chapter covers some basic principles of the organisation of a typical energy trading operation. Given the complexity and fast evolution of the markets, it is not an easy task.

- 1 To many of us the pace of progress may seem slow and uneven, but the progress is undisputable. Lion Feuchtwanger, in his novel *Die Füchse im Weinberg*, compared observing the march of history to watching a clock: if we look at the clock every 30 seconds, it seems to be at a standstill. But it marches on.
- 2 It is highly recommended that readers do not skip the Introduction (we almost always do when reading a technical book). The Introduction explains in more detail the philosophical outlook on energy markets underlying this book, the choice of topics and organisation of material.
- 3 It is, of course, possible that some producers and end users will choose to rely primarily on the spot markets. This is usually a risky strategy, both from the point of view of price exposures and operational risk.
- 4 Vincent Kaminski, 1997, "The US Power Markets. Restructuring and Risk Management," in *The US Power Markets* (London: Risk Books).
- 5 The elasticity of supply measures sensitivity of supply to the changes in prices.
- 6 One example of such regularities can be observed in the oil markets, when the spare OPEC production capacity drops below what markets see as a critical level. The perceptions of what may represent a critical level will vary with market conditions, but typically the spare capacity levels that may be wiped out by the removal of one critical supplier from the market (Iran, Nigeria) can trigger speculative buying.
- 7 Price gouging after natural disasters is prohibited in most jurisdictions. See Matt Zwolinski, 2008, "The ethics of price gouging," *Business Ethics Quarterly*, 18(3), pp 347–78, for more information and discussion of the related moral and ethical issues.
- 8 The outright dismissal of the possibility of market manipulation is often the consequence of strong beliefs in market efficiency, based more on conclusions drawn from the model of a perfect market than from practical experience on a trading floor. Such arguments remind us of the proof of God's existence by Saint Anselm of Canterbury (1033–1109): the definition of a perfect being contains existence, therefore God exists because God is perfect. This proof was devastated by a number of philosophers, including Immanuel Kant, who based his criticism on the fundamental distinction between analytic and synthetic judgements. In the author's view, the *a priori* dismissal of market manipulation as inconsistent with an efficient market mechanism is an affliction that even sequestration for a few hours in an elevator trapped between floors during the California energy crisis, caused by massive market manipulation, will cure.
- 9 Chapter 24 contains further clarification of the author's view on market manipulation and a discussion of the differences between market manipulation and speculation, which are often conflated in the popular press. Speculation is a necessary component of any market with an imbalance between natural longs and natural shorts, and a critical source of liquidity.
- 10 Friedrich Hayek, 1944, *The Road to Serfdom* (Chicago, Ill.: The University of Chicago Press).
- 11 Friedrich Hayek, 1960, *The Constitution of Liberty* (Chicago, Ill.: The University of Chicago Press).
- 12 IOSCO, 2012, "Functioning and oversight of oil price reporting agencies consultation report," March (see <http://www.iosco.org/library/pubdocs/pdf/IOSCOPD375.pdf>.)
- 13 See Footnote 12.
- 14 Neil F. Johnson, Paul Jefferies, Pak Ming Hui, 2003, *Financial Market Complexity: What Physics Can Tell Us about Market Behaviour* (Oxford, England: Oxford University Press).
- 15 When the Clinton administration proposed the so-called Btu tax (and later dropped the idea), many US residents did not object. Who would object to taxing the British?
- 16 <http://www.legislation.gov.uk/uksi/1995/1804/schedule/made>.
- 17 Letter M corresponds to the Latin numeral for 1,000. MM = 1,000 × 1,000.
- 18 [http://www.mitenergyclub.org/assets/2008/11/15/Units\\_ConvFactors.MIT\\_EnergyClub\\_Factsheet.v8.pdf](http://www.mitenergyclub.org/assets/2008/11/15/Units_ConvFactors.MIT_EnergyClub_Factsheet.v8.pdf).
- 19 An alternative visualisation is given and the reader has to decide which one is better. We find both of them to be very useful.

- 20 Adjustments, losses and unaccounted for.
- 21 The fossil fuel category also includes 0.01 quadrillion Btu of coal coke net exports.
- 22 As explained in the release notes, energy flow statistics are quoted from the US Department of Energy's Energy Information Administration. These statistics can be found in the "Annual Energy Review 2010," report number DOE/EIA-0384(2010). The original document can be found at: <http://www.eia.gov/totalenergy/data/annual/pdf/aer.pdf>. The graph can be found at: <http://www.eia.gov/totalenergy/data/annual/diagram1.cfm>.
- 23 Based on the post by Nate Hagens in "The Oil Drum": <http://www.theoildrum.com/story/2006/9/2/171333/1102>.
- 24 Sepulchral inscriptions for a rich noble family, the Statilii, list a total of approximately 428 slaves and freed persons from 40 BC to AD 65. Seneca, a man of extraordinary wealth, believed he was travelling frugally when he had with him one cartload of slaves (most likely four or five). (Ep. 87.2) Statilia Messalina, the wife of Nero, had four or five slaves (information obtained from: <http://www.worldhistoryblog.com/2005/06/slavery-in-roman-empire.html>).
- 25 Joseph A. Tainter and Tadeusz W. Patzek, 2012, *Drilling Down: The Gulf Oil Debacle and Our Energy Dilemma* (New York, NY: Copernicus Books). A US industrial worker in a relatively big town commands 28,000 people by a similar calculation.
- 26 When the deregulation started, many executives who had reached positions of responsibility in the sheltered environments committed serious mistakes and proved unfit to function in the new brave world of deregulated markets, with catastrophic consequences for them and their companies.
- 27 Wall Street Reform and Consumer Protection Act, H.R.4173. One challenge faced by the author at the time of writing is that many related regulations (by the CFTC and SEC) were a work in progress and likely to be challenged in courts.
- 28 Ronald Dore, 2000, *Stock Market Capitalism: Welfare Capitalism: Japan and Germany versus the Anglo-Saxons* (Oxford, England: Oxford University Press).
- 29 One has to recognise that national income statistics measure the use of resources in the financial industry and not so much its contribution to social welfare. A significant body of economic research points to positive correlation between the size of the financial system and economic growth. We could ask, however, if there is an optimal size of the financial sector relative to the rest of the economy. A 2011 paper (Jean Louis Arcand, Ugo Panizza, Enrico Berkes, 2011, "Too much finance?" *Vox*, March) concludes that "the marginal effect of financial development on output growth becomes negative when credit to the private sector surpasses 110% of GDP." The paper contains also a summary of economic research on this topic (see <http://upanizza.blogspot.com/2011/03/too-much-finance.html>.)
- 30 In the case of Enron, a significant portion of profits in some quarters was derived from tax oriented strategies. It is obvious that the use of such strategies had to stop one way or another.
- 31 John Kay, 2008, "Taxpayers will fund another run on the casino," *Financial Times*, September 17 (see <http://www.johnkay.com/2008/09/17/taxpayers-will-fund-another-run-on-the-casino>.)
- 32 The rush to jump into the new deregulated energy trading business quickly reached surrealistic and comic dimensions. "Every day we don't participate in the trading and marketing business, we are losing ground" was the mantra of energy company executives across the US.
- 33 A common practice in the industry in the late 1990s was the recognition of huge "first day mark-to-market" gains by both counterparties to the same transaction. It is obvious that both counterparties in this case could be wrong – they could not both be right at the same time.
- 34 <http://www.globaled.org/issues/176.pdf>.
- 35 [http://articles.businessinsider.com/2011-05-05/markets/29982186\\_1\\_china-s-gdp-michael-pettis-global-economy](http://articles.businessinsider.com/2011-05-05/markets/29982186_1_china-s-gdp-michael-pettis-global-economy).

- 36 One of the paradoxes of history is that the first decade of rapid globalisation (the 1990s) not only reintegrated billions of people from the former communist nations into the world economy, but also brought about low prices for certain critical commodities. The Russian economy collapsed in the early 1990s, forcing the nation to liquidate strategic reserves of certain commodities. The decline in prices created a sense of false security and delayed investments in extractive industries. The process of privatisation of the Russian economy contributed further to the commodity price declines. This is not the topic of the book, but, nevertheless, it is a good example of how geopolitical developments affect the commodity markets and how important it is to understand the proverbial "big picture." The price trends due to unique circumstances do not last forever. For more on this, see Marshall I. Goldman, 2003, *The Piratization of Russia: Russian Reform Goes Awry* (London: Routledge).
- 37 President Evo Morales of Bolivia had tried to reduce them, but eventually had to reestablish food and energy price subsidies after facing mass protests and riots.
- 38 In the extreme, it is the use of natural resources to further political aims.
- 39 There is often a tendency to accuse the most prominent politicians engaging in the policies inspired by resource nationalism to be part of a diabolical conspiracy. We are very sceptical of such claims where simple incompetence is a sufficient explanation.
- 40 One example of this is the initiative of a number of German companies at the time of writing. "Twelve German companies have joined the new German alliance aimed at securing raw materials supplies in the face of growing competition for key commodities, the Federation of German industry BDI said on Monday. In October 2010, Germany's government approved a new commodities strategy aimed at helping German industry secure supplies in the face of intense competition from China and other newly-industrialised countries which will include partnerships with supplier countries and greater cooperation between German commodity consumers" (see <http://af.reuters.com/article/metalsNews/idAFL5E8CU3EV20121030?pageNumber=2&virtualBrandChannel=0&sp=true>).
- 41 The growing risk of military conflicts as developed countries try to protect their access to increasingly scarce vital natural resources has been analysed by Michael Klare, 2001, *Resource Wars: The New Landscape of Global Conflict* (New York, NY: Metropolitan Books), and Michael Klare, 2012, *The Race for What's Left: The Global Scramble for the World's Last Resources* (New York, NY: Metropolitan Books).
- 42 These factors are extensively discussed in many financial blogs and publications. Several books, including basic investment advice publications, and academic papers have covered this topic. See, Parantap Basu and William T. Gavin, 2011, "What explains the growth in commodity derivatives?" *Federal Reserve Bank of St. Louis Review*, 93(1), January / February, pp 37–48; John Stephenson, 2010, *The Little Book of Commodity Investing* (Toronto, Canada: Wiley).
- 43 Gary Gorton, K. Geert Rouwenhorst, 2005, "Facts and fantasies about commodity futures," working paper, February 28; Ibbotson Associates, 2006, "Strategic asset allocation and commodities."
- 44 The authors constructed an equally weighted index of commodity futures. The hypothetical portfolio was collateralised in the following sense: Treasury bills were acquired for the positions held in the commodity futures and the interest earned on these holdings was included in the calculations of returns.
- 45 Inflation was measured by the Consumer Prices Index (CPI).
- 46 Gary Gorton, K. Geert Rouwenhorst, *op. cit.*
- 47 *Op. cit.*
- 48 Ibbotson Associates, *op. cit.*
- 49 Parantap Basu and William T. Gavin, *op. cit.* See also Parantap Basu and William T. Gavin, 2011, "Negative correlation between stock and futures returns: An unexploited hedging opportunity?" *Federal Reserve Bank of St. Louis*, working paper 2011-005A, January.
- 50 The growing importance of hedge funds whose strategies contribute to an increased correlation between different asset classes was documented by Bahattin Büyüksahin and Michel A.

Robe, 2011, "Speculators, commodities and cross-market linkages," working paper, February 12. The increasing correlations between prices on non-energy commodities and oil was discussed by Ke Tang and Wei Xiong, 2011, "Index investment and financialization of commodities," working paper, March 2012.

- 51 Gross market value is a superior measure of the market size. A contract with a very big notional value may have a negligible market value and represent a small overall risk.
- 52 M.W. Masters and A.K. White, 2008, "The accidental Hunt brothers: How institutional investors are driving up food and energy prices," p 1 (see <http://accidentalhuntbrothers.com/>.)
- 53 A few recent books on this topic include James E. Hansen, 2009, *The Storms of my Grandchildren: The Truth about the Coming Climate Catastrophe and Our Last Chance to Save Humanity* (New York, NY: Bloomsbury Press USA); Naomi Oreskes, 2010, *Merchants of Doubt: How a Handful of Scientists Obscured the Truth on Issues from Tobacco Smoke to Global Warming* (New York, NY: Bloomsbury Press). The position of the global warming sceptics is represented by Patrick J. Michaels (Ed), 2005, *Shattered Consensus: The True State of Global Warming* (Lanham, MD: Rowman & Littlefield). A book by US Senator James Inhofe published in 2012 contains a very strong criticism of the global warming theory: *The Greatest Hoax: How the Global Warming Conspiracy Threatens Your Future* (WND Books).
- 54 It may be a paradox that the global warming trend can cause colder winters. Melting of ice cover over the Arctic Ocean creates an area of high pressure pumping cold air into Europe. This mechanism was predicted by scientists at the Alfred Wegener Institute for Polar and Marine Research and by Vladimir Petoukhov of the Potsdam Institute (see <http://www.independent.co.uk/news/science/science-behind-the-big-freeze-is-climate-change-bringing-the-arctic-to-europe-6358928.html>).
- 55 Some opponents of the Keystone XL argued that the pipeline will make it possible to expand the production of oil from tar sands and that opening this source of hydrocarbons with high CO<sub>2</sub> intensity will tip the scale, making global warming all but inevitable. The latter may be true or not, but cancellation of the Keystone XL will not stop expansion of the Canadian tar sands activities.
- 56 Admiral Hyman Rickover, 1957, "Energy resources and our future," May 14 (see <http://www.energybulletin.net/node/23151>).

## *Energy Trading: The Organisation*

Running an energy trading operation is a fascinating and usually very lucrative occupation. However, running an energy trading operation well can often be an inhuman task. This is due to a number of factors, primarily information asymmetry and the social dynamics of trading floors. Information asymmetry is involved because of the peculiar properties of the energy markets, which form a closely coupled, integrated system, but still a system of connected, highly specialised niches. After a few months of immersion, a trader or analyst often becomes a world-class expert in a given area. They know more about a specific state retail electricity auction or pipeline flows around the Houston Ship Channel than any other person in the firm. Challenging an expert is always difficult, especially when the expert is young, successful and brash. Unquestionable but narrow expertise often leads to excessive, illusory perceptions of trading and deal-making skills and inflated self-assessment (it is sufficient to read annual self-evaluations to see this). This trait of human personality is known in psychology as the Dunning–Kruger Syndrome, which was summarised as follows.<sup>1</sup>

- *Prediction 1.* Incompetent individuals, compared with their more competent peers, will dramatically overestimate their ability and performance relative to objective criteria.
- *Prediction 2.* Incompetent individuals will suffer from deficient metacognitive skills, in that they will be less able than their more competent peers to recognise competence when they see it – be it their own or anyone else's.
- *Prediction 3.* Incompetent individuals will be less able than their more competent peers to gain insight into their true level of performance by means of social comparison information [...].
- *Prediction 4.* The incompetent can gain insight about their shortcomings, but this comes (paradoxically) by making them more competent, thus providing them with the metacognitive skills necessary to be able to realise that they have performed poorly.

The danger we have witnessed many times on trading floors when decisions have to be taken under stress and time pressure is that the incompetent people come across as decisive leaders and save the day (or so it seems). They have no doubts. As William Butler Yeats would say:

Things fall apart; the centre cannot hold; [...]  
The best lack all conviction, while the worst  
Are full of passionate intensity.

Given the laws of probability, incompetent people may be successful for a period of time and move up in the organisation. However, they will eventually cost an organisation a lot of money and may even threaten its existence. Hopefully this chapter and the rest of the book will provide useful information to counter such a toxic mix of highly specialised, concentrated knowledge and inflated egos (we are absolutely, positively confident this chapter is worth reading ...).

We will now discuss the basic principles of the organisation of an energy trading business, starting with the sources of profits and the ways in which it can obtain a competitive advantage. Extensive use of fundamental and competitive information is a key to successful commodity trading.

## **ORGANISATION OF AN ENERGY TRADING ORGANISATION AND THE SOURCES OF PROFITS**

### **Organisation**

The main question we address in this section is the definition and organisation of an energy trading operation. A very good description of an entity engaged in energy trading can be found in the now rescinded EITF 98-10 document.<sup>2</sup> According to this, the following features, repeated almost *verbatim*, would define an energy trading operation:

- the primary business of a business entity is not production, transportation and delivery of energy;
- the primary assets and liabilities are contracts;
- the energy products are offered by an entity acting as a dealer and not as a producer or transporter;
- the major competitors are other traders;
- the portfolio of contracts is characterised by a high turnover;
- the contracts for delivery/receipt are often outside the entity's primary geographical location;

- the entity does not have the resources to produce and gather energy in the primary environment;
- compensation of employees is based on trading profits calculated using a mark-to-market approach;
- the organisation communicates internally in terms of “trading strategy” or other similar concepts;
- employees are referred to as traders or were recruited from other trading operations;
- positions are tracked on an overnight basis;
- infrastructure of the operation is similar to that of a trading operation of an investment or commercial bank (front office, middle office, back office), with the separation of responsibilities dictated by prudent risk management practices;
- an infrastructure exists to track the price, credit and other risks in real time;
- the activities are managed on a book (portfolio) basis;
- the trading strategy is based on exploitation of arbitrage opportunities, unusual price relationships or outright directional bets;
- the contractual provisions allow for pairing off contracts with the counterparties (known as net cash settlements, offsetting, booking out, netting); and
- the contracts contain provisions for liquidated damages (ie, settling claims related to failures to perform based on the observed market prices).

The definition based on the criteria listed above is wide enough to include specialised trading operations in the big financial firms, hedge funds as well as specialised trading operations of large integrated energy producers and utilities. It is general enough to capture the operations of firms that deny that they are in the energy trading business (given that energy trading may be a dirty term to some people, and sometimes may be seen by shareholders, regulators, stock analysts and creditors as too risky).<sup>3</sup> If one is shown a picture of a large energy trading floor, one cannot say if the picture was taken on Wall Street or in Houston (unless one is familiar with a specific location). To be more specific, energy trading organisations may be embedded in a number of different companies, including;

- commercial and investment banks;
  - regulated utilities;
-

- big integrated oil/natural gas majors;
- independent oil and natural gas producers;
- industrial firms; and
- hedge funds.

The trading operations described above often perform dual functions as providers of services to other units of their companies (hedging, marketing, fuel acquisition, etc) and also as providers of risk management instruments to the rest of the energy industry and energy end-users. Speculative, directional bets are an important part of their activities (although this is not mentioned in polite society). Market participants with extensive physical operations (such as large, integrated oil companies) that are active in the financial markets often enjoy a significant competitive advantage – they have privileged access to information flows about the conditions of the physical systems, and at the same time they combine strong balance sheets with smaller (compared to their competitors in the financial industry) regulatory overheads.

In addition, most of the big producers and end-users of energy have clusters of employees responsible for marketing their production flows and acquisition of fuels and feedstock. They act as traders, analysts and risk managers, but they do not consider making directional bets as their mandate. Of course, the demarcation line between speculation, marketing and hedging is often very fuzzy in practice.

The design of an energy trading organisation closely follows the principles established for financial trading desks in commercial and investment banks, with some differences dictated by the unique features of the commodity markets. Historically, merchant energy companies learned the trading skills from the banks, either through joint ventures or by hiring experienced traders from Wall Street firms.<sup>4</sup> The biggest challenge was always blending the distinct cultures of physical and financial traders. The situation today is not much different; at some point, the tide turned and many financial institutions started drawing on the experience of merchant energy companies such as Enron, El Paso, Dynegy and many others. After the energy trading crisis of 2002–03, the DNA of the US merchant energy companies was dispersed across the industry. Energy trading in the US and Europe has become largely dominated, for better or worse, by the veterans of the merchant leaders of the late 1990s.

An energy trading organisation follows a standard design of a financial trading operation, with the separation of responsibilities enforced through the division of front, middle and back offices, although the boundaries between these units are often quite fuzzy.

The front office includes traders, sometimes with supporting personnel. The second group of employees typically included in the front office is the origination team (sometimes called the sales force), who are responsible for long-term structured transactions.

The middle office includes risk managers and analysts that support traders as well as risk management. The middle office sometimes also includes computer programmers supporting both the front and middle office. The analysts are typically divided into quants, who develop quantitative models for the valuation of derivatives and risk management, and “fundies” – ie, analysts responsible for fundamental analysis. Both groups often help in the calibration of valuation models, an activity that can be as important as the development of the quantitative algorithms and procedures. Middle office also includes the so-called structuring group responsible for the development of valuation and hedging models for complex, structured transactions, which may include multiple commodities and instruments. The models are deal-specific, require frequent updates and evolve over time as a transaction goes through a long negotiation process.

The back office includes clerical employees who are responsible for trade entry, updating and the marking-to-market of trading portfolios, running of risk and position reports, distribution of information to the management and responding to *ad hoc* queries about the portfolios, position levels and related information (price, volatility and correlation curves). There are many departures from this standard model. The most frequent variations on this theme are discussed below.

Before turning to a detailed discussion of the responsibilities of different groups, we first discuss the main differences in design of the organisational structures, related to reporting lines for different employees. A frequent controversy is related to the location of quantitative modellers, both in terms of sitting arrangements (ie, physical location) and the organisational charts. The latter issue is whether the quants should work in one centrally managed group or whether they should be assigned to different units (the trading desks, risk manage-

ment, IT, structuring). Under one solution, both front and middle offices are supported by a single independent quantitative analysis group. Under another, each business unit relies on its team of dedicated analysts. Traders strongly support the latter solution, ie, support by a group of quants who specialise in different markets, develop models at short notice for the valuation of new transactions and have multiple reporting lines (possibly a direct line to a trader and a dotted line to the chief quant). Irrespectively of reporting lines, such quants are controlled directly by traders, who determine their compensation. This model has a number of advantages, including direct exposure of quants to market realities, shortened communication lines, frequent feedback, the ability to react on short notice to market developments and the ability to troubleshoot if and when necessary.

However, what is often lost is the potential for cross-training and cross-pollination between different groups of quants, the advantages of centralised management and quality control. This arrangement also creates a potential for decentralised model development with multiple modelling philosophies, software packages and multiple versions of the same basic model propagating across the company. The groups of desk quants develop into information silos, with the loss of economies of scale and modelling transparency. One widely recognised shortcoming of this arrangement is the potential for quants “going native” – ie, becoming the alter egos of traders and being unable to resist pressures to produce models distorting valuations, used opportunistically to justify the trading strategy *du jour*. An alternative solution is to develop a centralised research group with resources shared by all the business units. The obvious economies of scale and team work are often negated by isolation from the trading desks and the tendency to treat the quants as a free resource (and to abuse and overexploit them).<sup>5</sup>

As in any human organisation, there are no perfect solutions, and the decision including how to utilise the quants depends on many factors including the culture of the organisation. As always, it is important to understand the pros and cons of different solutions, and to take proactive steps to avoid potential adverse consequences of the selected solutions. One frequent compromise (the one we personally favour) is to assign quants to different desks on a semi-permanent basis, allowing them to develop specialised skills

and acquire experience only obtained through immersion in a specific business over a long period of time. However, in many cases such quants may develop double loyalty and be used as pawns by a trader with their own agenda. This arrangement works only if the research group is run by someone with a strong personality and a clear sense of the mission. But even such a person (if they exist at all) will eventually get tired of navigating the conflicting political pressures inevitable in a big organisation.

Sometimes the solution that eventually emerges is a troubled co-existence between multiple quantitative analysis groups, supporting the trading desks, risk management, the structuring group and one centralised group responsible for vaguely defined corporate level projects. Sometimes, the Balkanisation of quantitative research can be taken to extreme levels, with different groups of quants and dedicated programmers supporting, for example, market risk and credit risk, although these exposures are two sides of the same coin and should be modelled in a unified framework.

The dilemma related to the organisation of quantitative modelling efforts can be seen in the different solutions used with respect to the placement of the structuring group and risk managers. The structuring group is sometimes organised as a central unit servicing all the trading and origination desks, but the analysts may be also allocated permanently to support specific desks or even specific traders/originators. Some companies have parallel risk organisations. The in-business risk group is responsible for alerting traders and originators to the potential problems in a contemplated transaction before it is submitted for final review and approval by the corporate (independent) risk group. These solutions suffer from the same benefits and shortcomings we discussed above in relation to the quant group.

Keeping in mind that in practice we may see a number of different solutions, we can proceed now to discuss the specifics.

### *The front office*

The front office executes two related types of transactions. One set of transactions includes highly standardised instruments, which can be bought or sold with a click of the mouse or a short telephone call. Other transactions are highly customised and require long negotiations and the involvement of employees with different skills, ranging

---

from quants to lawyers. The traders are responsible for execution of transactions which are reasonably standardised and repetitive. These transactions may be undertaken for a number of reasons (to be discussed in more detail later), including:

- speculation;
- market making;
- trading on behalf of a customer; and
- hedging existing positions.

The trading desks are organised typically by different commodity markets, including natural gas, oil and refined products, electricity, coal and emissions. The physical arrangement of the trading floor is developed to facilitate interactions between traders covering related commodity and geographical markets. For example, the natural gas and electricity traders covering Texas are likely to be placed next to each other, as they are likely to be exchanging information and trading ideas all the time, as both markets are closely related. The organisation of each specific commodity desk may vary from company to company but one can identify certain commonalities. Using the example of natural gas, one dimension along which the traders can be organised is physical versus financial trading. Another important dimension is regional specialisation, with typical units including Eastern US, Midcontinent, Texas and the West. Another possible specialisation is related to the transaction types, with financial transactions separated into such specialisations as Nymex trading (with a further distinction between futures and options trading), basis trading, trading of swaps and financial options at different geographical locations. Combinations of all those different dimensions create potential for multiple organisational and spatial arrangements. Specific solutions depend on the size of the organisation, the scope of coverage of different markets and profit margins available in different markets. In smaller organisations, employees have to wear, by necessity, many hats, and the extent of specialisation is limited.

Commercial personnel, responsible for structured long-term transactions, are often referred to in the industry jargon as “originators”. They are responsible for transactions that are typically long-term, involve significant volumes, require long and complex negotiations, and often represent significant legal and regulatory

risks. Unlike trades in standardised instruments which can be executed in a few seconds, structured transactions can keep a group of professionals representing many disciplines busy for months, require going through multiple levels of approvals and are sometimes associated with legal documents resembling the telephone book of a big city. These transactions also require the involvement of a dedicated group of analysts from a structuring and sometimes quant group, given the complexity of valuation and the complexity of embedded options. Unlike the traders, who tend to be more opportunistic and behave like hunter-gatherers, originators should in principle be more attentive to developing an understanding of the needs of their clients and helping them to address their problems, bringing to the table the skills of the entire organisation.

The distinction between traders and originators corresponds to the distinction we make between trading and marketing. The difference between the two is motivation and the way the counterparties are engaged. Trading is associated either with directional bets or hedging. Marketing is related to participation in the supply chain and requires the establishment of long-term, close relationships with the clients. Unlike trading, which can involve either physical or financial transactions, marketing always requires physical commodity flows. Trading and marketing are mutually co-dependent. Marketing transactions require hedging; marketing helps traders to increase the size of the trading book and gives them order flow and valuable market intelligence. The demarcation line between these two activities may be fuzzy at times, but it is still useful. The members of these two groups have a different time horizon, objectives, even personalities. One would know if they sat next to a trader or a marketer.

The volumes of commodities bought/sold under structured transactions are transferred to the trading books to be managed within the context of the entire portfolio. This creates a potential for conflict between the traders and originators. In most energy trading groups we have worked for, we heard complaints about the traders intercepting a portion of origination profit and loss (P&L) by temporarily adjusting the less liquid part of the price curve. For example, just before a huge sale happens, the forward prices and volatilities are adjusted, with reversion to the previous levels as soon as the big transaction is booked.<sup>6</sup> This issue can be easily addressed by

combining the traders and originators covering the same markets into one group, under the umbrella of the same management.

In addition to the quants, the traders and originators are often supported by clerical personnel and associates responsible for trade entry, downloading and uploading (from, and to, the trading system) of price information, and the handling of *ad hoc* requests.

#### *The middle office*

The responsibilities of the middle office are defined differently from place to place, but the emerging consensus identifies this operation with risk management which is broadly defined. The responsibilities of the middle office include:

- ❑ formulation and enforcement of a trading policy – a trading policy is a detailed document, typically subject to approval by the board of directors or senior management, containing a number of provisions, such as:
  - risk limits, expressed in terms of allowed volumetric position levels, sensitivity levels or market and credit value-at-risk (VaR) limits;<sup>7</sup>
  - risk concentration guidelines related to maximum positions and exposure levels allowed by individual traders, trading desks and business units, across maturities and geographical locations; and
  - procedures for notifying senior managers and the board about trading policy violations and losses exceeding certain levels.
- ❑ oversight of the risk models' development and the design of risk reports;
- ❑ oversight of production of regular (usually daily) model runs, review and analysis of the daily risk reports; and
- ❑ validation of the models used by the front desk for transaction valuation, including validation of the inputs and techniques used for model calibration.

#### *The back office*

Back office is responsible for:

- ❑ updating the information contained in the database of the software system supporting trading, such as:
  - price curves;
  - volatility curves;

- correlation curves; and
  - other information required to capture the trading positions (volumes, settlement calendar).
- producing and distributing position reports, risk reports and daily P&L statements;
  - confirming the details of transactions through exchange of appropriate information with the counterparties;
  - managing collateral calls and cash transfers with the brokers and other counterparties;
  - interacting with the accounting groups to settle derivative transactions (swaps and options);
  - handling physical flows of commodities (pipeline nominations, transmission scheduling, scheduling of tankers, etc).

Most energy trading operations rely on multiple systems for financial and physical transactions, with poorly designed interfaces between them. For example, financial transactions in gas and oil may reside in the main trading platform, but physical transactions, with associated scheduling modules, may be stored in specialised systems designed to capture commodity specific details, which often elude standardisation. This often happens because the logistic arrangements, market conventions, grade details, geographical definitions and units of measurement vary across markets. The need to use different systems to handle physical commodities results in a tangled web of systems communicating through *ad hoc* patches and ubiquitous spreadsheets. In some organisations, back-office personnel spend most of their time chasing information dispersed across thousands of spreadsheets, contributing to a very high average cost of processing transactions. Given the difficulty of validating calculations and formulas, and their sheer numbers, spreadsheets are one of the main sources of risk in big corporations.

### **The sources of profits**

What are the sources of profits of an energy trading organisation, ignoring the obvious prescription of buying low and selling high (although this is easier said than done)? This section will offer a classification of ways of making money. This is somewhat artificial, as a single transaction may contain elements related to different strategies. It helps, however, to enumerate the options available to

management and to identify the skills required to engage in different strategies.

It may seem a banal statement but is one well worth repeating: a trading organisation makes money by taking risks, and the main function of the front office is to decide which risks should be laid-off to third parties and which risks should be retained in-house. In practice, the risks that are warehoused are the risks that are difficult or impossible to divest – and which are associated with the widest profit margins. The risks that are transferred to third parties are not magically annihilated: rather, they are transformed into other, more manageable or more acceptable risks. On many occasions, the author has heard excessively optimistic statements about the ability of an energy trading organisation to break up efficiently any transaction into component atomic blocks (standardised instruments such as options, swaps, swaptions), with all the individual risks being effectively priced and managed.<sup>8</sup> Such pronouncements are true only with respect to the most basic transactions and in the most extreme forms are either a manifestation of a poor understanding of the energy business or of the worst form of corporate propaganda.

One has to recognise that many complex deals are not decomposable into simple atomic transactions and cannot be effectively hedged because there are no corresponding transactions available in the market place, or the underlying is not traded, or is not traded in a sufficiently deep market, which allows for active hedging transactions. Energy markets are often incomplete and shallow; one has to be humble enough to admit the obvious.<sup>9</sup> Also, the cost of hedging all component risks would be prohibitively high. The transfer of all imaginable risks to third parties, after the incorporation of transaction costs, leaves no profit margin for a trading organisation.

An energy trading operation makes money in many different ways, including:

- market making;
- directional trading;
- trading around assets; and
- asset management.

We shall comment briefly on some of the points above.

### *Market making*

This book uses a narrow definition of market making, understood as the readiness to show both bid and offer prices at the request of a customer and take either side of the trade. The revenue is derived from capturing the bid–offer spread, and is driven by two critical factors: the volume of transactions and the size of the bid–offer spread. The spread is typically larger for less liquid markets and longer maturities. One has to recognise, however, that the frequency of long-term transactions may be low, and that a firm may have to warehouse risks when engaging in transactions in a low liquidity market. It may require considerable time to close a position related to a large, long-term transaction in an illiquid market, with considerable risk of a significant price move in the meantime.<sup>10</sup>

Some traders are willing to address this risk by using a stack-and-roll strategy, which was very popular in the early 1990s, before the dangers of this way of hedging were exemplified by the drama of MG Marketing and Refining.<sup>11</sup> This strategy consists in using short-term instruments to cover the entire volumetric exposure, distributed over time. As the short-term hedges move closer to expiration, they are liquidated and the position is renewed using the next available sets of short-term instruments. For example, the entire hedge may be concentrated in the first available futures contract. The position is liquidated a few days before contract expiration, and a hedge is then reinstated using the next available contract. The risks of such strategy are obvious: the cost of rolling over the hedges may be significant, especially if the existence of a big position is revealed to the market. The traders working for other firms will extract their pound of flesh, by front-running the hedger.

The profit margins related to market making in highly liquid markets are relatively low, as the following calculation can demonstrate. EnronOnLine (EOL) was once the biggest electronic trading platform, with trading in natural gas dominating its business. Assuming a bid–offer spread of US\$0.005/MMBtu and 5,000 transactions a day, the annual revenue would be equal to:

$$5000 \times 5000 \times 250 \times 0.005 / 2 = \text{US\$}15,625,000$$

The natural gas contract traded on EOL had underlying volume of 5,000 MMBtus: this explains the number in the calculations above. Division by two reflects the need to make a round trip, ie, buy and

sell (or vice versa) in order to capture the bid–offer spread. The total revenue is equal to about US\$16 million, not an insignificant amount, but a far cry from the costs of maintaining the platform. Of course, moving along a forward curve to longer maturities is associated with wider bid–offer spreads but the potential volumes decline.

Additional revenues can be generated in the form of so-called rebates by directing business to different exchanges. A more significant revenue stream is generated from the market-making activity that is often described as customer-oriented business or trading on behalf of customers, and treated as a separate activity. An entity operating as a hedge provider inserts itself between the producers and end-users of a given commodity, usually through long-term transactions. The producers are willing to sell forward at a price below expected spot price in return for elimination of the price risk:

$$E(S_T) - f_p$$

where  $E(S_T)$  denotes the expected spot price at time  $T$ , and  $f_p$  denotes the discount that the producers are willing to accept (the use of letter  $f$  for fear seems to be appropriate in this context). The end users are willing to pay a price above the expected spot price:

$$E(S_T) + f_e$$

The margins captured by the hedge intermediary are equal to  $f_p + f_e$  (we can call it the hedging spread).

It is obvious that it is in the collective interest of the institutions serving as intermediaries to keep the information about the level of the hedging spread confidential. This may be accomplished by transacting in the OTC markets or using customised transactions, with the levels of transaction prices treated as commercial secrets. Making this information public could create incentives for producers and end users to deal directly among themselves (and potentially a new group of intermediaries willing to accept lower margins entering the industry, in the same way as discount brokers invaded the brokerage industry). The margins can be successfully buried in the complex transactions, with multiple embedded options and complicated and convoluted pricing formulas.

It is also in the collective interest of the hedge providers to steer the commercial hedging entities into long-term transactions, benefitting not only from the increased volumes but also from migration to the

illiquid segments of the forward price curves. The marketing tools that can be used range from emphasising commodity price volatility (a cynic could say that a terrified customer is a satisfied customer) to insisting on hedging as a precondition for obtaining a loan. Under US law, tying loans to certain conditions is illegal, but many borrowers will use the lender as a provider of hedges for convenience and for cultivation of the relationship. The strong opposition of the financial industry to the provisions of the Dodd–Frank Act requiring trading of “standardised” swaps on exchanges, and reporting of executed swaps and prices to a central data repository, can be explained as a rational effort to defend an important source of profits and an effort to keep the markets as non-transparent as possible. What is more curious is the opposition to mandatory exchange trading and clearing of derivatives coming from the end-user community. This strong reaction cannot be dismissed as the actions of poorly informed and uneducated hicks ready to carry water for the financial industry. The reality is complex and cannot be explained in black and white terms. The end users draw many benefits from relying on bilateral arrangements with hedge providers, although these benefits do come at a cost.

Further discussion of this apparent puzzle will be provided in Chapter 6 in the section covering organised exchanges and collateral management. At this point, however, it is important to recognise that the hedging counterparties often offer an important service that justifies the compensation they receive by acting as financial intermediaries (which does not mean, as argued above, that they would not like to make a little extra by keeping the markets more opaque).

#### *Directional trading*

Making directional bets is typically associated with proprietary trading desks at hedge funds and other financial institutions,<sup>12</sup> although it does take place in many other trading organisations – including large oil companies, utilities and other merchant energy companies. Sometimes, it happens under the guise of hedging or engaging in “customer-related business” (as explained above), with or without the direct knowledge of senior management. The line between hedging and speculating is a fuzzy one, as everyone in this business knows. In many cases, the merchant energy companies

choose not to advertise aggressive trading to their shareholders, who may have doubts regarding the ability of many companies outside the financial industry to engage in speculative trading for too long. In general, the market does not reward merchant companies that engage in successful trading (with or without the fig leaf of hedging), and punishes those who fail.

Directional trading is seen as a zero-sum game and, in a very general sense, this is true. Since the beginning of the 2000s, we have seen several cases of hedge funds with opposite positions facing each other like two duellists, with an obvious and inevitable finale: the ruin of one player and the riches for the manager with more staying power or a more reckless character.<sup>13</sup> It does not explain, however, how hedge funds can continue to exist in the long run, given that less-skilled fund managers will eventually run out of money and will have to leave the industry, and the successful managers may choose to retire and enjoy the rewards of their labours and peace of mind. Without constant inflows of new funds, the successful funds would have to close their doors.<sup>14</sup> As always, the answer is more complicated than an assumption that the gains of one speculator come from the losses of another.

The most skillful proprietary traders engage in strategies that do not necessarily generate profits from the losses of other speculators. The gains come from a number of sources, such as the following.

- ❑ *Risk-transfer services.* As explained above, the producers and end users of energy commodities are willing to trade lower expected return for risk reduction. The speculators step in between them and make profits by reducing the returns of both sides in exchange for the services they are providing.<sup>15</sup>
- ❑ *Exploitation of the behavioural biases of other market participants or market imperfections.* In the chapters on different market segments, we shall attempt to identify such opportunities reflecting unique features of different parts of the energy complex.
- ❑ *The ability to identify relative value opportunities.* In other words, a skilled trader does not usually place bets on prices of a specific commodity going simply up or down. The bets are made on relative prices of different commodities or relative prices of the same commodity at different locations or at different points in time. Some of the bets are exceptionally uninspiring (for example,

making bets on relative prices of natural gas and oil on an MMBtu basis,<sup>16</sup> as if a physical mechanism restoring some presumed equilibrium between the two existed currently).<sup>17</sup> The bets that work are usually based on hard work and diligent accumulation of supply and demand data or information coming from other sources (see the section on competitive intelligence in this chapter).

- ❑ *Strategies based on the exploitation of market imperfections over very short-time periods*, irregularities that last sometimes for just a few seconds. This so-called algorithmic trading is likely to become more popular in the future.
- ❑ *Superior weather prediction skills*. Weather information is critical to energy trading, as explained above, and many well-financed trading operations can afford top meteorological talent. Organisations with better fundamental skills are effectively front-running other market participants.

The reliance on profits from directional trading in many financial institutions has been negated by inclusion in the recent Dodd–Frank Act (The Wall Street Reform and Customer Protection Act, or WSRCPA, referred to as the Act in this book), a provision against proprietary trading defined as:<sup>18</sup>

“[E]ngaging as a principal for the trading account of [a] banking entity or [systemically important nonbank financial company] in any transaction to purchase or sell, or otherwise acquire or dispose of, any security, any derivative, any contract of sale of a commodity for future delivery, any option on any such security, derivative, or contract, or any other security or financial instrument that the appropriate Federal banking agencies, the [SEC], and the [CFTC] may, by rule ... determine.”

The financial institutions falling under this provision of the Act would have to spin off or sell their proprietary trading desks. The consequences of the implementation of this provision referred to as for the Volcker Rule for energy are at this point difficult to discern as the detailed rules are still being written (at the time of writing). The challenge consists in the ambiguity of the Act, which contains a long list of permitted activities, including:<sup>19</sup>

- ❑ market-making-related activity;
- ❑ risk-mitigating hedging,<sup>20</sup>
- ❑ underwriting;

- transactions on behalf of customers;
- transacting in government securities;
- certain insurance activity;
- investments in small business investment companies, public welfare investments and certain qualified rehabilitation expenditures under federal or state tax laws;
- certain offshore activities; and
- other activities that agencies determine would promote and protect the safety and soundness of banking entities and US financial stability.

These activities are carried out through the same actions that traders take to make directional bets (a trade is a trade), and the line separating them is very fuzzy. Market making requires keeping an inventory of different instruments and warehousing risk. One of the potential outcomes, however, is the migration of energy trading to hedge funds and merchant energy companies.

#### *Trading around assets*

Asset optimisation has many different definitions in the US energy industry and is often used as an excuse to participate in trading activity without incurring the risk of receiving complaints from the regulators and shareholders about engaging in speculation.<sup>21</sup> This is especially true of utilities. Many stakeholders of regulated utilities, including investors, community leaders and regulators, are often unhappy about any involvement of the regulated utilities in energy trading. Many investors are income-oriented and quite conservative and do not want to assume exposures related to activity they see, with some justification, as very risky. The regulators are afraid that potential losses in energy trading will undermine the ability of a public utility to maintain the assets and invest in generation and transmission in order to satisfy future load growth in the service area. This explains why many utilities engage in energy trading using the shield of asset optimisation or asset management.

It is difficult to find a precise definition of trading around assets in the existing literature. This term is used frequently in a rather careless manner. We define trading around assets as tactical decisions being taken every day in order to improve the financial results of a portfolio of physical assets, including associated hedges and supply

and sales contracts. This definition can be illustrated with an example of an integrated oil company operating a refinery supported by dedicated oil fields. The composition of the basket of refined products is a function of refinery design and configuration, and the chemical and physical characteristics of crude oil used as a feedstock (this will be discussed in detail in later chapters on crude and refined products). A company has the choice to rely on dedicated crude supply and sell the refined products in the local market. The alternative is to engage in a strategy of shipping crude oil to other refineries (better suited to processing the specific type of oil and possibly facing demand with different characteristics) and buying crude for its own refinery from a different source. The imbalance between the basket of outputs and local demand and supply commitments can be addressed by shipping some products to different markets and buying other products from different refiners. For example, a European refinery may ship gasoline to the US and import diesel fuel. This strategy is dictated by relatively greater demand for diesel fuel in Europe (where close to 50% of cars run on diesel) relative to, for example, the US. The market transactions around existing assets and contractual positions can be seen as an alternative to transforming the physical assets through costly and time-consuming investments.

Asset optimisation transactions are likely to have a relatively modest, but low-risk and recurring, contribution to the bottom line for two basic reasons:

- ❑ trading around assets as defined above is related to marginal adjustments to the existing positions and not sweeping and bold strategies – by definition, the profits cannot be excessive; and
- ❑ the execution of this strategy requires building an efficient trading operation – which is quite expensive, due to:
  - the high compensation required by skilled energy traders and risk managers;
  - investment related to the IT and communications systems, data feeds, regulatory and legal compliance; and
  - capital required to protect against market, credit and operational risks.

To offer an example, our experience, supported by inquiries with traders and risk management experts, provides some reference profit

---

targets related to trading around assets. For power trading, a consensus profit margin from trading around assets is about US\$0.25/MWh. To put this into perspective, an independent power producer with a generation fleet of 10,000 MW and a load factor of 70% (a somewhat generous assumption) can expect profits of around US\$15 million (this assumes a 7 days/24 hours operation). Assuming a more realistic operation (250 weekdays and profits realised only in the on-peak hours)<sup>22</sup> reduces the profits to US\$7 million. These numbers are before cost allocations. A few assumptions about the staffing and IT needs of a trading operation allow estimating the cost of providing this service. One can assume that the operation will require six traders at an average all-in cost of US\$500,000 and 20 support staff at US\$150,000 (on average). Space, computer systems and additional costs may be equal to US\$5–10 million.<sup>23</sup> This translates into a total cost in excess of US\$10 million.

Of course, this opinion has to be qualified in case of special circumstances. In some cases, trading around assets may be a euphemism for taking advantage of superior understanding of the market trends based on daily exposure to the supply chain and reliance on directional trades. It may also represent the ability to take advantage of local market power, due to control of the physical infrastructure around a trading hub, or the ability to profit by inserting a trading operation between a regulated utility and the wholesale market.

#### *Asset management*

Asset management is an operation under which a financial institution or an energy marketer takes over the control of assets of a producer or an end user of energy and commits to achieving certain financial objectives. The entity providing this service may be compensated by:

- a profit-sharing arrangement (above a certain negotiated threshold);
- a fee (with a possibility of a sliding scale related to financial results); or
- profits extracted from trading after meeting negotiated objectives related to physical operations – such obligations may include, for example, filling the natural storage facility at a certain cost by certain time; sometimes an asset manager pays the client for the privilege of managing their assets, hoping to meet contractual

requirements at a lower cost than that which is implied by the contract.

Asset management transactions can be very profitable from the point of view of an asset manager who runs an efficient operation, reaches a critical mass of assets and contracts under management, and receives additional benefits from participating in certain markets through order flow and competitive intelligence. On the other hand, such transactions often prove to be labour-intensive and based on an unrealistic assessment of trading and risk management skills by the providers of asset management services. On some occasions, providers are outwitted in negotiations by owners of assets who have superior understanding of local conditions and the characteristics of their assets. We will provide an example of asset management transactions in the chapter on the power markets: the so-called full requirements transaction.

## FUNDAMENTAL ANALYSIS

The philosophy of this book is based on the assumption that a successful energy markets professional has to combine skills in financial engineering with an understanding of the fundamental drivers of supply and demand. This section will therefore offer some basic observations on the importance of fundamental analysis to energy trading, while the details specific to each individual set of commodities will be presented in subsequent chapters. The discussion of financial engineering is deferred to Chapter 4 (the basics) and the follow-up book.

One of the dichotomies of commodity-related applied research is a gap between two distinct groups of analysts that support traders. The pure quants, working on risk management and valuation models, operate at a high level of sophistication and often use the most recent analytical techniques developed within academia. The so-called fundamental research often has different reporting lines and the level of mathematical complexity is often quite low. Fundamental research mostly revolves around the collection of data related to the drivers behind supply and demand for different commodities and dissemination of this information to the traders. However, there are exceptions to this general rule, which will be noted below. The quantitative models used by fundamental analysts

vary from simplistic (basic arithmetic) with an emphasis on display and visualisation of data, to extremely complex representations of physical reality (for example, power flow models).

The most important aspects of fundamental analysis, which will be discussed in detail for specific commodities across this book, include:

- ❑ analysis of the drivers behind supply and demand:
  - load (demand) prediction;
  - production forecasting;
  - analysis of the inventory levels;
  - the conditions of the physical infrastructure; and
  - weather conditions.
- ❑ modelling of commodity flows in the transportation network:
  - modelling of pipeline flows (natural gas); and
  - modelling of electricity flows in the transmission grid;
- ❑ regulatory developments affecting energy markets;
- ❑ strategies of major competitors; and
- ❑ geopolitical developments affecting energy markets.

The supply and demand analysis involves the identification of drivers behind the supply of different commodities and assessment of the potential price impact. This is the most critical effort for a commodity trading operation and a precondition to successful price forecasting. The data collected includes:<sup>24</sup>

- ❑ information about reserves of different hydrocarbons and the rate at which newly discovered resources can be brought to the market; equally important is the rate at which current production flows from existing reservoirs and mines are declining;
- ❑ data about existing and projected production and transportation capacity; actual flows of commodities over the transportation system;
- ❑ data about storage facilities and the level of inventories; in the markets in which official inventory levels are made available on a regular basis (natural gas, crude oil and certain refined products in the US), the research units are often responsible for forecasting periodic storage data releases;
- ❑ prediction of demand for a specific commodity at a global, national and local level over different time horizons;

- ❑ potential supply/demand disruptions (for example, outages of the power plants, pipeline shutdowns);
- ❑ strategies and position of competing traders in other companies; and
- ❑ regulatory and political developments that may affect production flows and market operations.

The collection of supply data (reserves, production capacity, inventories) is carried out on different geographical scales, from global to highly local levels. In each case, this is a task of enormous complexity and difficulty. On a global scale, one is facing the challenge of analysing statistical data compiled from different sources, using different definitions, multiple collection and aggregation techniques, and biases resulting from political and economic motivations.<sup>25</sup> On a local scale, it is difficult to obtain information at a low level of granularity (which may only be available to local players) about rigidities of the production and transportation system that can create pockets of special market conditions. The skills required to interpret the local data are not easily transferable to other locations and/or commodities. For example, a fundamental team supporting natural gas traders in the Western US has to keep an eye not only on production and natural gas pipeline-related data, but also on electricity production (some power producers are important buyers of natural gas), weather forecasts and, especially, the hydro conditions in the US Northwest and Western Canada.

One has to develop a very good understanding of the local infrastructure and the behaviour of local players. Such knowledge is typically not codified and can be accumulated only through years of immersion in a specific local market. Trading natural gas in Texas or the Northeastern states will require similar skills but acquisition of information will require a considerable amount of time. This explains why energy trading often tends to be labour-intensive, both in terms of trading and support talent, and local strategies are often not scalable across different markets. A key to creating a successful trading operation is often the ability to leverage location specific fundamental research techniques by making them more general and standardised across the organisation. Sharing models and analytical approaches across the organisation is a critical task for management. The trading desks often prefer to monopolise analytical resources, and this approach results sometimes in an excessively large and frag-

mented fundamental support group with no cross-pollination of ideas and sharing of models.<sup>26</sup>

One important aspect of fundamental research is the identification of constraints and dislocations in the production or distribution system and, once a disruption happens, forecasting how long it will last. Some dislocations happen due to weather patterns or recurring weather events, such as hurricanes. For example, certain natural gas pipelines swing from being partly empty to being fully utilised (due to seasonal variations of demand), with the actual transition point dependent on a configuration of many factors that change from year to year. A disruption following a weather event such as a hurricane can sometimes require a long time to resolve uncertainty around the scale and duration of the infrastructure destruction and resulting curtailment of commodity flows. A trading organisation that is more successful in collecting this information will have an obvious edge over its competition. Some catastrophic events, such as a natural gas pipeline explosion, are unpredictable – but once they take place, it is imperative to obtain information about the time required to fix the problem. The ability to improvise and allocate resources to the task of acquiring information can be critical. Catastrophic events in the commodity markets often create legendary traders because of the size of profits and losses.

## **COMPETITIVE INTELLIGENCE**

Competitive intelligence is a critical and highly specialised component of fundamental analysis. This activity can be defined as the use of legal sources and methods to gather information regarding:

- ❑ the state and operations of physical infrastructure in the energy industry;
- ❑ the policies, strategies, methods, processes and reputation of competing firms;
- ❑ the policies of the governments, government agencies and international organisations that impact energy markets;<sup>27</sup> and
- ❑ due diligence on, and screening of, potential counterparties to avoid entering into business relations with non-creditworthy, disreputable or criminally implicated partners; this also includes background checks on employees in critical positions with access to sensitive information, proprietary technology, controlling bank accounts, etc.

Most such information is available in the public domain and is simply dispersed across multiple documents, databases and human brains. Access to this information is often free, almost always legal<sup>28</sup> and frequently requires low levels of sophistication (but a lot of patience). The main challenge is collecting, organising and interpreting information in a cost-effective way. In other words, the challenge is not in collecting information but in connecting the dots. Because legal and ethical boundaries may be crossed, collecting information often should be left to professional groups that understand the complex rules of the game, not only in the US but also in other jurisdictions.

It is also necessary to draw the distinction between “information” and “intelligence.” “Information” refers broadly to any kind of data, whether or not it is necessarily relevant to one’s commercial needs. Such data is widely and readily available to all consumers, is not differentiated, and is therefore of limited commercial and competitive value. Distinct from information, “intelligence” is by nature targeted, differentiated and relevant to decision-making. It is intended to address a specific set of knowledge requirements that, if properly understood, would allow a decision-maker to proceed with certainty. Intelligence collection efforts identify data from relevant, credible sources in a manner directly responsive to commercial needs. Users of commercial intelligence do not make decisions with more information, but with better information than their peers.

A study conducted by the Central Intelligence Agency<sup>29</sup> concerning decision-making processes show that the human mind becomes increasingly frail when overloaded with information. As the mind works to incorporate and make sense of a growing dataset, particularly one that is circular and self-reinforcing, confidence in a “correct” judgement tends to increase in parallel. Conversely, however, researchers found that the subjects’ decision-making accuracy remained static or actually decreased in spite of the additional informational inputs. In other words, an executive presented with hundreds of spreadsheets reinforcing the same conclusion generally will commit more resources to a decision – including a wrong decision – than if they had thoroughly analysed a more limited, focused dataset. The over-build of gas-fired generation in the late 1990s and early 2000s is a great example of this phenomena (see the chapters on the US electricity markets).

---

Data collection about the conditions and operation of physical infrastructure is critical in markets where trading strategies evolve around logistical, operational and/or regulatory constraints and bottlenecks in commodity flows. The industry has a long tradition of monitoring such activities. In the early 20th century, oil scouts would travel to the oil fields on clear nights during the full moon to gauge the production rate by counting how many times per minute moonlight was reflected by the revolving parts of a rig.

More recently, competitive intelligence firms have been able to monitor a variety of factors impacting infrastructure – shipping channel blockages, regulatory permitting delays, severe weather, etc – to accurately predict disruptions in the supply of commodities. The same infrastructure-focused monitoring methodology holds true whether one is monitoring LNG, aluminium or timber.

Today, the collection of data can require extensive networks that process information in real time. For example, in the mid-1990s Enron developed a system for monitoring power flows on high-voltage transmission lines by measuring the intensity of the electromagnetic field. The objective of the data collection effort was not the transmission system *per se*, but the condition of the power plants. Specifically, it is possible to infer accurately information about the disruptions to a generation unit from perturbations of the power flows. A similar system was developed independently a few years later by Genscape and offered as a commercial service.<sup>30</sup>

Modern technology can be used to produce similar types of information using technologies such as satellite imaging, aerial photography, visual monitoring of plants and transportation hubs using closed-circuit TV cameras and the Internet. For example, planes may be used to overfly:

- ❑ coal-fired power plants to monitor the inventory of coal piled outside plants;
- ❑ storage tank farms to monitor the volume of oil, gasoline or heating oil available at a given location;<sup>31</sup>
- ❑ hydro reservoirs to monitor water levels; and
- ❑ construction status of power generation and transmission lines.

In the natural gas industry, several companies, such as BentekEnergy and Ventyx, monitor natural gas flows on the interstate pipeline

system. The information about nominations is cobbled together from the various websites of different pipeline operators and, after aggregation and validation for potential errors, is distributed to subscribers. The chapters on natural gas contain an extensive review of this technology and of the potential uses of data.

Genscape, BentekEnergy and similar service providers are reshaping the energy markets. In the past, physical market participants had a competitive advantage due to their access to non-price information about the state of the physical systems. Today, however, the ability to quickly compile and distribute large volumes of data creates an even playing field for companies with no physical assets, and sometimes even gives them a distinct advantage. A trading operation affiliated with a regulated pipeline cannot access, for regulatory reasons, the information flows produced in the regulated entity and has to rely on the same market information that is available to its competitors.

Data-gathering networks such as Genscape or BentekEnergy typically collect perishable data that has a short-term impact on the market. Such information must be acquired on a continuous basis and assessed against historical trends. To accomplish this, subscribers need to develop internal systems for processing, storing and interpreting the data. In most cases, the acquisition of data feeds from these service providers is a precondition for participating in the physical market. It is the ground floor of market knowledge and serves as a preemptive and defensive move designed to match the knowledge base of competing firms, not exceed them.

### **Event-based disruptions**

Many highly effective trading strategies revolve around event-based disruptions to physical infrastructure that have the potential for an immediate and significant market impact. Such events include, for example, pipeline explosions and outages,<sup>32</sup> permanent or long-term shutdowns of nuclear power plants, mine closures and acts of violence in unstable parts of the globe with significant energy production.<sup>33</sup> When such events happen, the market typically reacts in a highly volatile and often hectic manner, with trading decisions being taken under time pressure and often based on rumours, innuendo and intuition; access to reliable information in these circumstances is critical to avoiding costly mistakes and monetising market disruptions.

---

One example is provided by the developments following major hurricanes. In August and September 2005, Hurricanes Katrina and Rita inflicted unprecedented damage on the Gulf Coast. Apart from the horrendous human toll resulting from collapsed levies and the mismanagement of the mediation efforts by a number of federal and state agencies, the first physical casualty of the storms was the single most important US energy infrastructure system, responsible for producing just under 10 Bcf/day of natural gas, approximately 18% of domestic US production at the time, as well as the associated transportation, processing, refining and storage facilities.

After Hurricane Rita compounded the damage inflicted by Katrina, the energy trading community faced the difficult task of balancing two opposite trends. On one hand, the volumetric flows of natural gas and oil were curtailed across the entire supply chain. Damage was inflicted not only on the drilling rigs and underwater pipelines,<sup>34</sup> but also on the processing facilities on-shore, including natural gas processing plants and refineries. At the same time, demand was also partially destroyed due to the physical damage to industrial facilities and population displacement.

The final direction of prices depended on the relative impact of these two trends. A significant amount of information was provided on a daily basis from federal, state and municipal government agencies, but much of the crucial time-sensitive data was unavailable, seemingly contradictory or clouded in secrecy. Table 2.1 contains an example of the type of daily updates made publicly available by the Department of Energy regarding the conditions of the energy-related infrastructure in the Gulf after the hurricane. This information was available to all market participants and, therefore, offered no competitive advantage.

An example of creative information gathering is provided by the developments following Hurricane Katrina, when information about the state of physical infrastructure was at a premium. In our long career in merchant energy business, this was one of the most interesting and traumatic events. Katrina changed its course in the late afternoon on a Friday, after active trading stopped and the energy trading floors in Houston and New York City emptied. When the change of trajectory was announced around 5pm Central time, it was too late to re-trade the positions. Physical and emotional devastation was added to by the devastation of hedging and speculative portfo-

**Table 2.1** LA/MS/AL Gulf Coast Refiner Impacts

<b>Operator</b>	<b>Location</b>	<b>Capacity</b>	<b>Status</b>
XOM	Baton Rouge, LA	493,500 b/d	Reduced runs
Valero	Baton Rouge, LA	80,000 b/d	Reduced runs
Placid Oil	Port Allen, LA	48,500 b/d	Reduced runs
COP	Belle Chasse, LA	247,000 b/d	Shutdown
Marathon Oil	Garyville, LA	245,000 b/d	Temporary shutdown due to minor damage
Motiva (Shell)	Convent, LA	235,000 b/d	Shutdown
Motiva (Shell)	Norco, LA	226,500 b/d	Shutdown
Shell Cam	St Rose, LA	55,000 b/d	Shutdown
Chalmette	Chalmette, LA	187,200 b/d	Shutdown
Valero	St Charles, LA	185,000 b/d	Temporary shutdown 1 to 2 weeks due to water damage
Murphy Oil	Meraux, LA	120,000 b/d	Shutdown
Chevron Texaco*	Pascagoula, MS	325,000 b/d	Shutdown
Shell Cam	Saraland, AL	80,000 b/d	Unknown

\*Refineries with company owned and on-site power generators.

lios. Information was at a premium. One source of updates was information collected by the competitive intelligence firms. On the morning of August 30, 2005, immediately following the storm, TD International (TDI), a competitive intelligence firm based in Washington, DC and Houston, Texas, began a series of reconnaissance sorties over the impacted area using a chartered aircraft equipped with high-resolution cameras and sophisticated communications equipment. The company based its operations from a private airfield in Panama City, Florida, many hundreds of miles away from the storm's path.

Multiple reconnaissance sorties had been strategically planned over the previous five days, while the storm's path was still somewhat uncertain. Targets were prioritised in the order of major production platforms, gathering facilities, storage locations, onshore support infrastructure and refineries. TDI's aircraft was on site, in many instances, long before the asset owners, first responders or the Federal Agency Management Service (FEMA). The information collected in these sorties was made available to TDI's clients on a practically real-time basis. The intelligence proved market-moving; high-resolution images of the devastation showed a "ground truth" that differed sometimes dramatically from reports provided by media and corporations struggling to understand and/or "spin" what actually happened.

A review of the photographs taken during these aerial reconnaissance sorties gave considerable advantage to the recipients. The damage to a specific platform could be translated into days and weeks of reduced flows and this, in turn, led to an accurate assessment of the storm's true market impact well ahead of most market participants and media outlets. This effort required the collective judgement of specialists in many different fields, including fundamental analysts, engineers and traders. The cost of this effort, however, was minuscule compared to the benefit of having accurate, differentiated, unique intelligence on the storm's impact.

Other examples of the uses of competitive intelligence are too detailed to summarise. What is important to understand is that the information-collection effort described here is a critical component of the fundamental analysis effort, and for capitalising on opportunity while managing commercial risk. Used effectively, intelligence becomes just as critical and valuable a resource as the other more traditional inputs that firms seek to maximise, such as people, capital and technology.

## CONCLUSIONS

Energy markets are not only very complex but also very tightly coupled with the rest of the global economy. I can hardly think of any major economic, financial or geopolitical development that will not impact energy markets, sometimes in the most unexpected ways. This is why designing an efficient trading operation, capable of instantaneously processing huge volumes of information and rapidly reacting to market shocks is so critical to successful trading. A well-honed organisation requires the seamless integration of front, middle and back offices and the effective collaboration of professionals with different skills and backgrounds. Trading energy is a very competitive business with huge overheads (given the complexity of the underlying business), and keeping transaction costs under control requires well-designed software platforms.

It is obvious that the importance of fundamental information in energy trading gives an edge to physical players who are handling physical commodity flows and benefit from the information that can be acquired in this process. A financial institution may offset this informational handicap through creative solutions in information gathering and the ability to automate data processing. The

subsequent chapters will provide more detail regarding specific commodities.

- 1 Justin Kruger, David Dunning, 1999, "Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments," *Journal of Personality and Social Psychology*, 77(6), pp 121–1134.
- 2 The Emerging Issues Task Force of the Financial Accounting Standards Board (FASB). The referenced issue, "Accounting for contracts involved in energy trading and risk management activities," rescinded in 2003 (EITF Issue No. 02-03), applied to accounting for energy-related transactions with a physical component.
- 3 In most cases, the profits from energy trading are not reported separately. An inadvertent disclosure of energy trading profits offers a rare glimpse into the scale of such operations. "The midday email, inadvertently sent Friday to media organizations, contained spreadsheets, tables and charts that gave a breakdown of Chevron's profits, losses and exposures trading crude and refined products such as gasoline. They offered an unusually detailed look at how one of the world's largest energy producers was able to leverage its size and geographic reach to profit to the tune of \$360 million this year to date through trading activities." See: Brian Baskin and Ben Lefebvre, 2011, "Chevron's email 'oops' reveals energy giant's sway over markets," *Wall Street Journal*, July 16.
- 4 Enron followed both strategies by entering into a joint venture with Bankers Trust, and then, when the relationship soured, by hiring traders and risk managers directly from Wall Street firms. Many successful traders started as natural gas schedulers and developed into physical traders, benefitting from excellent knowledge of the pipeline grid.
- 5 Anybody who read Garrett Hardin, 1968, "The tragedy of the commons," *Science*, (162)3859, December 13, pp 1243–48, will not be surprised by this conclusion. This problem is often discussed in the context of the overgrazing of common pastures. The Swiss peasants found a solution to this problem 800 years ago. One can bring cows to the Alpine meadows in the summer, only if one carried them over the winter (see, <http://www.openmarket.org/2009/10/13/2002-economics-nobel-prize-winner-vernon-smith-on-2009-winner-elchinostrom/>). The trading desks should be ready to invest in the professional growth and well-being of the quants through good and bad times, and not treat them as an expendable resource (ie, hire and fire them at will).
- 6 One could argue that this is an issue of minor importance, as it is a matter of the distribution of rewards inside a trading organisation, not the issue related to overall profitability. One cannot ignore, however, that such internal front-running has a long-term corrosive impact on the ability to generate profitable transactions. The most promising originators choose to migrate to other firms, with better governance.
- 7 A detailed discussion of the principles underlying trading policy is contained in the chapter on risk management in our next book.
- 8 "We shall break up the position into swaps and options, and hedge the basis risk with options. The volumetric risk will be managed with weather derivatives. We shall cover counterparty risk with credit default swaps. We can manage all the component risks through a seamless matrix of dynamically adjusted positions." Any executive in energy trading should fire such an employee on the spot.
- 9 Unfortunately, wishful thinking and the drive to book big, complex transactions at any cost leads to many abuses. As Satyajit Das put it somewhat cynically: "There are salespeople – they lie to clients. Traders lie to sales and to risk managers. Risk managers? They lie to people who run the place – correction, think they run the place. The people who run the place lie to shareholders and regulators." Satyajit Das, 2006, *Traders, Guns & Money: Knowns and Unknowns in the Dazzling World of Derivatives* (Harlow, England: Financial Times Prentice Hall), p 54.

- 10 The author was involved in structured transactions (in the early days of market development) that would require about a year to close a position by transacting on Nymex and in the over-the-counter (OTC) markets.
- 11 This topic will be covered in greater depth in the next book from the author.
- 12 The scope of proprietary trading by the US financial institutions will be limited following adoption of the Dodd–Frank bill.
- 13 In the commodity markets, we see from time to time two trading organisations with mirror positions. Such situations are often compared to the Texas chicken game: two drivers on a collision course, with one of them swerving at the last moment, earning the reputation of a coward. Older readers undoubtedly remember a similar competition in the 1955 movie *Rebel Without a Cause*. This type of competition was used as inspiration for game-theoretic models applied to a wide range of problems, including mutually assured destruction in the case of a nuclear war (see Bertrand Russell, 1959, *Common Sense and Nuclear Warfare* (London: Allen & Unwin)).
- 14 One could argue that there always will be a supply of new investors and excessively confident young traders willing to take the risk of confronting established informed traders. It is said that the second marriage is a triumph of optimism over experience; the investment game is not much different.
- 15 Historically, the same objective was accomplished by both sides (a producer and an end-user of energy) negotiating risk-sharing agreements. Excessive profits received by one side in a given time period would be transferred to a special account and shared at some point in time. This was the standard way of managing risk prior to the development of liquid forward markets. The speculators make this process much more efficient and flexible.
- 16 One barrel of crude oil contains about 5.8 MMBtus. Oil is not trading, however, at heat parity to natural gas (ie, roughly 6 MMBtus of natural gas trading at the price of one barrel of crude). This does not stop some traders from proposing strategies based on the price convergence of natural gas and oil in terms of heat parity.
- 17 Winston Churchill remarked once that the best argument against democracy is a five-minute conversation with an average voter. One could paraphrase his words and say that the best argument against speculation is listening for five minutes to some fund managers on CNBC.
- 18 12 USC § 1851(h)(4).
- 19 Financial Stability Oversight Council, 2011, "Study & recommendations on prohibitions on proprietary trading & certain relationships with hedge funds & private equity funds," Completed pursuant to section 619 of the Dodd–Frank Wall Street Reform and Consumer Protection Act, January.
- 20 It is debated whether the definition of hedging contained in the proposed implementation rules (at the time of writing) effectively watered down the language of the Act to make prohibition of proprietary trading ineffective (especially the provisions allowing for *portfolio hedging*). See Edward Wyatt, 2012, "JPMorgan sought loophole on risky trading," *The New York Times*, May 12.
- 21 It is always difficult to draw a well-defined line between aggressive trading around existing financial and physical positions and speculation. One can argue, however, that if one sees an obvious case of speculation, one will recognise it. One could agree that if – hypothetically – a CEO of a large exploration and production company personally makes daily trading decisions, it is closer to speculation than not. In the more conservative days, their job would be to poke holes in the ground in the right places and at a low cost.
- 22 The market for off-peak electricity remains relatively anaemic in the US: the surplus of cheap generation suppresses profit opportunities.
- 23 We assume roughly a 50–50 split between personnel and non-personnel cost.
- 24 Certain types of information are collected through activities known as competitive intelligence, covered later in the chapter.
- 25 One obvious example is information about oil reserves that is tainted by efforts to inflate

reserves for political and nationalistic reasons or to qualify for a higher production quota in an organisation such as the Organization of the Petroleum Exporting Countries (OPEC).

- 26 This is easier said than done, because in big organisations some desks compete against each other. Quite often, the volumes of natural gas turned over by a power desk exceed positions of the natural gas desk. There are many reasons why this happens, including hedging transactions and also the limited ability to trade power, given the relative immaturity of the electricity market.
- 27 A thriller could be written to describe information-gathering efforts that have accompanied many OPEC meetings. We shall leave this task to a person with better command of the English language.
- 28 Any organisation should avoid like the plague the temptation to collect information using illegal means. It is not only improper but also irrational. Given that there so much information in the public domain, breaking the law is foolish.
- 29 Richards J. Heuer, Jr, 1999, "The psychology of intelligence analysis," Central Intelligence Agency.
- 30 This system will be covered in more detail in the chapter on the power markets.
- 31 Most modern storage tank have floating roofs (to avoid accumulation of vapours in the empty space above the liquid) and one can calculate easily the available volume from the level of the top.
- 32 Examples of such events are the El Paso pipeline explosion near Carlsbad, New Mexico, on August 19, 2000, and the April 2008 outage of the pipeline connected to the Anadarko Independence Hub in the Gulf of Mexico, which sent natural gas futures soaring to 27-month highs.
- 33 One example is the impact of the news about the Movement for Emancipation of the Niger Delta (MEND) activities roiling the oil markets from time to time.
- 34 Many pipelines were affected by boats dragging anchors along the bottom of the Gulf of Mexico.



## *Weather Information in Energy Trading*

Some of our academic friends who read drafts of this book were somewhat flabbergasted by the amount of space devoted to weather information and the depth of detail provided. One has to spend some time on the energy trading floors to understand the level of interest, bordering on obsession, that traders have in weather forecasts. We had the privilege to work with a group of very talented professionals at Enron who produced daily weather reports, and had extremely appreciative consumers of this information on the trading floor. The traders developed an excellent understanding of the weather modelling process, sometimes shocking climate scientists invited to deliver lectures with the sophistication and detailed nature of their questions. They would download weather forecasts early in the morning to their PDAs and analyse these reports on tiny screens while driving to work (trading is an occupation for risk takers). It is easy to understand this level of interest: weather is a critical driver of most energy prices on a daily basis, as will be shown in this chapter.

We will start with a review of the channels of transmission from weather conditions to market prices. The channels of transmission in this context are defined as the totality of market processes and interactions between producers, end users and traders, through which weather conditions ultimately find reflection in the market prices of energy commodities. The next topic will be a review of the weather forecasting models used by energy trading organisations and how information generated by these models is used in practice. Weather anomalies are predictable weather patterns, usually unfolding over long time periods (measured in months, years and sometimes decades). The best-known examples of such anomalies are the famous El Niño and La Niña events. Anomalies are important because they help occasionally (if they occur on a sufficiently large

scale) to form longer-term seasonal weather outlooks. The chapter will end with a discussion of hurricane events and the reasons why they are so important to energy trading.

### **WEATHER AND COMMODITY MARKETS: CHANNELS OF TRANSMISSION**

The progress of civilisation has been associated with a reduction in the impact of weather on 19th century economic activity. In the 19th century, one of the popular explanations of the mechanism behind the business cycle were theories linking fluctuations in agricultural output, driven by anything from weather cyclicalities through sun spots, to the changes in the level of economic activity. In the early 1800s, even the short-term monetary policy of the Bank of England was weather dependent.<sup>1</sup> A dial in the court room linked to a weather vane on the bank's roof alerted staff to changes in wind direction. Winds blowing from the east would bring the cargoes into the port of London and increase the demand for cash balances from the traders buying the goods unloaded in the docks. Western winds would have the opposite effect: the supply of goods would drop and the demand for money would decrease, calling for an adjustment in monetary policy. Technological progress reduced the dependence of national economies on weather, except for the commodity markets, where the weather impact is still of primary importance in the short and medium term.<sup>2</sup> It is not a coincidence that most morning traders' meetings in the energy business start with a weather report. What follows here is a very basic introduction to the weather-related information an energy trading professional should follow on a daily basis.

Weather information has always been critical to commodity trading. One of the best books in the US literature about speculation describes the extent to which a grains trader depended on weather reports and how sophisticated and data-intensive their analytical efforts were more than a 100 years ago:

The spring had been backward, cold, bitter, inhospitable, and Jadwin began to suspect that the wheat crop of his native country, that for so long had been generous, and of excellent quality, was now to prove – it seemed quite possible – scant and of poor condition. He began to watch the weather, and to keep an eye upon the reports from the little county seats and “centres” in the winter wheat States.  
[...]

From Keokuk, in Iowa, came the news that winter wheat was suffering from want of moisture. Benedict, Yates' Centre, and Douglass, in southeastern Kansas, sent in reports of dry, windy weather that was killing the young grain in every direction, and the same conditions seemed to prevail in the central counties. In Illinois, from Quincy and Waterloo in the west, and from Ridgway in the south, reports came steadily to hand of freezing weather and bitter winds. [...]<sup>3</sup>

Weather affects the commodity markets through an impact on supply and demand. The supply impact in the agricultural markets unfolds usually over longer time periods, except for certain events – such as hail or freeze affecting critical production areas.<sup>4</sup> In the case of energy markets, the supply impact is often related to traumatic events such as hurricanes and the resulting destruction of physical infrastructure. In most cases, the supply impact is highly localised and temporary. This does not mean that such relatively isolated impacts can be ignored: sometimes even a local disruption leads to highly profitable trading opportunities.<sup>5</sup> Trading organisations accumulate knowledge about such events through (sometimes bitter) experience. A few examples are listed below, although this is by no means exhaustive.

- ❑ Fog in a specific sea-lane that handles intensive energy-related ship traffic can cause short-lived supply disruption (the Houston Ship Channel is an obvious example).
- ❑ High sea levels off the coast of California may push kelp into the water intakes of nuclear power plants, reducing electricity production while a clean-up operation takes place.
- ❑ Volcano eruptions may disrupt historical weather patterns by injecting dust and ash into the stratosphere.<sup>6</sup>
- ❑ A drop in water levels in rivers during a drought (below the level of water-intake pipes) or a significant increase in the temperature of water in the cooling ponds of coal power plants can hamper electricity production.
- ❑ Water pollution may have a similar impact on the refineries or power plants (the 2010 oil spill in the Gulf of Mexico raised fears of disruption in the operations of industrial plants along the coast that were dependent on water for production processes).
- ❑ Protracted drought can reduce the availability of water required

for hydraulic fracturing (see the chapters on natural gas for more details related to water management in the production of oil and natural gas).

- ❑ Floods or low water levels can interfere with barge movement and impede deliveries of coal, gasoline and crude oil on some important waterways (Mississippi, Rhine). These disruptions may also affect other markets. Shortages of coal may force power generators to run gas-fired power plants more intensively at nights and on weekends.

Supporting an energy trading team takes more than good weather forecasting skills. A meteorologist has to understand how specific weather conditions affect energy markets, given the market conditions and the state of the energy infrastructure. The same weather forecast may have dramatically different impact on the markets, depending on the circumstances. Weather forecasters employed by energy trading operations learn from experience which weather patterns to look for and alert the traders about. It is important for any energy trading operation to develop a weather team with a good understanding of the fundamentals of the energy industry. Even the best weather forecaster who does not understand the complexity of the transmission channels from weather to market prices is unlikely to provide effective support for energy trading.

In some cases, the supply impact at a specific location materialises, due to weather-related demand change, in a different and sometimes very distant market. For example, a drought in France and Spain in the summer of 2005 reduced the level of hydropower production (and also nuclear power due to the low levels of many rivers in France) and increased demand for liquefied natural gas as a fuel for thermal power plants. In turn, this resulted in the diversion of LNG flows from the US market. This takes us on to the importance of weather impact on demand. One can expect that growing globalisation and integration of energy markets will create a need for monitoring weather conditions worldwide, even for local trading.

In the case of the US, the weather impact on demand primarily affects the electricity, heating oil and natural gas markets. In the US and other markets electricity demand tends to increase during the summer due to air conditioning load. In Europe, the weather-related impact is more pronounced in winter (due to the heating load), but

the proliferation of air conditioning is likely to change this seasonal demand pattern in the future. In the case of natural gas markets in the US and Europe, a spike in demand in winter is related to the heating load. In the US markets, natural gas demand spikes in the summer during heat waves due to demand from peaking gas-fired power plants that are critical to satisfying the air-conditioning load. The demand for heating oil is strongly affected by winter temperatures in the US, Europe and Asia.

The responsibilities of the weather team in most trading organisations include:

- ❑ acquisition, interpretation and dissemination of short- and medium-term forecasts;
- ❑ development of the seasonal outlook; and
- ❑ creation and support of a historical weather information database for demand forecasting and pricing weather derivatives.

## **WEATHER FORECASTING MODELS**

Atmospheric models are a key source of information for any successful weather forecast. These models are very complex and require powerful computers and huge databases, which is why they are usually designed and operated by government and quasi-government entities. The public agencies that run the weather models include the National Center for Environmental Prediction in the US, Environment Canada and The European Center for Medium-Range Weather Forecasts (ECMWF) and the UK Met Office. The results of the models are made available to weather consulting firms and sometimes the general public. A weather team typically receives the forecast from a number of weather consulting firms and its responsibility is limited to the interpretation, packaging and dissemination of information. Given that practically everybody in this field has access to the same datastreams and models, one can compete by acquiring and processing the information in a more effective way. This includes delivering the information to traders faster than competitors, and also making judgement calls when different models produce conflicting forecasts or when the consulting firms offer different forecasts. One of the critical in-house weather skills is an ability to assess the relative skills of different outside forecasters. Given that weather forecasting is a combination of science and art,

with a very important component of experience, meteorologists develop skills that vary across seasons and locations. The ability to assess relative skills of various outside consultants is critical to the success of a trading operation. This skill can be taken to the next level, and in some cases in-house meteorologists try to anticipate the market-moving forecasts that are made available according to a fixed time schedule.

Some weather groups acquire access directly to the output of a few existing large-scale numerical models for weather prediction in order to beat the competition. This means that the task of interpreting model output is transferred from outside consultants to the in-house weather group or is supplemented/modified internally.

The industry uses several key numerical weather prediction models, with models gaining or losing popularity depending on their recent performance. Existing models may be classified either as global, or models covering the entire globe, or regional, covering certain parts of the world. Weather forecasting models are imperfect for a number of reasons. First, any numerical model has to be initialised with actual data collected through international cooperation from many different sources, including weather satellites, planes, ships, buoys and land weather stations. However, the coverage of different regions is unequal, with relative data voids extending across some continents and oceans. The second fundamental problem is related to the non-linearity of weather and the chaotic nature of models describing complex dynamic systems. This property, a sensitivity to initial conditions, was observed for the first time by Edward Lorentz in 1963 and manifests itself in large differences in model outputs caused by even very small changes in the initial conditions. This, in combination with uneven and imperfect initialisation, reduces the quality of the forecast and can cause, on some occasions, significant divergence of forecasts produced by different models. In such cases, a meteorologist has to make a call, based on their experience and intuition – this is when an investment in a weather forecasting group bears the most fruit.

One technique used widely to assess the sensitivity of weather forecasts to initial conditions is based on running the same forecast many times, with slightly modified initial conditions.<sup>7</sup> This approach is a form of Monte Carlo analysis, consisting of repetitive runs of the same model with different initial assumptions, based on small

perturbations of the inputs. If the plots of different runs, known as ensemble plots, are converging to the same forecast, it increases the confidence in the quality of the forecast. If the plots look like a messy bowl of spaghetti, the reliability of a forecast is therefore questionable. The ensemble plots may be used to produce probabilistic statements about future weather conditions by counting the number of simulation runs producing the same (or similar) outcomes.

There are several numerical models that receive the most attention in the energy community.

*Global Forecast System (GFS)*<sup>8</sup> is run by the National Center for Environmental Prediction (NCEP) – a division of National Oceanic and Atmospheric Administration (NOAA) – four times a day, with diminishing spatial and temporal resolution (ie, bigger and bigger space and time blocks are used as the horizon increases) for up to 16 days. The spatial resolution is either 35 (zero to seven days) or 70 (between seven and 16 days) kilometres; the atmosphere is divided into 64 vertical layers. The model produces a forecast for every third hour for the first seven days, switching to 12 hours for the longer horizon. The consensus is that the model provides useful trading information five to seven days into the future. The results of this model are available at no charge from NCEP (as the development and maintenance cost is borne by the taxpayers) and can be obtained through consulting firms or directly from government agencies.

*The Global Environmental Multiscale Model (GEM)* was developed as a joint venture of the Recherche en Prévision Numérique (RPN), Meteorological Research Branch (MRB), and the Canadian Meteorological Centre (CMC). The model is run as a data assimilation system and global/regional short/medium term forecasting system. The model is very popular among energy trading professionals and is among the most closely watched information sources.

*European Models.* Two widely used European models include the UKMET model, developed in the UK by the Meteorological Office. The other model is the ECMWF (European Center for Medium Range Weather Forecasts), referred to as The European Model, which runs out to ten days<sup>9</sup> and provides 10-day 'Deterministic' forecasts of Mean Sea Level Pressure, of Wind Speed and Temperature at low levels, and of the Height of the 500-hPa (hectopascals – VK)

isobaric surface.<sup>10</sup> The ECMWF is supported by 28 countries and is located in England.<sup>11</sup> Information about these and other models offered by NOAA, NCEP and other institutions is available from the links in this end note.<sup>12</sup>

These models are often referred to in the industry jargon as the *American model* (GFS), *Canadian model* and *European model*, and these terms are often heard during morning weather presentations at the energy trading operations. From time to time, one of the models becomes more popular than the others and dominates the thinking of traders. Understanding this is as important as knowing what is the technical analysis tool *du jour*.

Given the preponderance of model-based weather forecasts and their availability to the public, one can ask whether the potential contribution of weather analysts working on energy trading floors and their cost is justified. The answer is “yes”, but the range of solutions used in practice is very wide.

The most sophisticated energy trading organisations evolve towards running proprietary models developed in-house or acquired from research firms. There are persistent rumours in the industry about some promising new quantitative techniques acquired on the basis of exclusivity by some hedge funds. Most likely, such proprietary systems process the outputs of multiple prediction models and come up with a synthetic aggregate forecast. It is doubtful that any private firm, even with large resources, could run an equivalent of a European or Canadian model, given IT processing requirements and the need to acquire massive datasets for model initialisation.

Most weather forecasting groups focus on the value chain between model outputs and the dissemination of final conclusions to the traders. The objective is to reduce the time required for the review and interpretation of the model runs, especially if different models point in different directions. A difference of a few minutes, or even a few seconds, could be critical.<sup>13</sup>

The contribution of weather forecasters to energy trading is typically based on two types of reports.

- ❑ Daily updates of short- and medium-term weather forecasts available from a number of models (as mentioned above).
- ❑ Discussion of seasonal winter and summer outlooks, which are influenced to a great extent by different types of “anomalies,”

defined as a departure of current weather patterns from the “climatology,” the average or typical state of climate based on many years of historical observations. The weather anomalies include patterns such as the Pacific Decadal Oscillation (PDO), El Niño Southern Oscillation (ENSO), North Atlantic Oscillation (NAO) and the Madden–Julian Oscillation (MJO).

## WEATHER ANOMALIES

Attention paid to different types of anomalies surprises any outside visitor to an energy trading floor. When we started our career in energy trading in Texas in 1992, the mention of El Niño would produce blank stares from most people, or at best would invoke the memories of Christmas decorations in a Hispanic barrio of Houston. Today, one can hardly attend a morning meeting of traders without being exposed to a barrage of acronyms such as PDO and NAO, which will be explained shortly. A word of warning is required at this point. The challenge related to reliance on weather anomalies in developing a seasonal outlook is that many different patterns may overlap and interact at the same time.<sup>14</sup> The most important anomalies that have become household words are La Niña and El Niño, which are covered at the end of this section.

The NAO is an anomaly discovered in the 1920 by Sir Gilbert Walker (although there is some evidence that the seafaring Viking were aware of the phenomenon). NAO is defined as an index of fluctuations in the difference of atmospheric pressure at sea elevation level (sea level pressures, or SLPs) between Iceland and the Azores.<sup>15</sup> As explained by James W. Hurrel and his co-authors:

[O]ver the middle and high latitudes of the Northern Hemisphere (NH), especially during the cold season months (November–April), the most prominent and recurrent pattern of atmospheric variability is the North Atlantic Oscillation (NAO). The NAO refers to a redistribution of atmospheric mass between the Arctic and the subtropical Atlantic, and swings from one phase to another produce large changes in the mean wind speed and direction over the Atlantic, the heat and moisture transport between the Atlantic and the neighboring continents, and the intensity and number of storms, their paths, and their weather. Agricultural harvests, water management, energy supply and demand and yields from fisheries, among other things, are directly affected by NAO.<sup>16</sup>

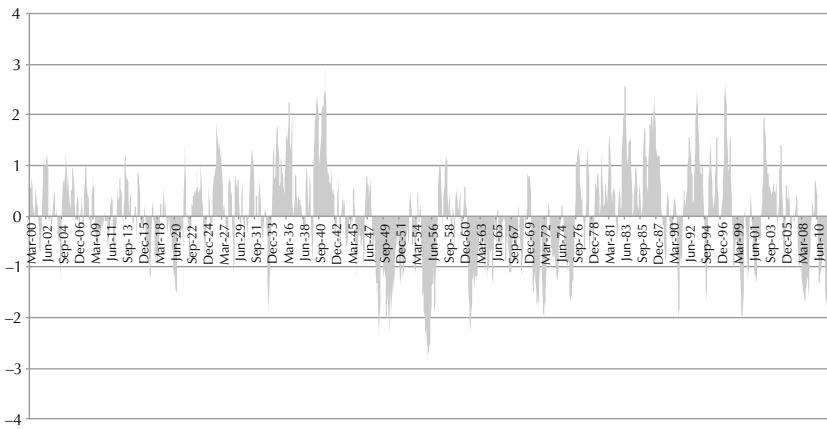
Positive NAO is associated with inflows of relatively warm and moist air into Europe, resulting in the warming of most of North America

and the cooling of North Africa and the Middle East. The temperatures over the northwest Atlantic tend to decrease. High NAO index winters are associated with drier conditions over the Canadian Arctic, Greenland, central and southern Europe, the Mediterranean and parts of the Middle East. More precipitation may be expected from Iceland through Scandinavia, with obvious consequences for the supply of hydropower in Norway and the rest of Scandinavia. A low NAO index is associated with reduced precipitation in Scandinavia. For example, in the summer of 1995, the summer inflow peak was sufficient to generate 10,000 GWh and 5,000 GWh per week in Norway and Sweden, respectively. In 1996, with a highly negative level of NAO (between December 1995 and March 1996), the corresponding levels were only 7,000 and 4,500 GWh per week.<sup>17</sup>

The second anomaly that has captured the attention of the industry in recent years is the MJO,<sup>18</sup> a disturbance that takes place in the tropics.<sup>19</sup> Its main characteristic is that it has a relatively short term (30–40 to 50–60 days), involving variations in wind, sea surface temperature (SST), cloudiness and rainfall. The reason why the MJO is important is that there is mounting evidence that it contributes to (and may actually be the trigger of) the ENSO, increasing the speed of development and intensity of both La Niña and El Niño. The US impacts of the MJO include higher and more frequent precipitation along the West Coast and colder weather during winter. As explained by Jon Gottschalck and his co-authors (see footnote 19):

The MJO is characterized by eastward propagation of regions of enhanced and suppressed tropical rainfall, primarily over the Indian and Pacific Oceans. The anomalous rainfall is often first evident over the Indian Ocean, and remains apparent as it propagates eastward over the very warm waters of the western and central tropical Pacific.

One of the best-known phenomena associated with the MJO is the so-called Pineapple Express. As explained by the Climate Prediction Center, “[d]eep low pressure located near the Pacific Northwest coast can bring up to several days of heavy rain and possible flooding. These events are often referred to as “pineapple express” events, so named because a significant amount of the deep tropical moisture traverses the Hawaiian Islands on its way towards western North America.”<sup>20</sup> Power traders in the US Northwest should familiarise themselves with this anomaly.

**Figure 3.1** PDO index, three-month average by month (1900–April 2012)

Source: <http://jisao.washington.edu/pdo/PDO.latest>

The most important impact of the MJO, in addition to West Coast rainfall, from the point of view of energy trading is the modulation of tropical cyclone activity. A good short summary of the MJO and a number of critical references are available in a very concise paper by Dennis L. Hartmann and Harry H. Hendon.<sup>21</sup>

The PDO is a pattern identified in 1996 by Steven Hare, who was studying the relationship between Alaska salmon production and the Pacific climate. PDO cycles are of a long-term nature and persist in the same phase for 20–30 years (see Figure 3.1).<sup>22</sup> The warm (or positive) phase of the PDO is associated with above-average water temperatures along the western coast of North America and the central Pacific Ocean. The cold phase of the PDO is associated with the opposite pattern. A very good description of the PDO is available in an article by Nathan Mantua, who provides a summary of PDO-related weather patterns for the North American continent.<sup>23</sup>

The PDO is of particular interest to electricity traders in the Western US and Canada: as mentioned, the positive (warm) phase of PDO leads to below-average precipitations in the Pacific Northwest – with obvious consequences for electricity output from the hydropower plants in this region. This is an illustration of information that can be obtained from following this anomaly. A detailed discussion of all these anomalies would be prohibitively long: one must rely on a friendly weather forecaster for continuous updates.

**Table 3.1** Summary of Pacific and North American climate anomalies associated with extreme phases of the PDO

Climate anomalies	Warm Phase PDO	Cool Phase PDO
Ocean surface temperatures in the northeastern and tropical Pacific	Above average	Below average
October–March northwestern North American air temperatures	Above average	Below average
October–March Southeastern US air temperatures	Below average	Above average
October–March southern US/Northern Mexico precipitation	Above average	Below average
October–March Northwestern North America and Great Lakes precipitation	Below average	Above average
Northwestern North American spring time snow pack and water year (October–September) stream flow	Below average	Above average
Winter and spring time flood risk in the Pacific Northwest	Below average	Above average

Source: Nathan Mantua, "The Pacific Decadal Oscillation," ([http://www.atmos.washington.edu/~mantua/REPORTS/PDO/PDO\\_egec.htm](http://www.atmos.washington.edu/~mantua/REPORTS/PDO/PDO_egec.htm))

The ENSO anomaly is the mother of all global climate patterns, a household term and a topic of daily discussion on energy trading floors. The El Niño/La Niña (baby boy or baby girl, and can also be translated as baby Christ) phenomenon was given its name by Peruvian fishermen who observed that, sometimes around November/December, the temperatures of coastal waters would increase, with the daily catch of anchovies decreasing dramatically, leaving them with plenty of time to fix their boats and nets. Sometimes such conditions would persist into late spring. These developments coincided with the Christmas season, which explains the other name of El Niño. The mechanism causing this phenomenon and its impacts on global climate is fairly well understood, although it is not well known what causes the turning point in the cycle.

Under normal conditions, the easterly trade winds (the winds blowing from the coast of the Americas in the western direction) push ocean water westward, causing the waters in the western Pacific to raise about 50 centimetres above the level of water in the eastern Pacific. This, in turn, allows cold waters that are rich in nutrients to rise to the surface off the west coast of South America (known as upwelling). This explains the abundance of fish. A reader is encouraged to examine the graph<sup>24</sup> illustrating this mechanism, which

shows a cold finger extending from Peru across the Pacific toward Indonesia. The so-called thermocline, separating the layers of cold and warm water, is deflected and rises close to the surface in the east and is depressed deep below the surface in the west. During the El Niño phase, the trade winds weaken and the process is reversed. The sea level drops in the west and increases in the east Pacific; the thermocline in the east is pushed deeper below the surface and the supply of the nutrients to the surface is cut off. This is why the Peruvian fishermen become idle during the El Niño phase and the economies of the Latin American countries dependent on the fishing industry suffer.

La Niña can be seen as a reversal of El Niño, with the global weather system returning to normal and overshooting in the process. Given that it is seen as the opposite of El Niño (child), it also known as "old man" (el viejo in Spanish). As can be seen from the history of the Southern Oscillation Index (SOI) and sea surface anomalies, it is associated with colder sea surface temperatures and positive SOI.

Climate scientists have developed special indexes that allow for the measurement of the intensity of El Niño. Two of the most popular indexes are SOI and SST average anomalies. SOI is calculated in a number of different ways. One formulation available from the Australian Bureau of Meteorology<sup>25</sup> is given by:

$$SOI = 10 \times \frac{Pdiff - Pdiffav}{SD(Pdiff)} \quad (3.1)$$

where:

$Pdiff$  = (average Tahiti mean sea 1level pressure (MSLP) for the month) – (average Darwin, Australia, MSLP for the month)

$Pdiffav$  = long-term average of  $Pdiff$  for the month in question

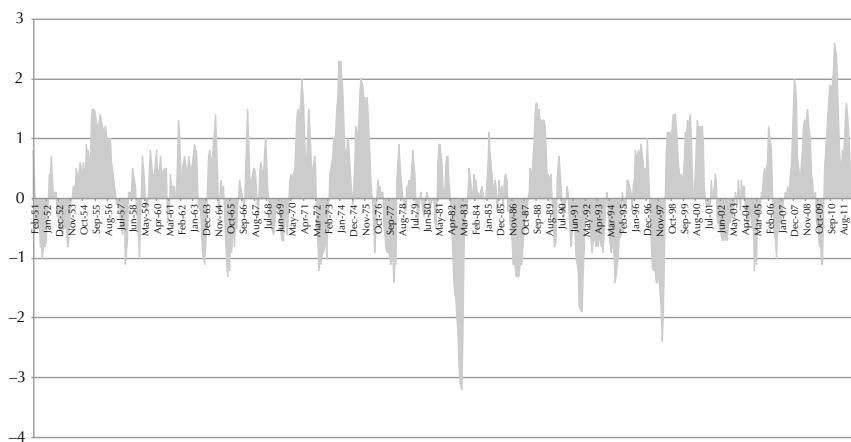
$SD(Pdiff)$  = long term standard deviation of  $Pdiff$  for the month in question

The negative value of the index corresponds to periods when SLP in the French Polynesia area is low relative to Darwin and the trade winds are weak. This corresponds to the El Niño period. Positive values are associated with La Niña (described above). Another index is the averaged sea surface temperature anomaly given by 5° N, to 5° S, 90° W to 150° W,<sup>26</sup> known as the El Niño 3.4 region. Large positive values of this index define El Niño conditions. Figure 3.2 shows the history of SOI. An updated version of these graphs should end up

every month on the desks of all energy traders and energy risk managers across the globe.

The ENSO influences weather in many parts of the world, but from the point of view of an energy trader one important influence is related to the development of hurricanes in the Atlantic. El Niño causes intensification of the subtropical jet stream which shears the tops of the Atlantic tropical storms before they have an opportunity to organise or grow stronger.<sup>27</sup> Other impacts include a drier-than-normal autumn and winter in the US Pacific Northwest, a wetter-than-normal winter in the Gulf States and central and Southern California, and a warmer-than-normal late autumn and winter in the Northern Great Plains and upper Midwest.<sup>28</sup> The US consequences of La Niña include colder winters in the northern US, and warmer and drier conditions in the southern and southeastern parts of the US. The number of Atlantic storms during La Niña is likely to increase. The consequences to the energy markets are obvious. The paper by Michael McPhaden (see footnote 26) lists other global consequences of El Niño and La Niña, which should be of interest not only to energy traders but also financial traders because of their often-dramatic impact on human economic activity.

**Figure 3.2** Southern Oscillation Index, Tahiti–Darwin, sea-level pressure, (three-month average)



Source: <http://www.cpc.ncep.noaa.gov/data/indices/soi.3m.txt>

## HURRICANE: PREDICTION AND IMPACT ASSESSMENT

The monitoring of hurricanes and predicting hurricane paths and potential consequences to the energy markets is one of the most important functions of a weather group. This task goes well beyond the mere prediction of a hurricane's trajectory and strength.

The hurricane season in the North Atlantic extends from June 1 until November 30, but the industry starts paying attention to the prediction of hurricane activity a few months before the starting date – early predictions of the hurricane season usually become available in December of the preceding year. The most prominent forecast is available from Colorado State University forecasters Philip Klotzbach and William Gray, and also from the Mark Saunders group at UCL,<sup>29</sup> who provide an expected number of named North Atlantic storms, hurricanes and major hurricanes, as well as probability of a hurricane striking the US Gulf Coast.<sup>30</sup> This forecast is likely to evolve over the following months, prompting some of the more aggressive traders to position and reposition their portfolios for the expected level of turbulence.

Among other data, information released in the Colorado State University forecasts includes (with the forecast for 2011 in parenthesis):

- named storms (16);
- named storm days (80);
- hurricanes (9);
- hurricane days (35);
- major hurricanes (5); and
- major hurricane days (10).

The scientists behind this forecast caution the recipients of this information, pointing out the probabilistic nature of the data:

We issue these forecasts to satisfy the curiosity of the general public and to bring attention to the hurricane problem. There is a general interest in knowing what the odds are for an active or inactive season. One must remember that our forecasts are based on the premise that those global oceanic and atmospheric conditions which preceded comparatively active or inactive hurricane seasons in the past provide meaningful information about similar trends in future seasons. *This is not always true for individual seasons* [emphasis added by the author]. It is also important that the reader appreciate that

these seasonal forecasts are based on statistical schemes which, owing to their intrinsically probabilistic nature, will fail in some years. Moreover, these forecasts do not specifically predict where within the Atlantic basin these storms will strike. The probability of landfall for any one location along the coast is very low and reflects the fact that, in any one season, most US coastal areas will not feel the effects of a hurricane no matter how active the individual season is.<sup>31</sup>

This message is often lost on the trading community, and the forecast sometimes has an impact out of proportion with its historical accuracy.

During the hurricane season, every tropical wave developing off the coast of Africa is tracked and the probability that it will evolve into a tropical depression, tropical storm and, eventually, a hurricane is constantly assessed. At some point the focus changes to assessing a probability that a hurricane, or potential hurricane, will enter the Gulf of Mexico and, conditional on this event, what is its potential trajectory, strength, point of landfall and likely impact on energy infrastructure.

The trajectory of a hurricane is important from the point of view of its impact on offshore natural gas and oil installations. Drilling platforms are usually shut down and evacuated ahead of a hurricane if it is likely that it will reach dangerous strength. This is a supply impact. The location of landfall is also important for assessing the impact of a hurricane on land-based energy installations and on population, as well as commercial and industrial operations. This is a demand impact. As the information about hurricane progress becomes available, traders often make hectic adjustments to their portfolios. Sometimes the change of hurricane trajectory or strength is so sudden that no adjustment can be made in time.<sup>32</sup>

It is important to recognise that a hurricane is a highly unstable dynamic system. The strength of a hurricane will vary over its path, and it may either intensify or weaken as it hits patches of warmer and colder water over the Gulf of Mexico.<sup>33</sup> Hurricanes in the Atlantic are classified according to a five-level scale known as the Saffir–Simpson Hurricane Scale. The categories are:

- ❑ Category 1: Winds of 74–95 mph; damaging winds are expected
- ❑ Category 2: Winds of 96–110 mph; very strong winds will produce widespread damage

- ❑ Category 3: Winds of 111–130 mph; dangerous winds will cause extensive damage
- ❑ Category 4: Winds of 131–155 mph; extremely dangerous winds causing devastating damage are expected
- ❑ Category 5: Winds over 155 mph (these are VERY rare); catastrophic damage is expected<sup>34</sup>

The Saffir–Simpson Scale is routinely mentioned in the media during the hurricane season and is well known to the public. However, it has been criticised for its discrete nature – a 2–3 mph difference can move the hurricane from one category to another, with a big change in perceptions of its strength.<sup>35</sup>

An alternative index that has been developed to characterise the destructive force of a hurricane is the Carvill Hurricane Index (CHI),<sup>36</sup> used for the settlement of the hurricane futures and options traded on the Chicago Mercantile Exchange (CME). The CHI index is calculated as

$$CHI = \left( \frac{V}{V_0} \right)^3 + \frac{3}{2} \left( \frac{R}{R_0} \right) \left( \frac{V}{V_0} \right)^2 \quad (3.2)$$

where  $V$  denotes the maximum wind speed of a hurricane in miles per hour and  $R$  the radius of hurricane force winds (winds above 74 mph) in miles around the centre. The reference values  $V_0$  and  $R_0$  used by the CME are 74 mph and 60 miles, respectively.

The actions taken by traders after a hurricane makes landfall are often more critical and complicated than predicting the hurricane's path. Market prices evolve under conditions of very limited and conflicting information, in an atmosphere of panic, with the trading teams being sometimes dislocated and dispersed. Some consequences of a hurricane, such as shutdowns and damage to offshore platforms and underwater pipelines, damage to onshore energy facilities such as natural gas processing plants and compressor stations, may point towards the reduction of supply and higher prices. The shutdown of industrial plants and power plants consuming energy, and the dislocation of population, may reduce demand and prices. Some developments may also produce a Balkanised and disconnected market, with shortages and higher prices at some locations and surpluses and lower prices at others. For example, some compressor stations and pipelines may be

non-operational and natural gas may become stranded in some pockets, without the ability to put it into storage or transport it to market.

The shutdowns of offshore platforms may be quite protracted, even if there is no serious damage. Evacuation of personnel takes time. After a hurricane passes, the platforms have to be inspected for damage and brought back online in a systematic, time-consuming way. Industrial plants, such as refineries, take a long time to shut down and restart. This is related not only to the complexity of the chemical processes taking place, but also the high temperatures and pressures required for many processes. The installations have to be cooled down and then reheated at a slow pace, otherwise the connecting parts made of different materials (such as metals and alloys) may experience thermal contraction and expansion at different rates, producing a risk of subsequent leaks and breakdowns if the shutdown/restart happens too fast.

The shutdown of many industrial plants, commercial buildings, damage to transportation infrastructure and dislocation of population may result in a drop of demand and a counter-intuitive drop in prices at some locations. In some cases, the power transmission and distribution grid may be damaged, with power being unavailable to many customers for weeks. This reduces the demand for power and, indirectly, for natural gas.

This tug-of-war between forces leading to either higher or lower prices makes the task of traders and fundamental analysts very difficult. There are several sources of information, such as the reports made available in the past by Materials Management Service (MMS),<sup>37</sup> exploration and production companies and local utilities that can be of high value to analysts in this area.

## CONCLUSIONS

This chapter has documented the importance of weather forecasting to energy trading. Most energy traders, and especially those responsible for short-term physical transactions, have to develop a good understanding of the channels of transmission – from weather conditions to market prices – and learn about the strengths and limitations of different forecasting models. This is not the last time weather-related data and weather impact on energy prices will be discussed in this book. Weather information is an important input to many

quantitative models used for forecasting demand for natural gas, electricity and heating oil. The output from these models is an important feed to systems developed for predicting market prices and the evolution of inventory levels. Last, but not least, many merchant energy companies and financial firms offer specialised market instruments for managing weather risks. Modelling and pricing such instruments requires a good understanding of historical weather data, their statistical properties and potential shortcomings.

- 1 "Winds of change: A new gust of inflation recalls an older era", *Financial Times*, Monday, May 14, 2007, p 8.
- 2 Certain industries (airlines) and sectors (agriculture) of the economy remain sensitive to weather fluctuations, but the overall vulnerability of national economies has been greatly reduced.
- 3 Frank Norris, 1903, *The Pit: A Story of Chicago* (Cambridge, MA: R. Bentley).
- 4 Coffee in Brazil and orange juice in Florida are obvious examples.
- 5 Events in the Texas power markets illustrate the impact of weather transmitted through both the supply and demand channels. The Electric Reliability Council of Texas (ERCOT)'s electricity grid broke its winter peak demand record between 7 and 8 pm on February 2, 2011, at 56,334 MW. At the same time, extreme (for Texas) cold knocked out about 7,000 MW of capacity statewide, resulting in price spikes and rotating outages (see <http://www.star-telegram.com/2011/02/02/2818600/more-rotating-electricity-outages.html>).
- 6 An excellent analysis of the impact of volcanoes on weather conditions and analysis of the implication for agricultural and energy markets can be found in the *Browning Newsletter*, a monthly publication (available at <http://browningnewsletter.com/>).
- 7 Joel K. Sivillo, Jon E. Ahlquist and Zoltan Toth, 1997, "An ensemble forecasting primer," *American Meteorological Society Journals On Line*, (124), December.
- 8 This model replaced two older models: AVN and MRF.
- 9 The operational (deterministic) model runs out to 10 days. There is also an ECMWF Ensemble, which runs out to 15 days, like the GFS Ensemble.
- 10 The model results are available from <http://weather.unisys.com/ecmwf/index.html>.
- 11 <http://www.ecmwf.int/>.
- 12 <http://products.weather.gov/PDD/NCEPMAF.pdf> (see also <http://www.nco.ncep.noaa.gov/pmb/nwprod/analysis/>).
- 13 In one firm we worked for, a power trader demanded that he receive the morning weather report ahead of other trading desks in the same company. Electricity traders were transacting with natural gas desks a few yards away and wanted to have competitive advantage. Given such attitude, it is not surprising the company is no longer in business.
- 14 For example, the summer of 2010 was characterised by high precipitation in San Diego, due to the interaction of two weather patterns (La Niña and PDO).
- 15 There are several competing definitions of NAO, related to the specific stations used in compiling historical data or the numerical techniques involved. The most widely accepted definition is based on rotated principal component analysis (RPCA), used by Barnston and Livezey (see A. G. Barnston and R. E. Livezey, 1987, "Classification, seasonality, and persistence of low frequency atmospheric circulation patterns," *Monthly Weather Review*, 115, pp 1083–126).
- 16 James W. Hurrell, Yochanan Kushnir, Geir Ottersen and Martin Visbeck, 2003, "An overview of North Atlantic oscillation," in James W. Hurrell, Yochanan Kushnir, Geir

- Ottersen and Martin Visbeck (eds), *The North Atlantic Oscillation. Climatic Significance and Environmental Impact* (Washington DC: American Geophysical Union).
- 17 The correlation between the winter NAO index and the amount of precipitation over western Norway (Bergen) is equal to 0.77 (see James W. Hurrell, 1995, "Decadal trends in the North Atlantic oscillation: Regional temperatures and precipitation," *Science*, 269, pp 676-79.; this paper contains a tabulation of such correlations for other locations in Europe).
  - 18 The anomaly was discovered by Roland Madden and Paul Julian, hence the name.
  - 19 See Jon Gottschalck, Wayne Higgins, "Madden Julian oscillation impacts," NOAA/NWS NCEP Climate Prediction Center; Jon Gottschalck, Vernon Kousky, Wayne Higgins, Michelle L'Heureux, "Madden Julian Oscillation (MJO)" ([http://www.cpc.ncep.noaa.gov/products/precip/CWlink/MJO/MJO\\_summary.pdf](http://www.cpc.ncep.noaa.gov/products/precip/CWlink/MJO/MJO_summary.pdf)).
  - 20 [http://www.cpc.ncep.noaa.gov/products/intraseasonal/intraseasonal\\_faq.html](http://www.cpc.ncep.noaa.gov/products/intraseasonal/intraseasonal_faq.html).
  - 21 Dennis L. Hartmann and Harry H. Hendon, 2007, "Resolving an atmospheric enigma," *Science*, (318)5857, December 14, pp 1731-32 (<http://www.sciencemag.org/content/318/5857/1731.summary-off-2>).
  - 22 See <http://jisao.washington.edu/pdo/> for additional information and links.
  - 23 Nathan Mantua, 2002, "The Pacific Decadal Oscillation," ([http://www.atmos.washington.edu/~mantua/REPORTS/PDO/PDO\\_egec.htm](http://www.atmos.washington.edu/~mantua/REPORTS/PDO/PDO_egec.htm)).
  - 24 <http://www.pmel.noaa.gov/tao/elnino/gif/ElNino.gif>.
  - 25 Commonwealth of Australia 2007, Bureau of Meteorology (ABN 92 637 533 532).
  - 26 Michael J. McPhaden, 2006, "El Niño and La Niña: Causes and global consequences," Ted Munn (ed), *Encyclopedia of Global Environmental Change* (New York, NY: Wiley). This is a highly recommended, up-to-date review article.
  - 27 Given that many natural gas and, increasingly, oil traders define their summer strategies around the hurricanes, ignoring the effect of El Niño can be devastating. The history of the US natural gas markets in the US and some hedge fund fortunes would be completely different if it were not for the El Niño.
  - 28 NOAA Magazine Online (story 24).htm.
  - 29 University College London (see: <http://www.tropicalstormrisk.com/>).
  - 30 An alternative forecast is available from the NOAA National Weather Service Climate Prediction Center. NOAA's seasonal hurricane outlook does not project where and when any of these storms may hit.
  - 31 "Extended range forecast of Atlantic seasonal hurricane activity and landfall strike probability for 2011" (see, <http://www.news.colstate.edu/content/documents/apr2011.pdf>).
  - 32 The change in the direction of the hurricane Katrina became obvious after the close of open-outcry session at Nymex on Friday afternoon, too late to make significant adjustments to the portfolio. Some hedge funds that did not make investments in high-quality weather teams suffered significant losses.
  - 33 Many sophisticated weather groups will constantly monitor the information coming from the buoys about the surface sea temperatures along the projected hurricane path. This is very important as the hurricanes may behave in a highly unpredictable way. For example, Katrina changed its course and intensified as it hit a patch of warm water known as the Louisiana loop.
  - 34 <http://www.enchantedlearning.com/subjects/weather/hurricane/classification.shtml> (see also <http://www.nhc.noaa.gov/aboutsshs.shtml>). Earlier versions of this index were based on additional criteria, such as central pressure and storm surge. The dynamic nature of the hurricanes and the potential for conflicting rankings led to simplifications of this classification.
  - 35 See "A detailed overview of the CME Hurricane Index (Chi)," for more criticism, <http://www.cmegroup.com/trading/weather/files/Hurricane-CHI-Whitepaper.pdf>.
  - 36 The index is based on the paper by Lakshmi Kantha, 2006, at the Department of Aerospace Studies at the University of Colorado in Boulder, Colorado, "Time to replace the Saffir-

- Simpson Hurricane Scale?,” *Eos*, (87)1, January 3. The idea was adopted by the insurance broker RK Carvill.
- 37 The MMS information is available at <http://www.gomr.boemre.gov/homepg/what-snew/hurricane/index.html>. The website contains damage assessments for past storms, which may be used as a way to learn from past experience of hurricane events. “On October 1, 2011, the Bureau of Ocean Energy Management, Regulation and Enforcement (BOEMRE), formerly the Minerals Management Service (MMS), was replaced by the Bureau of Ocean Energy Management (BOEM) and the Bureau of Safety and Environmental Enforcement (BSEE) as part of a major reorganization” (see <http://www.boemre.gov/>). The industry still refers to this website as MMS.



## Section 2

# **Participants and Instruments**



## *Energy Markets: The Instruments*

This chapter will provide a review of the different instruments that can be used by energy markets participants to market or trade different energy commodities, hedge their exposures related to energy prices, or to make directional bets on market directions. The objective is to concentrate on common features shared by the various instruments, irrespective of differences with respect to underlying physical commodities. The review of such instruments is necessary in order to follow the discussions of different markets contained in subsequent chapters. The instruments covered here include forwards and futures, swaps and options. More complex types of transactions will be covered in the next chapter.

Understanding derivative instruments is a critical skill that any participant in the energy markets requires. Energy-related transactions are saturated with options: any contractual flexibility with respect to delivery timing and location, quality of the underlying commodity, volume, ability to discontinue or restart production, can be translated into some type of a derivative. These derivatives tend to interact and overlap, and are contingent on each other and on non-market variables (for example, weather) – making pricing and risk management devilishly complicated. For many years, these options were often given away for free, and sometimes actively sought by those who intuitively understood their value. We had the good fortune to be hired by a company (Enron) that had decided to reinvent itself, and dive into the trading of energy-related derivatives at a time when this field was still a relatively clean slate. One of our responsibilities was to help in introducing employees of the company and its customers to the brave new world of derivatives. After 20 years (almost to the day), the depth of collective knowledge of financial instruments by energy industry professionals is quite

amazing. Irrespective of the sad end of its story, Enron did make one positive contribution to the industry.

## SPOT, FORWARD AND FUTURES MARKETS

### Definitions

All textbooks on the commodity markets start with a mandatory discussion of the spot, forward and futures markets – and this one will be no exception. Most textbook presentations are based typically on a number of abstract and very general statements that apply to many different markets, without any discussion of how specific markets are organised and how they function, or what is really meant in practice by a spot and a forward price. We shall follow this convention but with many warnings to readers that different energy markets vary with respect to their microstructure and the ways the spot and forward markets interact. It is critical for any practitioner to realise that the world is usually more complicated than the models – which, by necessity, represent a simplified and stylised representation of reality – might imply. The general observations made in this part of the book will be exemplified in the chapters that follow. The narrative in this chapter assumes that a reader is familiar with standard financial instruments, such as forwards, futures and options, and with the uses of these instruments for speculation and hedging. Our objective will be to focus on the issues that are specific to the energy (and in general, commodity) markets.

This chapter is important for a number of reasons. No practitioner can survive in the modern commodity markets without a solid understanding of derivatives. First, complex commodity contracts are usually valued and managed by decomposing them into a portfolio of forwards, swaps and options. In some cases, a schema of a complicated transaction looks like a diagram describing the mother-board of our laptop. However, a word of caution is required – many of these conceptual valuation blocks do not have market equivalents (for example, some complex options may not be traded in the markets) or have readily available prices but still can be valued and managed using standard financial engineering tools. Additionally, the derivative markets are critical for the price discovery process, either directly or as a source of inputs into quantitative models used for trading or decision support.

A spot market represents transactions for immediate delivery. In

the following chapters, we will attempt to identify and describe the spot markets for different energy commodities. In reality, there are only a few true spot markets, with real-time power pools arrangements being the best examples. Commodity trading happens for all practical purpose in the forward markets, which often have very short maturity but still meet all the criteria of the definition of a forward contract.

A forward contract is a bilateral transaction that provides for delivery of the underlying commodity at a future date, at the price determined at the inception of the contract, with cash being paid at delivery, or after an agreed number of days following the delivery. Each specific commodity market has its own set of conventions and definitions that any practitioner has to understand in depth. Many trading losses have been caused by inexperienced traders and originators extending the rules applying to one commodity to a similar market with significantly different procedures.

The rigidities of the physical distribution infrastructure require that a time lag exists between placing an order and taking delivery and this lag may vary from as little as a few minutes in the power markets, a day in the natural gas markets, to a few days.<sup>1</sup> Given that there is always a time lag between transaction and delivery, one can argue that most transactions in the commodity markets are really forward transactions. A pure spot transaction happens at a retail gasoline station or a local gasoline distribution terminal. Without going into a discussion resembling a debate about the number of angels that may dance on the head of a pin,<sup>2</sup> We think that everybody can agree with the following:

- ❑ the demarcation line between spot and forward transactions is, at best, blurred;
  - ❑ from the practical point of view, the forward markets are critical for energy trading – and this is what counts;
  - ❑ the price formation process in what one can call spot markets is increasingly dependent on the price discovery in the forward and futures transactions;
  - ❑ practical difficulties of identifying true spot prices lead most researchers to using the futures prices of the shortest maturity as the proxy for spot; and
  - ❑ the elusive nature of spot prices does not stop many modellers
-

from building models using spot prices as the cornerstone of their philosophical approach.

To reinforce the first bullet point above, it is important for any practitioner in the energy markets to recognise that the demarcation line between the spot and forward prices is, at best, a bit fuzzy. One should accept as a fact of life that every market has its set of conventions and definitions and some transactions will be referred to as spot, even if they clearly have some forward flavour. Instead of wasting time trying to come up with perfect definitions, time can be better spent developing insights into how specific prices are formed, discovered and communicated. This is where the profit potential is.

#### *Forward contracts: pricing and convergence to the spot*

The market value of a forward contract at inception is equal to zero. A moment later, as an underlying forward price changes under the impact of market forces, the contract turns into an asset for one counterparty and a liability for the second counterparty. Over the life of a contract, prices will usually fluctuate and the marked-to-market value of a specific position from the point of view of a given counterparty will oscillate between positive and negative values. A forward contract is a zero-sum game: the gains of one counterparty are equal to the losses of the other side. The counterparty benefiting from a higher price is called a long; a counterparty losing from a higher price is called a short. An equivalent definition is that a long has to take delivery and make a payment. The opposite is true of a short. Some forward transactions may be cash-settled, so this definition of a long does not apply. The most general definition of a long/short position is based on who benefits from a higher price. This discussion may give an impression of splitting a hair, but in practice it may be important. Some structured transactions are so complex that determining which counterparty can be called long and which can be called short may be quite difficult.

Forward prices are denoted in this book as  $f(t,T)$ , with  $f$  denoting the price level,  $T$  the maturity (expiration date) of the forward contract and  $t$  being today's date. In other words,  $f(t,T)$  denotes the forward price observed at time  $t$  for a contract expiring at time  $T$ . The spot price is denoted as  $S(t)$  or as  $f(t,t)$ . The last notation is based on

the convention under which the spot price is a forward price with zero time to maturity. In other words, the forward price at maturity converges to the spot price. The convergence is a feature of the stylised, textbook version of a forward contract. In practice, convergence is more complicated due to market frictions and imperfections, and the basic difficulty of deciding in many cases what the spot price is. More on this subject shortly.

Under the stylised definition of a forward price, the maturity of the contract, its financial settlement and the cashflow event coincide. In practice, all these dates may diverge and one should think carefully about the timing of different events and the consequences for valuation and the profitability of any transaction. It is important to distinguish between the timing of all these events in the design of a software platform for commodity trading. Future cashflows have to be discounted back to the portfolio valuation date using correct dates. This information should be captured in the database. A physical forward transaction does not necessarily mean that a short has an absolute obligation to make a delivery: there is always an option to reverse this transaction by entering into an offsetting transaction or through book-outs (ie, offsetting a given transaction with other transactions with the same counterparty). The futures are forwards traded on organised exchanges. A detailed discussion of the futures contract is offered in the next section.

The convergence of forward prices to cash prices is a fuzzy concept, because the forward/futures prices are often defined more precisely and are more transparent than the spot prices, which are sometimes difficult to identify. One can take the CME natural gas futures contract, discussed in more detail in the chapters on natural gas, as an example. The contract for 10,000 MMBtus of natural gas expires three business days before the end of a calendar month and requires ratable delivery over the course of the next calendar month, known as the delivery month. The delivery location is the Henry Hub in Erath, Louisiana. What is the spot price that the futures contract is supposed to converge to? One could argue that the proper reference point is the price (or the set of prices) of natural gas traded for next day delivery through the calendar month. The problem is that these prices are not known when the futures contract expires,<sup>3</sup> and they will be determined by supply and demand conditions that cannot be precisely predicted. In the case of natural gas, the daily

prices are influenced primarily by weather that cannot be predicted with any degree of accuracy for more than a few days (see Chapter 3).

One can also argue, more correctly in our view, that a better reference point for the futures price is the price of the base load physical gas for next month ratable delivery at the Henry Hub location (the so-called index price).<sup>4</sup> Such prices become known (as described in detail in the chapters on natural gas) on the first business day of a delivery month. We are back to the concept of convergence of the futures to the expected monthly prices. This means that at best one can take a position that the futures prices of natural gas converge to the anticipated prices of the Henry Hub physical monthly gas. To make things even more convoluted, the monthly index price is technically a forward price (more precisely, the fixed price of a short-term swap with a one-month tenor; a swap is a package of forwards while a forward contract can be looked at as a bullet swap), so we are really talking in this case about convergence of the futures to the forward price (the concept of the swap price is explained below, see also the chapters on natural gas for more details). One has to also recognise that the prices set in physical transactions that determine the next day/next month index are influenced by the prices of transactions executed on exchanges and reported in real time to the trading floors. In practice, the prices in physical and derivative markets are being set through a simultaneous, general equilibrium-type process where everything depends on everything else.

Given all the practical complications, the convergence of forward prices and the spot prices should be considered as a tendency that is not perfectly realised due to inevitable frictions, asynchronous trading, imperfect information and limitations of the physical infrastructure limiting the potential for arbitrage transactions.

The difficulty of identifying and monitoring a spot price for many energy commodities is another reason why most widely traded contracts are the forwards and futures. The physical flows are very important but a significant percentage of volumetric flows is priced using formulas based on the forward and futures prices. The true spot markets are often very opaque, based on conventions and personal relationships that are difficult to discern from the outside. In many cases, one would say that a pure spot market simply does

not exist. This is another reason why we are dealing primarily with the forward markets in this book.

Energy companies often use different versions of forward transactions to meet their diverse needs. A forward transaction can be looked at as a purchase with deferred delivery and deferred payment. A deferred delivery with an accelerated payment is called a *pre-paid forward* and is often used by the energy producers to finance the development of new producing properties. A producer receives an amount of money upfront in return for the promise to deliver *volumetric flows* in the future (we use the term volumetric flows to emphasise that a transaction is settled through deliveries of physical commodity, as opposed to cash settlement). Some companies used prepay's in the past to hide loans – ie, pretending that loans represented cashflows from commodity related transactions. Examples of such transactions will be provided in the chapters on natural gas.

### Futures

Futures are forward contracts traded on an organised exchange that offer credit guarantees of counterparty performance, provide trading infrastructure (typically an open outcry system or screen-based electronic trading) and develop the definitions of standardised contracts. The exchanges offer a number of safeguards, which reduce credit risk to a minimum (as explained below).

Technically, credit protection is offered by a clearinghouse, which may be associated with an exchange or may be an independent entity. In the US, commodity exchanges historically owned clearinghouses, but this is not the norm elsewhere. We shall discuss this topic in detail when we cover operations and the design of clearinghouses. After a futures transaction is executed, it is novated: an original transaction is broken up, with a clearinghouse being inserted between the original buyer and seller. The clearinghouse takes the other side of each of the two new transactions created through the novation process. At inception, each counterparty posts the so-called initial margin that is followed by the variation margin deposited by a counterparty incurring a loss on the contract, and credited to the account of a counterparty that profits from the transaction. The call for a variation margin happens when the equity in the account of a given counterparty falls below the maintenance margin. The initial

margin may be compared to the security deposit required when opening an account with a utility or renting an apartment. Initial margin provides a buffer for losses which a clearinghouse may incur in the case of significant market move or the inability of a customer to post variation margin in a timely manner, or at all. If the customer does not maintain an adequate level of margin, the clearinghouse may close the customer's position and return the initial margin net of transaction costs and unmargined losses (if anything is left). In practice, the mechanics of calculating the required margin may be quite complicated as an exchange may recognise that a given client may hold multiple positions in different contracts and different maturities of the same contract. The potential losses in one commodity may be partially or fully offset by gains in another futures position. Chapter 6 describes the margining system used by the CME and licensed to a number of other exchanges worldwide.

To illustrate the concept of initial and variation margins further, we can use the following example. Suppose that the initial margin for a futures contract is equal to US\$10,000 and the maintenance margin

### Futures contract identifiers

The specific futures contracts are identified using a combination of letters and numbers. For example, the natural gas Nymex contract for January 2013 delivery will be identified as NGF13. NG is a ticker for natural gas; F denotes January, 13 stands for the year. The contract month codes are listed below:

- F – January
- G – February
- H – March
- J – April
- K – May
- M – June
- N – July
- Q – August
- U – September
- V – October
- X – November
- Z – December

The most important energy contract tickers are: CL – crude oil; HO – heating oil; HU – unleaded gasoline; NG – natural gas; and RB – RBOB gasoline.

to US\$8,000. If a market participant incurs a loss of, for example, US\$2,500, their equity in the account (assuming they are long or short no other contracts) drops to US\$7,500 and a variation margin call of US\$2,500 is initiated, restoring equity in the account to the level of the initial margin. It is also important to remember that variation and initial margins are generic names. Different exchanges can use various terms to describe margins. For example, the CME uses the term performance bond.

The practice of marking-to-market the contract and posting (receiving) additional margin is the most important difference between futures and forward contracts. This critical distinction, which is often missed in structuring transactions, impacts the way futures- and forward-related transactions are accounted for and carried in the books. Forward transactions represent cashflows expected sometime in the future and should therefore be valued and carried on the books at their net present value (NPV), calculated using an appropriate discount rate.<sup>5</sup> Futures positions are associated with immediate cash settlements, can be converted into cash (or settled with cash) at the price levels readily available from the electronic screens, and discounting is not required. This distinction carries to swaps, which are just portfolios of forwards carried on the books at a single levelised price.<sup>6</sup> Ignoring this crucial distinction between accounting for futures and forwards, as obvious as it is, may (and often has) led to many structuring snafus. One may add that this distinction between forwards and futures is further complicated by the fact that most bilateral transactions are now collateralised based on the mark-to-market present value of the position.<sup>7</sup> It is important to observe that the counterparties to an OTC transaction may negotiate a waiver of collateral (especially for counterparties with a strong balance sheet and a good credit rating) or a partial waiver if the mark-to-market credit exposure is below a certain level (known as the threshold).

The exchanges often modify the level of initial margin, depending on market conditions. The increase in margins happens often following time periods of historically high price volatility. An increase in margins may provide a jolt to the markets as some weaker hands, unable to post additional margin, may be forced to reduce their positions. This explains why many periods of market turbulence produce predictions of margin changes, contributing further to

instability. One obvious example was an increase in initial margins for the silver contract leading to a dramatic price drop. This is a short summary of this event:

In the 2 week period starting April 26th, margins on one silver futures contract rose from \$8,700 to \$16,000 in just two weeks – an 84% increase. That is a huge amount of capital for a trader to have to find in such a short space of time. A lot of traders couldn't afford hence there was a huge sell off. The market corrected and as a result silver fell by 28% in around a week. [...]<sup>8</sup>

The irony is that raising margin is often seen as a remedy against excessive speculation and price volatility. The outcome is even more volatility.

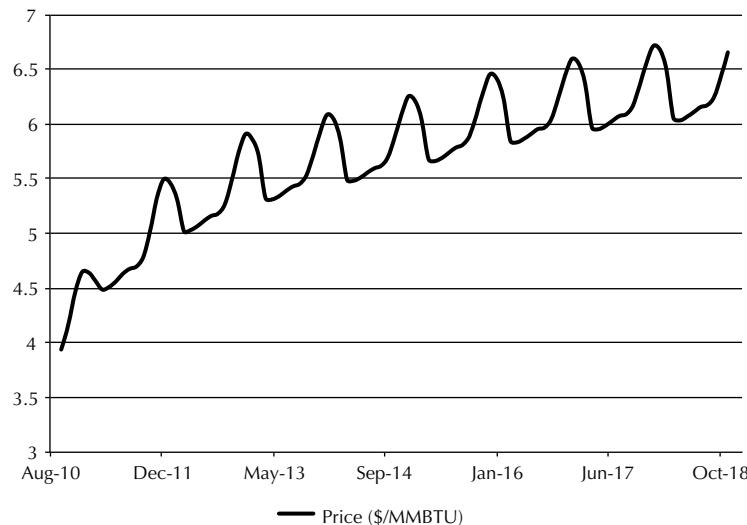
#### *Forward price curve*

It is easy to observe that at any time  $t$ , one can enter into multiple forward transactions for the same underlying commodity, with different expiration dates ( $T_1, T_2, \dots, T_n$ ). The collection of such forward prices as of time  $t$  is known as the forward price curve. There are several practical and theoretical issues related to the construction of a forward price curve.<sup>9</sup> One obvious question is the frequency at which the price data is reported and stored in the database. Each market has its own unique conventions regarding the frequency at which the forward price curve is specified. For example, in the oil and natural gas markets, forward prices are quoted on a monthly basis. In other markets, such as electricity, one can start with quotes for hourly prices, switching at some point to daily, and then to monthly prices. Sometimes, the conventions for quoting the forward price curve are quite intricate and can be explained only by the historical evolution of a specific market. In some markets, forward contracts are structured as cascading transactions, with some contracts covering longer time periods, and being eventually broken into shorter contracts closer to the maturity date. For example, in some European electricity markets, such as Norway, forward prices are quoted at varying frequencies, starting with annual and quarterly prices at the back of the curve. As a contract slides down the curve towards shorter maturities, a price is eventually broken up into higher frequencies. In the US power markets, summer trades initially as a block, with forward prices being broken into July and August forwards when the market for single months

becomes more active. The calendar defining cascading constitutes part of a contract specification. Figure 4.1 shows an example of a forward price curve for natural gas (Nymex NG contract, settlement prices as of September 3, 2010), with a pronounced seasonality.

The question traders and risk managers have to answer is how to save price curves in the database. In the case of cascading contracts, one solution is to carry original prices. An alternative is to report all the prices at the lowest common frequency. For example, one could break the prices for annual and quarterly forward contracts into monthly prices using a mathematical algorithm.<sup>10</sup> The advantage of this solution is a simpler database design and the ability to price transactions corresponding to sub-periods of currently traded contracts (for example, one could price a July transaction when only a summer block power is traded). The downside is that the original source data may be lost and that the algorithm used for power curve construction may be somewhat arbitrary and produce unrealistic prices.

One of the basic responsibilities of a trader is posting a forward price curve at the end of the trading day for the market they are responsible for, keeping track of the evolution of the curve during the day and reacting when a forward price or a segment of the forward price curve undergoes a rapid change. In liquid markets this task may be relatively easy, especially if the entire curve or its most liquid segment is formed on an exchange with the information displayed on electronic screens in real time (or close to real time). In less transparent markets, a trader receives during the day bits and pieces of data about the evolution of the market. At the end of the day, the trader has to come up with the curve as of the close of business, relying on the prices observed during the day, settlement prices reported by the exchanges and daily pricing sheets from brokers and specialised consulting firms collecting and disseminating forward prices to the subscribers. In many cases, a trader has to use their judgement to fill the gaps in the data, relying on their experience. For example, forward prices may be available from third parties on a seasonal or annual basis and a trader has to apply their seasonality coefficients to spread them over shorter time periods. It is easy to underestimate the practical importance of this feature of the commodity markets. Traders often become experts in some niche and illiquid markets, and a firm becomes critically dependent on

**Figure 4.1** Forward price Nymex natural gas contract

Source: Nymex

their expertise. It is difficult to challenge them if there are reasons to believe they made a mistake or misrepresented market conditions for some illicit ends. In our experience, every big trading operation has to deal every year with several cases of questionable price curves, with some traders departing the organisation with surprising regularity. One of the key responsibilities of the middle office (ie, risk management and compliance) is the validation of the forward price curves posted by traders.

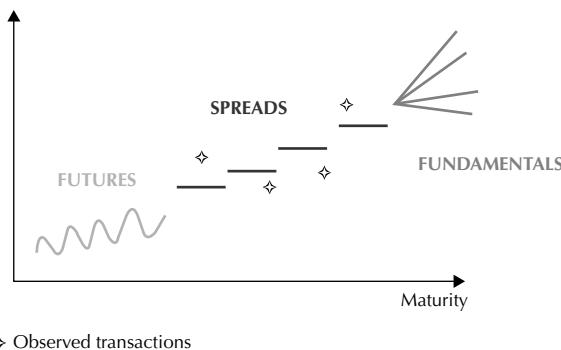
Figure 4.2 illustrates how a forward price curve is constructed in practice. This is a generic graph and the details will vary from market to market and commodity to commodity. In the case of commodities for which a futures market exists, the front of the curve corresponds to the prices available from the exchange. The industry ignores the potential difference between futures and forward prices (this is addressed below in more detail). At some point, a trader switches to forward market observations using information obtained from their transactions, bids and offers from their counterparties and information acquired from brokers and consulting firms, who send price sheets to their customers at the end of the day. The information obtained from brokers comes in the form of direct prices (for

example, as monthly forward prices) or as swap prices (ie, prices representing a bundle of several forwards).

An example of a swap would be electricity prices for July–September 2013 for a specific trading hub. Sometimes, the pricing information may arrive in the form of calendar spreads. The responsibility of a trader is to use this information to build a forward curve at a higher level of granularity. If a July–September swap price is provided, the trader has to distribute the swap price over specific months using seasonality assumptions and any information that may be applicable. At longer maturities, the information becomes more fuzzy: long-term OTC transactions are private, the pricing information is not available to the market participants, the transactions may be complex, highly structured deals bundling multiple risks and multiple commodities. Even if the terms of the specific transaction were fully known to the market, extracting a specific price would require complex modelling. At some point, even this limited information arrives at such long time intervals and is so unreliable that it becomes unusable from the point of view of forward price curve construction.

What are the choices if this is the case? One solution is to rely on the judgement of a trader, supported by some *ad hoc* improvised rules. For example, an average expected rate of inflation may be built into the forward prices, although one could complain that this is a statement of faith, not supported by any reliable information. There is no reason to expect that a forward or future spot price of a specific commodity should change in tune with the expected general

**Figure 4.2** Construction of a forward price curve



inflation (assuming one can come up with a reasonable forecast of inflation). The second solution is to use an algorithmic approach based either on applications of stochastic calculus or modelling a physical system (future supply and demand curves for a given commodity). Some models, known as hybrid models, combine both approaches to construct the forward price curve.

### **The differences between futures and forward prices**

A discussion of the difference between forward and futures prices often evolves around the comparison of stylised definitions of these contracts, ignoring rapid evolution of the market, the pace of financial innovations and the progress taking place in risk management. We shall list a number of statements reflecting conventional wisdom and then qualify them in view of the complexity of the energy markets.

*Standardisation.* Futures are standardised contracts, while OTC forwards, swaps and options are customised. The more correct statement would be that the OTC contracts *can be customised* and, in practice, there are no restrictions on the variability of contractual arrangements embedded in privately negotiated contracts. In order to trade futures with a click of a mouse or a few screams and hand gestures, the contracts have to be customised with respect to such defining features as expiration date, the specification of the underlying commodity or the financial instrument, the delivery location, quality of the commodity acceptable for delivery, maximum daily price change, etc. One has to recognise, however, that futures have many built-in escape hatches that allow for modification of some critical provisions of the contract through mutual consent (through, for example, alternative delivery procedures, ADPs, and exchange for physicals, EFPs),<sup>11</sup> and also embed some optionality (to be explained shortly). We explain EFPs in the chapters on natural gas. An ADP is a “provision in a futures contract that allows buyers and sellers to make and take delivery of the underlying product under terms or conditions that differ from those outlined by the contract.”<sup>12</sup>

At the same time, most OTC contracts are standardised for an obvious reason: if every OTC deal were to be negotiated from scratch, the burden on end-users and market makers, in terms of time required for bargaining, booking, settling, monitoring and

measuring the risk, would be prohibitively high. The beauty of the market is that participants gravitate, in their best interests, towards the transactions that represent a compromise between the need to target certain risk exposures and minimise the extent of customisation of transactions in order to reduce transaction costs. The trend towards standardisation explains to a large extent major productivity gains realised since the dawn of the industrial revolution, as pointed out by countless economists.<sup>13</sup> The commodity markets are no exception. Nonetheless, the trend towards the standardisation of the OTC contracts does not have to destroy the potential for innovation: predefined OTC contracts can be used as the building blocks of very complex hedging and trading strategies.

*Credit risk.* One very important feature of futures contracts is the daily marking-to-market and margining of existing positions (as explained above). A clearinghouse is inserted as the counterparty between the two sides of a transaction executed on an exchange. This is often characterised as futures having no credit risk. The more precise statement would be that mutualisation of risks through clearing reduces the probability of a default to a very low level but does not completely eliminate this risk. Although there was no default of a clearinghouse and an exchange like the Chicago Mercantile Exchange has multiple layers of defence against non-performance by one of the clearing members and their customers, the potential for insolvency always exists, as anybody old enough and still blessed with a good memory can point out:

On October 19, 1987, financial Armageddon nearly struck. The stock market plunge of 508 points on Black Monday is seared into market observers' memories; what is less well remembered is the credit crisis that ensued and threatened to destroy two of the United States' largest clearinghouses. More than a dozen clearing members of the Chicago Mercantile Exchange (CME) fell out of compliance with capital requirements, and half a dozen more faced margin calls that exceeded their capital. [...] Both CME and CBOE temporarily halted trading, lest the financial condition of the markets – and their clearinghouses – deteriorate further.<sup>14</sup>

The issue of the credit quality of clearinghouses is of critical importance to the operations of financial markets. The Dodd–Frank Act passed by the US Congress in 2010 mandated the exchange trading and clearing of standardised derivatives (with an exception for end

---

users). This increases the importance of clearinghouses to the stability of the financial system and underlines the importance of monitoring the credit quality and financial governance of any exchange and clearinghouse. This point has been made repeatedly in the context of discussions about the reform of the financial system, the most succinct statement being made by René Stulz:

As long as the clearinghouse is well capitalized and manages its risks well, there is no material counterparty risk with the clearinghouse. This fact explains the widely held belief that requiring clearing for over-the-counter derivatives will significantly reduce systemic risk. It is important, however, to understand that we have much experience with exchange clearinghouses and little experience with over-the-counter clearinghouses. Over-the-counter clearinghouses have not been tested in a financial crisis.<sup>15</sup>

At the same time, it is important to recognise progress in managing credit risk in bilateral transactions. One has to be aware that bilateral credit risk has become much better understood, monitored and mitigated than in the late 1990s. Bilateral transactions happen under credit agreements that establish the rules for posting collateral and other critical issues, such as the maximum acceptable credit exposure and the exposure threshold below which collateralisation is not required. The industry, supported by several initiatives undertaken by the Committee of Chief Risk Officers (CCRO), the Edison Electric Institute (EEI) and the International Swap and Derivatives Association (ISDA), has made significant progress in rationalising the legal framework and procedures for handling credit risk.

In view of these observations, the claim that credit risk disappears in the case of futures is too extreme. One advantage that futures contracts have over OTC transactions is derived from the more efficient management of capital required for margining/collateralising the outstanding positions. Migration from the OTC markets to exchanges saves the overall amount of working capital required for credit support. This is not necessarily true with respect to a specific market participant. This explains why the early proposals to force the end users of standardised derivatives to trade on the exchanges and clear the transactions were met with a loud protest from the industry. Most end users prefer to hedge with financial institutions, primarily due to the ability to negotiate a waiver or relaxation of collateral requirements below certain levels of exposure (this issue is discussed in more detail in Chapter 6).

*Price transparency.* Futures prices are available to the public at zero or nominal cost, whereas the prices negotiated in OTC transactions represent private information. This statement would be generally true with respect to the US markets only a few years ago. One has to recognise, however, two developments that require drawing a much finer distinction between these two markets. Forward price information, in aggregate form, has been available for a long time from price-reporting agencies and consulting firms. Forward price curves are compiled from the information collected from brokers, exchanges and market participants and disseminated to subscribers. The Dodd–Frank Act contains requirements that the swap transaction prices be reported to the so-called swap data repository (SDR). The detailed procedures are still being outlined at the time of writing but the major provisions of the rules can be divined from the proposals published by the CFTC. In addition to reporting the transaction data to the SDR, the Act contains provisions for dissemination of the data to the public.<sup>16</sup> What requires clarification are the requirements for the swap market participants to report the data to the SDRs or disseminators, and the SDRs' obligation to make public certain swap data, "as soon as technologically practicable." Block trades on swap execution facilities or exchanges like the CME or subject to the end-user exception are likely to be reported after a 15-minute time delay.

One can expect that the degree of transparency of the OTC prices will change dramatically for the US markets if the provisions of the Act are fully implemented. Given the challenges this law is still facing in Congress and the uncertainty regarding the final rules, the advice the author can give to the reader is: "Don't take any sweeping statements about the differences between the OTC and futures contracts for granted. Pay attention to the market developments."

*Market participation.* The OTC markets admit as a rule or by tradition only institutional players and have no retail participation. In the case of the energy markets, the participants are energy companies (producers, transportation companies and utilities), industrial firms, investment and commercial banks, hedge funds and energy marketers. This has a number of consequences for the market structure, including the following.<sup>17</sup>

- ❑ *Limited market activity.* The number of transactions over a given time period in the OTC markets is small compared to the number of trades taking place on exchanges.
- ❑ *Large individual transactions.* A typical transaction size in the OTC markets is much bigger than a typical trade executed on an exchange. Transaction costs on an exchange are very low compared to the OTC deals. Higher transaction costs have to be spread over a larger volume to make the OTC market viable.
- ❑ *Limited participation.* The number of OTC market participants is small compared with the number of participants trading on the exchanges. This is explained by the large size of the OTC transactions and the balance sheet requirements for the market participants.

*Forward prices versus futures: numerical differences.* As explained in the previous section, many forward price curves start with a segment based directly on futures prices (to the extent that a corresponding futures contract exists). In theory, the forward and futures prices are not necessarily identical. The difference between the two sets of prices can result from the practice of daily marking-to-market futures prices and margining. The counterparty that is posting margin has to borrow funds or divert cash from other corporate uses. In either case, the cost of margining depends on interest rates.

Suppose that interest rates are positively correlated with the commodity prices. If the prices go up, a trader who is long a futures contract has a gain in their account (margin will be credited to their account), which can be reinvested at higher interest rates. This translates into an additional gain for the long and makes them willing to pay a higher price for the futures contract, compared to a corresponding forward. For the short, increasing prices will require raising cash (to post margin) in a higher interest rates environment. This additional cost will induce them to ask for a higher price compared to the forward. This, in conjunction with the willingness of a long to pay a higher price, will lead to the futures prices being above corresponding forwards. The opposite is true in the case of negative correlations between commodity prices and interest rates.

Empirical verification of this proposition is difficult, as it is not easy to find corresponding time series of the futures and forward prices which are strictly comparable. A dissertation by Viola Markert

contains a study of correlations between commodity prices and interest rates.<sup>18</sup> From our point of view, the results reported by Ms. Markert regarding energy commodities are of primary interest. As one can see in Table 4.1, the correlations are low and statistically insignificant. We can conclude, following Ms. Markert and industry practices, that one can ignore the distinction between the forwards and futures. However, a word of caution is required. We have been operating in the world of low interest rates and relatively subdued inflation for a long time. It is likely that the relationship described above will be reversed if commodity price inflation accelerates and starts feeding into general price indexes and into the level of interest rates. One of the messages we are trying to convey in this book is that one can never step twice into the same river when it comes to commodity trading. One should avoid treating any statement about the markets as a universal, eternal truth.

### The relationship between spot and forward prices

The critical question every trader has to address in planning their trading and hedging strategies is related to the forces behind the shape and evolution of a forward price curve. Two basic shapes of a forward price curve are distinguished in the literature: backwardation and contango. The origins of these terms are somewhat shrouded in mystery. Most authors track their use back to John Maynard Keynes, but it stands to reason that Keynes did not coin these words himself but borrowed from the jargon used by traders in the City of London.<sup>19</sup> The definitions of backwardation and contango used in practice and in most textbooks are somewhat simplified. Backwardation corresponds to the case of a downward sloping

**Table 4.1** Correlations between commodity prices and interest rates

Commodity	Correlation	t-statistics
Crude Oil	0.01	0.1
Gas Oil	0.05	0.7
Gasoline	0.06	0.8
Heating Oil	0.07	1.0

Source: Viola Markert, 2005, "Commodities as assets and consumption goods: Implications for the valuation of commodity futures," PhD dissertation, May, p 16.

Note: Monthly futures log-returns and the 3-month Treasury T-Bill, 1/86 – 12/03.

forward price curve, with the spot price exceeding the forward prices, while contango is the opposite case, a forward price is sloping upward, with the spot prices below the forwards.

In practice, one can encounter many different shapes of a forward price curve. For example, a forward price curve may slope downward in the front, reversing into contango for longer contract maturities. Many economists have formulated slightly different definitions, and one has to be very careful when using these terms. Spot prices are sometimes not directly observable, and some authors replace the spot prices with forward contract prices of the shortest maturity, treated as proxies for the cash prices. Most academic studies of contango/backwardation rely on this approach. Backwardation is sometimes defined by comparing spot prices (prompt forward prices) with discounted forward prices – ie, forward prices corrected for the time value of money.<sup>20</sup> An alternative definition forms the foundation of the so-called normal backwardation theory, which will be explained shortly. There are many competing theories explaining the empirically observable shapes of the forward price curves. In addition, there are many competing heuristic techniques for simulating evolution of a forward price curve treated as a unified object. A detailed review of competing theories and subtle differences in alternative formulations is beyond the scope of this book, so we will concentrate on reviewing the two most important threads: the normal backwardation theory and the storage theory.

### **Normal backwardation**

Normal backwardation theory, sometimes called risk premium theory, was formulated by John Maynard Keynes in his usual casual way in a newspaper article and further elaborated upon in a book published in 1930.<sup>21</sup> A more systematic exposition was offered by John Richard Hicks and a number of other economists. If one were to summarise this theory in one sentence, one would say that the hedgers are net short – or, in other words, there is an imbalance between natural longs and natural shorts, with the latter dominating the market.<sup>22</sup> Natural shorts are those market participants who sell forwards (futures) to offset long positions resulting from their primary business. For example, a natural gas producer is long this commodity given their underlying business and they short forwards

to neutralise price risk (to transfer it to a third party). A natural long would be, for example, an industrial company using natural gas as feedstock. In other words, the definition of natural long (short) is based on the positions taken in the forward market. For example, the desire of the oil producers to hedge may not be perfectly offset by the hedge positions of the processors of oil and the users of refined products. The gap is filled by speculators, who need incentives to enter the forward markets to compensate them for risks they take. They will require a differential between a forward price and expected spot price for the time period corresponding to the expiration of the contract. In other words:

$$f(t, T) = E_t[S(T)] - P \quad (4.1)$$

where  $P$  is the risk premium required by the speculators.<sup>23</sup>

According to both John Maynard Keynes and John Richard Hicks, the risk premium is positive, making the forward price lower than the anticipated spot price. The speculators make money buying forwards at prices that are, on average, lower than the spot prices at contract expiration. Hicks supported this hypothesis with an additional observation that the processors are less sensitive to price risks than the producers.<sup>24</sup> The producers of commodities are, therefore, ready to sacrifice a greater percentage of returns to reduce risk compared to the market participants who acquire their products. One can observe that this assumption reflected the conditions of the time. When Keynes and Hicks were observing the forward markets, their thinking was shaped by the experience of agricultural futures (in which Keynes participated, sometimes to his regret). It is not unreasonable to assume that farmers, who were not as well capitalised as manufacturers and were at the mercy of the markets characterised by low price and income elasticity of demand, were more risk averse.

Under normal market conditions, when the expected spot price does not diverge from the current spot price (ie, when the markets are stable, the expected spot prices are roughly equal to the current spot prices), the difference between forward price and the current spot price is negative: we are in a situation of normal backwardation. This leads to an alternative definition of backwardation: forward prices being below *expected* spot prices. Under conditions of excessive inventories, the current spot price does not necessarily

correspond to the expected spot price. However, the expected spot price is still above the forward price: we may have a simultaneous condition of contango (with expected spot price being above the current spot price) and backwardation (forwards being below the expected spot price). This argument points to the very important role played by inventories. In the following chapters, we will elaborate on the sources of inventory information for each specific market and review studies linking inventories to the spot/forward price differential.

It is important to recognise that Keynes did not believe that “it is necessary that there should be an abnormal shortage of supply that a backwardation should be established.”<sup>25</sup> His theory of normal backwardation is often simplified to the upward spot price pressure resulting from the temporary short-term demand pressures hitting limited supply. “If supply and demand are balanced, the spot price must exceed the forward price by the amount which the producer is ready to sacrifice in order to “hedge” himself.”<sup>26</sup> The spot price, Keynes argued, incorporates the remuneration for the risk of price fluctuations, while the forward price excludes this. Keynes also added that, if the cost of hedging is excessive, many producers will prefer to run the price risk rather than pay it.<sup>27</sup> Excessive stocks lead to contango, but the forward price established by the cost-of-carry (a concept that will be explained shortly) argument tends to be lower than the expected future spot price, as explained above.

The tests of normal backwardation theory can be classified either as indirect (ie, relying on economic theories of asset pricing, such as the capital asset pricing model (CAPM) to prove the existence of systematic risk associated with holding a portfolio of futures) or direct (ie, based on conclusions drawn exclusively from tested theory without making reference to more general theoretical concepts). The direct tests revolve around the following implications of normal backwardation.

- ❑ Given that the forward prices are biased predictors of future spot prices (the difference between them being the risk premium), the forward price should gravitate towards the spot price, as the forward contract approaches maturity.
- ❑ Given that the theory of normal backwardation depends on structural disequilibrium between natural longs (speculators)

and natural shorts (producers), the net positions of these two groups of market participants should be mirror images.

- ❑ Risk premium should be real, not phantom – the speculators should make money as a group.

A detailed review of the results of these tests is beyond the scope of this book.<sup>28</sup> We shall limit our comments to general remarks outlining the difficulty of executing any tests in practice, given the limitations of the available data.

- ❑ The tests are applied to many distinct commodities with different underlying market structure; market conditions are not static and may evolve over time as new participants enter the market.
- ❑ The calculations of returns usually ignore the role of leverage in the forward/futures markets.
- ❑ The impact of price seasonality is ignored or is not addressed in a satisfactory manner.
- ❑ The market is analysed as if there were only two different homogeneous groups of agents: producers/end users and speculators. The possibility that the producers may engage in speculation on a massive scale is ignored (or perhaps such behaviour was unheard of in the days of Keynes and Hicks when the theory was formulated).

### **Storage theory**

Storage theory, sometimes referred to as a convenience yield theory, is focused on the critical role that inventories and availability of storage facilities play in the commodity markets. The forward and spot prices are linked through the formula that ties together the fair price of a forward contract, the spot commodity price, interest rates and the storage cost. Market mechanisms behind this equation are the potential arbitrage transactions that will be triggered whenever the identity shown below is violated. Specifically, the forward price ( $f$ ) is equal to the spot price ( $S$ ) adjusted for the financing and storage cost(s), or:

$$f(t, T) = S(t) \times (1 + r + s)^{(T-t)} \quad (4.2)$$

where  $T - t$  represents the life of the contract ( $t$  is the today's date,  $T$  is the maturity date, both measured in years),<sup>29</sup> and

$S(t)$	spot price as of time $t$ ;
$f(t,T)$	forward price of a contract expiring at time $T$ , as of time $t$ ;
$r$	annualised interest rate; and
$s$	annualised unit storage cost.

The logic behind this equation is based on the arbitrage argument that defines the upper and lower bounds for a forward price; one can prove that both bounds are equal and define a unique forward price. A forward transaction can be replicated by borrowing funds at rate  $r$ , buying the commodity at spot price ( $S$ ), storing the commodity and taking a short position<sup>30</sup> in the forward contract (selling). At contract expiration, the transaction is reversed: the commodity is delivered into the forward contract at price  $f$  and the financing and storage costs are recovered. The arbitrageurs will engage in such transactions, as long as the inequality  $f > S \times (1 + r + s)^{(T-t)}$  holds, or as long as one can make a profit by selling a forward contract and delivering the physical commodity from storage on the contract maturity date.

Alternatively, the lower bound can be established by recognising that one can sell the commodity in the spot market and invest the funds at rate  $r$ , rent the storage facility they own, taking at the same time a long position in the forward contract. At maturity, the delivery is taken and the position in physical commodity is restored. The transaction is profitable as long as we have  $f < S \times (1 + r + s)^{(T-t)}$ , or as long as the revenues from investing the proceeds of a short sale and renting the storage facility exceed the cost of buying back the physical commodity in the forward market. In a competitive and transparent market, any deviation from strict equality in Equation 4.2 will trigger the arbitrage transactions described above, restoring the postulated relationship between the spot and forward prices. Of course, in practice, the transaction costs and market imperfections will result in Equation 4.2 being realised only as a general market tendency.

The logic behind Equation 4.2 is often referred to as the replication mechanism – ie, recreation of the economic outcome of a forward contract through physical operations involving real assets. Equation 4.2 also demonstrates that, as time to delivery gets shorter ( $T - t \rightarrow 0$ ), the forward price will converge to the spot price. Under the notation for forward prices used in Equation 4.2, one can write the spot price as

$$S(t) = f(t, t) \quad (4.3)$$

This, as explained above, is equivalent to treating the spot price as a forward price with zero time to delivery.

The mechanism explaining the level of forward price formation can be treated as a general market tendency, but has to be qualified to recognise the influence of many market frictions and imperfections, including:

- bid–offer spreads;
- transaction costs;
- unequal access to the physical infrastructure and the markets; and
- restrictions on short selling.

### Bid–offer spreads

Bid–offer spreads correspond to the difference between prices at which one can buy (offer) and sell (bids). The bid–offer spread represents part of the profit of a market maker who posts bid–offer prices – ie, prices at which they are willing to transact (buy and sell). Bid–offer prices vary from market to market, location to location, and generally increase with the tenor of a contract and the size of a transaction. The bid–offer spreads are typically negligible in the front of a forward price curve in a liquid market but increase significantly for longer-term contracts. The existence of a bid–offer spread places a forward price curve within two bands, corresponding to the bid and offer sides of the market, with the transactions taking place at the boundaries or inside the band. A related issue is the difference between the interest rates at which one can borrow and lend. Bigger, more established and better-capitalised entities enjoy access to cheaper credit and can invest on better terms. The difference in borrowing and lending interest rates has the same effect as the existence of the bid–offer spread.

The bid–offer spread should be distinguished from the market impact of a given transaction. In any market there is a level of transactions over a given time period which will affect the market price. As a matter of fact, a liquid market is often defined as a market in which one can transact significant volumes without affecting the prices or widening the bid–offer spreads. Any forward price curve has implicit transaction volume assumptions behind it. An analyst

working on a large structured transaction should always consult a trader to recognise the impact of the proposed volumes on the profitability of a transaction.

### **Transactions costs and access to infrastructure**

Transaction costs cover a wide range of charges one incurs engaging in market transactions, such as commissions, fees, taxes, communications costs, the costs of processing transactions, etc. These costs may vary from one market participant to another and can have a significant impact on the profitability of different trading operations. A related issue is the access to the physical<sup>31</sup> and informational infrastructure of the markets. The author of this book may know, in principle, all the steps one has to take to arbitrage the differences in price levels of refined products between North America and Europe. It does not mean, however, that we could structure this transaction together on a short notice from our home study or university office, without the benefit of constant presence in the markets, cultivation of personal contacts or constant monitoring (or direct control) of information about the availability of storage tanks and tankers. Access to physical storage may be sometimes limited. Physical storage facilities are often controlled by market participants who prefer to keep a reserve of unused capacity to take advantage of emergencies and market opportunities.<sup>32</sup> Whether we like it or not, access to physical infrastructure is unequal. We live in a society based on the division of labour and specialisation, and this sometimes produces inequality of outcomes. In principle, differences in transactions costs, inequality in access to the markets and unequal borrowing costs could push the forward price towards the levels determined by the most efficient and lowest-cost market participant. This is, however, unlikely to happen, given that other market participants would be unwilling to transact with a single counterparty and make them captive customers of a single entity. The ability of a single participant to monopolise a forward market will be also limited by the internal risk management considerations of the market maker. Even the biggest financial institutions limit the size of the positions they are willing to take.

### **Convenience yield**

Given that both the storage cost and interest rate are positive, the arbitrage argument implies that the forward price is always greater than the spot price and that the forward price curve (a collection of

forward prices for different contract horizons) is an increasing function of time to expiration (the curve is always in contango). In reality, forward price curves for many commodities, including natural gas and crude, are often backwardated (the spot and the forward prices in front of the curve exceed the prices in the back), as explained in the previous section. This arises from the fact that for many market participants holding physical commodities has a higher value than controlling a forward contract. The stream of benefits that accrue to the owner of a physical commodity is called a convenience yield (introduced by Nicholas Kaldor, 1939). Assuming that the convenience yield can be stated per unit of commodity,  $d$ , the full cost-of-carry formula for an energy forward contract is given by the following equation:

$$f(t, T) = S(t) \times (1 + r + s - d)^{(T-t)} \quad (4.4)$$

The concept of a convenience yield is very popular among academics and generally ignored by practitioners. The position of this book is that convenience yield represents a doubtful practice of explaining one unknown variable through another unknown and that it is more productive to concentrate on forward prices that represent an observable reality. The convenience yield is typically estimated from the forward prices by assuming that the first observable forward represents the spot price and then by going through the process of bootstrapping the convenience yields from the remaining forward prices, under some assumptions of borrowing and storage costs. This comes across as a somewhat circular and dubious procedure: no additional information is created.

The arbitrage argument often breaks down in the case of commodity-related forward contracts for a number of reasons.

- ❑ In some cases, the existing market framework does not allow for shorting of the underlying physical commodity (there are no established procedures for borrowing the underlying commodity), eliminating one of the arbitrage bounds. In most commodity markets, the framework for such transactions does not exist or the transactions are associated with significant costs. Borrowing a commodity requires significant search time and engaging in physical operations, which may be expensive. Shorting financial instruments is relatively straightforward.<sup>33</sup>

- ❑ If this is the case, one has to revise the traditional paradigm used to obtain insights into the behaviour and properties of the forward prices. One can think of the forward markets not as mechanisms to move the existing supplies of commodities in time and space, but as a discovery market, a system for aggregating and processing information, used by the market participants willing to make bets in the same way bets are made on the outcome of a baseball game.
- ❑ Another complication arises from the fact that many energy commodities are not pure financial assets but are also consumption goods and inputs to the industrial processes undertaken to produce consumer goods and services. This means that the replication approach to pricing forwards breaks down as one has to compare at any point in time the current value of energy commodity as a consumer good with the replication-based value implied by Equation 4.4. Once the complications described above are recognised, the properties of forward prices can be derived from more complex equilibrium models that rely on the analysis of intertemporal supply and demand of a given commodity and the preferences of the market agents. Such models available in academic literature are looked at by the practitioners (those who care to read such papers) as somewhat simplistic and stylised, especially when compared to approaches used in practice and the massive volumes of fundamental information collected by the analysts<sup>34</sup> in the commodity trading operations, with millions of records pouring in every morning into the trading floors. We believe that such models represent a useful pedagogical tool and allow us to develop useful insights onto the operations of commodity markets. In most markets, depending on the circumstances, price formation is dominated sometimes by the forward replication mechanism and sometimes by the consumption value. A skilled trader who understands the market mechanism can assess the fundamental information flows, and especially the data on inventory levels and availability of storage, to guess which mechanism will dominate.
- ❑ Finally, some energy commodities may be unstorables (like electricity) or the supply of storage may be limited at times. We defer discussion of this important topic for now. The futures markets

in which the cost-of-carry mechanism breaks down and does not work effectively (for the reasons enumerated above) require the development of an alternative paradigm. This takes us back to the expectation hypothesis.

### What do the practitioners do?

In reading academic papers or attending industry conferences, one cannot fail to notice a deep gap between some theoretical approaches and the practice of energy trading. The traders think in terms of forward prices. Some theoretical approaches are based on modelling the evolution of a spot price over time and deriving the forward prices from the expected levels of the spot prices. This approach is not very popular with the practitioners for a number of reasons that will soon become clear. Also, many option-pricing models found in the academic literature are stated in terms of the spot prices. We do not subscribe to this philosophy for a number of reasons.<sup>35</sup>

First of all, most transactions in the energy markets are defined in terms of forward prices. If one uses a model based on a spot price, it is necessary to build a theoretical bridge to the forward price and this is often difficult in view of the complex and evolving relationships between these two markets. Using the spot price as a starting point to value a transaction referencing forward prices seems to be a detour fraught with dangers related to model imperfections and calibration errors, and a violation of the basic requirement of making a model as parsimonious as possible. One should use aggressively the Occam's razor<sup>36</sup> principle and gravitate towards a simpler model.

Second, using less parsimonious theoretical frameworks not only increases the model risk but also imposes significant costs in terms of the complexity of software systems and market data requirements. More information has to be captured, stored and processed, and this additional cost is seldom justified by the benefits of this approach.

Third, most spot prices found in the energy markets, once one digs deeper into the contract details, turn out to be forwards (although often very short-lived forwards). In the chapters on specific commodities, we shall point out which prices deserve the label of true spot prices. For example, in the electricity markets, even real-time prices can be interpreted as forward prices, given the process through which they are determined. Also, in many cases, the causality runs not from the spot towards forwards but in the

opposite direction. For example, the crude oil prices determined under the so-called P+ transactions (see the chapters on oil markets), which can be interpreted as true spot prices, are derived from the prompt Nymex West Texas Intermediate (WTI) contract prices.

Practically all the traders we have worked with tend to think in terms of forward price curves treated as unified objects. The questions they ask the quantitative and fundamental analysts revolve around the following.

- ❑ The best analytical techniques used to describe the evolution (upward and downward shifts and reshaping) of a forward price curve.
- ❑ The extent to which the current forward price curve deviates from the future spot price. There is a subtle distinction here. The traders do not ask the question about the expected spot price to come up with a better guess of what a forward price should be. For them, a forward price at which they can transact is the objective and observable reality. They want to know if they can put on a profitable trade, given the likely direction in which the market is going to evolve. In the decision-making process, the traders use eclectically the elements of different theories explaining the forward prices but also rely to a large extent on experience and intuition. We can recall the marching orders the head of the trading operation at Enron was giving to the traders trying to launch a new market: "Trade at the levels at which the market would trade if there was a forward market." This statement more or less summarised how traders think about the ways the market operates.
- ❑ The best fundamental tools to answer the questions listed above.

To summarise, it is difficult to find a generalised justification for using energy derivatives models relying on the modelling of the spot price, irrespective of circumstances.<sup>37</sup> Of course, we live in a society defined around the principle of the pursuit of happiness. If somebody is made happy developing certain types of models, one should by all means persist. There is nothing to be lost but the attention of the traders.

## ENERGY COMMODITY SWAPS AND OPTIONS

In a very general sense, a swap<sup>38</sup> is the exchange of two streams of cashflows: one of them is based on a fixed (predetermined) price and the other on a floating (variable) price. Swaps structured as the exchange of two floating (variable) cashflows are called basis swaps. Swaps can also be looked at as baskets of forwards that are traded as one package (no cherry-picking is allowed). Swaps can be either physical or financial. Financial swaps are settled in cash. The physical commodity swaps require actual delivery of a physical commodity. We elaborate on these very general statements below.

### The mechanics of a swap

A standard commodity swap transaction requires specification of the following items:

- the underlying commodity;
- the notional quantity (volume);
- the fixed price;
- the published price benchmark for floating price;
- the settlement date(s); and
- cashflow dates.

What follows is a stylised example of a swap with a notional amount of 10,000 MMBtu of natural gas with a fixed price of US\$4.00/MMBtu and a floating price equal to the Houston Ship Channel monthly prices (the chapter on natural gas explains how such prices are formed and discovered). The swap settles financially (ie, no physical commodity actually changes hands) on a monthly basis. The floating price becomes known on the first business day of the month. A few days after the floating price settles (the actual delay depends on specific market conventions and/or negotiated terms), the cash-flows are exchanged. For example, if the floating price is equal to US\$4.5, the floating payment is equal to US\$45,000 and the fixed price payment is equal to US\$40,000. In practice, these two payments are netted. The counterparty paying the floating price will send US\$5,000 to the counterparty paying the fixed price. If the floating price is equal to US\$3.5/MMBtu, the payer of the fixed price sends an amount of US\$5,000 to the counterparty paying the floating price.

Under market conventions, the payer of the fixed price is called a buyer of the swap and, by definition, is considered to have a long position in the swap. The payer of the floating price is short the swap (ie, a seller). The rationale for this convention is obvious. A short position benefits from a lower price. Given the fixed price, a payer of the floating price is better off if the current market prices fall. In some more complicated swap structures, identifying a buyer/seller is more difficult but, as always, the guiding principle is who benefits/loses from a higher price.

The term “pricing a swap” refers to the calculation of the fixed price of the swap. The calculation is based on the principle of equality of the value of the two legs of a swap (the floating leg and the fixed leg) at inception. In other words, the initial value of a swap is zero. This condition is obvious: if the value of one leg of a swap exceeded the value of the other leg, one counterparty would have no reason to enter into such a transaction. In practice, one can encounter swap transactions with a fixed price agreed at the inception diverging from the price calculated according to the principles stated above (off-market swaps). There may be different reasons for this practice, sometimes legitimate, sometimes not. One example is so-called tilted swap, which produces initially greater cashflows to one party than a market level swap. Such a swap may be simply a loan in disguise.

The value of the fixed leg is equal to the sum of the fixed payments, calculated by multiplying the notional volumes of the underlying commodity by the fixed price, and discounting these payments back to the starting date. The value of the floating leg is equal to the sum of the volumes multiplied by the floating price, discounted back to the starting date. The floating price, at inception, is obtained from the forward price curve, a collection of forward prices for contracts with different maturities. There are two reasons to use the forward price curve for the valuation of a swap. First, a swap is hedged using the forward markets and second, one can argue that the forward price curve offers the best information about the future level of the market prices.

In other words, the fixed price of the swap is derived from the following equality:

$$\sum_{i=1}^n \frac{V_{t_i} \times F_{t_i}}{(1+r_{t_i})^{t_i}} = \sum_{i=1}^n \frac{V_{t_f} \times F}{(1+r_{t_i})^{t_i}}, \quad (4.5)$$

where  $t_i$  denotes the time of the cash flow ( $i = 1, 2, 3, \dots, n$ ), expressed in years. The interest rate  $r_{t_i}$  denotes a bullet annualized interest rate, used to discount a single cash flow taking place at time  $t_i$ . For example, if a swap has 12 settlement periods,  $n = 12$ , and  $t_6$  denotes the time at which the cash flow corresponding to the 6th settlement will take place. The time is measured, by convention, in years. If  $t_6 = 2.5$ , the sixth cash flow will happen 2.5 years from the valuation date.  $F$  denotes the fixed price,  $F_{t_i}$  – the floating price for the settlement taking place at time  $t_i$ .  $V_{t_i}$  denotes the volume underlying cashflow taking place at time  $t_i$ . We assume, for simplicity that the cashflow dates and the settlement dates are the same. In practice, this is not usually the case. A cashflow is determined at settlement, but the actual transfer of cash may happen at a later date.

$$F = \frac{\sum_{i=1}^n V_{t_i} \times F_{t_i}}{\sum_{i=1}^n (1 + r_{t_i})^{t_i}} \quad (4.6)$$

The denominator in Equation 4.6 is often called the total discounted volume, and many daily trading position reports show discounted volumes in addition to nominal volumes underlying the swap positions. The interest rates used in pricing swaps are the so-called zero (called also spot)<sup>39</sup> rates, used to discount bullet cashflows occurring in the future. This means that, for different future time points, different discount rates are used. The zero rates should not be confused with bond yields used in pricing coupon-paying bonds, with the same discount rates applying to every future cashflow.

An obvious question is what interest rates to use. The merchant energy industry has agreed since the early 1990s to use the set of zero interest rates corresponding at the time to the AA credit quality interest rate curve. These rates reflected the cost of funds of large international banks transacting in the eurodollar markets and typically being assigned (in the good old days) the credit rating of double A. This interest rate curve is typically referred to as the Libor<sup>40</sup> curve as it is constructed from Libor rates and Libor-related instruments. The wide use of the Libor rate curve by the industry can be explained by two reasons. First, it was argued frequently in the early 1990s that an emerging market for energy and other commodity derivatives would be dominated by large financial institutions, and the rates representing their cost of funds would be a critical input to determine the price level of commodity-based financial products. Second,

the merchant energy companies that eventually came to dominate the market embraced the concept of pricing the cashflows, which often represented the IOUs of BBB or weaker companies, using the AA discount rates. Under mark-to-market accounting, this approach allowed some firms to manufacture instant earnings out of substandard material. Only the practice of pricing junk bonds at the US Treasury rates would be more aggressive. Of course, some merchant energy companies, but not all, followed the practice of creating a reserve for the potential credit losses, based on the difference between the AA and respective counterparty's credit ratings.

A Libor curve can be built using routines available from different software platforms built for the energy trading companies or for interest rate risk management. A trading desk embedded in a big financial institution can easily obtain the Libor spot rates (or corresponding discount factors) from a fixed income desk. Any energy trading organisation should make an investment in skills related to interest rate analytics. We believe that there is still a knowledge gap between the merchant energy industry and the financial industry that must be closed.<sup>41</sup>

Equation 4.5 implies that portfolios of commodity swaps have interest rate sensitivity. A standard job interview question is how this risk can be managed. In our experience, in more than 50% of cases the answer is that there is no exposure to interest rates at all, another illustration that there is still a knowledge gap between financial and physical commodity traders.

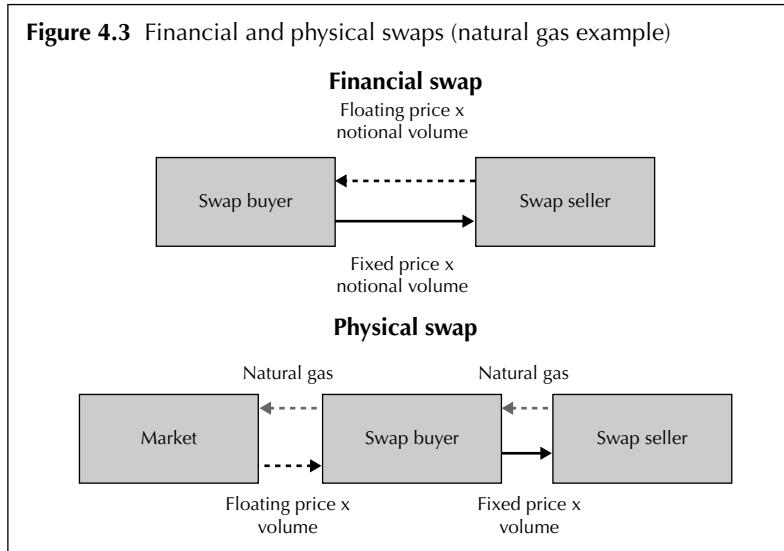
### **Physical, financial and basis swaps**

In practice, a swap may be physical or financial. A sale of an energy commodity at a fixed price, with deliveries distributed over time, is referred to and priced as a swap. As mentioned before, a swap may be a purely financial transaction, with one counterparty sending periodically to another the cheque for the fixed–floating price difference (multiplied by the underlying volume), with the direction of the cashflow dependent on the sign of the difference. In a physical swap, one counterparty receives molecules or barrels that, ignoring the transaction and operating costs, can be taken immediately to the market and sold at the current price, and are, therefore, seen as being equivalent to receiving a floating price. Figure 4.3 illustrates this concept.

One of the special features of the commodity markets is that in many markets one can trade directly the spreads between two different prices. In other words, a spread trades a single underlying and a position in a spread can be established in one transaction. Of course, a spread position can be established by independently entering into two separate transactions. One can, for example, establish a long position in the January/October spread by buying the contract expiring in January and selling a contract expiring in October. Once a position in a spread is established, it can be maintained to expiration or it can be closed or modified by eliminating one of its legs. For example, a trade can transform a January/October spread into a direct January position by buying back the October leg.

The spread position can be established with respect to location, time, quality and commodity type. In the US markets, natural gas markets locational spreads are referred to as basis transactions and trade as a single underlying.<sup>42</sup> The chapter on natural gas describes in more detail how this market functions and how basis swaps are priced. Spread positions created by buying/selling contracts expiring at different times are known as calendar spreads and are very popular for a number of reasons, including risk reduction and the ability to lower the cost of financing a given position, due to the

**Figure 4.3** Financial and physical swaps (natural gas example)



natural offsets in combined margin required for the two legs of a spread position (if the spread is established on an exchange). Another less obvious reason for the popularity of spread positions is the ability to circumvent risk management systems, often by creating stealth positions that are largely invisible, in spite of their massive size, to the risk management system. A long/short position of the same volumetric size in a March/April (famous or infamous H/J spread) natural gas contracts are likely to be invisible to the risk system based on the notional physical quantity limits.<sup>43</sup> This spread between March and April natural gas contracts is known as a widow maker, given its volatility and implied risk. More than one trader and some big hedge funds went down betting on this spread.

Of course, one can observe that spreads do not take down hedge funds. Some traders do. The reason why this position may be ignored by a risk system is that, if a trader is long 10,000 MMBtus in one contract month and short 10,000 MMBtus in the next month, this may be interpreted as a zero overall position. All risk management systems require aggregating positions to some extent in order to reduce the dimensionality of the problem. If March and April contracts are thrown into the same risk bucket, the risk magically disappears but is not in reality removed.

Risk managers try to remove this imperfection of the system by limiting net positions calculated over a moving time window, which is typically 12 months. The spread established for the adjacent months would still evade these control measures. A risk system using VaR could still be tricked to ignore the risk of this position as the price return correlations assumed in the model for the adjacent months would often be very high, sometimes close or equal to one (even though price returns for March and April tend to have weaker correlations than other adjacent contracts). One should not be surprised that the March/April natural gas Nymex spread enjoys a bad reputation in the industry.

### **Customised swaps**

The commodity markets developed a number of customised swaps through two different types of financial engineering. One approach consists of packaging swaps together with other financial instruments, such as options, while the second consists of modifications to the structure of a standard swap in order to meet a client's require-

ments related to risk management or speculation. A few examples follow.

A basis swap denotes the exchange of two cashflows based on floating (ie, variable) prices. For example, one can enter into a transaction consisting of the exchange of two monthly index prices of natural gas at different locations. One cashflow is based on the Houston Ship Channel monthly price,<sup>44</sup> the second on the monthly price of natural gas at Chicago Citygate. If these two prices for a given month are equal to US\$4.00/MMBtu and US\$4.50/MMBtu, respectively, and the notional volume is equal to 10,000 MMBtus, the two cashflows will be equal to US\$40,000 and US\$45,000. In practice, these two cash payments will be netted and one counterparty will send a cheque for US\$5,000 to the other party. Pricing of the basis swaps is based on the same principle that applies to the fixed-for-floating swap: equality of the present values of the two legs constituting a swap at the inception of the transaction. The only difference is that the swap has two floating legs instead of one.

In practice, meeting the requirement of equality of the values of the two legs of a basis swap at inception requires that one of the prices is adjusted, up or down, by a fixed differential. It is a matter of convenience which price is adjusted: a positive adjustment differential for one price is a negative adjustment for another price. Expanding the example above, the terms of a swap may require adjusting the Chicago Citygate floating price down by 10 cents. This means that the payments of US\$40,000 and US\$44,000 would be exchanged. Pricing a basis swap is equivalent to calculating the level of the differential at the inception of a transaction. Figure 4.4 illustrates the concept of a basis swap.

There is an alternative way of looking at the basis swap. Let us assume that a basis swap involves exchanging a payment based on  $\text{Price}_1$  versus a payment based on  $\text{Price}_2 + \text{Differential}$  (which may be positive or negative). This is equivalent to exchanging two cashflows:

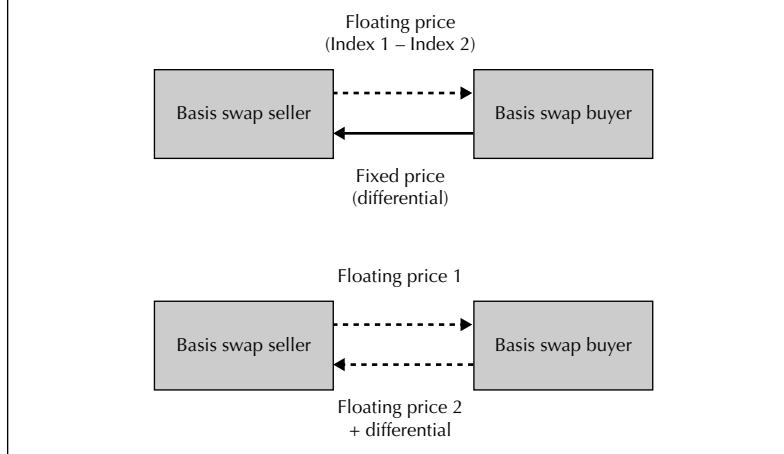
- cashflow based on the difference of  $\text{Price}_1$  and  $\text{Price}_2$  (ie, basis);  
and
- cashflow based on the fixed differential.

This means that a basis swap can be looked at as a fixed-for-floating swap, with the floating leg corresponding to the basis between two

prices (varying over time) and the fixed leg based on the constant differential. If a separate forward market for the price differential (ie, the forward market for basis) exists, the pricing of a basis swap can be carried out using a basis forward curve. Such curves can be constructed from the forward basis transactions observed in the market or derived from the two forward price curves defining the basis.

An example will help to explain the equivalency of these two ways of looking at a basis swap. Suppose that, under a basis swap contract, the swap counterparty A pays monthly basis ( $\text{Index}_1 - \text{Index}_2$ )<sup>45</sup> and receives a fixed price of F from counterparty B. Suppose that, for a given month, the  $\text{Index}_1$  is equal to US\$4/MMBtu and  $\text{Index}_2$  is equal to US\$3.75/MMBtu, and the fixed differential (F) under the swap agreement is US\$0.10/MMBTU. This means that counterparty A pays US\$0.25/MMBtu (= 4 – 3.75) and receives US\$0.10/MMBtu, for the net cash inflow of negative US\$0.15/MMBtu. This is equivalent to the swap counterparty A paying  $\text{Index}_1$  and receiving  $\text{Index}_2 + \text{Differential}$ . This translates into paying US\$4/MMBtu and receiving US\$3.75/MMBtu + US\$0.10/MMBtu. In both cases, the swap counterparty A has a negative cashflow of US\$0.15/MMBtu, which is a positive cashflow of 15 cents for B. In the first case, we can look at this transaction as the fixed-for-floating swap, in the second case, as the floating-for-floating swap (ie, basis swap).

**Figure 4.4** Basis swap



Identifying the long position hinges on determining who benefits from a higher price (in this case, a higher basis). In our example, counterparty A loses if the basis increases (think of Index<sub>1</sub> going to US\$5/MMBtu, with Index<sub>2</sub> remaining unchanged). This means that counterparty A is short basis. This is consistent with the convention under which the buyer of a swap is the payer of the fixed price.

An Asian swap is defined in terms of a variable price equal not to a settlement price at a specific point in time, but an average of settlement prices calculated over a certain time window. The economic rationale for Asian swaps transaction and pricing procedures are discussed in the section on Asian options.<sup>46</sup>

## Options

Energy-related options are among the most complicated instruments traded in the markets. This section provides definitions, a classification of available options and discusses the basics of option valuation.

An option is a contract that gives the right, but not the obligation, to buy (a call option) or sell (a put option), a certain instrument, good or commodity (called the underlying) at an agreed price (known as the strike or exercise price)<sup>47</sup> by or on a certain date (known as the maturity or expiration date). The last part of the definition is important because this is an explicit recognition of the fact that an option is a wasting asset, with a finite life. The seller of an option is called an option writer and the compensation for their troubles is called an option premium. The buyer of the option is typically referred to as an option holder. Once the option is sold, its value changes with the passage of time and the fluctuations of the market conditions.

The options available in the market can be classified according to many different criteria.

I. *Exercise rights.* Options that can be exercised at maturity only are known as European options. Owning a European option does not mean that one is stuck with it: one can always sell the option in the market, sometimes at a significant discount to its fair value, if the market for the option is thin. Options that can be exercised at any time during their life are known as American options. Options that can be exercised during a sub-period of their life are called

Bermudan options (for obvious reasons). Many executive stock options are of this type, as they have a vesting period after which they change from contingent to binding contracts.

II. *Contract initiation.* Option contracts are classified based on the way they are transacted. The most basic distinction is between exchange-traded options, which can be acquired on Nymex or another exchange, and OTC options negotiated bilaterally between two counterparties. OTC options are often highly customised, in the same way that OTC forward contracts can be tailored to the needs of both sides. A special case of OTC options are options embedded in the contracts that often are not even explicitly identified as options. They typically arise from flexibilities and conditionalities included in energy-related contracts. Any flexibility in terms of volume, price, place of delivery or time of delivery can be restated as an option and valued using the technology developed for pricing standalone options (the ability to price such options does not mean that this task is easy). Such options have existed in the energy business from time immemorial and correspond to the practical requirements of doing business in highly volatile and shifting environments. The revolution in energy risk management that took place since the mid-1990s centred around increasingly sophisticated techniques to identify such options, capture them in a risk-tracking system and value and risk manage them using technology adapted from the financial markets. Finally, one can identify so-called real options (options embedded in physical assets that correspond to the flexibilities (or rigidities) in the way a physical asset can be operated). Such options are acquired by buying an asset or by making modifications to an existing asset through investment decisions. In the most extreme applications of this approach, one can look at an asset as a portfolio of options and other derivatives, and value and manage the asset accordingly.

III. *Payout profile.* The payout profile of an option at expiration (or at the exercise time) can typically be illustrated with a familiar hockey stick diagram (see the next chapter for illustration). The mathematical definition of the payout profile for a call on a forward is given by the function *max* which returns the greater of the two arguments:

$$\text{Payout(call)} = \max(F_T - K, 0) \quad (4.7)$$

where  $F_T$  denotes the forward price at option expiration time  $T$  (or option exercise time), and  $K$  denotes the strike price. The payout is always greater than or equal to zero, as the holder has always the right to let the option expire as worthless. The payout for the put is defined as:

$$\text{Payout}(put) = \max(K - F_T, 0) \quad (4.8)$$

The progress in the option markets since the 1990's has been associated with the development of options with more complicated payouts. This trend was largely a reflection of the needs of the options' end users, who required instruments better matching their economic exposures. In some cases, the pendulum swung too far and many options brought to the market were rejected by end users as excessively complicated and impossible to value without a major investment in the risk management organisation.

One example of an option with a complex payout that proved useful and was embraced by the industry is an Asian option. Asian options are defined in terms of an average price calculated over an agreed time interval, known as the averaging window. This window may correspond to the entire life of the option or its subinterval. The payout of such an option is given by:

$$\text{Payout}(\text{Asian call}) = \max(\text{average}(F) - K, 0) \quad (4.9)$$

$$\text{Payout}(\text{Asian put}) = \max(K - \text{average}(F), 0). \quad (4.10)$$

The definition of the average price is described in detail in the option contract. A special type of Asian option is a floating strike option that has no predefined strike price: the strike price is calculated as an average over an agreed time window.

A spread option is defined as an option on price difference. The payout profile of a European spread option is:

$$\text{Payout}(\text{call on spread}) = \max((F_T^1 - F_T^2) - K, 0) \quad (4.11)$$

$$\text{Payout}(\text{put on spread}) = \max(K - (F_T^1 - F_T^2), 0) \quad (4.12)$$

where  $F_T^1$  and  $(F_T^2)$  denote the values of two forward prices at expiration. For example,  $F^1$  may denote the price of electricity and  $F^2$  the price of fuel, both expressed in the same units.

Binary options are defined as options with payouts that depend

on the realisation of a certain condition. The state of the market in which the condition has been realised is denoted by 1. In the opposite case, the state of the market is indexed with 0. This explains why such instruments are called binary (or digital) options. Some of these options come close to being pure bets. The payout is defined as a certain amount of cash received on exercise or as the underlying asset (instrument) itself. For example, a binary call option may pay US\$100 if the price of the underlying (let us say natural gas) exceeds US\$7/MMBtu on the day of option expiration. In this case, the payout will be the same if the price exceeds the strike by a penny or 10 dollars. Such an option is referred to as *cash-or-nothing*. An alternative option would pay the underlying itself (one MMBtu of natural gas, for example) if the price exceeded the strike (US\$7/MMBtu, for example). The payout of this option would increase with the value of the underlying (it is better to receive the underlying when it is worth US\$15 than US\$7.01). What is important in the definition of this option is that the payout is discontinuous at the price level of US\$7/MMBtu – it jumps from 0 to 7 dollars and then increases penny for penny with the price of the underlying.

The payout profiles for the first type of the binary option (cash-or-nothing) are:

$$\text{Payout}(\text{cash or nothing call}) = \max(\text{cash} | F_T > K, 0) \quad (4.13)$$

$$\text{Payout}(\text{cash or nothing put}) = \max(\text{cash} | F_T < K, 0) \quad (4.14)$$

where " | " means "conditional on". The payout profiles of the asset-or-nothing options can be defined in a similar way.

IV. *The underlying*. Another obvious classification of options is based on the nature of the underlying. The possible classification schemes depend on the commodity the option refers to and the nature of the underlying. The underlying may be either a physical commodity (spot) or a financial instrument, such as a forward, swap or option. Alternatively, the underlying may be a portfolio of physical commodities or financial instruments. Such portfolios are typically referred to as baskets. Options on baskets are increasingly important instruments, used for hedging or as building blocks of structured products. One can treat an Asian option as a special case of basket

option (both averaging and calculation of the value of the underlying basket reduce to summation over some predefined dimensions). Options on swaps are called swaptions and options on options are called compound options. One important distinction is the difference between strips of options and a swaption. A strip of options is a portfolio of options on the same underlying, with varying exercise dates. Such options can be defined as a sequence of daily, weekly or monthly options, for example, and they can be exercised separately. An option on a swap, when exercised, applies to the entire structure (the option is exercised into a swap), without the ability to cherry-pick some periods during the life of the swap.

Real options represent a special class of derivatives that do not have an independent existence separately from a certain physical asset. For example, the ability to adjust output of a power plant at short notice can be looked at as an option and valued and managed accordingly. At this point, we can just signal that the industry is using option-based techniques on an increasing scale in order to establish valuations of physical assets. This is another reason why one cannot survive in the modern energy industry without being familiar with derivatives.

*V. Sources of risk.* Most options have been designed to speculate on, or hedge against, the variability of the price of the underlying. One of the important sources of risk in the energy markets is uncertainty regarding volume. This risk is related to the difficulty of producing a perfect forecast of actual quantities of a commodity required in the future. The industry developed a special class of options designed to address this risk – and these options have become most popular in the natural gas markets, although they can be found in other markets, often as provisions embedded in the long-term contracts and asset-related transactions. In the natural gas markets, such options are known as swing options. A swing option is a contract that gives the right (but not the obligation) to acquire a given commodity  $n$  times during a certain time period. In addition, when the option is exercised, the volume taken can be modified within pre-established limits (with the lower bound being possibly zero). In order to offer protection to the seller, the holder is obliged to take a certain minimum volume during the life of the option (or over a certain time period) and cannot exceed a maximum total volume. In case the total volume limits are exceeded or the volumes taken fall

short of the contractual minimum, the holder of the option may have to pay penalties. These penalties may be calculated using, for example, agreed unit charges or a formula based on current market prices (typically, as of the option expiration date). The long-term contracts, which have the features described above and require the buyer to take a minimum volume of commodity in a given time period, are known as *take-or-pay*.

An example for a swing option is provided below. The user has the right to buy between 30,000 MMBtus and 50,000 MMBtus of natural gas 10 times during the month at US\$7/MMBtu. The minimum monthly volume is 250,000 MMBtus. If the minimum total quantity requirements are not met, the holder pays a penalty on the shortfall volume calculated using a unit charge of US\$8/MMBtu. One interesting aspect of this option is that it may be sometimes to the holder's advantage to exercise the option even if the strike price is below the current market price. For example, if the market price is US\$6.98/MMBtu, it may be profitable to pay the strike price, if it helps to avoid the penalty for volume deficiency.

## CONCLUSIONS

This chapter has covered the basics of derivative instruments used in the energy commodity business. These instruments can be thought of as building blocks of more complex trading strategies and transactions. The derivatives are important for other reasons, which may not always be obvious. Two important types of derivatives are stand-alone contracts and derivatives embedded in structured transactions and physical assets. Many physical assets can be operated and valued with the understanding that they can be looked at as portfolios of options. The pricing of energy commodities in the spot markets is often based on the prices established in the financial markets. A buyer of a cargo of crude oil may be an accidental participant in the derivatives market. In Molière's play *Le Bourgeois Gentilhomme*, good Monsieur Jourdain discovers one day that he was speaking prose all his life. In a similar way, energy commodity professionals have been in the derivatives business all their lives.

Familiarity with derivatives is not just an additional strong point on an energy professional's résumé. It is a critical ingredient in a package of skills required to survive in today's difficult markets. It is necessary to point out, however, that good training in derivatives

alone, without understanding the context in which they are used, is dangerous. Energy-related derivatives do not exist in an institutional and physical vacuum. Their value may be reduced or denied through constraints present in the industry infrastructure and through the actions of other market participants. Operating an asset with great embedded optionality is not a spectacular achievement if access is controlled by owners of transmission, or a pipeline or a hostile local government. Hopefully, the rest of the book will help the reader to develop necessary insights into the potential perils of the energy business, or at least equip them with the information necessary to identify potential problems.

- 1 Physical transactions in the electricity markets are sometimes arranged in 10–15 minutes. In the US natural gas markets, volumes transacted on Fridays flow over the weekend and into Monday of the next week.
- 2 One can use an arbitrary distinction between spot and forward transactions based on a number of days to delivery. In our view, this is an example of theorising through classification but the US legal standards support this approach. A forward contract is “(A) a contract (other than a commodity contract) for the purchase, sale, or transfer of a commodity...or any similar good, article, service, right, or interest which is presently or in the future becomes the subject of dealing in the forward contract trade, or product or byproduct thereof, with a maturity date more than two days after the date the contract is entered into, including, but not limited to, a repurchase transaction, reverse repurchase transaction, consignment, lease swap, hedge transaction, deposit, loan, option, allocated transaction, unallocated transaction, or any similar agreement[...]” 11 U.S.C.A. § 101(25) (2005).
- 3 The final settlement price of the expiring Nymex natural gas contract is determined as a volume weighted average of the prices of transactions executed during the settlement window between 14.00 and 14.30 Eastern Time (during a two-minute window prior to the last day of trading). The price of natural gas for delivery over weekdays (Tuesday to Friday) is determined through transactions that take place during the previous day. The Saturday through Monday price is set on Friday. A subset of these transactions reported to the so-called index publishers is used to derive the daily index price that becomes known in the evening when electronic copies of *Platts* and *Argus* newsletters are distributed (after 17.00 Central).
- 4 In practice, the market participants taking contracts to delivery will hold reasonably big positions that will result in physical flows distributed over the course of a month and flowing at commercially viable levels.
- 5 The discount rate should reflect the credit quality of the counterparty or a credit reserve (credit valuation adjustment, CVA) should be used as an adjustment for credit risk. We shall further discuss this topic in our next book.
- 6 Swaps are explained in the next section.
- 7 The management of credit risk and collateral rules will also be detailed in our next book.
- 8 <http://seekingalpha.com/article/309958-silver-and-margin-requirements-how-the-cme-system-increases-volatility>.
- 9 Again, the quantitative aspects of constructing forward price curves and modelling their dynamics are topics we shall delve into in our next book.
- 10 An example of an algorithm used in construction of a forward price curve is given in Fred Espen Benth, Steen Koekebakker, Fridthjof Ollmar, 2007, “Extracting and applying smooth

forward curves from average-based commodity contracts with seasonal variation," *Journal of Derivatives*, Fall.

- 11 See <http://www.cmegroup.com/rulebook/NYMEX/2/200.pdf>.
- 12 <http://invest.yourdictionary.com/alternative-delivery-procedure>.
- 13 "[If] the markets are free, standardization relieves him of many troubles of marketing which might weigh him down, and it enables him to give his energies to that work which is specially his. He looks to selling his standardized product without inordinate expenses on advertising, or in building up business connections with individual firms." See Alfred Marshall, "Industry and Trade," pp 166 (<http://socserv.mcmaster.ca/econ/ugcm/3113/marshall/Industry%26Trade.pdf>).
- 14 Jeremy C. Kress, "Credit default swap clearinghouses and systemic risk: Why centralized counterparties must have access to central bank liquidity" (<http://ssrn.com/abstract=1583912>).
- 15 "Testimony of René M. Stulz to The House Committee on Financial Services," October 7, 2009 ([http://www.house.gov/apps/list/hearing/financialsvcs\\_dem/prof\\_rene\\_m.\\_stulz.pdf](http://www.house.gov/apps/list/hearing/financialsvcs_dem/prof_rene_m._stulz.pdf)).
- 16 See Covington & Burling LLP, 2010, "CFTC proposes new rules on swap data reporting, recordkeeping and real-time dissemination of swap data," December 1.
- 17 ISDA, 2011, "Block trade reporting for over-the-counter derivatives markets," January 18 (<http://www.isda.org/speeches/pdf/Block-Trade-Reporting.pdf>).
- 18 See also Kenneth R. French, 1983, "A comparison of futures and forward prices," *Journal of Financial Economics*, 12, pp 311–42; Hun Y. Park and Andrew H. Chen, 1985, "Differences between futures and forward prices: A further investigation of the marking-to-market effects," *Journal of Futures Markets*, 5, pp 77–88.
- 19 The term backwardation is intuitively obvious. The term contango goes back to a fee paid in the 19th century by a buyer of securities on the London stock exchange to the seller for the privilege of deferring payment. This payment corresponded to the interest foregone by the seller. It is still prevalent in some exchanges such as the Bombay Stock Exchange, where it is referred to as *Badla*. The word contango is likely to be derived from the word "contingent" or "continuation" (see [www.dictionary.com](http://www.dictionary.com)). As W. S. Gilbert put it in *Utopia*, "A Company Promoter this with special education, Which teaches what Contango means and also Backwardation – To speculators he supplies a grand financial leaven, Time was when two were company – but now it must be seven."
- 20 As explained earlier, the forward contracts are carried on the books at discounted levels, while the futures are carried at the screen values.
- 21 John Maynard Keynes, 1923, "Some aspects of commodity markets," *Manchester Guardian*, reprinted in *Collected Writing of John Maynard Keynes, Volume XII* (London, England: Macmillan), and *The Treatise on Money, Volume II* (New York, NY: Harcourt, Brace and Co); John Richard Hicks, 1946, *Value and Capital: An Enquiry into some Fundamental Principles of Economic Theory* (2e), (Oxford, England: Clarendon Press). See also Yves Simon and Delphine Lautier, 2006, "Marchés dérivés de matières premières," *Economica* (3e).
- 22 One can extend this theory to include the case of normal contango, with the long being net short (see P. Cootner, 1960, "Returns to speculators: Telser vs. Keynes," *Journal of Political Economy*, 68, pp 396–404).
- 23 See, Yves Simon and Delphine Lautier, op.cit., pp 59.
- 24 According to J. R. Hicks, "technical conditions give the entrepreneur a much freer hand about the acquisition of inputs, (which are largely needed to start new process) than about completion of outputs [...]. The desire to hedge planned purchases [...] tends to be less insistent than the desire to hedge planned sales." Today we would say that the processors have more optionality to postpone or shut down temporarily the operations than, for example, farmers. This argument has some merits with respect to the producers of natural resources, although it cannot be elevated to the level of a general law.

- 25 John Maynard Keynes, 1930, *A Treatise on Money* (Vol 2) (London, England: Macmillan), p 143.
- 26 Ibid.
- 27 This should serve as a warning to hedge providers that insisting on excessive short-term profits may kill the goose that lays the golden eggs.
- 28 A detailed discussion of different tests of the normal backwardation theory can be found in the dissertation of Delphine Lautier (see, Delphine Lautier, 2000, "La structure par terme des prix des commodités: analyse théorique et applications au marché pétrolier," Thèse de doctorat, Université Paris IX).
- 29 By convention, the life of a contract is measured in years, with the unit of measurement being one year, and the current time denoted by  $t$  (maturity time is under the standard convention equal to  $T$ ).
- 30 A long position holder benefits from the price increase and has to take the delivery. The opposite is true in the case of a short position. Given growing complexity of market transactions and the provisions to cash-settle many of them, the first part of the definition of a long/short position is more useful in practice. Some energy-related transactions may be quite complex and the exposure may shift between long and short sides, depending on the levels of the underlying prices.
- 31 Robert Jacob Alexander Skidelsky, 2009, *Keynes: The Return of the Master* (New York, NY: Public Affairs), p 73, provides a somewhat hilarious example of limited access to storage: "Once, in 1936, he [J. M. Keynes] even had to take delivery of a monthly's supply of wheat from Argentina on a falling market. He planned to store it in the crypt of King's College Chapel, but found this was too small. Eventually, he worked out a scheme to object to its quality knowing that cleaning would take months. Fortunately by then the price had recovered and he was safe. There were loud cries that "infernal speculators" had cornered the market." One additional thought: complaining about the quality of physical commodity is a standard operating procedure if one wants to avoid taking delivery.
- 32 In addition, using a credit card to provide financing could inflate the costs.
- 33 This does not mean that it is always easy. When Salomon Brothers started doing business in Japan in the 1980s, one of the first stumbling blocks was the need to develop the legal and institutional framework for shorting securities (a precondition to offering many other market transactions and services) that did not exist prior to Salomon's arrival.
- 34 As one can see, the prospect of making or losing money motivates people better than the prospect of enhancing one's academic reputation.
- 35 Most models reviewed use the forward price curve (a collection of forward prices for different maturity horizons) as the starting point. Some models based on the concept of the convenience yield will be reviewed for the sake of completeness, given their popularity, but they are not recommended.
- 36 Occam's razor is a principle formulated by William Occam (Ockham), a 14th century English Franciscan friar. As one could expect, it is expressed in a really short and succinct Latin rule, *entia non sunt multiplicanda praeter necessitate*, or "the beings should not be multiplied beyond necessity."
- 37 We shall continue the discussion in the next book in which we cover the quantitative aspects of the energy markets.
- 38 Interest rate swaps were invented in the 16th century by bankers in the Italian city of Genoa. (See Michele Fratianni and Franco Spinelli, 2005, "Did Genoa and Venice kick a financial revolution in the quattrocento?", working paper, September.
- 39 Spot, in this context, means applying to a specific point in time.
- 40 Libor stands for the London Interbank Offered Rate and denotes the rates at which big banks transact in the London Eurodollar market (ie, lend to, and borrow from, each other). Libor rates are available daily from the British Banker's Association. This is the most widely used short-term interest rate benchmark worldwide.

- 41 When a physically oriented trading operation in Houston was taken over by a big financial institution a few years ago, one of the first questions in the due diligence process was: "How do you construct your Libor curve?" The answer was: "What is Libor?"
- 42 In most financial textbooks, basis denotes the difference between futures and spot prices.
- 43 If one is long 10,000 MMBtus in one contract month and short 10,000 MMBtus in the next month, a risk management system may interpret this as a zero overall position. The limits may sometimes be based on discounted volumes and in this case the same notional volumes would result in a small net position.
- 44 Pricing of natural gas is covered in subsequent chapters.
- 45 In the US natural gas market, the basis is defined with respect to the Nymex final settlement price of the expiring contract (ie, Basis = Local Monthly Index Price – Nymex settlement). This market will be discussed in detail in the chapters on natural gas.
- 46 See also Don M. Chance and Don R. Rich, 1996, "Asset swaps with Asian-style payoffs," *The Journal of Derivatives*, Summer, pp 64–77.
- 47 In the case of most options, the strike price is fixed in the option contract. In some cases, the strike may be unknown at the inception of the contract. For example, the strike may be defined as an average price of the underlying calculated over a certain time window (floating strike option).

## *Energy Markets: Structured Transactions*

This chapter will continue our coverage of energy-related instruments available in the marketplace. As with the instruments discussed in the previous chapter, those reviewed here can be used by individual and institutional investors to acquire an exposure to energy commodities in order to speculate or to invest (with the difference between the two objectives being usually somewhat fuzzy), or to hedge pre-existing risks. Structured transactions examined in this chapter differ in many important aspects from the simple instruments covered in Chapter 4. These differences include:

- ❑ they usually represent a package (a portfolio) of the instruments covered in the previous chapter;
- ❑ they offer exposure to multiple sources of risk by referencing prices of multiple commodities, or even non-energy-related prices or variables (exchange rates, interest rates, weather); and
- ❑ they are often subject to rules and regulations applying to financial instruments, such as bonds and equities.

As in the previous chapter, we will seek to cover the instruments that do not require knowledge specific to a particular commodity. Subsequent chapters will offer examples of the instruments discussed below which have been customised for use in specific markets, such as natural gas or electricity. We start here with a discussion of relatively simple structures that represent packages of plain vanilla swaps and options, before continuing with a look at hybrid instruments, which allow hedgers or investors to acquire exposure to multiple sources of risk, including risks not related

directly to energy prices. We use an example of weather derivatives as an illustration. We will conclude by exploring instruments used by investors and speculators to acquire an exposure to energy commodities (ie, the passive investment instruments: commodity indexes, exchange-traded funds (ETFs) and mutual funds).

### **STRUCTURED TRANSACTIONS**

Stand-alone swaps, forwards and options are referred to in this book as atomic instruments if they satisfy these criteria:

- ❑ they can be priced using generally accepted valuation algorithms;
- ❑ market data parameters required for valuation (ie, arriving at a fair market price of these instruments) are directly observable or can be easily extracted from market data; and
- ❑ there is a reasonably liquid market for these instruments.

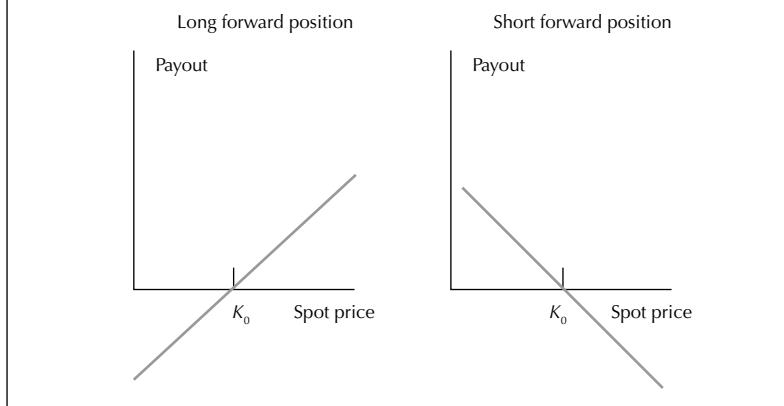
The rationale for using the term atomic is that these instruments can be used as building blocks of more complex transactions, representing portfolios of options, swaptions and swaps. Financial engineering consists to a large extent of breaking up complex transactions into building blocks; simpler transactions that can be easily priced and risk managed. However, a word of caution should be offered at this point. Many structured transactions that originators can dream up do not have the property of separability – ie, they cannot be broken up into portfolios of atomic instruments. For example, a structured transaction may contain an embedded, complicated Asian option for which there is no existing generally accepted valuation model, or the option is not actively traded in the market. We have often received requests to value options such as, for example, spread or Asian options with a maturity of 30 years. Of course, a model for pricing such options can always be developed and inputs can be pulled out of thin air, but the model risk and hedging costs would usually exceed any benefit of offering this product. It is unlikely that the underlying of a 30-year option will be actively traded and the ability to delta hedge will be at best limited, if not non-existent. Simplicity and reliance on the most straightforward and liquid instruments should be the order of the day.<sup>1</sup>

The development of a new market or satisfying customer's needs may sometimes require pushing out the envelope of existing analytical tools and taking significant business risks. There is nothing wrong with that, as long as it recognised that a firm is engaging in a qualitatively new type of business and is acquiring exposures that may be difficult to mitigate. In many organisations for which the author has worked, as long as a trader agreed to take a transaction into their book and manage it, the transaction would be magically transformed into the safest, plain vanilla transaction. There is something to be said about the wisdom of the old adage, "Out of sight, out of mind."<sup>2</sup>

The rest of this section will cover relatively simple structured transactions, combining some very basic atomic instruments. These transactions can be decomposed into portfolios of simpler instruments and valued directly through pricing their component blocks. More market specific examples will follow in subsequent chapters. The transactions will be illustrated with simple diagrams that describe payout patterns at the maturity of a transaction. A payout of a transaction at maturity means that we add up cash flows that happen when a transaction terminates and ignore the initial cost of entering into a transaction. For example, we illustrate how much money an option expiring in-the-money will throw off, but ignore the premium paid to acquire it. Figures 5.1–5.8 depict the transactions that we chose to illustrate structured deals.

Long (short) forwards were explained in Chapter 4. In this section, we shall provide a diagram of the payout profile at the maturity of a contract. A long (short) position is put on at price  $K_0$ . At maturity, the payout of the long forward position has positive (negative) value if the spot price (assuming perfect forward price convergence to the spot) is above (below)  $K_0$ , ie, if  $S > K_0$  ( $S < K_0$ ), with  $S$  denoting the spot price.<sup>3</sup> The opposite is true of a short position.<sup>4</sup> A simple example illustrates this relationship. If long forward position for natural gas at a given location was established (a contract was bought) at the price of US\$3.00/MMBtu and the spot price at maturity is US\$3.50/MMBtu, the holder of the contract may take delivery, pay US\$3 per MMBtu, turn around and sell the entire volume for US\$3.50/MMBTU, earning US\$0.50/MMBtu.

In Figure 5.1, the payout of a forward transaction at maturity is represented by straight lines. The grey line representing the value

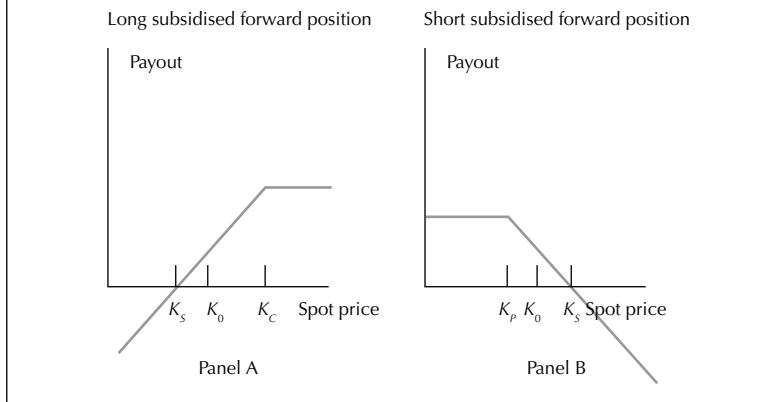
**Figure 5.1** Payout of a forward position

of a forward position at maturity is drawn at the angle of  $45^\circ$  ( $135^\circ$ ) with respect to the horizontal axis, for the long (short) position, respectively.

The subsidised forward combines an option and a plain vanilla forward. In Figure 5.2, a purchase of a forward contract (a long position is established in a forward) is combined with sale of a call option with strike  $K_C$  and with the same maturity as the forward.

The option is sold to the same financial institution that sells the forward contract. The premium received from the sale of a call subsidises the purchase of the forward. The forward contract is acquired at price  $K_S$ , lower than the going market price  $K_0$  for a stand-alone forward. The buyer of the forward contract loses the upside should the spot price increase above  $K_C$ . This is illustrated by the horizontal section of the grey line in Figure 5.2 (Panel A). In the case of a short position in the forward, the seller of a forward contract sells a put struck at  $K_P$ . The premium received from the shorted put improves the terms of the forward transaction from the point of view of the seller. The forward contract is sold at the subsidised price  $K_S$ , better than the current market price of the forward ( $K_0$ ). In case the spot price at maturity is below  $K_P$ , the upside from selling the forward contract is capped.

The subsidised forward is a very simple example of packaging two derivatives into a single structure. It is also a good way to illustrate how the payout diagram can be constructed for portfolios of derivatives. The payout in Panel A of Figure 5.3 is drawn by super-

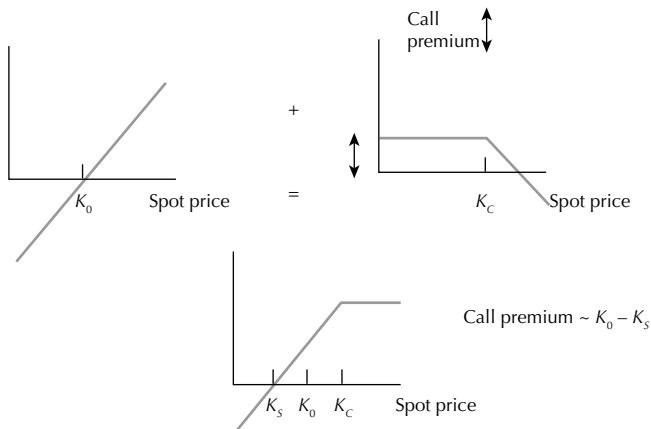
**Figure 5.2** Payout of a subsidised forward position

imposing the payout diagram of a short call onto the payout diagram of a long forward. Improvement in the forward price ( $K_0 - K_S$ ) is equal to the revenue from selling the call option (the option premium).<sup>5</sup> Of course, in practice bid–offer spreads and other transaction costs will reduce this benefit.

Option spreads (as distinguished from spread options) are transactions involving simultaneous sale and purchase of two calls (puts) with:

- the same maturity;
- the same underlying; but
- different strikes.

An example of a call spread is purchase of a call with strike price  $K_1$  and sale of call with strike price  $K_2$  ( $K_1 < K_2$ ), with everything else (the underlying, maturity, etc) being the same. The premium paid for a long call (strike  $K_1$ ) is higher than the premium received from sale of the call option with a higher strike. This means that this strategy produces a net cash outflow at inception. The rationale for engaging in this strategy is the expectation that prices of the underlying will move up but the upside potential may be limited. Selling a call with a higher strike reduces the overall cost of this strategy without giving much of the upside (if the perception that upside beyond  $K_2$  is limited is correct). Another reason for using this strategy is hedging against potential price increases in the commodity underlying the

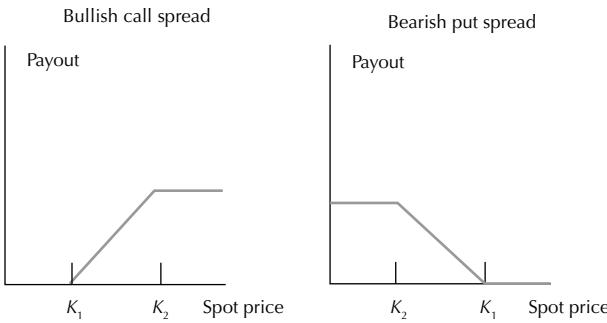
**Figure 5.3** Payout of a subsidised long forward position

call option, with the assumption that protection beyond  $K_2$  is not required. We shall elaborate on this topic in subsequent chapters in an analysis of the operations of natural gas processing plants.

The call spread described above is called a bullish call spread because the buyer of the spread expects prices of the underlying to go up. A bearish call spread would consist in selling a call with a lower strike ( $K_1$ ) and buying a call with a strike  $K_2$  ( $K_1 < K_2$ ). This strategy, associated with positive net cash inflow at inception, could be used in expectation of prices of the underlying being below  $K_1$  at maturity. Should the prices of the underlying end up below  $K_1$ , as hoped for, when the options expire, the seller of the call spread would pocket the net cash inflow (both options, a long and a short call, would expire worthless). Selling a naked call alone would be a very risky strategy. The concurrent purchase of a call with a higher strike is a risk-reduction strategy.

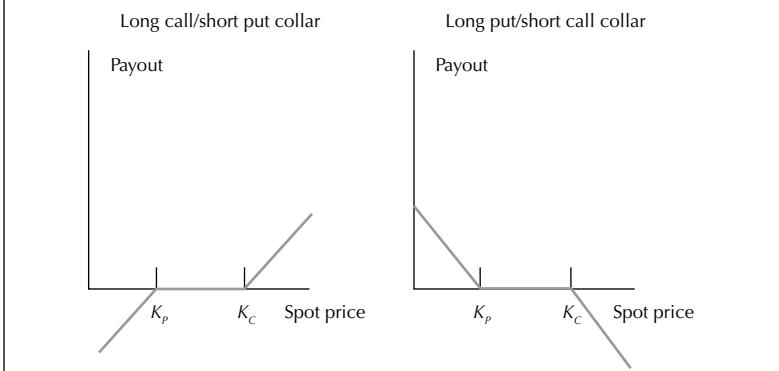
One can engage in similar bullish and bearish strategies through put option spreads. The examples of a bullish call spread and bearish put spread are shown in Figure 5.4.

Both figures ignore initial option premiums (the payout at maturity is shown). I am sure the reader could come up with other examples of option-spread strategies. However, as with other strategies described in this section, one has to keep in mind that the payout diagrams presented here apply to market conditions prevailing at

**Figure 5.4** Examples of option spreads

option/forward expiration. An early termination of these strategies involves additional risks related to the evolution of bid–offer spreads over time and changes in the shape of the volatility smile (volatility levels used to price options with different strikes but otherwise identical). A long option position may make money but buying back a short option may prove to be quite expensive if implied volatility for these options spikes. One may face the necessity of taking the option strategy to expiration.

A collar strategy is one of the most popular strategies used in the commodity markets. A collar consists in the simultaneous purchase of one option (such as a put) and the sale of another option with the same underlying and maturity (a call in this case), to offset the cost of a long option position. The strike prices of the two options are denoted by  $K_P$  and  $K_C$ , respectively. A natural candidate for such a transaction is a producer (ie, a generic producer of, for example, natural gas who can go out of business if the price of natural gas drops), who seeks protection against a drop in prices of the commodity they produce and wants at the same time to reduce the cost of the hedge. The strikes of two options ( $K_C$  and  $K_P$ ) may be calibrated in such a way that the net cash flow at inception is zero (hence the name costless collar) or even positive, which explains the popularity of this transaction. Another example is the sale of a put and the purchase of a call. An example of the use of this collar structure is provided by the case of an airline selling a put and buying a call on crude oil (as a proxy for hedging jet fuel exposure). The rationale behind this strategy is that a jump in oil prices will increase fuel costs and a long call position provides a hedge against this risk. Should oil

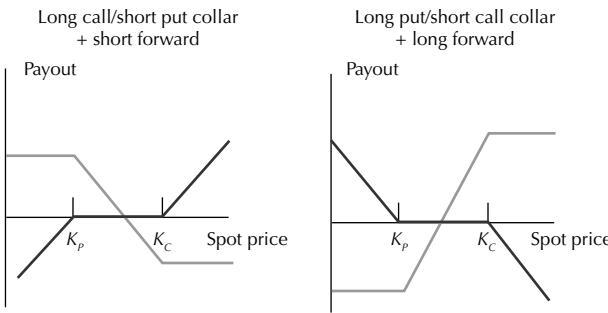
**Figure 5.5** Collar structures

prices drop, an airline may lose money on the short put position but will benefit from the lower fuel prices it buys to operate its planes. The cost of buying a call is partially or fully offset by selling a put. Two types of collars are illustrated in Figure 5.5.

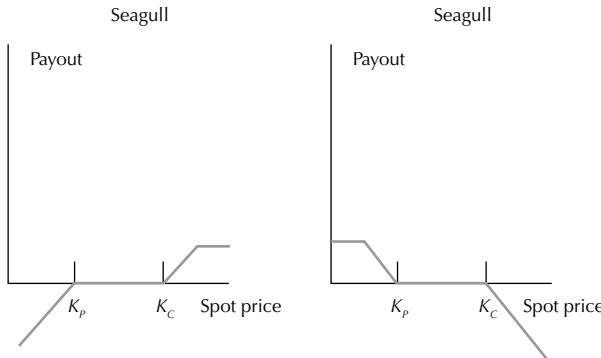
Figure 5.5 illustrates a collar structure in isolation from other positions in the portfolio. It makes more sense to evaluate a collar in the context of a natural long or short forward position (a typical situation for a producer or end-user of energy commodities) that is being hedged (see Figure 5.6). A producer is naturally long the underlying and their unhedged position can be illustrated as being long the forward. Buying a collar (long put, short call) creates a floor (in case market prices drop) but truncates gains (in case market prices go up). This explains why this structure is sometimes called a fence.

Two strategies related to an option collar are a three-way collar and an extendable collar. A three-way collar is, for example, a short put, long call combination associated with the sale of an additional call option with a strike  $K_C^1$  above the strike of a long call. We can use the example of an airline again to explain the rationale for this strategy. The sale of a call helps to reduce the overall cost of the strategy and can be justified if it is unlikely that the price of oil increases beyond  $K_C^1$  at the horizon of this strategy. This strategy is also known as a seagull (see Figure 5.7). Of course, the danger of this strategy is that the price of oil spikes, leaving the airline without adequate price protection.

An extendable collar contains an embedded option to extend the structure at the discretion of the seller for an additional time period.

**Figure 5.6** Collar + forward structure

In other words, the hedging entity sells the option to extend to a provider of the hedge in order to reduce the overall cost of price protection. This may be a very tempting proposition when the hedge is contemplated, but it may be a very costly strategy in the long run. Options to extend are very difficult to model, and a typical commercial user of hedges does not have the apparatus to determine a fair value of this structure. In the worst case, an extendable collar becomes the economic equivalent of an extendable dog lead, trapping a hedging entity in a financially precarious situation for a long period of time. An oft-observed sequence of hedging decisions in the case of a producer is the purchase of a strip of collars (long put, short call positions of different maturities) followed by an upward shift in prices along the entire forward price curve. The overall position has a high overall negative mark-to-market value and the hedge

**Figure 5.7** Seagull

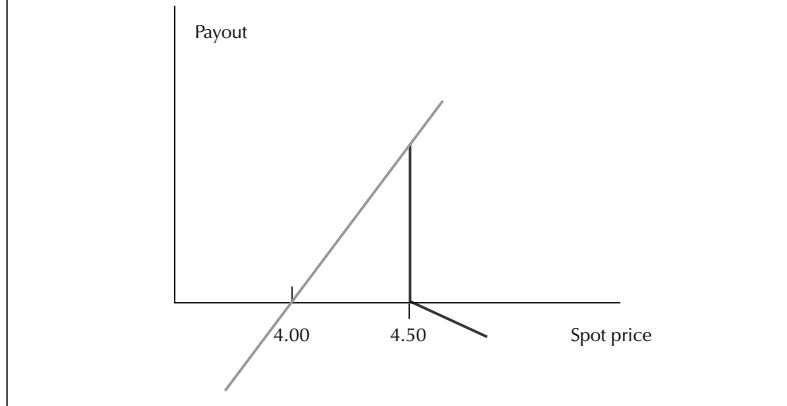
manager (with their reputation and job at stake) starts looking for a solution. A friendly hedge provider comes up with a solution: restructuring of the overall transaction with modification of strikes, extension of maturity and an additional embedded option to extend the entire structure by an additional time period (on top of the firm maturity extension). The outcome for the hedger is in many cases years of misery, a large liability related to an underwater option position and loss of flexibility.

These observations make it clear that a “costless” collar is a seductive but sometimes a very risky proposition. Let us use an example of a producer. Selling a call to offset the cost of a put may be regarded as writing a covered option: the risk of call is covered with future production flows.<sup>6</sup> What is often overlooked is that a shift of a forward price curve upward will reduce the value of long puts and increase the value of short call position. The value of the entire strip of collars will be negative and a producer may be required to post collateral corresponding to the entire hedge position that may extend years into the future. At the same time, they receive the benefits of higher commodity prices one day at a time.<sup>7</sup>

A cancellable forward (see Figure 5.8) is a transaction under which a forward position is extinguished under certain market conditions.

These conditions may be, in principle, related to any market. In practice, they apply to the same market that underlies the forward. An example is a forward contract for which the underlying is the Nymex June 2013 contract. The forward transaction is established at

**Figure 5.8** Cancellable forward



the price of US\$4.00/MMBtu, with the condition that it is cancelled if the settlement price of the underlying Nymex contract on the last day of trading (in May 2013) is equal to, or exceeds, US\$4.50/MMBtu. We have a position which has a potential market value at expiration of US\$4.499(9)/MMBtu per unit (if the settlement price <US\$4.50/MMBtu) or nothing (if the settlement price => US\$4.50/MMBtu). This transaction may be valued by considering a hypothetical portfolio that has the same payout at expiration and is associated with the same cashflows over its life. This portfolio is equivalent to a cancellable forward and, therefore, has the same value, consists of:

- a long forward;
- a short binary option (call) with a strike of US\$4.50/MMBtu and payout of US\$0.50; and
- a short call with a strike of US\$4.50/MMBtu.

Both options have the same underlying and expiration date as the forward. When the settlement price approaches US\$4.50/MMBtu (from the left), the market value of a forward position approaches US\$0.50/MMBtu. When the settlement price hits or exceeds US\$4.50, we want to wipe this value out. This is accomplished by being short a binary option (a bet option) with a lump sum payment of 50 cents when it is in the money. What we accomplished was wiping out the value of the forward at a specific point. The forward position included in the portfolio will, however, continue to accumulate value at prices over US\$4.50. This additional value is wiped out with a short call struck at US\$4.50.

More complex structured transactions will be covered in subsequent chapters.

## WEATHER DERIVATIVES

Weather is a very important risk factor in the energy markets, as was repeatedly pointed out in Chapter 3. A joke from a 2011 book by David Graeber illustrates that there are no easy ways of dealing with this risk:

There was a small town located along the frontier between Russia and Poland; no one was ever quite sure to which it belonged. One day an official treaty was signed and not long after, surveyors arrived to draw a border. Some villagers approached them where they had

set up their equipment on a nearby hill. ‘So where are we, Russia or Poland?’ ‘According to our calculations, your village now begins exactly thirty-seven meters into Poland.’ The villagers immediately began dancing for joy. ‘Why?’ the surveyors asked. ‘What difference does it make?’ [...] ‘It means we’ll never have to endure another one of those terrible Russian winters!’<sup>8</sup>

Unfortunately, we cannot have our town in Texas to secede and join Vermont for the summer and California for the winter. Changing jurisdictions is not a solution. However, the merchant energy business did find a partial solution to weather risk, by designing financial instruments known as weather derivatives.

Weather derivatives are among the best-known examples of hybrid structures – ie, structures combining features of different financial instruments. Weather derivatives became very popular in the late 1990s and later occupied a rather specialised and relatively small niche of the energy markets. In the interest of full disclosure, we have to comment on our contribution to this market. Around 1996–97, this author spent a lot of time working at Enron together with an associate, Ding Yuan, on the development of a new framework for the management of weather risk. Enron executives recognised that the company combined skills in weather fundamentals and trading energy commodities with insights into the impact of weather on the operations and profitability of energy companies. In 1997, Enron executed the first weather derivative trade with Koch Industries, an event that gave both companies the unmistakable right to claim the honour of inventing a new class of financial instruments. We subsequently spent much time introducing the concept to the industry at different speaking venues and in publications. Over time, we noticed a troubling trend: very limited repeat business. Most clients would test the waters by making small transactions but would seldom come back to hedge on a larger scale. A few success stories, such as a pub in London seeking protection against weather conditions driving its patrons away, or the producer of long johns hedging the risk of warm weather around the time their catalogues are sent out, were cited over and over again as proofs of a vibrant market. At the time of writing, this business, as envisaged originally, remains relatively small, as illustrated by the volumes of weather derivatives traded on the Chicago Mercantile Exchange.

What has not materialised, contrary to our hopes and expectations, was a market of actively traded and re-traded contracts,

characterised by high liquidity, narrow bid–offer spreads and low transaction costs, used both by hedgers and speculators. What continues, however, and enjoys healthy growth is a specialised market supported by a cluster of companies which found a profitable niche at the interface of the insurance and energy businesses and succeeded through the introduction of highly innovative products. Many of these companies are run by the veterans from the early days of this business and usually operate as units or affiliates of insurance companies. Another exception is the growing use of options with an energy commodity as the underlying, with exercise contingent on the occurrence of certain weather conditions. The firms active in this space include energy companies operating as hedgers or principals (Constellation, Mercuria, Direct Energy and E.ON), financial companies (JP Morgan, Goldman Sachs and Morgan Stanley), reinsurers (RenRe, Swiss Re and MSI Guaranteed Weather), hedge funds (Tudor, DE Shaw, Cumulus) and brokers (Evolution Markets, ICAP, VCM Partners, MarexSpectron, The Dow Corp).<sup>9</sup> However, the market for weather is still very small by the standards of other derivative markets. According to information posted on the website of the Weather Risk Management Association, “there is now around US\$12 billion of risk being transferred annually in the weather derivatives market.”<sup>10</sup> This market is likely to grow and deserves attention. The rest of this section will cover the reasons why this market failed to take-off as expected, and why it may grow in the future.

### **Weather derivatives: The design**

Weather derivatives have been envisaged as instruments combining the features of derivatives and insurance products, used to hedge weather risk or make directional bets on weather. The instruments would be structured as swaps or options,<sup>11</sup> with environmental variables (temperature,<sup>12</sup> precipitation, wind speed over a defined point or area in space) being used as the underlying. The difference between weather derivatives and weather insurance is that one does not have to prove that actual loss has been sustained (as in the case of weather derivatives); one has to document only that the weather event took place.<sup>13</sup> In other words, a buyer of weather derivatives does not have to own an insurable interest. Another difference between weather derivatives and weather insurance was related to the nature of risks covered by the two instruments. Weather insur-

ance applies to low probability, high-severity exposures, whereas weather derivatives were conceived to include protection against high probability, low-impact events. The industry chose to rely on application of actuarial techniques<sup>14</sup> for the valuation of weather derivatives, making them closer to insurance products than to standard financial instruments. Weather derivatives are cash-settled, as nobody has yet developed a technology for delivering weather.<sup>15</sup>

Weather derivatives may be used as a hedge against a number of different adverse developments, including:<sup>16</sup>

- ❑ operational risk and potential disruptions in commercial operations (for example, airplanes cannot take off if the runway is covered with ice or snow);
- ❑ higher operating costs (for example, low temperatures require frequent de-icing of airplanes and this translates into higher payroll costs and more equipment and materials use);
- ❑ higher volume of inputs a company has to buy and higher volume of outputs a company has to deliver; for example, a local utility has to acquire more natural gas to run peaking units during a heat wave and has to satisfy higher local demand – the use of weather derivatives for management of volumetric risk was promoted as one of the main advantages of weather derivatives; and
- ❑ higher volumes of inputs and outputs may be associated with higher market prices of inputs and higher costs.

### **Weather derivatives: Examples**

#### *Exchange-traded weather derivatives*

Different weather derivative structures can be best summarised by looking at the different instruments offered by the Chicago Mercantile Exchange.<sup>17</sup> Temperature derivatives which were listed (as of February 2012) include:

- ❑ US cooling monthly;
- ❑ US cooling seasonal;
- ❑ US heating monthly;
- ❑ US heating seasonal;
- ❑ US weekly weather;
- ❑ Canada CAT monthly;
- ❑ Canada CAT seasonal;

- Canada cooling monthly;
- Canada cooling seasonal;
- Canada heating monthly;
- Canada heating seasonal;
- Europe CAT monthly;
- Europe CAT seasonal;
- Europe heating monthly
- Europe heating seasonal;
- Asia-Pacific monthly;
- Asia-Pacific seasonal;
- Australia cooling monthly;
- Australia cooling seasonal;
- Australia heating monthly; and
- Australia heating seasonal.

These products include both futures and options. Cooling (heating) derivatives are based, respectively on cooling degree days (CDDs) (heating degree days, HDDs). The difference between monthly and seasonal derivatives is that the first category applies to individual calendar months, the second to entire seasons. Monthly cooling derivatives in the US are available for: May, June, July, August, September plus April and October; monthly heating derivatives for November, December, January, February, March plus October and April.<sup>18</sup> The contracts settle on the CME Degree Days index, calculated as the cumulative total of respective degree days over a calendar month. Seasonal contracts settle based on cumulative degree days over the entire season. For cooling degree days contracts, a season is defined as a minimum of two, and maximum of seven, consecutive calendar months, April–October. For heating degree days, the season is defined as a minimum of two, and maximum of seven, consecutive calendar months, October–April. CAT contracts settle based on cumulative average temperature (the term average applies to a daily average of maximum and minimum temperatures). Of course, the lists of months, definitions of seasons and the list of cities will vary across different locations (Canada, Australia and Europe).

Other weather contracts offered by CME include hurricane, rainfall, snowfall and frost derivatives. Hurricane derivatives available for nine regions include four types of contracts (event, seasonal,

seasonal max, second event) and three types of structures (futures, options, binary options). A standard hurricane contract has payout of US\$1,000 times the respective CME Hurricane Index (the index is calculated by MDA).<sup>19</sup> The details of these and other contracts can be found on the CME website.

The mechanics of settlement of a weather derivatives futures contract can be illustrated with the example of the heating degree day contract for Chicago, for the month of February (of a given year). Suppose that the contract trades at 1,100 HDDs on February 1 and settles at the end of the month at 1,200 (100 HDDs colder weather than initially expected). The buyer would receive 100 HDDs  $\times$  US\$20/HDD = US\$2,000. The amount of US\$20 is the value of one heating degree day, per contract specification.

### **OTC weather derivatives**

OTC weather derivative contracts are structured more like typical swaps or options. An example is provided below:<sup>20</sup>

Structure: Cumulative HDD swap

Term: November 1, 2008 through March 31, 2009

Seller: HeatCo

Buyer: Renaissance Trading Ltd (RTL)

Weather Station: Boston (WBAN #14739)

Strike: 4,342 HDDs

Notional Amount: US\$25,000/HDD

Maximum Payout: US\$10,000,000

(RTL pays below 4342 HDDS; HeatCo pays above 4342 HDDs)

In this example, a company distributing heating oil (HeatCo) enters into a hedge against warm winter weather, which would reduce the volume of fuel they sell to their customers. If the cumulative number of heating degree days is above 4,342 (winter is cold), the hedge provider receives (HeatCo pays) US\$25,000 for each heating degree day above the threshold level. Hopefully, the benefits of cold weather more than offset this expense. In the opposite case, HeatCo receives payments from the hedge counterparty.

The industry offers a number of creative instruments targeting volumetric risk.

Quanto structures<sup>21</sup> offer protection against two risk factors (such as volume and price) that enter into the payout definition in a multi-

plicative way. An example is a contract that hedges the exposure of a natural gas local distribution company to two potential adverse scenarios:

- ❑ a cold winter that increases throughput (ie, volumes sold to the customers), although it may be a mixed blessing if the utility turns out to be under-hedged and has to procure additional supplies of natural gas at high prices; and
- ❑ a warm winter that reduces throughput and forces the utility to dispose at low prices some of the natural gas acquired in anticipation of more normal weather.

The payout of a quanto contract is given as:<sup>22</sup>

$$\text{Volume} \times (HDD_{var} - HDD_{fix}) \times (P_{var} - P_{fix})$$

where

$HDD_{var}$	stands for the realised number of heating degree days
$HDD_{fix}$	the fixed strike for heating degree days, agreed in the contract
$P_{var}$	the floating price of natural gas
$P_{fix}$	the fixed price (strike price) of natural gas

This contract is effectively a swap based on two variables. The payout is illustrated in Figure 5.9. A utility receives payments for the following events.

- ❑ The realised number of heating degree days is above the strike ( $HDD_{fix}$ ). This means that there were more cold days than expected, resulting in the throughput exceeding planned levels. Concurrently, natural gas is expensive ( $P_{var} > P_{fix}$ ), as one could expect during cold winter. The utility has to buy more gas in the market and the additional expense is offset through cashflows from the weather contract. Both differences in the contract payout formula are positive, the result is positive as well.
- ❑ The realised number of heating degree days is below the strike  $HDD_{fix}$ . This means that there were fewer cold days than expected and the utility sells less gas than expected. The utility may have to sell excess amounts of gas (the utility bought too much gas for the winter) in the over-supplied market. A utility

may incur a loss on these transactions and the regulators may refuse to allow recovery through higher rates. After all, natural gas prices were low, so it would be difficult to justify a surcharge on rates to the ratepayers. Both differences in the payout formula are negative and this produces a positive result.

In Figure 5.9, the utility receives payments in the North West and South East quadrants. The utility makes payments to the seller of protection in the other cases (South West and North East).

Quanto structures are probably the fastest growing segment of this market. As reported in *Energy Risk*:

Standard vanilla weather derivatives volumes remain relatively lack-luster. However, more and more end-users are waking up to the value of the quantity-adjusting option (quanto) derivative, where weather acts as a trigger to an underlying commodity price payout. As such, there have been significant strides in the market for these tailor-made quanto products. A recent quanto deal was reported to be worth US\$100 million.<sup>23</sup>

Another popular structure is a swap/option on an energy commodity, contingent on a weather event. For example, an electric utility in California or in Chile (both transactions actually happened) could buy a call on fuel (natural gas or LNG) which is knocked-in (the option becomes alive) if precipitation at a specific location drops below a certain level. The rationale for this transaction is the dependence of the utility on hydropower. Low precipitation levels reduce the levels of water in the reservoirs and river flows, increasing dependence on thermal generation units. A bad hydro year is likely to coincide with high prices of natural gas (as demand for this fuel

**Figure 5.9** Quanto weather swap

		$P_{\text{var}} - P_{\text{fix}} > 0$	$P_{\text{var}} - P_{\text{fix}} < 0$
$\text{HDD}_{\text{var}} - \text{HDD}_{\text{fix}} > 0$	Utility receives cold winter, expensive gas	Utility pays	
	Utility pays		Utility receives warm winter, cheap gas

Source: Author's own

increases), and a long-option position offers a protection against this risk exposure. Conditioning the option payout both on precipitation and natural gas prices lowers its price. Figure 5.10 illustrates how this specific option applies only to the subset of the potential states of the world, resulting in lower price, compared to an option defined with respect to natural gas or precipitation only.

An example of such a structure is shown below:<sup>24</sup>

Structure: Weather-contingent, natural gas calls

Term: June 1, 2008

Weather Location: Grand Island (Coop #253395, WBAN #14935), NE

Weather condition: Accumulation of less than 3.25 inches of cumulative precipitation

Gas strike: US\$8.50/MMBtu

Daily volume: 20,000 MMBtu per 1/10th inch

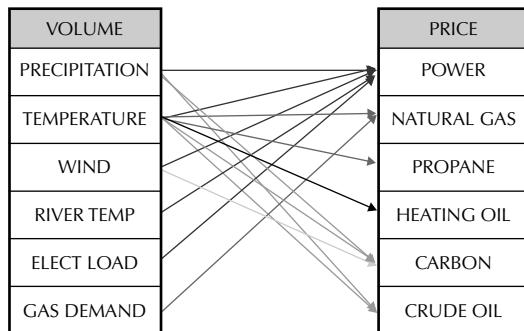
Settlement: If the weather condition is struck, AgCo owns US\$8.50/MMBtu natural gas calls on 20,000 MMBtu per 1/10th of an inch of rain.

Figure 5.11 illustrates how different environmental variables translate into price impacts for different energy commodities. This can be seen as a road map for different double-trigger options, referencing both prices and weather conditions.

An alternative reason for weather derivatives is related to the accounting treatment of different instruments. A weather-related contract with an insurance company can be transformed into a derivative by inserting another company (a transformer) between an

**Figure 5.10** Natural gas option contingent on precipitation

		Natural gas prices	
		Low	High
Precipitation	High		
	Low		Call option is triggered

**Figure 5.11** The impact of weather on prices of energy commodities

Source: Galileo.

insurance company and its client. There may be many reasons for the use of a transformer, including the varying treatment of derivative instruments in different jurisdictions, different tax laws or accounting rules.

### Weather derivatives: Pros and cons

One could ask the question why weather derivatives did not fulfill its initial promise and what specialised niches remain profitable. In the interest of fairness, we should mention that weather derivatives have many enthusiastic supporters (including the author and we continue to see the merits of customised weather protection). Of course, our observations may be tainted by a sense of disappointment and wishful thinking, but we think a few useful points can be made.

After they were launched in the late 1990s, weather derivatives enjoyed a period of rapid growth fuelled by the success of a number of energy companies in this field. Some market participants made significant investments in corrections of historical temperature and precipitation data, which gave them a significant advantage over the counterparties relying blindly on raw historical time series provided by government agencies. Some historical data contained errors and omissions (missing data points, misplaced decimal points) distorting historical distributions used in the calibration of weather derivatives' pricing models. A related issue is the correction of the data for the so-called "human island" phenomenon, describing the local climate impact of economic activities or growing population density. A local

airport might have grown from a sleepy backwater into an airline hub, with hundreds of planes landing and taking off every day. The readings from the weather stations located at such an airport could be affected over time, and even small deviations from historical norms could translate into huge differences in valuations and financial outcomes of weather-related contracts. In the late 1990s, a few weather derivatives traders became aware of these potential problems and invested time and resources in studying and fixing the historical time series. They made a lot of money trading against less sophisticated counterparties, and their clever games are now legend in energy trading. At some point in time, this business went away as numerically savvy MBAs separated numerically challenged MBAs from their money (yes, it helps to study mathematics). Some merchant energy companies took a directional view of weather and went against insurance companies that were relying on 50–60 years of data. The trends towards global warming and strong weather patterns in the late 1990s (El Niño), favoured less conventional approaches, relying on shorter time series. Over time, such sources of profitability evaporated as the money-losing counterparties withdrew from the market and the initial enthusiasm for weather derivatives subsided.

Weather derivatives were originally designed to hedge high probability/low severity risk – and this was probably a mistake. We tend to spill coffee on ourselves every day but we do not call on a highly priced Wall Street banker to manage this problem. We treat it as a cost of doing business and learn to live with it. It is more economically rational to incorporate such costs in the price structure and treat them as a recurring and predictable expense. Using derivatives to cover highly predictable expenses puts the traditional insurance principle (“The many will pay for the losses of the few”) on its head. It turns into an unsustainable proposition of “The few will pay for the losses of the many.” In the interest of fairness, one has to recognise that the industry identified this problem and many weather traders were looking for counterparties with the opposite exposures to the same weather conditions in order to act as intermediaries, inserting themselves between them.<sup>25</sup> A weather desk would cover the losses of one group with profits of another group, capturing commissions and bid–offer spreads. In practice, finding such opportunities of offsetting exposures and consenting adults willing to trade proved fairly difficult.

The industry settled relatively early on pricing models for weather derivatives based on an actuarial approach. With this method, the price of the derivative is an expected payout (calculated using historical data to derive empirical probability distribution) plus a risk premium. The differences of opinions regarding the number of observations to be used, specific probability distributions used to fit historical data, and the need to fix the data given their imperfections, supported trading activity. As long as both counterparties believed they were taking the other side for a ride, the market flourished. Pricing of weather derivatives based on this approach usually clashes with a willingness to pay for the protection offered by weather derivatives. Given the recurring nature of the risks they are supposed to hedge against, potential buyers have less expensive solutions as a fall-back. Instead of incorporating weather risk in their basic pricing structures, they can use borrowing to smooth cashflow fluctuations. The cost of borrowing may be much lower than the risk premium required by the weather derivative seller. Those companies that cannot easily borrow are not likely to be treated by financial firms as desirable clients to whom one would sell complex financial instruments. To make this example more specific, a big hotel chain with operations on many continents may easily absorb the losses related to a rainy summer in one location. Even without geographical diversification, a hotel chain may use a short-term loan or fall back on cash reserves when business slows down. This is especially true in the low interest rate environment. A small hotel in Colorado may go out of business after a low precipitation winter, but is unlikely to court Wall Street to buy weather protection.

Given the many reasons for being skeptical with respect to weather derivatives, one can ask the questions – what types of weather derivatives survived the test of time and why? There are two important reasons for using weather derivatives:

- they are seen as a hedge against volumetric risk; and
- the demand for energy commodities may be contingent on the realisation of certain weather conditions.

Volumetric risk is a significant exposure faced by most energy companies, as the majority of contracts contain embedded explicit or

implicit volume optionality that a counterparty may take advantage of. Unpredictable fluctuations in volumes demanded by customers are usually related to weather conditions. If abnormal weather conditions persist, volumetric impacts can accumulate, leading to significant losses. Using weather derivatives for hedging combined volumetric/price risk may often be a fairly ineffective strategy. On the one hand, extreme weather conditions may have no impact on market conditions. A heatwave when all the generating units are available may not translate into a spike in electricity prices. On the other hand, a price spike may happen during normal weather when many outages take place at the same time. Given the risk of the decoupling of weather and volumetric impacts, one can ask why we should bother to use weather contracts at all. The answer is that it may be often difficult to verify private volume information and this information may be distorted on purpose. Weather information is objective and publicly available. It is interesting that, in the cases where volume information is available and can be relied on, contracts based directly on quantity, without the weather detour, have developed. Such contracts cover risks related to the overall system load in different power pools.

Conditioning the payout of a derivative on weather has two basic advantages to hedgers. The cost of protection is reduced, as the contract applies to a smaller set of the future states of the world (states characterised by certain weather conditions). In contracts with physical delivery of the underlying, a hedging entity does not receive the commodity it does not need (due to weather conditions). For example, as a Chilean utility can fall back on hydropower, it does not need natural gas for the thermal power plants. Using an option defined with respect to one source of risk may result in logistical complications (the need to dispose of the commodity received under the contract when it is not needed) or may require unwinding the contract before delivery is taken. Unwinding a contract under duress is often quite expensive. An option to buy natural gas, which can be acted on only if precipitation over a defined time period is low, turns out to be a better solution.

## PASSIVE INVESTMENT INSTRUMENTS

We will now discuss instruments used by investors and speculators to acquire an exposure to energies. These are known as passive

investment instruments and come in the following forms: derivatives settling on commodity indexes (such as the Goldman Sachs Commodity Index), ETFs and commodity-oriented mutual funds. A discussion of financial instruments that investors can use to acquire exposure to commodity prices would not be complete without examining these alternatives. These instruments are sometimes very controversial for several reasons. As will be explained later in this chapter, critics believe that they contribute to upward price pressures in many markets and also increase price volatility.

### **Commodity indexes**

A commodity index tracks returns on a basket of commodities or commodity-related financial instruments over a specified time period. Most indexes used in the energy markets represent portfolios of futures contracts and share many common design features. In practice, they offer the ability to invest in commodity futures without worrying about the mechanics of trading and margin management. The index can be used as a measure of commodity price trends and as a benchmark for settlement of different derivative OTC and exchange-traded contracts. A few indexes gained wide recognition and are used extensively by most market participants, although in principle one can design highly customised indexes to meet the needs of institutional investors with unique objectives or funding constraints. The definition of any index has to specify (at a minimum) the following items:<sup>26</sup>

- index composition;
- index weights;
- frequency of rebalancing;
- use of leverage; and
- return calculation.

Specific examples of different designs are provided below. A leveraged index is a basket of futures, with the returns reflecting significant leverage offered by these instruments (an investor has to come up with the initial margin and potential variation margins, as opposed to paying the full price of the underlying basket). An unleveraged index deploys cash equal to the value of a commodity basket in risk-free fixed income instruments, with the interest payments reinvested

in the index. The return calculation can be based on an arithmetic or geometric average. The geometric average return is defined as:

$$r_g = -1 + \left[ \prod_{t=1}^n (1+r_t) \right]^{1/n} \quad (5.1)$$

where  $r_t$  denotes returns realised in period  $t$ .

### *Goldman Sachs Commodity Index*

The most widely followed commodity index is the Goldman Sachs Commodity Index (S&P GSCI), which was introduced in 1991.<sup>27</sup> Three S&P GSCI indexes are available: excess return, total return and spot return. The excess return index measures the returns from an uncollateralised basket of nearby commodity futures, the total return index measures the returns from a fully collateralised portfolio of nearby futures and the spot index measures the level of nearby commodity prices.<sup>28</sup> The indexes are based on a basket of 24 liquid commodity-related futures, rolled forward from the fifth to the ninth business day of each month (at the rate of 20% of contracts per day).<sup>29</sup> Of course, the commodities with the futures expiring on a quarterly, and not monthly, basis do not have to be rolled every month. The roll is necessary as futures are instruments with a finite life and an expiring contract has to be replaced with the next available one. Rollovers distributed over a few days are used to avoid market disruptions. The weights used in the construction of the index correspond to global production volumes.

The weights are generally updated every year. Table 5.1 provides information on index composition and the rollover rules for the S&P GSCI. CPW stands for the contract production weight, ACRP for the average contract reference price, PDW for percentage dollar weight, RPDW for reference percentage dollar weight and TDVT for total dollar value traded. Table 5.2 shows the dollar weights for 2011 for the S&P GSCI. As one can see, the index is heavily weighted towards energy (70.5% as of December 2011).

In addition to the basic S&P GSCI index, there are any sub-indexes and specialised indexes available from the same source, including S&P GSCI Reduced Energy, Light Energy and Ultra-Light Energy, as well as forward month versions (from one to five months).<sup>30</sup>

**Table 5.1** S&P GSCI details, contracts included in the S&P GSCI for 2012

Trading Facility	Commodity	Ticker	2011 CPW	2012 CPW	2012 ACRP (\$)	Unit	2011 PDW	2012 RPDW	2012 TDVT (USD billions)	2012 TVM	Designated at Month Begin	Contract Expiration at Month Begin
CBT	Chicago Wheat	W	18188.56	18217.58	7.466	bu	3.28%	3.23%	919.4	124	H H K K N N U U Z Z Z H	
KBT	Kansas City Wheat	KW	4134.2	5004.071	8.333	bu	0.83%	0.99%	281.9	124	H H K K N N U U Z Z Z H	
CBT	Corn	C	28210.87	29648.15	6.6	bu	4.50%	4.64%	2771.1	259.7	H H K K N N U U Z Z Z H	
CBT	Soybeans	S	7708.699	8037.317	13.38	bu	2.49%	2.55%	2913.6	496.8	H H K K N N X X X X F F	
ICE-US	Coffee	KC	16710	17406.22	2.472	lbs	1.00%	1.02%	474.9	202.3	H H K K N N U U Z Z Z H	
ICE-US	Sugar #11	SB	340773.4	344724.8	0.278	lbs	2.29%	2.28%	870.6	166.3	H H K K N N V V V H H H	
ICE-US	Cocoa	CC	4.015306	4.116321	3085.75	MT	0.30%	0.30%	139.3	201	H H K K N N U U Z Z Z H	
ICE-US	Cotton #2	CT	51632.55	53411.21	1.41	lbs	1.76%	1.79%	435	105.9	H H K K N N Z Z Z Z H	
CME	Lean Hogs	LH	70271.76	72823.44	0.865	lbs	1.47%	1.49%	325.3	94.7	G J J M M N Q V V Z Z G	
CME	Live Cattle	LC	91458.23	92591.82	1.102	lbs	2.44%	2.42%	571.2	102.7	G J J M M Q Q V V Z Z G	
CME	Feeder Cattle	FC	13417.1	13596.46	1.279	lbs	0.42%	0.41%	97.4	102.7	H H J K Q Q U V X F F	
NYM/ICE	Crude Oil	CL	14314	13557.23	94.111	bbl	32.59%	30.49%	22038.9	316.7	G H J K M N Q U V X Z F	
NYM	Heating Oil	HO	72571.85	71569.8	2.802	gal	4.92%	4.80%	3463.8	316.7	G H J K M N Q U V X Z F	
NYM	RBOB Gasoline	RB	72504.78	73694.1	2.714	gal	4.76%	4.78%	3454.5	316.7	G H J K M N Q U V X Z F	
ICE-UK	Brent Crude Oil	LCO	6262.977	6959.701	105.134	bbl	15.93%	17.14%	12639	316.7	H J K M N Q U V X Z F G	
ICE-UK	Gasoil	LGO	313.6761	359.2745	879.063	MT	6.67%	7.36%	5455.4	316.7	G H J K M N Q U V X Z F	
NYM/ICE	Natural Gas	NG	28797.24	28984.31	4.273	MMBtu	2.98%	2.94%	5275.1	781	G H J K M N Q U V X Z F	
LME	Aluminium	MAL	41.288	42.53	2512.938	MT	2.51%	2.53%	3351.3	575	G H J K M N Q U V X Z F	
LME	Copper	MCU	16.62	17.14	9194.146	MT	3.70%	3.74%	7477.9	870.1	G H J K M N Q U V X Z F	
LME	Lead	MPB	7.574	7.872	2514.708	MT	0.46%	0.47%	631.7	585.1	G H J K M N Q U V X Z F	
LME	Nickel	MNI	1.286	1.352	24796.583	MT	0.77%	0.80%	1138.6	622.8	G H J K M N Q U V X Z F	
LME	Zinc	MZN	10.68	11.04	2336.917	MT	0.60%	0.61%	1174.7	834.9	G H J K M N Q U V X Z F	
CMX	Gold	GC	78.12632	76.58309	1476.492	oz	2.79%	2.68%	7224.8	1171.6	G J J M M Q Q Z Z Z G	
CMX	Silver	SI	649.4452	665.5205	34.085	oz	0.54%	0.54%	3573.1	2888.3	H H K K N N U U Z Z Z H	

Source: S&P GSCI index methodology (May 2012) <http://www.standardandpoors.com/>

**Table 5.2** S&P GSCI index component index weights  
(December 30, 2011)

<b>Commodities</b>	<b>Dollar weights (%)</b>
<i>Energy</i>	
Crude oil	34.90
Brent crude oil	16.50
Unleaded gasoline	4.70
Heating oil	5.20
Gasoil	7.10
Natural gas	<u>2.10</u>
	<b>70.50</b>
<i>Industrial metals</i>	
Aluminium	2.00
Copper	3.10
Lead	0.40
Nickel	0.60
Zinc	<u>0.50</u>
	<b>6.60</b>
<i>Precious metals</i>	
Gold	3.00
Silver	<u>0.50</u>
	<b>3.50</b>
<i>Agriculture</i>	
Wheat	2.90
Kansas wheat	0.70
Corn	4.50
Soybeans	2.30
Cotton	1.20
Sugar	2.00
Coffee	0.90
Cocoa	<u>0.20</u>
	<b>14.70</b>
<i>Livestock</i>	
Feeder cattle	0.50
Live cattle	2.70
Lean hogs	<u>1.50</u>
	<b>4.70</b>

Source: [www.spindices.com](http://www.spindices.com) Fact Sheet

### *Dow Jones–UBS Commodity Index*

The Dow Jones–UBS (DJUBS) Commodity Index was created by AIG International Inc in 1998 and acquired by UBS Securities LLC in May 2009. UBS signed an agreement with Dow Jones to market the index through a joint venture. Dow Jones assigned its interest to CME Indexes. The index (or more precisely the family of related indexes) has many similarities and some significant differences with the S&P

GSCI. The DJUBS Commodity Index is based on exchange-traded commodity futures contracts. Many different sub-indexes representing major commodity sectors are included within the broad index: energy, petroleum, precious metals, industrial metals, grains, livestock, softs and agriculture. The futures are rolled approximately every other month, with a schedule defined for each commodity in the index. The roll happens over a five-day period, starting with the fifth business day of the month.

The weights used by the DJ-UBS index are constructed using a two-stage procedure. In June of each year, the commodity liquidity percentage (CLP) is determined using a five-year average of the product of the trading volume multiplied by the US dollar value of the futures contract selected to represent a given commodity. This product is divided by the sum of similar products for other commodities. The commodity production percentage (CPP) is calculated in a similar way, using production data in place of the trading volumes. The commodity index percentage (CIP) is established by combining the two indicators described above (CLP and CPP) using a ratio of 2:1. The CIP is adjusted further to better achieve the objective of diversification. On the fourth business day in January, the CIPs are combined with the prices of the commodities included in the index, forming the commodity index multiplier (CIM), which remains unchanged for the entire year. The daily value of the index is calculated by multiplying the CIMs by the US dollar prices of the designated contracts chosen to represent the commodities included in the index. The current weights for the DJUBS Commodity Index are shown in Table 5.3.

As in the case of S&P GSCI, there are many sub-indexes available that track single commodities and commodity groups.

#### *Thomson Reuters/Jefferies CRB Index*

The CRB (Commodity Research Bureau) Index is the oldest index around, established initially as a research and statistical tool in 1957. It includes 19 commodities (aluminium, cocoa, coffee, copper, corn, cotton, crude oil, gold, heating oil, lean hogs, live cattle, natural gas, nickel, orange juice, silver, soybeans, sugar, unleaded gas and wheat) and trades on the ICE futures exchange. The index mutated into the Reuters/Jefferies CRB Index in 2005, when it was revised through a collaboration with Reuters (now Thomson Reuters) and Jefferies

**Table 5.3** Dow Jones UBS Commodity Index  
(2011 target weights)

Commodity weight	(%)
Gold	10.26%
Natural gas	9.12%
WTI crude oil	8.92%
Soybeans	8.59%
Copper	7.50%
Corn	6.23%
Aluminium	6.09%
Wheat	5.54%
Brent crude oil	5.16%
Live cattle	3.89%
Unleaded gasoline	3.63%
Soybean oil	3.52%
Zinc	3.41%
Sugar	3.39%
Heating oil	3.31%
Silver	2.90%
Lean hogs	2.48%
Nickel	2.41%
Coffee	2.02%
Cotton	1.62%

Source: Dow Jones–UBS Commodity Index, Fact Sheet, [http://www.djindexes.com/mdsidx/downloads/fact\\_info/Dow\\_Jones-UBS\\_Commodity\\_Index\\_Fact\\_Sheet.pdf](http://www.djindexes.com/mdsidx/downloads/fact_info/Dow_Jones-UBS_Commodity_Index_Fact_Sheet.pdf)

Financial Products, LLC. The index was renamed the Thomson Reuters/Jefferies CRB Index in 2009 after the 2008 merger of the Thomson Corporation and the Reuters Group PLC. The most recent revision, in 2005, established the following weights for the most important commodity groups:

- ❑ energy: 39%;
- ❑ agriculture: 41%;
- ❑ precious metals: 7%; and
- ❑ base/industrial metals: 13%

Details of the index weights are shown in Table 5.4.

The Thomson Reuters Equal Weight Continuous Commodity Index comprises 17 commodity futures that are continuously rebalanced (cocoa, coffee, copper, corn, cotton, crude oil, gold, heating oil, live cattle, live hogs, natural gas, orange juice, platinum, silver,

**Table 5.4** Thomson Reuters/Jefferies CRB Index weights (percentages)

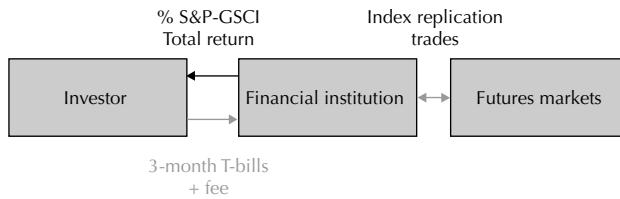
<i>Energy</i>	
RBOB gasoline	5
Natural gas	6
Heating oil	5
Crude oil	<u>23</u>
	<b>39</b>
<i>Agriculture</i>	
Cocoa	5
Lean hogs	1
Orange juice	1
Wheat	1
Coffee	5
Cotton	5
Sugar	5
Live cattle	6
Soybeans	6
Corn	<u>6</u>
	<b>41</b>
<i>Precious metals</i>	
Gold	6
Silver	<u>1</u>
	<b>7</b>
<i>Base metals</i>	
Copper	6
Aluminium	6
Nickel	<u>1</u>
	<b>13</b>

Source: [http://www.jeffries.com/pdfs/TRJCRB\\_Index\\_Materials.pdf](http://www.jeffries.com/pdfs/TRJCRB_Index_Materials.pdf)

soybeans, sugar No. 11 and wheat), using equal weights (5.88%). The index is sometimes referred to as the “old” CRB Index.

Investors have several different choices when it comes to acquiring an exposure to an index. The most popular solution is entering into an OTC commodity index swap with a financial institution. Here, an investor receives total return of a specific index over a defined time period, and typically pays a three-month Treasury bill rate and an agreed management fee.<sup>31</sup> The financial institution hedges its exposure under an index swap actively trading the futures contracts, to offset its exposure under a swap (see Figure 5.12). Other choices include, for example, exchange-traded instruments based on commodity indexes.

**Figure 5.12** Commodity index swap (assumes a swap based on the S&P GSCI)



### Exchange-traded funds and notes

ETFs are important to energy trading for a number of reasons. They have increased the interest of individual and institutional investors in the commodity markets and allowed them to acquire exposure to certain markets at a low cost and in a time-effective way. Certain unique features of the ETF market design support-trading strategies carried out on a recurring basis by some of the most sophisticated market participants. We cover the general principles underlying ETF design first, and then use the example of the United States natural gas fund (UNG) to put them in the context of a specific energy-related instrument.

ETFs are securities designed to track market performance of a specific basket of securities, an index (such as S&P 500 or a commodity index), or a specific commodity (such as natural gas or oil).<sup>32</sup> Some commodity ETFs represent portfolios of futures (with some marginal positions in other instruments), such as, for example, United States Oil (USO),<sup>33</sup> or represent mostly physical holdings of a specific commodity (for example, gold (GLD) or silver (SLV)).<sup>34</sup> ETFs are hybrid instruments sharing certain features of mutual funds and closed-end funds. Like close-end funds, they trade continuously on exchanges, like any listed stock, at prices fluctuating from one transaction to another. This is different in comparison with a mutual fund. A mutual fund investment or redemption is based on net asset value (NAV) calculated at the end of the trading day. Like mutual funds, ETFs are operated by professional managers and created by sponsors representing the most experienced and trusted financial institutions. The purchases of shares from, and sales of shares to, the fund manager can be executed only by authorised participants (APs), who exchange creation units (large blocks of shares, typically in tens or hundreds of

thousands) for basket of underlying assets. The APs provide liquidity and maintain close relationships between the market prices of ETFs and the baskets of instruments or commodities that these ETFs represent. This can be explained using an example of a market situation in which the ETF market price drops below the value of the instruments they represent. In this case, the APs will buy the shares in the open market, form the creation units, redeem the units that they receive from the ETF custodian for the underlying instruments and sell them to lock-in profit. Understanding the role of APs and the details of mechanics of the transactions they engage in is key to the design of ETF-related trading strategies.<sup>35</sup> ETF investors trade the shares created through this process in the secondary markets.

ETFs offer a number of benefits to the investors, including low transaction costs and low structuring costs. They are generally not actively managed and this feature translates into lower marketing, distribution and accounting expenses (generally the annual costs are around 50 basis points, although some actively managed ETFs may be more expensive). An investor can acquire exposure to a given sector within seconds or minutes, and can rebalance their portfolio structure in a very efficient way. ETFs can be sold short or purchased on margin, increasing an investor's flexibility in asset-allocation decisions. ETFs are also tax-efficient, as they are characterised by low turnover and do not have to sell securities to meet redemptions. This means that an investor, in general, will not be surprised by unexpected capital gains and losses: a potential risk faced by mutual fund investors (except for those related to their own transactions).<sup>36</sup>

UNG is a ticker for a natural gas ETF. The design of UNG is described as follows:

The United States Natural Gas Fund, LP ("USNG") is a Delaware limited partnership organized on September 11, 2006. [...] USNG is a commodity pool that issues limited partnership interests ("units") traded on the NYSE Arca, Inc. [...]. [...] The investment objective of USNG is for the changes in percentage terms of its units' net asset value ("NAV") to reflect the changes in percentage terms of the spot price of natural gas delivered at the Henry Hub, Louisiana, as measured by the changes in the price of the futures contract for natural gas traded on the New York Mercantile Exchange (the "NYMEX"), that is the near month contract to expire, except when the near month contract is within two weeks of expiration, in which case it will be measured by the futures contract that is the next month contract to expire, less USNG's expenses.<sup>37</sup>

**Figure 5.5** UNG holdings (as of June 12, 2012)

<b>Security</b>	<b>Quantity</b>	<b>Price</b>	<b>Market value</b>
Natural gas futures and other natural gas interests			
Total return swaps	US\$149,989,564.50		
NYM Nat.Gas Future NG JUL12	20,722	US\$2.232	US\$462,515,040.00
NYM Nat.Gas Future NN JUL12	17,952	US\$2.232	US\$100,172,160.00
ICE LOT Nat.Gas Clrd Swap JUL12	34,367	US\$2.232	US\$191,767,860.00
US Treasuries			
US T BILL ZCP 6/21/12	250,020,000	US\$100.00	
	US\$250,018,333.32		
<b>Cash</b>			
US dollars	799,665,288	US\$1.00	US\$799,665,288.36

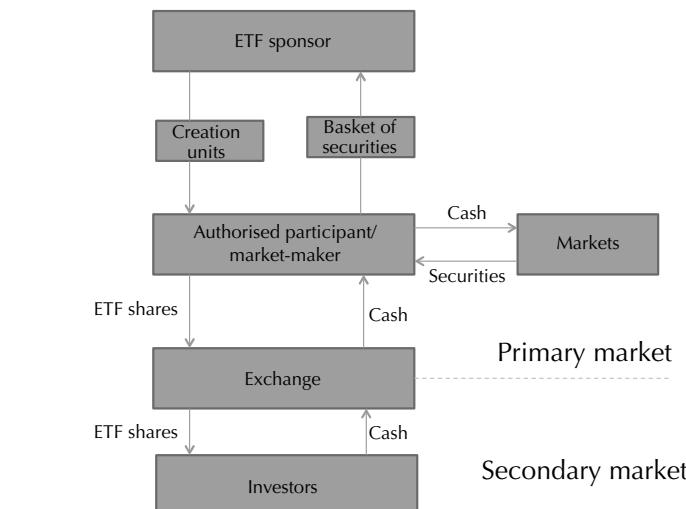
Source: <http://www.unitedstatesnaturalgasfund.com/ung-holdings.php><sup>38</sup>

The USNG is managed by its general partner, USFC. The net assets of the USNG include primarily natural gas futures, but may include investment contracts for other energy commodities such as crude oil, heating oil, gasoline, acquired on Nymex, ICE and in the OTC markets (including forwards, swaps and options). For example, daily holdings (as reported on June 12, 2012) on the fund's website included the instruments listed in Table 5.5.

The natural gas futures positions of the fund are concentrated in the first available contract and in the instruments settling on this contract. This is not true of the rollover period, when the fund holds a mix of the first and next available futures, as positions in the expiring contract are liquidated and positions in the next available contract are established. The details of rollover process can be found in the UNG SEC filing.<sup>39</sup>

The units created by the USNG can be acquired only by APs in blocks of 100,000 units (called creation baskets). The payment for a basket is equal to the aggregate NAV of units included in a basket (plus a US\$1,000 fee), calculated at the end of the business day. An analogous procedure applies to redemptions of the units. Whenever a basket is sold, USFC purchases natural gas interests with the value approximating the amount of cash received upon the issuance of a basket.

Understanding the rollover schedule is important because the trading community often pre-positions itself for the rollover of a fund's holdings. It may happen on occasion that a fund may choose

**Figure 5.13** An ETF unit creation mechanism

Source: Srichander Ramaswamy, 2011, "Market structures and systemic risks of exchange-traded funds," BIS Working Papers No 343.

not to roll over the positions in the expiring futures contract on Nymex but may engage in an exchange for swaps (EFS) transactions. This, in turn, produces an unexpected jolt to the prices when a fund does not roll its holdings in the expected manner (ie, does not sell and buy natural gas futures and lookalikes on Nymex and ICE outright, without recourse to EFSs).

It is important to recognise that a growing number of rapidly proliferating ETFs uses a different mechanism, which exposes the investors to additional risks (of which they are often ignorant). Instead of investing in actual assets behind the index, a replication mechanism based on derivative contracts is used. Such synthetic ETFs expose the investors to the credit risk of the derivative counterparty.<sup>40</sup>

Other commodity-related ETFs available in the US include (with their respective tickers in parentheses):

- United States 12 Month Natural Gas Fund (UNL);
- United States 12 Month Oil Fund (USL);
- United States Brent Oil Fund (BNO);
- United States Gasoline Fund (UGA);

- United States Heating Oil Fund (UHN);
- United States Short Oil Fund (DNO); and
- United States Oil Fund (USO).

Exchange-traded notes (ETNs) are debt instruments that offer returns related to the performance of certain assets or market indexes. The first ETNs were launched in June 2006 by Barclays. One ETN was based on the Dow Jones–AIG Commodity Index Total Return (ticker DJP), the second on the S&P GSCI Index Total Return (ticker GSP). However, it is important to recognise that there is no pool of assets supporting an ETN. These instruments represent the unsecured debts of the sponsoring financial institutions. Another important difference is that ETNs are not supported by the same mechanism that keeps their market prices in line with the value of the referenced assets or indexes – ie, there are no authorised participants who can redeem the existing shares or trigger the creation of new ones. The corresponding feature of ETNs is the early redemption provision that keeps market valuations in line with the referenced assets. Early redemption, however, requires a notice of a few days, reducing the efficiency of this mechanism. The redemption value (or indicative value) is given as the difference between formulas 5.2 and 5.3.<sup>41</sup>

$$\text{Principal} \times \frac{\text{Index Value at Redemption Date}}{\text{Index Value at Issue Date}} \quad (5.2)$$

$$\sum_{t=1}^m \text{Annual Investor Fee (\%)} \times \text{Principal} \times \frac{\text{Index Value}_t}{\text{Index Value at Issue Date}} \times \frac{1}{365} \quad (5.3)$$

where  $m$  is the number of days since inception.<sup>42</sup>

ETNs, like ETFs, offer several benefits to investors, but should be treated with caution. The tax treatment is one obvious problem: the notes do not pay interest but may be subject to taxation on accrued income. Assessment of credit risk is another challenge, as is the potential for an ETN value to be affected adversely by hedging transactions undertaken by the issuer. Of course, there is always a possibility that an issuer can hedge its exposures under an ETN in more creative ways. Many ETNs allow the investor to benefit from negative returns (so-called short ETNs). If a sponsoring financial institution issues roughly equal dollar amounts of long and short ETNs, it will have a natural hedge on its books and may not need to

transact in the market to protect itself (except for some small amounts corresponding to imbalances between the short and long notes outstanding).

### Performance issues

The performance of commodity indexes is highly sensitive to the shape of the forward price curve. A contango results in losses related to the cost of rolling from one contract to another, known as the negative roll yield. As explained in the prospectus of the UNG ETF:

The price relationship between the near month contract to expire and the next month contract to expire that compose the Benchmark Futures Contract will vary and may impact both the total return over time of USNG's NAV, as well as the degree to which its total return tracks other natural gas price indices' total returns. In cases in which the near month contract's price is lower than the next month contract's price (a situation known as 'contango' in the futures markets), then absent the impact of the overall movement in natural gas prices the value of the benchmark contract would tend to decline as it approaches expiration.<sup>43</sup>

The opposite is true in the case of backwardation.

An investor should beware. A period of persistent contango may turn an ETF or index-based investment into a money-destruction machine. Investors should be also realise there are potential hidden costs related to the index roll. The roll schedules are well known to the trading community and can be exploited for short-term gains. As explained by John Dizard:

As the volume in indices began to grow, the speculator community began to game the index managers by running up the prices just when the managers needed to roll their positions. This didn't hurt the managers, who were committed to delivering the prices on a certain date, whatever the prices were. But it does hurt the investing public. Jonathan Spencer, the president of Gresham Investment Management, which manages over US\$2bn in commodities funds that are not passively indexed, says the firm's research indicates that commodity index investors lose between 100 to 150 basis points a year of yield thanks to the professionals taking advantage of the traffic jam at index roll dates.<sup>44</sup>

Of course, as indicated in the quoted article, an investor should compare the implicit rollover costs (in the case of a passive index investment) against management fees in the case of an investment in actively managed commodity funds.

Front-running of commodity index rolls is a staple of many hedge fund commodity-related strategies. This practice has also attracted the attention of a number of academics.<sup>45</sup> What remains unexplained, in our view, is why this strategy – widely known and widely practiced – fails to eliminate excess returns it can extract. A discussion of this topic is beyond the scope of this book, and all we can offer is advice based on an old adage: “What cannot go on forever, eventually has to stop.”<sup>46</sup>

As signalled in Chapter 1, index funds and commodity-related ETFs are highly controversial. Inflows of investment funds into commodity-related funds and ETFs are blamed for price spikes of crude oil and agricultural commodities. The position taken by the index fund critics can be best summarised through the words of one of the most vocal opponents of this investment approach:

Institutional investors, with nearly US\$30 trillion in assets under management, have decided *en masse* to embrace commodities futures as an investable asset class. In the last five years, they have poured hundreds of billions of dollars into the commodities futures markets, a large fraction of which has gone into energy futures. While individually these investors are trying to do the right thing for their portfolios (and stakeholders), they are unaware that collectively they are having a massive impact on the futures markets that makes the Hunt brothers pale in comparison. In the last 41/2 years assets allocated to commodity index replication trading strategies have grown from US\$13 billion in 2003 to US\$317 billion in July 2008. At the same time, the prices for the 25 commodities that make up these indices have risen by an average of over 200%. Today’s commodities futures markets are excessively speculative, and the speculative position limits designed to protect the markets have been raised, or in some cases, eliminated. Congress must act to re-establish hard and fast position limits across all markets.<sup>47</sup>

The confrontation between critics of the index funds and their supporters are likely to escalate as inflationary pressures intensify in the future and the public starts looking for an easy explanation of price trends.<sup>48</sup>

### **Mutual funds**

An alternative way of acquiring exposure to commodities, offered primarily to the individual investors, is participation in specialised mutual funds. This may come as a surprise to investment professionals, as mutual funds operate under restrictions with respect to the sources of their income.<sup>49</sup> Commodity-oriented mutual funds

circumvent these rules by establishing offshore shell subsidiaries, with no real physical presence, which invest in commodities. Mutual funds hold the shares of these subsidiaries and this is considered by the Internal Revenue Service as an acceptable investment.<sup>50</sup> It remains to be seen if this conduit to commodity markets for individual investors remains open in the future (see footnote 50 for more information).

## CONCLUSIONS

The first decade of the 21st century was a period of rapid proliferation of different instruments offering investors many choices in establishing exposure to energy markets and/or in hedging energy-related risks. In this chapter, we have reviewed instruments that are created through packaging simple instruments (like futures and options) into more complex portfolios, hybrid derivatives, index-related investments and ETFs. A reader is encouraged to follow developments in this area: the new tools made available to investors affect the dynamics of energy prices.

Structuring complex energy transactions brings to mind building vastly impressive structures using Lego blocks. The blocks are simple and standardised; the edifices are monuments to creativity and persistence of children. Many structured energy transactions are extremely complicated, with the block diagram of a deal resembling the schema of a motherboard of a personal computer. Each individual component is usually a very simple instrument. All the instruments combined together may create a devilishly complicated system. A few words of caution are worth remembering.

- ❑ Even the best diagram will not capture all the risks of a complex transaction. In the early days of our career in energy business we worked on a transaction that was losing a lot of money from month to month but nobody was quite sure what the reason was. Reading and re-reading the contract and staring for hours at the deal schema were not very productive. After looking at the multi-dimensional graph capturing the contractual provisions across many different market scenarios, it took about 30 seconds to come up with the answer.
- ❑ The transactions tend to be complex enough and there is no need to complicate them any more by using fancy building blocks. The

component blocks should be plain vanilla instruments and the temptation to use exotic derivatives should be resisted. According to a quote attributed to Albert Einstein, one should “make things as simple as possible, but not simpler.”

- No transaction exists in an institutional and economic vacuum. There are many potential risks that cannot be captured in a diagram. The bottom line is that there is no substitute for common sense and mature judgement. Hopefully, this book will help to identify and explain many soft risks that cannot be captured in formal, stylised models.

Familiarity with passive commodity-based investment vehicles is important to a trader, even if they are not investing in such instruments. Exchange-traded funds affect the dynamics of energy prices and should be closely followed.

- 1 Of course, it is conceivable that a very complex transaction could be designed to take advantage of superior financial engineering skills when dealing with a counterparty lacking quantitative sophistication. If such motivation is detected, the best advice we can offer is to discontinue the business relationship.
- 2 The author's follow up book to be published by Risk Books will address the issue of risk management of non-standard instruments.
- 3 As explained before, a forward contract has a zero value at inception.
- 4 There is an assumption in the background that a forward price converges to a spot price at maturity ( $f(t,t) = S(t)$ ). In many commodity markets, convergence may be imperfect due to transaction costs, limited availability of information and other factors (such as the market power of some market participants).
- 5 In practice, the proceeds from the sale of a call option to a financial institution will be split between the counterparties. A hedge provider may capture the lion's share of the benefits if a hedging counterparty is not a sophisticated, well-informed market participant.
- 6 Writing covered calls is a popular strategy used in the equity markets. Portfolio managers and individuals often sell calls against the stocks they own to enhance overall returns.
- 7 The risk of a strategy based on the use of collars may be catastrophic if one does not manage the overall risk correctly. A good case study is provided by Christopher Gilbert, 2001, “Has the Ashanti Goldfields Loss Discredited Collar Hedges?” Vrije Universiteit, working paper.
- 8 David Graeber, 2011, *Debt: The First 5,000 Years* (Brooklyn, NY: Melville House Publishing).
- 9 “Introduction to weather markets 1.0; broker's perspective,” Evolution Markets (<http://www.wrma.org/weather-markets-webinar.html>).
- 10 Peter Brewer, chief investment officer of Cumulus Funds, 2012, “Weather markets 1.0 webinar: history, acronyms and terminology and overview,” 22 March.
- 11 One reason behind using derivative structure for weather contracts was avoiding regulation by the US state insurance commissioners.
- 12 Following the conventions used by the utilities, temperature was usually translated into heating or cooling degree days. These are defined in the US with respect to a cut-off temperature of 65°F. An average daily temperature of 80°F (35°C) translates into 15 (30) cooling (heating) degree days. In Europe, a temperature of 18°C would be used by convention. Most

- contracts applied to a cumulative number of heating or cooling degree days over a defined time period.
- 13 Most contracts relied on the weather data provided by governments, recorded at a specified weather station (typically located at an airport or an astronomical observatory).
  - 14 Some authors proposed valuation models based on the option-pricing framework but the industry did not embrace this approach.
  - 15 One could argue that delivering a commodity under a weather-related contract represents effectively “delivery of weather” in disguise.
  - 16 Many energy companies disclose weather risk in their regulatory filings. “Warmer weather can lead to lower margins from fewer volumes of natural gas being sold or transported. Colder weather that increases the volumes of natural gas sold to weather-sensitive customers can result in the inability of some of our customers to pay their bills. Either warm or cold weather that is outside the normal range of temperatures can lead to less operating cash flow, thereby increasing short-term borrowings to meet current cash requirements.” Piedmont Natural Gas Co Inc., Form 10-Q, March 13, 2002.
  - 17 Weather data to the CME is provided by MDA Federal.
  - 18 The following cities are included: Atlanta, Chicago, Cincinnati, New York, Dallas, Philadelphia, Portland, Tucson, Des Moines, Las Vegas, Boston, Houston, Kansas City, Minneapolis, Sacramento, Detroit, Salt Lake City, Baltimore, Colorado Springs, Jacksonville, Little Rock, Los Angeles, Raleigh Durham, Washington D.C.
  - 19 MDA Information Systems.
  - 20 <http://www.tegma.org/Documents/Windle.pdf>.
  - 21 Quanto options have payouts denominated in a different currency than the underlying. Foreign exchange rate enters into definition of payout in multiplicative way. Multiplication in the contract definition is the reason why this term is used with respect to some weather derivatives.
  - 22 Thomas Kammann, “Exposed to the elements: developing customised weather risk” ([http://www.worldcommercereview.com/publications/article\\_pdf/110](http://www.worldcommercereview.com/publications/article_pdf/110)).
  - 23 Alex Davis, 2010, “A new direction for weather derivatives,” *Energy Risk*, June 4.
  - 24 RenRe (<http://www.wrma.org/pdf/WRMA%20Industry%20Applications.pdf>).
  - 25 For example, heavy snowfall could benefit ski resorts and the hurt construction industry in a given area.
  - 26 See Maria Katharina Heiden, 2006, “Commodities as an asset class,” diploma thesis, Technical University Munich Department of Financial Mathematics. The thesis contains an excellent review of commodity-related financial instruments.
  - 27 S&P purchased the index in 2007.
  - 28 <http://www2.goldmansachs.com/services/securities/products/sp-gsci-commodity-index/tables.html>.
  - 29 As explained on the Goldman Sachs website, “Taking the first day of the roll as an example, just before the roll takes place at the end of the day, the S&P GSCI consists of the first nearby basket. That portfolio, constructed the night before and held throughout the fifth business day, has a dollar value. For the roll, that dollar value is distributed across the first and second nearby baskets such that the *number of contracts or the quantity* of the first nearby basket is 80% of the total and the *quantity* held of the second nearby basket is 20% of the total.” (<http://www.goldmansachs.com/what-we-do/securities/products-and-business-groups/products/gsci/roll-period.html>).
  - 30 More information is available from [www.indices.standardandpoors.com](http://www.indices.standardandpoors.com).
  - 31 It is estimated that 85–95% of institutional investors rely on the OTC index swaps. See M. W. Masters and A. K. White, 2008, “The accidental Hunt Brothers: how institutional investors are driving up food and energy prices,” *Special Report: A blog dedicated to discussing the topic of index speculation*.
  - 32 Technically, an ETF is a security certificate that documents the legal right of ownership of a certain basket of securities or commodities.

- 33 USO is a ticker for one of the crude oil ETFs.
- 34 SLV is the ticker for iShares Silver Trust. GLD is the ticker for SPDR Gold Trust.
- 35 Ben R. Marshall, Nhut H. Nguyen and Nuttawat Visaltanachoti, 2010, "ETF arbitrage" (<http://ssrn.com/abstract=1709599>).
- 36 To be more specific, the ETFs are generally more tax efficient than mutual funds but in some cases an investor may be surprised with unexpected capital gains distributions. See, Ari I. Weinberg, "Most ETFs Are Tax-Smart. But Others..." *The Wall Street Journal*, January 9, 2012.
- 37 USNG, 10-K 2010.
- 38 NG stands for the Nymex Henry Hub futures contract. NN is the Globex contract which is cash settled based on the final settlement price of the NG contract (NN has notional volume of 2,500 MMBtus daily). For the full specification of this contract, see [http://www.cmegroup.com/trading/energy/natural-gas/henry-hub-natural-gas-swap-futures-financial\\_contract\\_specifications.html](http://www.cmegroup.com/trading/energy/natural-gas/henry-hub-natural-gas-swap-futures-financial_contract_specifications.html). The specification for the ICE position is available from [https://www.theice.com/publicdocs/productSpecs/PSpec\\_OTC\\_Natural\\_Gas.pdf](https://www.theice.com/publicdocs/productSpecs/PSpec_OTC_Natural_Gas.pdf).
- 39 USNG, 10-K 2011.
- 40 See Srichander Ramaswamy, 2011, "Market structures and systemic risks of exchange-traded funds," BIS Working Papers No 343, April.
- 41 Colby Wright, Dean Diavatopoulos and James Felton, 2009, "Exchange Traded Notes: An Introduction," working paper, August 10.
- 42 This is the difference between the compounded principal amount of the note less an accumulated investor fee.
- 43 "Prospectus, United States Natural Gas Fund, LP," August 12, 2009.
- 44 John Dizard, 2007, "How the speculators profit from investors in commodities," *Financial Times*, January 23.
- 45 See, for example, Yiqun Mou, 2011, "Limits to arbitrage and commodity index investment: front-running the Goldman roll," Columbia University, July 15.
- 46 Some standard references explaining why arbitrage transactions fail to eliminate anomalies, including Andrei Schleifer and Robert W. Vishny, 1997, "The limits of arbitrage," *Journal of Finance*, pp 35–56; Dilip Abreu and Markus Brunnermeier, 2002, "Synchronization risk and delayed arbitrage," *Journal of Financial Economics*, pp 341–60; Darrell Duffie, 2010, "Presidential address: asset price dynamics and slow moving capital," *Journal of Finance*, pp 1237–67.
- 47 Michael W. Masters, 2009, "Testimony before the Commodity Futures Trading Commission," August 5.
- 48 The best summary of the conflicting positions taken in this debate can be found in: M. W. Masters and A. K. White, 2008, "The accidental Hunt Brothers: how institutional investors are driving up food and energy prices," *Special Report*: A blog dedicated to discussing the topic of index speculation. The rebuttal is offered by Scott H. Irwin and Dwight R. Sanders, 2010, "Testing the Masters hypothesis in commodity futures markets," October.
- 49 As explained by the 2011 *Investment Company Fact Book*, "[M]utual funds are subject to special tax rules set forth in subchapter M of the Internal Revenue Code. Unlike most corporations, mutual funds are not subject to taxation on their income or capital gains at the entity level, provided that they meet certain gross income, asset, and distribution requirements. To qualify as a regulated investment company (RIC), under subchapter M, at least 90 percent of a mutual fund's gross income must be derived from certain sources, including dividends, interest, payments with respect to securities loans, and gains from the sale or other disposition of stock, securities, or foreign currencies." (see, [http://www.icifactbook.org/fb\\_appa.html](http://www.icifactbook.org/fb_appa.html) for more information).
- 50 According to Senator Levin, "[T]his blatant end-run around the 90/10 restriction has nevertheless been blessed by the IRS which has issued dozens of private letter rulings, listed in Exhibit 7(d), deeming the offshore arrangements to be investments in securities rather than

commodities, since the parent mutual funds hold all of the stock of their offshore subsidiaries. The IRS has recently put a moratorium on those private letter rulings while it studies the issues. In addition, the offshore shells are currently exempt from CFTC registration requirements, despite operating as commodity pools, a situation the CFTC is reviewing as a result of a petition filed by the National Futures Association.” “Opening statement of Senator Carl Levin before US Senate Permanent Subcommittee on Investigations hearing on excessive speculation and compliance with the Dodd-Frank Act,” November 3, 2011.

## *Energy Markets: Exchanges*

This chapter starts a review of different energy market participants. We begin with exchanges, and explain the general principles of their business models and operations using examples of the US- and Canada-based exchanges that are of critical importance to the energy markets. We shall also revisit this topic in discussions of the exchanges that specialise in certain markets such as, for example, European natural gas and electricity. We shall also cover a number of topics related to trading on exchanges, such as:

- ❑ position limits (ie, upper bounds on the size of positions held in the futures contracts) and hedge exemptions (ie, waivers given to commercial users of futures);
- ❑ the SPAN system used for determination of initial margin;
- ❑ comments on the organisation and mechanics of clearinghouses;
- ❑ the contents and potential uses of the commitment of traders (COT) reports, which are followed closely by the trading community; and
- ❑ controversy surrounding proposals regarding mandatory exchange trading and clearing of “standardised” derivatives.

The exchanges have moved into the eye of the storm and face a number of challenges, some of which may be good problems to have in the long run. The most important developments to watch are related to the evolving regulatory landscape in the US and Europe. The Dodd–Frank Wall Street Reform and Consumer Protection Act (H.R. 4173), signed into law by President Barack Obama on July 21, 2010, faces multiple court and legislative challenges and may be at some point revised to a significant degree. The CFTC and SEC are still working on some of the final rules related to the implementation of the Act.

In Europe, the ongoing economic crisis may trigger similar policy initiatives that will affect the mechanisms of financial markets.

- ❑ Mandatory exchange trading and clearing of standardised derivatives<sup>1</sup> (with some exceptions for end users, ie, commercial hedgers) required under the Dodd–Frank Act (the Act) benefits the exchanges in the long run. This will lead to an increase in volumes but, in the mean time, involves the need to overhaul IT systems and business processes. The transition happens in the climate of uncertainty, given many legal and political challenges to the Act and the delays in promulgating the final rules by the CFTC and the SEC.
- ❑ Some provisions of the Act will affect exchanges negatively. Restrictions on proprietary trading under the Volcker rule, higher capital requirements and stricter position limits will reduce the trading activity by exchanges' customers and – this will translate into lower revenues.
- ❑ The way exchanges interact with the OTC markets will change. The trend towards clearing of bilateral transactions will continue, with competition between different exchanges and clearinghouses intensifying. The Act created a new type of market participant, the swap execution facility (SEF),<sup>2</sup> but the final rules and the way these new entities will interact with the existing exchanges are still unclear.

The regulation of derivatives has a long history going back to the Hammurabi's code.<sup>3</sup> What represents a challenge to the exchanges is not regulation itself but uncertainty regarding its final shape. (As changes in this area are frequent and rapid, please forgive the author any inaccuracies that are based on the information we had at the time of writing.)

The reader should keep in mind that references to exchanges in popular parlance include a very heterogeneous group of entities. The main distinction between the classes is the extent of regulatory oversight. In the case of the US, there are two primary types of institutions:<sup>4</sup>

- ❑ *Designated contract markets (DCMs)*. As explained by the CFTC, "DCMs are boards of trade (or exchanges) that operate under the

regulatory oversight of the CFTC, pursuant to Section 5 of the Commodity Exchange Act (CEA), 7 USC 7. DCMs are most like traditional futures exchanges, which may allow access to their facilities by all types of traders, including retail customers.”

- *Exempt commercial markets (ECMs)* were created by the Commodity Futures Modernization Act of 2000. As explained by the CFTC, ECMs are “exempted from most requirements of the Commodity Exchange Act (CEA) and most CFTC regulatory oversight. There are two kinds of exempt markets – exempt commercial markets (ECMs) and exempt boards of trade (EBOTs).” The Dodd-Frank Act eliminated the ECMs but (as of the time of writing) they continue to operate under grandfathered rules.

A discussion of the convoluted legislative history leading to the present solution is best left to the lawyers.

## **ENERGY EXCHANGES**

This section will review the three most important exchanges from the point of view of energy trading: CME Nymex, Intercontinental Exchange and NGX. These exchanges are crucial as they establish some of the most widely used price benchmarks. We review their business models and explain briefly the differences in their legal status. Other exchanges that specialise in certain niche geographical and commodity markets will also be covered in subsequent sections of this book.

### **CME Nymex**

Nymex history can be traced back to 1872 when the Butter and Cheese Exchange of New York (later The Butter, Cheese and Egg Exchange) was created by a group of dairy traders. The name of the Exchange was changed to the New York Mercantile Exchange, with an expansion into additional commodities such as dried fruits and poultry. The name of the exchange was shortened to Nymex (first informally and then as the official name), and industry veterans still use this term in spite of many organisational changes. In 1993, Nymex merged with Comex,<sup>5</sup> retaining its name. In 2006, Nymex demutualised and converted into a publicly owned company listed on the New York Stock Exchange as Nymex Holdings Inc, with Nymex and Comex

being two wholly owned subsidiaries. In March of 2008, Nymex Holdings was acquired by the Chicago-based CME Group.<sup>6</sup>

Energy trading on Nymex is one of the greatest success stories in the history of the US markets. The first energy-related futures contract introduced by Nymex was one for heating oil (1978), followed by crude oil (1983) and gasoline (1984). A natural gas contract was opened in April 1990. All these contracts grew very quickly both in terms of open interest and volume, with crude and natural gas becoming critical cornerstones of the US energy industry. Energy-related contracts account for 24% of revenues of the CME group (the first quarter of 2012), more than any other asset class, including interest rates (21%).

Electronic trading started on Nymex with the creation of the electronic platform known as ACCESS. This platform was launched in 1993, to allow for trading following the termination of the open outcry session. Nymex ACCESS migrated to the CME Globex platform on August 6, 2006 (beginning with the August 7 trades). This migration coincided with the introduction of parallel trading of certain contracts: through an open outcry system on Nymex and electronically through Globex. The CME Direct platform was later launched to enable the side-by-side electronic trading of CME futures, including all Nymex benchmarks, and the electronic trading of OTC markets, supported by selected, independent brokers.<sup>7</sup>

One important development in the US commodity markets was the development of a platform that allows for the transformation of bilateral OTC forward contracts into the economic equivalent of the futures contracts. This platform, known as ClearPort, was created in 2002 by Nymex in response to the credit-related slump in energy trading caused by bankruptcy or near-bankruptcy of many merchant energy companies, as well as the exit of many market participants. ClearPort allows for a breakdown of a bilateral transaction into two futures transaction through a process known as “novation”. In novation, Nymex is substituted as a counterparty to each side of the original transaction, which is nullified. ClearPort, created as a platform for energy trading, currently covers a number of different asset classes, including agricultural commodities, energy, green products and metals. Other asset classes covered include credit default swaps, equities, FX and interest rates instruments, with weather derivatives also expected to become available. The average daily volume

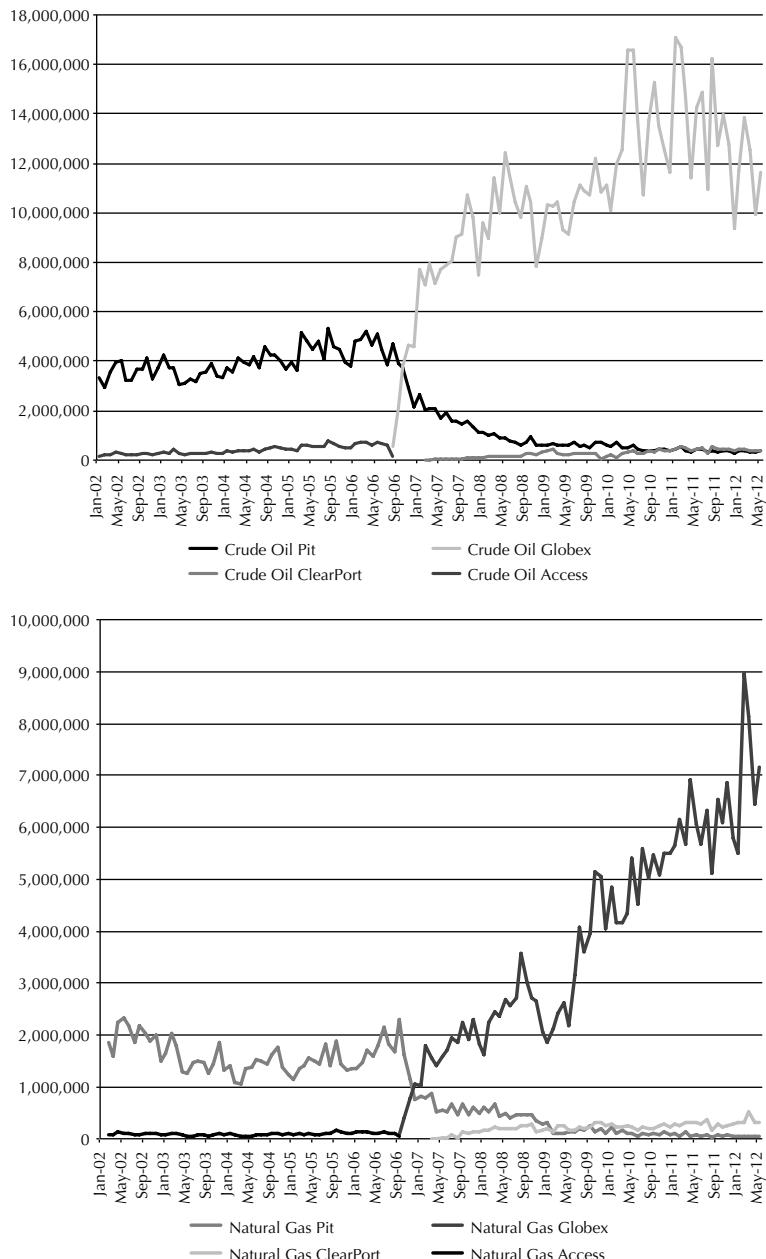
increased from 24,137 contracts in 2003 to about 459,500 in 2011 (the first quarter). The ClearPort product slate includes over 17,000 different listed contracts, 10,000 registered users and 500,000 contracts processed daily.<sup>8</sup>

There are several different ways a trade can be processed through ClearPort. One solution is to negotiate a transaction on the OTC market through interdealer brokers (IDBs), dealers or on electronic communications networks (ECNs).<sup>9</sup> After the trade is concluded, the broker submits the transaction (with the consent of both counterparties) to ClearPort, using a web-based graphical user interface (GUI) or an application programming interface (API). The submission of a transaction takes place through the clearing members. Another procedure uses the CME facilitation desk. One of the counterparties may contact CME and submit a transaction for clearing. The transaction is processed when the other side confirms its approval. The third trade entry method starts with a transaction executed on a third-party matching confirmation platform, which writes the trade directly into ClearPort using an API.

As of the time of the writing, the trading of natural gas and WTI is dominated by Globex. Figure 6.1 illustrates the evolution of trading volumes in the pit, on ACCESS, ClearPort and Globex.

One aspect of Nymex history was its long commitment to trading based on the open outcry system and its own unique set of rules and unique culture, after many other exchanges started migrating over to screen-based electronic trading. This system is being effectively phased out through a switch to two parallel trading platforms (electronic trading through GLOBEX and pit-based trading, as explained below). We watch with certain nostalgia and melancholy as pit-based trading rapidly fades away. However, it does make sense to devote some space to the open outcry system, as it is still used in many countries and remains more than a pure historical curiosity.

Under the open outcry system, traders stand in the pit, a set of concentric, descending rings (or polygons), in which floor brokers execute transactions through voice and hand signals. The hand gestures are standardised and well understood by all traders. The floor brokers, or floor traders, also use abbreviated phrases to rapidly make “bids” to buy and “offers” to sell (one hits a bid and lifts an offer). The traders wear identifying exchange badges. They also wear jackets with different colours that help to identify their affiliations

**Figure 6.1** CME trading volumes: crude oil and natural gas

Source: CME

with different companies. The traders from different brokerage houses or firms receive orders through phone clerks, who are located at workstations outside the pit area. Other employees known as runners deliver the buy/sell orders to the pit, using small slips of paper on which the details of the transactions are written (volume, contract, buy or sell designation and other requirements specified by a client). Some telephone floor clerks are close enough to the pit to hand orders or speak to their executing brokers directly.

The offer price is the price at which a market participant is willing to sell; the bid price is the price at which a market participant is willing to buy. The difference between these two prices is the bid–offer spread. Some floor traders, known as locals, trade in their own name for their own accounts, using their own (or borrowed) capital. They make money by speculating, exploiting small irregularities in prices or capturing bid–offer spreads. The locals perform a useful function of providing liquidity to other market participants and making the market more efficient.

The different categories of Nymex traders are described in an article by Victoria Woolley, as follows:

At the Nymex, there are three types of broker–traders. The first is a broker who's employed by a company to execute customer orders. [...] [B]rokers have more financial security, as they are equipped with a set salary plus commission. The second type is a loosely associated local, a professional who trades with his own money for his own account. [...] The third type is known as a market maker. [...] The difference between locals and market makers is that the latter are better capitalised and financed, which enables them to trade larger.<sup>10</sup>

In a simplified example, a trading firm ABC Co calls on the telephone to brokerage firm XYZ Co and speaks to XYZ Co's phone clerk and places an order. The phone clerk relays the order to the XYZ Co's pit trader, typically using a written order ticket. The ticket is time-stamped and preserved as part of the records of the transaction. The pit trader executes (or "fills") the order in the pit by transacting with another floor broker using hand gestures and voice. The selling pit trader records the details of the transaction on a different piece of paper called the pit card, and tosses it into the middle of the trading ring, where it is collected by Nymex clerks. The brokerage firm will also record the details of the transaction on the order ticket. Both the seller and the buyer will report their transaction details to the

clearing or brokerage firms for assignment to appropriate customer accounts. The audit trail for the transactions is created by: (i) the brokers' execution records of opposing broker, volume, and price; (ii) the exchange of pit cards; and (iii) the customer order tickets.

After a trade is executed through agreement by voice and hand signals, the pertinent details of a transaction are recorded by both the buying and selling floor brokers on a small card, including information such as quantity, delivery month, price, trader's badge name and badge name of the buyer. A seller must provide details to the exchange for price recording in essentially real time. The card has to be thrown to an exchange reporter located on the floor of the pit, where it is collected by a Nymex employee,<sup>11</sup> time stamped and entered into the computer system. Nymex rules require that this is done within one minute of the trade. One of the consequences of this system is that the time sequences in which the transactions are executed are sometimes reported slightly out of order. The time stamps are subsequently reviewed by back-office Nymex officials, and corrected if there is a reason to believe they are out of sequence. Of course, this problem will be eliminated over time when open outcry trading is entirely replaced with an electronic trading system. Currently, however, trades in the natural gas and other futures contracts can be executed through the open outcry system or electronically under the arrangement Nymex has with Globex. Globex volume accounts for between 95% and 100% of daily transactions in the main energy contracts (except for options where the open outcry share may be as high as 50%).

Prices of natural gas transactions executed on Nymex are displayed by the exchange with a small delay. Different financial information networks, such as Bloomberg, Telerate or CQG, display the prices in real time to their subscribers. Historical prices, tick-by-tick, are available from different vendors at a relatively small charge. The daily settlement prices (the end-of-day prices) are most important from the point of view of margining (the calculation of margins is based on these prices) and risk management (the traders and institutions use them to mark their positions to market for risk management and for P&L calculation).

On Nymex, settlement prices are calculated as volume-weighted averages of the prices of transactions executed during the last two minutes of each trading day. For the expiring contract on the last day

of trading, the settlement price is based on a 30-minute, volume-weighted average.

One of the advantages of the open outcry system is the ability of the market to absorb information in a very effective way. All transactions are executed in front of other traders. Experienced market participants develop over time the ability to observe the action in the pit and react very quickly. The other side of the coin is that the hyper-active and tense pit atmosphere makes it easy to deliberately inject information into the market that can create a “herd mentality” that results in large shifts in the market in a very short time. The injection of such information and/or the creation of this herd mentality may happen through aggressive, large volume and ostentatious trading, particularly if the trading activity was designed to influence the perceptions held by other traders with respect to market dynamics and the balance between supply and demand.

### **Intercontinental Exchange**

IntercontinentalExchange (ICE) was created in May 2000<sup>12</sup> as a joint venture of a number of financial and merchant energy companies, as an answer to the electronic trading platform started by Enron (EnronOnLine, EOL). EOL was set up as a one-to-many arrangement: Enron was the buyer for all sellers and the seller for all buyers. For nearly three years (1999–2001), Enron became a true dominant player in the energy space – offering about 1,400 different products, with screens, documentation and technical support in multiple languages.

After the company’s bankruptcy in 2001, EOL became history and ICE emerged as a major factor in the commodity markets in Europe and North America. The current business model of ICE, is defined around four clusters of closely related businesses:

- futures exchanges;
- OTC markets;
- clearinghouses; and
- market data and processing.

#### *Futures*

ICE Futures Europe was created through the acquisition of the International Petroleum Exchange (IPE), which was founded in 1980.

ICE Futures Europe operates as a recognised investment exchange in the UK and is regulated by the UK Financial Services Authority (FSA).<sup>13</sup> The exchange made a full transition to screen-based electronic trading and offers a number of very important energy-related contracts, including:

- Brent crude;
- WTI crude;
- ICE gas oil;
- ICE ECX CFI,<sup>14</sup>
- ICE UK electricity; and
- ICE unleaded gasoline blendstock.

Table 6.1 contains the data illustrating the growth of the European futures business.

Clearing services were initially provided by LCH.Clearnet Ltd. (London Clearing House, LCH). In April 2007, ICE announced a plan to establish a wholly owned European clearing service, parallel to services established in the US and Canada, where all ICE futures contracts are cleared by ICE Clear US and ICE Clear Canada, respectively.<sup>15</sup> The establishment of the ICE European clearing service means that the relationship with LCH was terminated by the end of 2008.

**Table 6.1** Main energy-related futures contracts

	Year Ended December 31					
	2011		2010		2009	
	Number of Contracts (In thousands)	Notional Value (In billions)	Number of Contracts (In thousands)	Notional Value (In billions)	Number of Contracts (In thousands)	Notional Value (In billions)
ICE Brent Crude Futures and Options	134,248	14,493	100,217	8,103	74,351	4,748
ICE Gasoil Futures and Options	66,184	6,111	52,283	3,565	36,253	1,963
ICE WTI Crude Futures and Options	51,936	4,878	52,790	4,241	46,412	2,969

Source: Intercontinental Exchange, 2011, 10-K, <http://files.shareholder.com/downloads/ICE/1231523476x0xS1193125-12-45255/1174746/filing.pdf>

The US-based ICE clients could initially access ICE Futures platform under the “no action” letter from the CFTC (effectively waiving CFTC jurisdiction over the European ICE futures business). This decision of the CFTC came under close scrutiny by a number of the US Congress committees during the financial crisis of 2007–08. As reported in the ICE Annual Report for 2010,<sup>16</sup> in “June 2008, the CFTC revised ICE Futures Europe’s no action letter for products that settle on the price of a US exchange’s futures contract to require ICE Futures Europe to adopt position limits and enhanced trader reporting equivalent to those required of designated contract markets.” In August 2009 the letter was amended further, with one additional condition related to the obligation of ICE Futures Europe to provide transaction data to the CFTC for the contracts linked to US markets (through the US delivery points or settlements based on contracts transacted on the US designated contract markets).

ICE futures operations in the US are not important from the point of view of this book as no energy contracts are offered on this platform.

#### *OTC markets*

ICE OTC business offers an energy trading electronic platform<sup>17</sup> that allows participants to enter into transactions, both cash-settled and involving physical delivery, over a wide spectrum of different commodities, including natural gas, power, natural gas

**Table 6.2** ICE OTC energy commodity-related contracts

	Year Ended December 31					
	2011		2010		2009	
	Number of Contracts (In thousands)	Notional Value (In billions)	Number of Contracts (In thousands)	Notional Value (In billions)	Number of Contracts (In thousands)	Notional Value (In billions)
North American natural gas	338,957	3,539	257,354	2,442	204,690	2,024
North American power	68,117	289	69,223	315	53,599	344
Global oil and refined products	8,720	5,462	5,722	2,420	2,232	811

Source: Intercontinental Exchange, 2011, 10-K Form, <http://files.shareholder.com/downloads/ICE/1231523476x0xS1193125-12-45255/1174746/filing.pdf>

liquids, chemicals, crude and refined oil products. Table 6.2 contains summary data about the growth of the ICE OTC business, demonstrating a growing acceptance of the OTC platform as one of the two main US energy markets.

The contracts transacted on ICE can be executed as bilateral or cleared transactions. The introduction of cleared OTC contracts was one of the most important innovations introduced by ICE, a concept it pioneered in 2002. Clearing was initially performed under an arrangement negotiated by ICE with the LCH. After November 2008, clearing was carried out through ICE Clear Europe, allowing ICE to capture the clearing revenues that were previously flowing to the LCH. The decision of using either form of an OTC contract (cleared and non-cleared) is left to the discretion of the counterparties. The benefits of clearing include reduced credit risk and the ability to lower the working capital required for credit support, due to an ability to cross margins related to the OTC and futures positions established under the ICE umbrella. Clearing is made available to the futures commission merchants (FCMs) who are members of the clearinghouse, or market participants who maintain accounts with such FCMs. The cleared contracts have been widely accepted by the market. Under the Dodd–Frank Act, ICE has to accept for clearing approved contracts executed on other platforms. The clearing services are made available to counterparties transacting outside the ICE platform through voice brokers. The so-called block trades<sup>18</sup> can be transmitted electronically for clearing and are subject to 50% of the standard ICE commissions.

The electronic trade confirmation system is an extension of the OTC trading platform. The service allows reviewing of electronic trade data received from different counterparties, comparing the terms electronically and then highlighting any discrepancies in a report to the traders' respective back offices. The service, available irrespective of the platform on which a trade was executed (other exchanges, voice brokers, etc), reduces transaction costs for the energy trading business.

As of December 31, 2011, 668 energy-related contracts were offered on ICE, including:

- 143 cleared natural gas contracts;
- 188 cleared power contracts; and
- 251 cleared oil contracts.

In 2011, 92% of the ICE OTC energy contracts traded on a cleared basis.

ICE OTC business operates as an ECM (as of the time of writing) under the Commodity Exchange Act and regulations of the CFTC. ECM is a new category of market participants created under the Commodity Futures Modernization Act of 2000 (more information on this topic will be provided later in this chapter). The passage of the Dodd–Frank Act in 2010 affected ICE in many ways. The ICE OTC platform will have to register as an SEF, which will be subject to regulation under core principles comparable to those applying to the futures markets. The US-based OTC ICE, a platform crucial for the energy markets, already meets many SEF requirements as a result of compliance with certain provisions the 2008 Farm Bill (as will be discussed in more detail in the section on position limits later in this chapter).<sup>19</sup> Also, many ICE customers may have to register as swap dealers or major swap participants. This requirement may create incentives for some clients to abandon the markets or reduce the scale of their activities.<sup>20</sup>

#### *Market data*

The ICE data business units were established in 2002 to address the growing appetite of energy markets for transaction-related data. ICE repackages data harvested from the trading activities on its platforms into price indexes and other types of information that can be used to validate internal pricing by the trading companies. The services provided by the market data division include:

- ❑ activity reports, including reports about ICE futures, indexes and contract records;
- ❑ real-time data, historical reports (OTC end-of-the-day), market commentary; and
- ❑ market price valuation for contracts with no directly observable prices.

Quantitative analysts can acquire detailed historical price data for model calibration.

#### **Natural Gas Exchange**

Canadian Calgary-based Natural Gas Exchange (NGX) with its associated clearinghouse is wholly owned by the TMX Group.<sup>21</sup> It

opened for business in 1994, funded initially by the Canadian pipeline operator Westcoast Energy. The platform was designed initially to trade physical natural gas in western Canada. In March 2000, OMHEX, the provider of transaction technology, acquired a 51% interest in NGX and, in January 2001, acquired the remaining 49%. In 2004, TMC acquired 100% of NGX. NGX has an alliance with ICE that allows it to utilise the ICE trading platform. NGX was a spectacular success after starting from a very low base, as illustrated by its transaction volume growth.

The products offered by NGX include:<sup>22</sup>

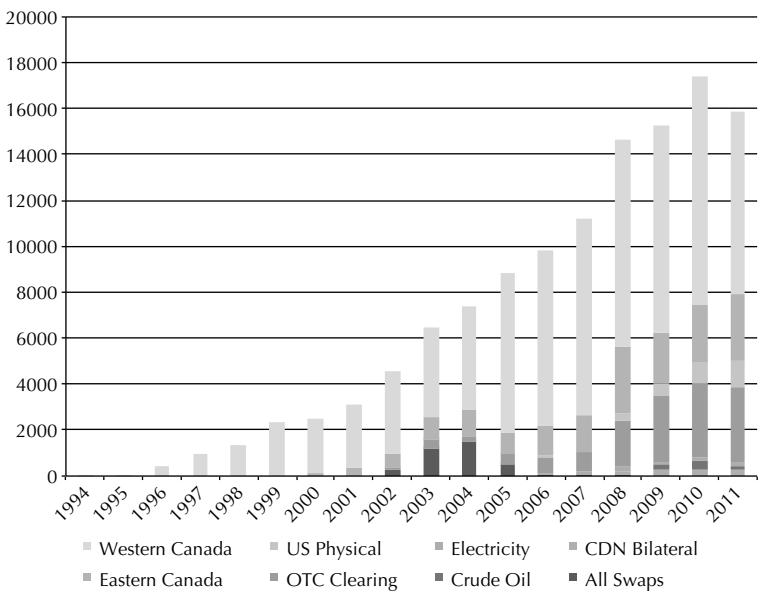
- physical and financial natural gas spot and forward contracts deliverable in Canada and USA;
- financial natural gas options for Alberta;
- physical crude oil spot and forward contracts deliverable in Canada and USA;
- financial power forward contracts for Alberta and Ontario; and
- ancillary services spot and forward contracts for Alberta (via the Watt-Ex subsidiary).<sup>23</sup>

The highlights of NGX operations include calculation and dissemination of the price indexes for Canadian natural gas and power. NGX posts daily about 15,000 settlement price indexes covering about 480 forward curves. The main points of the NGX business model include one-to-many design (NGX is a buyer to all sellers, a seller to all buyers) and hybrid transaction design (forward contracts based on futures principles, with respect to settlement and margining).

## **REGULATORY DEVELOPMENTS POSITION LIMITS AND MARGINS**

### **Position limits in the futures markets: Rules and regulations**

Position limits refer to maximum holdings of futures (and futures equivalents) held in a single trading account allowed by the regulations (subject to the rules regarding aggregation of accounts and certain exemptions). This section contains a review of the rules promulgated by the CFTC,<sup>24</sup> setting limits on derivative positions in 28 different commodities. The futures position limits in their current form can be traced back to the Commodity Exchange Act of 1936,

**Figure 6.2** Historical NGX trading activity

Source: <http://www.ngx.com/yearlyvolgraph.html>

which recognised that excessive speculation can impose an “undue and unnecessary burden” on interstate commerce, and authorised the Commodity Exchange Commission (CEC, the predecessor of the CFTC) to introduce limits on trading by market participants. Section 4a of the CEA reads as follows:

Sec. 4a. (1) Excessive speculation in any commodity under contracts of sale of such commodity for future delivery made on or subject to the rules of contract markets causing sudden or unreasonable fluctuations or unwarranted changes in the price of such commodity is an undue and unnecessary burden on interstate commerce in such commodity. For the purpose of diminishing, eliminating, or preventing such burden, the commission shall, from time to time, after due notice and opportunity for hearing, by order, proclaim and fix such limits on the amount of trading under contracts of sale of such commodity for future delivery on or subject to the rules of any contract market which may be done by any person as the commission finds is necessary to diminish, eliminate, or prevent such burden.

The CEC only had authority over agricultural futures and proceeded to establish, in December 1938, position limits and trading limits (although the law technically referred to trading only) with respect to

wheat, corn, oats, barley, flaxseed, grain sorghums and rye (collectively known as grains). CEC did not establish its authority with respect to the trading and position limits in many other contracts. The exchanges established their own voluntary limits for many contracts outside the CEC (and later, CFTC) jurisdiction.<sup>25</sup>

The Dodd-Frank Act amended section 4a(a)(2) of the CEA, mandating the CFTC to establish position limits for futures and options traded on DCMs. Additionally, the Act required the CFTC to establish position limits for swaps economically equivalent to such futures and options traded on DCMs. To summarise, the Commission was directed to establish also the position limits for:<sup>26</sup>

- (1) Contracts listed by DCMs;
- (2) swaps that are not traded on a registered entity but which are determined to perform or affect a “significant price discovery function”; and
- (3) foreign board of trade (‘FBOT’) contracts that are price-linked to a DCM or swap execution facility (‘SEF’) contract and made available for trading on the FBOT by direct access from within the United States.

The CFTC acted on October 18, 2011, promulgating the rule regarding position limits for futures and swaps. The rule is highly controversial and is being challenged in the court system and through the political process. Objections to the proposed position limits have come at a fast and furious pace from all directions, including:

- ❑ H.R. 2328, “End Excessive Oil Speculation Now Act of 2011,” introduced by Representative Hinchey (D-NY), and a similar proposed legislation in the Senate (S. 1200 sponsored by Senator Bernie Sanders (I-VT)) would tighten position limits;
- ❑ H.R. 3006 “Anti-Excessive Speculation Act of 2011,” introduced by Representative Peter Welch (D-VT), and a similar proposed legislation in the Senate (S. 1958 sponsored by Senator Bill Nelson (D-FL)) would, among other things, establish separate long position limits for the speculators;
- ❑ A suit filed in December 2011 by ISDA and the Securities Industry & Financial Management Association (SIFMA) challenging CFTC’s position limits rule for, among other things, failing to carry out adequate cost/benefit analysis or even to determine whether position limits were in fact necessary.

- ❑ H.R. 1840 (to improve consideration by the Commodity Futures Trading Commission of the costs and benefits of its regulations and orders) introduced by Representative Conaway (R-TX), would require the CFTC's chief economist to evaluate the costs and benefits of any proposed rules; and
- ❑ S. 1615, the "Financial Regulatory Responsibility Act of 2011," introduced by Senator Shelby (R-AL) imposes on the regulators a requirement to engage in rigorous economic analysis with respect to every new rule they propose.

It is difficult to determine at the time of writing whether these challenges will be successful. After the book went into production, the U.S. District Court for the District of Columbia vacated on September 28, 2012, the position limit rules and remanded them to the CFTC.

The most important change made by the Dodd–Frank Act to the CEA was an extension of the scope of position limits to "swaps traded on or subject to the rules of a designated contract market or a swap execution facility, or swaps not traded on or subject to the rules of a designated contract market or a swap execution facility that performs a significant price discovery function<sup>27</sup> with respect to a registered entity."<sup>28</sup> The CFTC was mandated under the Act to consider the following factors in determination if a swap performs significant price discovery function:<sup>29</sup>

- (A) Price Linkage. The extent to which the swap uses or otherwise relies on a daily or final settlement price, or other major price parameter, of another contract traded on a regulated market based upon the same underlying commodity, to value a position, transfer or convert a position, financially settle a position, or close out a position.
- (B) Arbitrage. The extent to which the price for the swap is sufficiently related to the price of another contract traded on a regulated market based upon the same underlying commodity so as to permit market participants to effectively arbitrage between the markets by simultaneously maintaining positions or executing trades in the swaps on a frequent and recurring basis.
- (C) Material Price Reference. The extent to which, on a frequent and recurring basis, bids, offers, or transactions in a contract traded on a regulated market are directly based on, or are determined by referencing, the price generated by the swap.
- (D) Material Liquidity. The extent to which the volume of swaps being traded in the commodity is sufficient to have a material effect on another contract traded on a regulated market.
- (E) Other Material Factors. Such other material factors as the Commission specifies by rule or regulation as relevant to determine

whether a swap serves a significant price discovery function with respect to a regulated market.

The limits would apply to futures and options on contracts traded on or subject to the rules of a designated contract market. The highlights of the new rules adopted by the Commission include:

Spot month<sup>30</sup> and non-spot month limits for all “referenced contracts,” defined as:

- core referenced futures contracts; or
- futures contracts, option contracts, swaps or swaptions linked to the price of a core referenced futures contract or the price of the commodity underlying a core referenced futures contract delivered at locations with the same demand/supply characteristics.<sup>31</sup>

The core futures contracts related to energy include:

- Nymex light sweet crude oil (CL);
- Nymex New York Harbor No. 2 heating oil (HO);
- Nymex New York Harbor gasoline blendstock (RB); and
- Nymex Henry Hub natural gas (NG).

The position limits will be implemented in two phases. In the first phase, all the spot month contracts and non-spot legacy contracts will be covered (60 days after the term “swap” is defined for the purpose of the DF Act). In the second phase, non-spot/non-legacy contracts will be affected (after 12 months of swap data is compiled, as explained below). The term “legacy contracts” corresponds to certain nine agricultural contracts for which the position limits were set historically by the CFTC and its predecessors.<sup>32</sup> In the case of non-legacy contracts, the position limits were set by the exchanges at their own initiative, even before this was mandated by the CFTC and its predecessors. There are nine legacy and ten non-legacy core agricultural contracts.

The limits have been defined differently for spot and non-spot months.<sup>33</sup> The position limits in the case of spot months are defined as 25% of the estimated spot month deliverable supply. Deliverable supply has been defined as “the quantity of the commodity meeting a derivative contract’s delivery specification that can reasonably be

expected to be readily available to short traders and saleable by long traders at its market value in normal cash marketing channels at the derivative contract's delivery points during the specified delivery period, barring abnormal movement in interstate commerce.”<sup>34</sup> Netting of physical and cash-settled contracts is not allowed (even if they are economically equivalent).

Position limits for cash-settled contracts, in the case of Henry Hub natural gas, are an exception to the general rules explained above. They are set at the level equal to five times the limit for physical delivery contracts.<sup>35</sup> The limits will take effect 60 days after the term “swap” is defined through the process of rulemaking carried out jointly by the SEC and the CFTC. The initial energy-related limits are shown below:

- Nymex light sweet crude oil: 3,000;
- Nymex New York harbor gasoline blendstock: 1,000;
- Nymex Henry Hub natural gas: 1,000; and
- Nymex New York harbor heating oil: 1,000.

The CFTC will have to revise the limits in the future in light of new information. The limits will expire on January 1 of the second calendar year after they become binding. The CFTC will update the limits using data available at this point.

The non-spot month positions limits apply to net long or short position across all contract months combined or with respect to a single month. Given the energy markets orientation of this book, we have to worry only about the so-called “non-legacy Referenced Contracts (ie, those agricultural contracts that initially were not subject to Federal position limits as well as energy and metal contracts).”<sup>36</sup> The position limits are set as “10 percent of the first 25,000 contracts of average all-months combined aggregated open interest with a marginal increase of 2.5 percent thereafter.”<sup>37</sup> The initial position limits will become available within a month after necessary data are compiled. This illustrates very well the circular nature of many problems faced by the CFTC. It is difficult to determine what the open interest is without the machinery to collect the OTC transaction data put in place and running. The specific numerical levels for the limits will be revisited by the CFTC after two years.

The CFTC rules contain very important provisions regarding hedge exemptions, ie, permissions to exceed the limits under certain circumstances defined for a position which:

- (i) Represents a substitute for transactions made or to be made or positions taken or to be taken at a later time in a physical marketing channel;
- (ii) Is economically appropriate to the reduction of risks in the conduct and management of a commercial enterprise; and
- (iii) Arises from the potential change in the value of one or several –
  - (A) Assets that a person owns, produces, manufactures, processes, or merchandises or anticipates owning, producing, manufacturing, processing, or merchandising;
  - (B) Liabilities that a person owns or anticipates incurring; or
  - (C) Services that a person provides, purchases, or anticipates providing or purchasing; or
- (iv) Reduces risks attendant to a position resulting from a swap that –
  - (A) Was executed opposite a counterparty for which the transaction would qualify as a bona fide hedging transaction pursuant to paragraph (a)(1)(i) through (iii) of this section; or
  - (B) Meets the requirements of paragraphs (a)(1)(i) through (iii) of this section.<sup>38</sup>

Hedge exemption is contingent on orderly establishment and liquidation of a position. There are many other detailed provisions related to hedge exemptions, which may be very important in practice. It is critical that any trader or risk manager becomes familiar with the detailed provisions of the final rule by following information available from the CFTC website and discussions at the industry conferences. Equally important are the provisions regarding aggregation of accounts promulgated by the CFTC in the rule regarding position limits. A trader is under obligation to combine the positions held in all the accounts in which they have at least 10% ownership (assuming the trader actually knows the positions and that aggregation would not violate federal laws (regarding insider trading, for example)).

### **Position limits and margins: Current practice**

Nymex position limits as of 2012 are shown in Table 6.3. Position limits tend to change and any trader should follow the developments in this area.

Current practice distinguishes between position limits for the prompt month (which are binding) and accountability limits (which are soft limits, which may trigger an action by an exchange if exceeded). Position limits apply during the last three days of contract

**Table 6.3** Position and accountability limits on Nymex (as of June 15, 2012); number of contracts

Commodity	Position limits prompt month	Accountability level (single month)	Accountability level (all months)
Natural gas (NG)	1000	6,000	12,000
RBOB gasoline (RB)	1000	5,000	7,000
Heating oil (HO)	1000	5,000	7,000
Crude oil (CL)	3000	10,000	20,000
Central Appalachian coal (QL)	200	5,000	5,000

Source: <http://www.cmegroup.com/rulebook/NYMEX/1/5.pdf#page=49>

trading. Accountability limits are defined as any-one-month and all-months limits.

Nymex, like other exchanges, specifies an initial margin and a maintenance margin (called also performance bonds). The initial margin is comparable to a security deposit posted when renting an apartment, maintenance margin is used to determine if additional margin (variation margin) is required. The Nymex website provides information about current levels of both margins (an example as of June 15, 2012 is provided in Table 6.4).<sup>39</sup>

A complication arises from the fact that the actual margin is assessed on the entire portfolio, taking into account possible risk reduction arising from diversification of different holdings. This will be covered in more detail in the next section.

A period of very high volatility in the commodity markets associated with gapping prices (ie, prices changing through discrete jumps and not through small, almost continuous, changes) resulted in frequent adjustments to the exchange margins. This was a direct

**Table 6.4** Nymex crude oil margin requirements

Start Period	End Period	Initial (USD)	Maintenance (USD) Scan
May-12	Jul-12	6,210	4,600
Aug-12	Oct-12	5,873	4,350
Nov-12	Apr-13	5,535	4,100
May-13	Dec-13	5,198	3,850
Jan-14	Dec-20	4,860	3,600

Source: <http://www.cmegroup.com/clearing/margins/#e=NYM&a=CRUDE+OIL&p=CL>

result of growing concerns that unprecedented price fluctuations in certain markets (such as, for example, silver) could result in defaults by some market participants. An example of such action is a decision by the CME – announced on May 9, 2011, and which became effective close-of-business May 10, 2011 – to increase performance bonds on a number of contracts. The changes in margins introduced by the exchange are not only replicated but often amplified by clearing members. The change of margins forced some over-leveraged longs to curtail positions, forcing prices down, potentially triggering execution of stop-loss orders, and resulting in a significant price move.

It is a paradox that an action undertaken to reduce volatility may actually increase it. Any risk manager (or trader) should always keep in the back of their mind the possibility of margin adjustments and should review their position in view of the potential market impact of such changes. The risk of this event increases in a period of market instability, raising the importance of monitoring exchange policies and trying to predict their reactions to market developments.

## CLEARINGHOUSE DESIGN

A clearinghouse provides a number of very important services, including the processing, matching, registering, guaranteeing and settling of trades executed on an exchange. In addition, a clearinghouse performs the following functions: it replaces a web of bilateral relationships between many trading counterparties into claims on a clearinghouse and liabilities to a clearinghouse. This has the beneficial outcome of reducing credit risk to a very low level and reducing the amount of working capital required for collateral management.

The regulatory developments taking place in the US related to the Dodd–Frank Act make clearinghouses (referred to in official documents as central clearing counterparties, or CCPs) one of the critical pillars of the overhauled financial system.<sup>40</sup> One of the most important and also very controversial provisions of the Act is the mandatory exchange trading and clearing of most derivatives. The specific rules defining the universe of derivatives that will fall under this provision will be determined through the process of rule making by the SEC and the CFTC (underway at the time of writing). The end users of derivatives (ie, the entities hedging commercial exposures) will receive exceptions (not exemptions, as is usually reported by the

press) from this requirement that in itself are very controversial (to be discussed below). Understanding the economic functions and different designs of clearinghouses becomes quite important in the changing landscape of the financial industry (not that they were unimportant in the past). Historically, different models of central CCPs evolved through trial and error, and the replication of successful solutions implemented elsewhere. After all, imitation is the subtlest form of flattery.

In the most general sense, a clearinghouse is an entity with two critical functions: settling trades for its customers and guaranteeing the fulfillment of contracts. The first clearinghouse was established in 1882 in Le Havre, France, as the Caisse de Liquidation des Affaires en Marchandise for traders importing cotton and sugar to France. This was undoubtedly one of the most important financial innovations in history and was quickly copied across the globe. In the interest of fairness, one has to mention there were some important forerunners, including the Dojima rice futures market established in Osaka, Japan, in 1697, and some New York City trading venues (the Coffee Exchange, founded in 1882).<sup>41</sup>

Clearinghouses can be structured as vertical or horizontal operations. A vertical structure (in its most straightforward version) is a clearinghouse closely associated or directly owned by an exchange, with all transactions taking place on the exchange being subject to mandatory clearing. Additionally, a clearinghouse accepts for clearing exclusively the transactions executed on the exchange it is associated with. An example of such a clearinghouse is Nymex prior to creation of ClearPort.<sup>42</sup> A horizontal structure corresponds to a clearinghouse which accepts for clearing transactions from multiple trading venues (for example, London Clearing House). Both solutions have some advantages. The vertical structure facilitates seamless processing of trades with each transaction executed on the exchange being immediately broken up with the clearinghouse being inserted between the two original counterparties. Horizontal clearing gives more choices to the market participants and stimulates competition.

However, in practice, these two stylised structures no longer exist and clearinghouses keep evolving to meet the needs of diverse market participants and under regulatory pressures. On one hand, one can detect a tendency to form clearinghouses owned by

exchanges, which relied in the past on services of independent clearinghouses. The rationale for this trend is the need to increase revenues through the provision of additional services.<sup>43</sup> On the other hand, the regulators have pressured existing clearinghouses to adopt a more open, flexible system. As explained by ICE in its most recent annual filing:<sup>44</sup>

MiFID<sup>45</sup> II contains provisions relating to open access for exchange-traded derivatives and the introduction of fungible clearing. Requiring a clearing house to accept and clear trades executed on an unrelated trading facility would expose our clearing house to trades that we may not have the same level of confidence compared to a trade that had been executed on our trading facility. Further, clearing trades from other trading facilities could make it more difficult to track positions and counterparty risk exposure, which will make the operation of our clearing houses riskier and more difficult since there will need to be common rules and margin requirements as well as more information sharing between competing clearing houses. Finally, the open access provisions could diminish the value of our OTC swaps execution platform by enabling competing electronic venues to submit trades to our clearing houses for clearing. Also, our clearing houses will have less control over the decision of whether to accept and clear trades from various execution facilities.

The second defining feature of a clearinghouse is the way it interacts with market participants. In practice, a clearinghouse interacts directly with a small group of financial institutions that have a status of clearing members. A distinction is sometimes made between the general clearing members (GCMs) and direct clearing members (DCMs). A GCM clears its own trades, the trades of its customers and the trades of non-clearing members (NCMs). A DCM may clear its own trades.<sup>46</sup> This means that clearing members function as a conduit to a CCP to other, non-clearing members and to the general public. Of course, the evolution of the financial system and frequent innovations may modify and complicate the clear distinctions between different types of CCP members.

### **The role of clearing members: Exchange safeguards**

Developments related to the bankruptcy of MF Global focused public attention on the role played by the clearing members and illustrate the importance of careful dissection of the somewhat simplistic understanding of credit risk of exchange trading. The conventional wisdom of many practitioners is that exchange trading

and clearing eliminates completely credit risk, but this is not entirely correct. Central counterparty clearing is like plumbing in a house: it usually works and is taken for granted, unless it does not. It works most of the time but it always breaks at the worst possible moment (typically just before the guests arrive), and the consequences are dismal. It is more correct to say that central clearing reduces credit risk to a very low level, but that a very low probability/very high severity exposure always exists.

There are several requirements that a clearing member has to meet, including the level of capitalisation.<sup>47</sup> Anybody trading on an exchange should be aware of the potential exposure to credit of a clearing member they use as a broker (or who is used by their broker, who has no clearing privileges). The CME representations, for example, make it perfectly clear that the clients of clearing brokers are exposed to credit risk:<sup>48</sup>

Customers face credit risk in doing business through any particular clearing member. Consequently, the selection process for a suitable clearing member is important. While the policies applicable to segregation of customer monies for products traded in regulated markets are specifically designed to protect customers from the consequences of a clearing member's failure, they do not always provide complete protection should the default be caused by another customer at that firm.

Historically, there were very few cases of clearinghouses that ran into problems and required some kind of a bailout.<sup>49</sup> The exchanges have developed a system of firewalls to provide protection against default resulting from the non-performance of an exchange client or a clearinghouse member (clearing or non-clearing). Examples of such safeguards can be found on the websites of many exchanges, including the CME.<sup>50</sup> The safeguards quoted in the CME document include the following.

- ❑ Two full settlement cycles, marking to the market once in the late morning and once in the late afternoon.
- ❑ Performance bond requirements that are good faith deposits to guarantee performance on open positions, and are often referred to as "margins."
- ❑ Cross-margining with other clearinghouses, such as LCH.Clearnet and the Fixed Income Clearing Corporation.
- ❑ Segregation of customer funds at the clearing brokers from the

clearing firm's own funds.<sup>51</sup> It seems at this point that MF Global failed to segregate properly the customers' funds and, possibly, its trading positions.

- ❑ Capital requirements for the clearing brokers.<sup>52</sup> CME clearing members that are subject to the CFTC regulation are required to maintain adjusted net capital (ANC) at prescribed levels. Effective as of January 1, 2009, all active clearing member were required to maintain the greatest of:
  - US\$5,000,000;
  - CFTC minimum regulatory capital requirements (see below); or
  - SEC minimum regulatory capital requirements.
- ❑ Financial surveillance. The CME Group audit department, in conjunction with other self-regulatory organisations, operates a sophisticated financial surveillance programme, including notification, inspection and information sharing.
- ❑ Intra-day monitoring.
- ❑ Market regulation.
- ❑ Clearing member risk reviews.

The actions the clearinghouse can take in the case of a default by a clearing member vary to some extent depending on the source of the problem (mismanagement by a clearing member of its own business or a customer's failure), but the remedies envisaged for both cases share many commonalities, including:

- ❑ transfer of all segregated and secured customer positions and monies to another clearing member;
- ❑ taking control of, or liquidation of, the positions in the house account;
- ❑ applying the clearing member's guarantee fund and house performance bond deposits to the obligation shortfalls;
- ❑ attaching all other assets of the clearing member that are available to the clearing house (eg, shares and memberships); and
- ❑ invoking any applicable parent guarantee.

The most important provision from the point of view of a clearing-house customer is related to separation of customers' funds at the house level:<sup>53</sup>

Under no circumstances will customer segregated or secured performance bond deposits held by CME Clearing for one clearing member be used to cover either a house or customer default of another clearing member. Customers doing business through a clearing member not involved in a default are insulated from losses incurred by the failure of another clearing member.

Performance of a clearinghouse is guaranteed by a number of reserve funds and emergency credit lines.<sup>54</sup>

The case of MF Global is a lesson that the best safeguards deployed by an exchange are not a substitute for vigilant risk management. A report on CCPs published in the late 1990s identified a number of risks related to central clearing.<sup>55</sup>

- ❑ *Defaults by clearing members*, producing:
  - a replacement costs risk (the cost of substituting new contracts for the contracts under default) – this cost varies with market fluctuations.;
  - a liquidity risk: shortages of immediately available funds even if no losses were incurred – this may in turn translate into replacement risk as time passes; and
  - physical (delivery) risk if physical delivery of the underlying is required.
- ❑ *Settlement bank failures*: This risk happens primarily due to the timing difference in crediting the bank account of the defaulting member prior to receiving the funds from him.
- ❑ *Operational risk*: Unavoidable delays in resolving conflicting claims and sorting out paperwork.
- ❑ *Legal risks*: Certain multilateral credit arrangements may receive conflicting treatment in different jurisdictions.
- ❑ *Investment of clearinghouse funds*: A clearinghouse invests its funds in the capital markets using instruments that may have credit risk associated with them.

We may add that there is an additional risk related to the nature of the bankruptcy process. Bankruptcy is like a black hole in which the laws of physics no longer apply. Each bankruptcy case results in unexpected developments and complications and the skills and firepower of one of many competing legal teams will be always a critical factor.<sup>56</sup>

It seems that the critical issue at the heart of the MF Global fiasco

was the proper separation of customers' funds from the company's own funds. The clearinghouses use two models to manage related cash flows: net or gross. The gross model (used by the CME) requires that the margin payments, posted by non-clearing members and customers, are passed by clearing members directly to the clearinghouse without netting and offsets, which apply under the netting regime. The customers make the same margin payments under both systems. The difference is who is effectively holding the funds and in what proportions: a clearinghouse or a clearing member. One can anticipate that one of the issues the courts will have to resolve in the aftermath of MF Global is what is actually meant by separation of funds. Is separation accomplished just by appropriate accounting entries or through actual physical separation of assets (different bank accounts, escrow accounts, etc). Nobody asks such questions when things are going well.

### THE SPAN SYSTEM

Daily marking positions to market and margining is the cornerstone of clearinghouse operations. In practice, the determination of initial margin is a fairly complicated process but should be well understood by any trader. Financial results of many trading strategies depend on large volumes and razor-thin unit profits, and the cost of funding margin may be significant to the overall outcomes.

The calculation of margin is based on the system known as Standard Portfolio Analysis of Risk (SPAN), developed by the CME in 1988 and used under licence by a number of exchanges,<sup>57</sup> including Nymex.<sup>58</sup> SPAN represents a version of a risk management system based on enhanced stress testing, with an attempt to recognise the benefits of diversification. Due to its transparency and flexibility (it allows the exchanges to adjust the parameters to their risk preferences), the SPAN system has survived the test of time in spite of being a fairly simplistic solution by the standards of modern risk management practices.

Understanding of the SPAN system is important for a number of reasons. A portfolio manager has to understand how different margining arrangements work for different components of their portfolio. In a complex portfolio, different pieces of the same strategy may be executed across different markets and platforms, and are therefore subject to different margining arrangements. For example,

the risk of an OTC transaction may be hedged using a Nymex position and the collateral received/posted may diverge, which may stress the firm's liquidity (the details of credit risk management for OTC transactions will be discussed in our next book project). Second, the US financial markets use two margining systems based on varying philosophies, and many equity and fixed income traders moving to the energy markets may sometimes ignore these differences. The US securities industry has historically used a strategy-based system<sup>59</sup> developed under Regulation T.<sup>60</sup> The term "strategy-based system" is not very descriptive; what it means is that two or more positions can be seen as mutually offsetting in terms of risk only if they are parts of the same predefined strategy. If this is not the case, the positions are treated as if there was no diversification effect. The SPAN system takes a portfolio approach, adding gains and losses for all the portfolio positions under different market scenarios. This reflects the policy of an exchange to liquidate the entire portfolio if there is a default on a single contract held by a customer. Another important difference is the treatment of the option positions under Regulation T.<sup>61</sup> Margin required for options is related to the value of the underlying instrument (through a fixed percentage requirement); SPAN uses the option-pricing model.

The main function of SPAN is to create a bridge from the maintenance margins established by an exchange to the estimate of a loss that may be incurred by the portfolio from one day to the next at certain probability levels, ranging from 95% to 99%. Of course, the loss estimate is not a measure of maximum potential loss, and both the exchanges and the futures brokers can address this by assessing surcharges to their members and clients,<sup>62</sup> in excess of the guidelines determined under SPAN.

The design of the SPAN system is based on a number of principles. Financial instruments based on the same underlying are grouped together for analysis and are treated as one combined commodity. The risk of a portfolio is analysed for each combined commodity, followed by an assessment of the benefits of diversification across different combined commodities.

The process of risk assessment is based on the simulation of potential market moves and calculation of gains and losses for individual portfolio positions. An exchange or clearinghouse defines the so-called price scan ranges and volatility scan ranges that are used to

generate potential market scenario. Scan ranges describe the magnitude of possible market price and volatility changes. The system accommodates unlimited number of scenarios; most organisations use a set of 16 scenarios called SPAN risk arrays. A risk array is a tabular representation of losses/gains under each of scenarios. In the case of linear instruments, such as futures, the calculation is quite straightforward; in the case of options, a specific option-pricing model is required. The risk arrays are next combined into a Risk parameter file, disseminated to the users of the system.

Risk arrays contain information about potential changes to the mark-to-market (MTM) value of a specific instrument under each of 16 specific scenarios. The changes in MTM values are calculated by moving the price of the underlying up or down in 1/3 increments of the price scan range. Losses are shown as positive numbers. Of course, calculations for options will be more complicated, as a specific option valuation model has to be applied for each strike. Next, for each product, the scenario with the maximum aggregate loss is chosen. For example, a natural gas position maximum loss would be calculated across all futures contracts by looking at the scenario with the biggest potential loss, not by summing the maximum losses for different contracts. A similar calculation is carried out for the short option positions, taking into account the short option minimum charge, mentioned above.

The 16 scenarios used by Nymex (as of the time of writing) are summarised below:<sup>63</sup>

- (1) Futures unchanged: Volatility up
- (2) Futures unchanged: Volatility down
- (3) Futures up 1/3: Volatility up
- (4) Futures up 1/3: Volatility down
- (5) Futures down 1/3: Volatility up
- (6) Futures down 1/3: Volatility down
- (7) Futures up 2/3: Volatility up
- (8) Futures up 2/3: Volatility down
- (9) Futures down 2/3: Volatility up
- (10) Futures down 2/3: Volatility down
- (11) Futures up 3/3: Volatility up
- (12) Futures up 3/3: Volatility down
- (13) Futures down 3/3: Volatility up

- 
- (14) Futures down 3/3: Volatility down
  - (15) Futures up extreme move:<sup>64</sup> Volatility unchanged
  - (16) Futures down extreme move: Volatility unchanged

This list illustrates how scenarios are constructed in practice. For example, scenario 8 is constructed moving current price of a given futures contract up 2/3 of the range, and moving volatility down. In this case, volatility is the input used in pricing an option with a given futures contract as the underlying.

The extreme scenarios are used to assess the risk of out-of-the-money options that may produce large losses in the case of extreme market moves. Options close to expiration have higher deltas and gammas (see the next book for more details). In the case of a significant market move, an option position with a very small pre-move market value can suddenly translate into a big gain or loss.

The example below shows SPAN calculations for a simple portfolio of one long Nymex WTI September 2010 contract and two short WTI December 2010 contracts. The risk array file was downloaded from the CME website and used in the calculations are shown in Table 6.5.

From this table, the largest potential loss results from scenarios 11 and 12 (prices up 100% of the scan range).<sup>65</sup> It is not a surprising conclusion given that the portfolio has a net short bias. It is also

**Table 6.5** SPAN calculations for a simple portfolio

Scenario	CL9	CL12	Gain/loss
1	0	0	0
2	0	0	0
3	-431	-428	425
4	-431	-428	425
5	431	428	-425
6	431	428	-425
7	-861	-857	853
8	-861	-857	853
9	861	857	-853
10	861	857	-853
11	-1292	-1285	1278
12	-1292	-1285	1278
13	1292	1285	-1278
14	1292	1285	-1278
15	-1279	-1272	1265
16	1279	1272	-1265

worth noting that the system recognises the benefits of diversification and valuation across the entire portfolio, as opposed to cherry-picking maximum losses for each separate portfolio position under different scenarios. Next, the applicable intra-commodity spreads, inter-commodity spreads or spot month charges are added/ subtracted.

The analysis is further complicated due to the consideration of spread risk. A default assumption used in SPAN is that the futures price curve changes through parallel shifts. This means that a spread position in the same commodity (for example, long the 5th nearby, short the 10th nearby, same volumes) would have zero margin requirements. Exchanges or clearinghouses using SPAN can capture the forward curve reshaping risk by charging an intra-commodity spread charge. Also, an additional charge may be added to a position in the first nearby contract in order to recognise exceptionally high volatility observed in the futures prices near expiration. SPAN recognises that the futures prices for the same underlying do not correlate exactly across different months and that the forward price curves do not reshape through parallel shifts. If the risk arrays use the same values for contracts of different maturities, the curve reshaping risk is accounted for through the intra-commodity spread charge, which is customised for specific groups of two or more contracts.

SPAN recognises the benefits of diversification resulting from high levels of correlation between offsetting positions through the inter-commodity spreads. Each exchange identifies combinations of different instruments that qualify as risk-reducing positions and determines the margin credit that applies to them (as a percentage reduction of initial risk estimates). The SPAN system takes the inter-commodity spread table defined by an exchange and searches for the defined spread formations that give margin credit.

An alternative way of recognising the benefits of diversification across calendar months is an approach called scanning-based spreading. The risk offsets are accounted for by scanning related commodities together. The system identifies the benefits of offsetting correlated positions, including option positions. SPAN has the capability to recognise spreads established on different exchanges.

The short positions in out-of-the money options may be found to represent a minuscule risk using a risk array of only 16 scenarios. Under extreme market conditions, such positions could produce

catastrophic losses as the options may mutate into in-the-money positions. SCAN captures this risk through the short option minimum that is applied in case the scan risk is lower.

The risks related to the position in the delivery month are handled with the delivery add-on charge. Such positions may represent a special risk from very high volatility due to potential squeezes resulting in big price moves.

## COMMITMENTS OF TRADERS AND RELATED REPORTS

COT reports, distributed weekly by the CFTC, are a useful window into current activity in the futures markets. These reports have many uses, including the development of trading strategies,<sup>66</sup> forecasting futures prices and academic research.

The level of activity on futures exchanges is measured using two statistics: open interest and volume. Open interest is the total of all futures contracts that have not been offset through delivery or closed. The volume is the number of transactions executed during a specific time period. An example may help to clarify this distinction. Suppose that a new contract is opened and X sells 10 contracts to Y during the first hour of trading. The open interest is 10 at this point, and this number is the same as the long or short open interest (there is a buyer for every seller). Before the end of the trading day, X and Y reverse their transaction and X buys their 10 contracts back from Y. Net open interest at this point is 0 and the volume for the day was 20 contracts (assuming there were no other trades).

The outstanding positions in over 90 futures contracts are reported by the CFTC in the COT report, and this publication (in spite of its limitations) is an important information source followed by commodity traders.<sup>67</sup> COT is available for Tuesday positions (published on Fridays at 3:30 pm EST) for contracts in which 20 or more traders hold positions exceeding the levels established by the CFTC.<sup>68</sup> These reporting levels depend on the overall maturity of a given contract and the overall open interest, and may vary from low tens to a few thousand.<sup>69</sup> If a single trader exceeds the established limits in just one of the delivery or option expiration months, their entire position will be reported. According to CFTC, 70–90% of outstanding positions are covered.

The reports are currently available in a number of formats. Prior to September 4, 2009, they were based on the classification of traders

into commercial and non-commercial categories. Commercial traders were defined as those who engage in hedging transactions,<sup>70</sup> as defined by the CFTC. The reports based on this distinction are still available from the CFTC website and are called legacy COT reports. The non-commercial category included everybody else who was reporting positions for the purpose of this report.

The disaggregated COT reports use classification of traders based on more granular categories:

- producer/merchant/processor/user*: an entity engaged primarily in production, processing, packaging or handling of a physical commodity; futures are used for risk management;
- swap dealer*: an entity that uses futures primarily for management of risk associated with swap transactions – the counterparties may be speculators (for example hedge funds) or commercial companies;
- managed money*: money managers who manage and trade futures on behalf of customers; and
- other reportables*: everybody else not fitting into three categories above.

The disaggregated COT report has been available since September 2009 and was reconstructed by the CFTC back to June 2006.

Both the legacy and disaggregated reports are produced in a number of different formats: futures only (long and short form) and futures and options (long and short form). The option positions are reported on the delta basis. For example, a position in 1,000 option contracts with the delta of 0.5 will be reported as a 500 contracts position.

The disaggregated COT report has greatly improved the quality of reported data compared to the previous format, but it still suffers from one basic shortcoming: it does not provide sufficient information about the activities and positions of the indexers – ie, financial companies offering financial instruments replicating performance of a certain commodity basket (see Chapter 5). The CFTC addressed this problem by producing another report, called supplemental commitments of traders, available since 2007. This report, available for 12 agricultural futures contracts, supplements the old format COT report by removing the futures positions held by indexers from

the aggregate commercial/non-commercial data, and putting them into a separate category. The shortcoming of this report is that a given financial institution identified as an indexer may hold futures positions related to many different types of activities (for example, speculative positions, index-related positions or positions established as hedges of swaps). The CFTC will treat in this case the entire portfolio of futures as index-related (“once an indexer, always an indexer”). This report may be somewhat important to energy traders: it is used sometimes to produce estimates of positions held by indexers in the metals and energy markets.<sup>71</sup> Producing such estimates requires making an assumption that the indexers display the same patterns of activity across a wide spectrum of commodities. This approach has been criticised in a number of studies. The major shortcoming of this estimation technique is that the extent of internal netting between long and short positions varies from commodity to commodity. It is *de minimis* in the case of agricultural markets but quite important in the case of the metals and energy.<sup>72</sup>

The shortcomings of the supplemental COT report were addressed in a report produced under a procedure known as “special call” under Commission Rule 18.05. The call was issued in June 2008 (initially to 43 institutions) and is still continuing. The data is published on a monthly basis (quarterly data prior to June 30, 2010). What is important is that the information is available for the entire book of business, not on a netted basis. An example of a monthly Index Investment Data report is shown as Figure 6.3, while the relationships between different reporting systems discussed in this section are shown in Table 6.6.

The distinction between different classes of traders is based on self-classification. An entity is classified as commercial (in the new producer/merchant/processor/user classification) or a swap dealer by filing the CFTC Form 40<sup>73</sup> – of course, the designation is subject to review by the staff of the CFTC. If a given corporate entity has multiple business units with different business models, these units may be subject to different classifications. Non-commercial traders (ie, anybody outside the producer/merchant/processor/user category) can be either short or long or may be classified as spreading, to the extent they hold offsetting contracts (futures and options) “in the same or different calendar months.”<sup>74</sup> A residual position is reported in the long or short column. Spread positions are very popular in the

**Table 6.6** Index investment data (May 2012)

US Futures Market <sup>1</sup> (Notional value > 0.5 billion US\$) <sup>2</sup>	Notional value (Billions US\$)			Futures equivalent contracts <sup>3</sup> (Thousands)		
	Long	Short	Net L (S)	Long	Short	Net L (S)
Cocoa	1.0	(0.3)	<b>0.7</b>	48	(16)	<b>32</b>
Coffee	4.5	(1.9)	<b>2.6</b>	74	(32)	<b>42</b>
Copper	9.1	(3.6)	<b>5.5</b>	108	(43)	<b>65</b>
Corn	16.2	(5.5)	<b>10.7</b>	596	(200)	<b>397</b>
Cotton	4.2	(1.6)	<b>2.6</b>	118	(44)	<b>75</b>
Feeder cattle	0.8	(0.2)	<b>0.6</b>	10	(3)	<b>7</b>
Gold	22.8	(6.5)	<b>16.4</b>	146	(41)	<b>104</b>
Heating oil	9.3	(2.8)	<b>6.5</b>	81	(25)	<b>56</b>
Lean hogs	5.3	(2.2)	<b>3.0</b>	150	(62)	<b>88</b>
Live cattle	8.9	(3.2)	<b>5.7</b>	185	(67)	<b>118</b>
Natural gas	15.5	(5.2)	<b>10.3</b>	596	(212)	<b>384</b>
Platinum	0.7	(0.2)	<b>0.5</b>	10	(2)	<b>8</b>
RBOB unleaded gas	10.3	(2.1)	<b>8.2</b>	93	(19)	<b>75</b>
Silver	5.8	(1.6)	<b>4.3</b>	42	(11)	<b>31</b>
Soybean meal	0.8	0.0	<b>0.7</b>	21	(1)	<b>19</b>
Soybean oil	5.1	(1.8)	<b>3.3</b>	171	(62)	<b>109</b>
Soybeans	18.6	(6.0)	<b>12.6</b>	285	(91)	<b>194</b>
Sugar	8.7	(3.0)	<b>5.6</b>	394	(139)	<b>255</b>
Wheat (CBOT)	12.4	(5.9)	<b>6.4</b>	377	(182)	<b>194</b>
Wheat (KCBT)	1.8	(0.4)	<b>1.5</b>	55	(12)	<b>43</b>
WTI crude oil	46.0	(14.0)	<b>32.0</b>	529	(161)	<b>368</b>
Subtotal (>0.5 billion US\$)	207.7	(68.1)	<b>139.6</b>			
Subtotal (<0.5 billion US\$)	0.8	(0.2)	<b>0.7</b>			
Total notional US mkts	208.6	(68.3)	<b>140.3</b>			
Total not'l non-US markets	61.3	(17.6)	<b>43.7</b>			
Total all markets	269.9	(85.9)	<b>184.0</b>			

<sup>1</sup>Each listed US futures market includes index investment for all futures and OTC markets related to or referenced to that US futures market. For example, the US market listed as "WTI Crude Oil" includes (with the NYMEX's Light "Sweet" crude oil futures market) investments held in the NYMEX "Crude Oil Financial" market and the ICE Futures-Europe WTI Light Sweet crude oil market, because both of those contracts' settlement prices are determined by reference to the NYMEX Light "Sweet" crude oil futures contract.

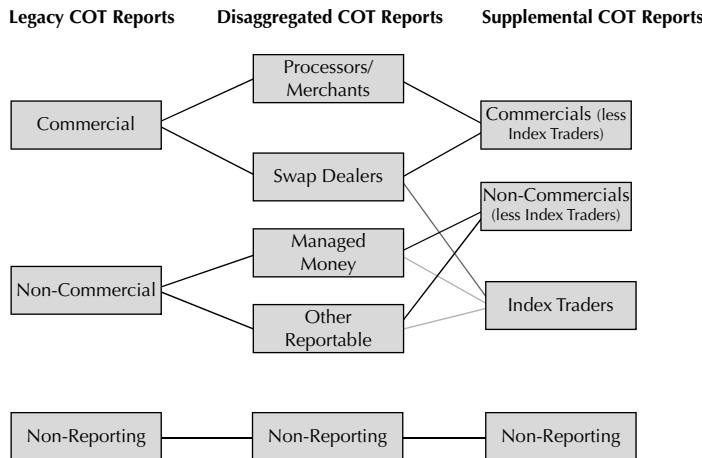
<sup>2</sup>US Futures Markets with 0.5 billion US dollars or more in reported net index investment notional value on the report date.

<sup>3</sup>Futures Equivalent Contracts: Futures plus delta-adjusted options, estimating what Futures Contracts would have been established absent offsets.

Source: CFTC, <http://www.cftc.gov/MarketReports/IndexInvestmentData/index.htm>

energy markets as seasonal plays or as ways to make directional bets on a specific month using an offsetting position in another month contract as a partial hedge.

The breakdown of traders into commercial and non-commercials corresponds to the pre-1982 distinction used in the COT between

**Figure 6.3** Legacy, Disaggregated, and Supplemental COT Reports

Source: Scott H. Irwin and Dwight R. Sanders, "Testing the Masters Hypothesis in Commodity Futures Markets," October 2010

speculative and non-speculative positions. The classification of traders is a controversial subject, and it is often discussed to what extent developments in the financial markets should be recognised in the modifications of COT reports. For example, one could argue that a big oil company may enter into a transaction with a hedge fund on one side and hedge this position on Nymex. The latter position could be classified as commercial but its economic essence would be a conduit to establish indirectly a speculative, directional position.<sup>75</sup>

The data available from COT reports are often used to divine the intention of the big speculative players associated with the hedge funds and commodity advisers. The data underlying COT reports were used in a number of studies on the impact of speculators on commodity prices.

The large trader reporting system is an additional data collection system about the activities of large traders whose positions exceed the limits determined by the CFTC. This information, gathered daily, is provided by clearing members of the exchanges, FCMs and foreign brokers, collectively identified as reporting firms.<sup>76</sup> The data collected include futures and option on futures positions. If a trader

exceeds a reporting threshold in a single contract or option expiration month, their entire position in a given commodity has to be reported. According to the CFTC, aggregate large trader positions represent 70–90% of the open interest in any given market. The thresholds range from 25 to over 1,000 contracts, depending on the overall size of the given market. The thresholds are determined by the CFTC based on the total open interest, the size of positions held by different traders, the history of trading in a given commodity and the availability of physical commodity for delivery in case of market disruptions. Positions held by a single trader across many different brokerage firms are aggregated based on the additional information collected by the CFTC. The accuracy of reported numbers is validated through cross-calibration of data received by the CFTC from different sources.

### The uses of COT reports

Energy markets analysts should be warned against excessive reliance on, and drawing far-reaching conclusions from, COT reports. These reports provide a relatively timely but partial snapshot of commodity-related derivative positions. Judging the shape and size of an iceberg relying on partial information is a dangerous practice, as an officer on the deck of Titanic must have concluded before hitting the water. These reports may also lead to wrong conclusions due to potential distortions resulting from self-selection bias. Many traders we have talked to believe that exchanges attract speculators relying on strategies tailored for a reliance on the futures markets, such as trend-following and algorithmic trading. Futures, especially when screen trading is available, allow for quick adjustment of positions. Traders that take a long-term view, try to optimise collateral and seek to hide their positions, and often choose to rely to a greater extent on the OTC markets. Their positions may be eventually reflected in COT reports, but may show up under different labels as the positions of swap dealers or merchant energy companies. Most studies of the impact of speculation on the commodity markets are based on COT data (or more granular data underlying the COT reports), and predictably find that speculators follow the markets, instead of leading them.

If our trader friends are correct in their characterisation of speculative strategies and the migration of certain types of speculators to

exchange-based instruments, the conclusions of quantitative studies of the impact of speculation on the commodity markets are easy to understand. If one studies a market dominated by trend followers, one will find that market participants are trend followers. Of course, this observation does not answer the basic question about the influence speculators have on the direction of commodity prices: the impact may be huge or *de minimis*. The point we are trying to make is that economic policy and regulatory solutions should not be driven by the studies of the impact of speculators on commodity prices based on incomplete data.

These remarks are not intended to send a message that COT reports are useless and should be ignored. One can often find in them the evidence of important shifts in market structure. One example of creative use of COT information can be found in the 2012 *Financial Times* "Alphaville" post about the WTI market.<sup>77</sup>

Figure 6.4 shows net positions of different groups of WTI market participants derived from COT reports. The futures contracts used for this figure are the Nymex sweet crude contract (contract symbol CL) and the corresponding ICE contract (contract symbol T). (Note that, we talk here about net short or long positions so the word net is not repeated for convenience.)

As one can see, the short positions of producers, offsetting in the past long positions of managed money (primarily index funds) have been dwindling in size since 2011. There are several explanations of this development one can find in the cited "Alphaville" post. Most likely, the reason for the observed structural shift is the migration of some market participants away from a contract based on WTI (most likely to the Brent contract). This is due to partial decoupling of WTI prices from the prices of other grades of crude oil that became chronic in 2011 (see more on this point in the chapters on the oil markets). On the other hand, short positions of swap dealers have been increasing, reflecting the growing scale of activities in which they were accommodating growing long positions held by managed money.

Another example of the use of COT reports is the assessment of the impact of speculators and speculative activity on commodity prices. The formula used to accomplish this is based on the contribution of Holbrook Working,<sup>78</sup> who came up with a measure of speculative activity (Working's T index) in a futures market. The

logic behind this index is very simple. In a market without speculators (it would be probably a very dull market), the commercial longs and shorts would be, by necessity, perfectly balanced. In reality, there is never perfect balance and the speculators have to step in.

The size of the gap they have to fill can be used to measure the level of speculative heat. The formulas for Working's T ratio are given below:<sup>79</sup>

$$T = 1 + \frac{SS}{HL+HS} \text{ if } HS \geq HL$$
$$T = 1 + \frac{SL}{HL+HS} \text{ if } HL > HS$$

where SS stands for the short positions of the speculators

SL – long positions of the speculators

HS – short positions of the hedgers

HL – long positions of the hedgers

All the position levels are in absolute numbers.

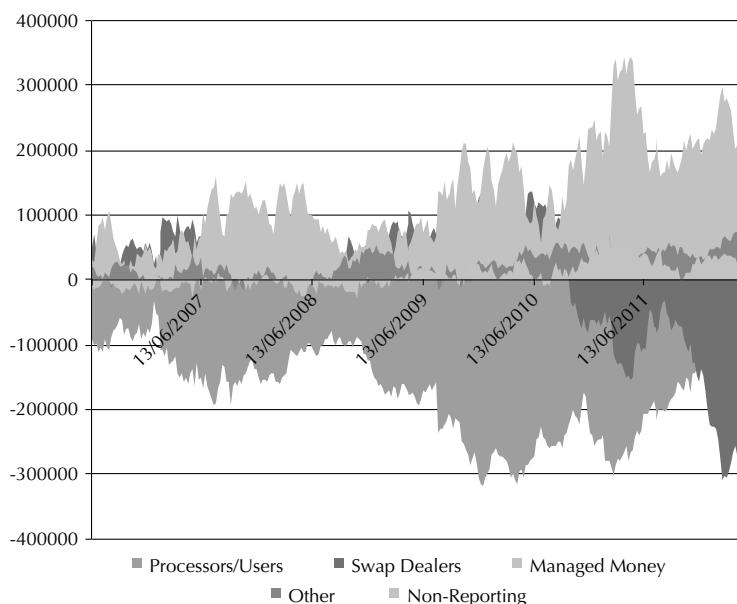
The history of the T-ratio shown in Figure 6.5 for the crude oil Nymex futures contract confirm the intuitive expectations of even a casual observer of the energy markets, but have to be used with caution. The CFTC data underlying the calculations are not flawless.

As mentioned above, the distinctions between speculators and hedgers are based on self-classification (through the CFTC Form 40) and fail to represent a very thin line separating these two types of trading, and the fact that in any trading organisation the speculative and hedging activities may be combined under one common umbrella. It is also important to recognise that speculative activity is typically concentrated in the first few most liquid contracts, whereas hedging positions may be extended over a longer range of maturities. The aggregate nature of the CFTC data does not allow for making this distinction. More information about the index of speculative activity for US energy markets can be found in a research paper by Bahattin Büyüksahin and Michel A. Robe (see footnote 79).

### **TO CLEAR OR NOT TO CLEAR?**

The Dodd–Frank Act requires that all standardised derivatives be traded on exchanges and cleared. This happened to be a very unpopular provision from the point of view of the end users of derivatives.

**Figure 6.4** Combined Nymex and ICE light sweet crude oil: Net positions of traders

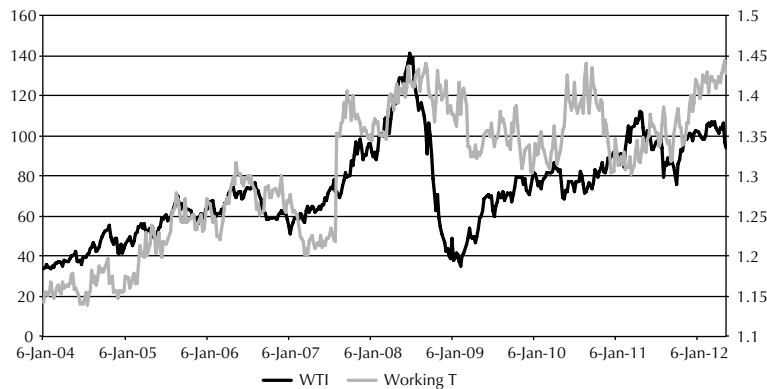


Source: CFTC

Protests against this part of the Act (and objections to a similar proposal in Europe) led to the “end-user exception”: a waiver of the requirement for *bona fide* hedgers. This section explains the reasons for this controversy.

The dilemma is very simple but the debate surrounding the issue is quite contentious.<sup>80</sup> An end user of derivatives (a commercial hedger) has the choice of transacting on an exchange such as Nymex (to the extent a contract satisfying their criteria of hedge effectiveness is available) or transacting in the OTC markets, with a financial institution or trading desk of a big oil company or utility acting as counterparty. In the first case, the hedging entity has to post an initial margin and possibly a variation margin. In the second case, the exposures under a bilateral transaction may or may not be collateralised. Under a frequently used arrangement, a hedge provider (typically a financial institution) waives the collateral requirements (ie, a deposit of cash, letter of credit, marketable securities serving as guarantee of performance under a derivative contract) by extending to the

**Figure 6.5** Speculative activity in the Nymex crude oil futures contract, Working's T ratio



Notes: CFTC COT Reports and author's calculations, futures only <http://www.cftc.gov/MarketReports/CommitmentsofTraders/HistoricalCompressed/index.htm>  
WTI – weekly prices

counterparty an automatic credit facility, secured with the company's assets, corporate guarantees or letters of credit. A financial institution may also accept unsecuritised exposure up to a credit threshold negotiated with counterparty. An example of such an arrangement was provided in one of the regulatory hearings on the subject:<sup>81</sup>

[On] June 30, 2008, when Chesapeake owed approximately US\$6.3 billion under our OTC derivative contracts, we had pledged collateral valued at more than US\$11 billion to our derivative counterparties. The collateral we pledged included both mortgages on our oil and gas properties – our underlying business assets – and letters of credit. While the security is not in cash, our counterparties were and continue to be well-secured. This is how most end-users utilise this market and, as a result, help alleviate systemic risk.

Financial institutions have a strong preference for such arrangements, as explained in a very lucid way by J. P. Morgan:<sup>82</sup>

Exposure to end-users that is either secured by assets that a clearing-house can't accept (eg, property, reserves, equipment, commodities) or is unsecured[,] does not pose a systemic risk. From the standpoint of the end-users, the ability to pledge this kind of collateral or to transact on an unsecured basis is very favorable, as it allows them to enter into risk management using the assets that they own in their businesses with no disruption to those businesses and without draining their liquidity to come up with cash. From the standpoint of

the dealers, exposure to these end-users often is what is referred to as “right way risk” in that the exposure moves in the same direction as the value of the collateral or the overall business.

Many bankers we talked to emphasise the benefits of offering hedges in the context of managing the overall business relationship with a customer. Better understanding of customer’s business results in better credit terms, both in terms of the cost and the levels of credit lines. They see the importance of accepting as collateral only well-defined, specific assets of a client, and avoid relying on the overall assets, which may be pledged under other contracts.

As mentioned above, during the political debates on reform of the financial system in the US (related to the Dodd–Frank Act) and in Europe, a proposal was floated to mandate the clearing of all “standardised” derivatives. This provision was met with a universal outcry from the end users of derivatives, who objected to this measure in an exceptionally strong way, forming a unified front to pre-empt its passage.<sup>83</sup> A good summary of the rationale for this position can be found in the testimony by Thomas C. Deas, Jr., Vice President & Treasurer, FMC Corporation:<sup>84</sup>

By forcing end users to post cash margin, the regulators will take the balanced structure I’ve just described and impose a new risk. Treasurers will have new and unpredictable demands on their liquidity. Swap dealers are market makers who take open positions with derivatives and we agree central clearing and margining is appropriate for them. However, since end users are balanced, with derivatives exactly offsetting underlying business risks, forcing them into the swap dealers’ margin rules adds the considerable risk for end users of having to fund frequent cash margin payments. This will introduce an imbalance and new risks onto transactions that are matched and will settle with offsetting cash payments at maturity.

The end users of derivatives contesting mandatory clearing of derivatives emphasise the macroeconomic consequences of this solution, as explained in Mr. Deas’ testimony:<sup>85</sup>

Forcing end users to put up cash for fluctuating derivatives valuations means less funding available to grow their businesses and expand employment. The reality treasurers face is that the money to margin derivatives has to come from somewhere and inevitably less funding will be available operate their businesses.

The unified front of end users of derivatives resulted in the extension of a waiver from this provision (referred to as an exception) in the

Dodd–Frank Act. The position taken by the commercial hedgers contrasts sharply with very critical comments coming from a few practitioners and derivatives experts, who believe that the position taken by the hedgers is an example of “the turkey voting for Thanksgiving or Christmas” and borders on delusion. The arguments and counter-arguments revolve around the explicit and hidden costs of hedging arrangements.

The critics point out that there is no free lunch. If a financial institution extends a credit line to a client (securitised or not), the opportunity cost of the loan has to be recovered somehow, with the most obvious source being an increased profit margin on the hedging transaction. Pricing of OTC transactions is often opaque and the hedgers do not have information about the bid and offer prices: a producer who sells natural gas forward does not know<sup>86</sup> what price is charged to an electricity producer when natural gas is sold down the line. An automatic credit facility extended to a company in lieu of collateral is effectively a loan, and a company that pledges its assets as collateral, or offers a corporate guarantee, encumbers some of its assets and has its explicit borrowing capacity reduced. Corporate guarantees and letters of credit are not free and should be explicitly priced. A hedging entity would be better off, critics say, hedging on an organised exchange and borrowing to post initial and variation margins. The overall cost of explicit loans and lower bid–offer spreads and transactions costs arising from trading on a transparent platform would bring net savings compared to an arrangement based on indirect loans and inflated and unknown bid–offer spreads. This point of view can be best summarised through the following statement from a post by John Parsons and Antonio Mello:<sup>87</sup>

The micro mistake is the delusion that absent a collateral requirement companies are able to trade derivatives at no cost to their balance sheet. This is plainly not true. If you don't back up your derivative trades with a cash collateral account, then you are backing them up with a promise that you are good for it, ie, with credit. Companies have limited debt capacity, so using credit is costly, too. A regulation that requires using cash instead of credit costs the company on one side, but loosens its constraints on the other. The net effect on the company's free cash flow is zero. [...] One could argue that the cash requirement is costlier than credit, but then you would have to figure out by how much. That would be an extra, very difficult step in the calculation, and any reasonable estimate for the differential would drive the headline number down enormously, possibly to zero.

Can these divergent points of view be reconciled? The end users of derivatives firmly believe that they would be better off without the mandatory clearing of derivatives. Companies with better credit ratings can negotiate with the hedge providers high credit thresholds (below which collateral is not required) and automatic credit facilities (secured or not) if the credit thresholds are exceeded. Companies with lower credit ratings rely on loans secured with assets. Any CFO has to understand that such credit facilities or liberal credit thresholds come at a cost and that the financial institutions have to be compensated for their largesse in a number of ways, primarily through higher bid–offer spreads.<sup>88</sup> Failing to recognise this would be equivalent to financial illiteracy. At the same time, end users believe that this additional cost is justified for a number of reasons.

- ❑ The credit arrangements with the providers of derivatives based on the waivers of collateral requirements and automatic credit lines leave the end users with more liquidity (“It’s all about cash,” I was told many times by corporate CFOs and risk managers).
- ❑ Bilateral arrangements with hedge providers allow for the construction of more precise, customised hedges, and the benefits of better hedges justify additional costs. The accounting rules in the US strongly favour effective hedges.<sup>89</sup>
- ❑ A significant shift in a forward price curve may trigger a liquidity crisis in case future volumetric flows are hedged with futures. A company has to borrow on a short notice and this may be difficult at times. A higher cost of hedging is a premium paid for avoiding financial distress.

Another important factor behind the strong preference for bilateral hedging arrangements seems to be, as always, accounting. A company hedging on an exchange and clearing through a clearing-house would have to raise additional debt or reduce its liquid assets in order to post margins. An automatic credit facility from a hedge provider or a waived margin requirement may not register on the firm’s balance sheet as debt. Of course, in a perfect world, a credit rating agency should see through the veil of accounting and recognise that a debt is a debt is a debt. We do not live, however, in a perfect world.

The comments made above can be summarised in term of the pros

and cons of different approaches to mandatory clearing. From the point of view of an end user of derivatives, an exemption<sup>90</sup> from mandatory clearing and a reliance on the bilateral arrangements described above has many advantages, primarily in terms of cash and liquidity management.

- ❑ An end user of derivatives can negotiate multiple bilateral credit arrangements with several providers of risk management instruments and avoid posting collateral by flying below the level of a threshold with each counterparty.
- ❑ A financial institution managing the collateral exposure of a client in the context of an overall business relationship can be more comfortable extending credit on better terms if it understands all their balance sheet and off-balance sheet liabilities.
- ❑ Access to collateral funding may be obtained at short notice without long negotiations with creditors, with potentially unpredictable consequences. A loan negotiated under duress is likely to be more expensive than credit extended under a revolving facility.

There is, however, downside to the bilateral approach to management of credit risk.

- ❑ An end user of derivatives may become a captive customer of a financial institution, extending to them an automatic credit in lieu of collateral collection. Trading on an exchange and relying on a central CCP offers flexibility – and sometimes flexibility is worth paying the necessary price.
- ❑ The benefits of automatic credit arrangements replacing collateral collection are driven to a large extent by an accounting arbitrage.<sup>91</sup> A company relying on exchange-traded derivatives and central clearing, and required to post initial and potentially variation margins, would have to reduce its cash holdings and/or to borrow. A bilateral arrangement with a hedge provider does not have the same accounting consequences and may not show up as a loan. It still leaves the hedging company with encumbered assets (reducing its overall borrowing capacity) and with a contingent liability. It is effectively a way of borrowing without making it explicit on the balance sheet. At the end of the

day, however, a liability is a liability. Some liabilities are contingent and will fluctuate in value, but they cannot be ignored.

In the long run, this debate should be settled as a purely empirical study of the costs and benefits of different credit arrangements. The provisions of the Dodd–Frank Act that require reporting of swap transactions to SDRs and dissemination of this information to the public will allow end users to assess better the true costs of hedging. Some end users of derivatives will most likely suffer from buyer's remorse. Sophisticated end users of derivatives that are capable of negotiating favourable terms will continue to rely on bilateral arrangements.

## CONCLUSIONS

This chapter covered a number of issues related to the organisation and operations of exchanges and clearinghouses. We have addressed a number of topics related to position limits, margining (the SPAN system), some provisions in the Dodd–Frank Act regarding mandatory clearing of standardised derivatives and the controversy surrounding this requirement. We have reviewed COT reports that provide a useful window into trading patterns in the futures markets.

Exchanges and clearinghouses are a critical part of the infrastructure of the energy markets, and this happens to be the part that is certain to undergo major transformation. Alignment of competitive and regulatory pressures combined with technological progress will change the landscape of this industry. One of the developments one can expect is the propagation of high-frequency trading in energy commodity markets. It is difficult to report on the developments in this area given the proprietary nature of algorithms currently under development in major financial institutions and hedge funds, but the work proceeds at a hectic pace. Another development will be increasing regulatory scrutiny and oversight of clearinghouses.<sup>92</sup> This will be a somewhat unexpected result of financial reforms which elevated clearinghouses to the status of systemically important institutions. This, in turn, is likely to lead to extension to clearinghouses of access to central bank emergency funding in the event of a liquidity crisis. Any market practitioner should stay tuned and follow the developments in this area.

- 1 Similar developments are likely to take place in Europe under the proposed European Market Infrastructure Regulation (EMIR). This is consistent with the commitment made by the leaders of G20 in Pittsburgh in September 2009. "All standardised OTC derivative contracts should be traded on exchanges or electronic trading platforms, where appropriate, and cleared through central counterparties by end 2012 at the latest. OTC derivative contracts should be reported to trade repositories. Non-centrally cleared contracts should be subject to higher capital requirements." (See <http://www.g20.utoronto.ca/2009/2009communique0925.html>.)
- 2 A SEF is defined by the Act to be "a facility, trading system or platform in which multiple participants have the ability to execute or trade swaps by accepting bids and offers made by other participants that are open to multiple participants in the facility or system, through any means of interstate commerce."
- 3 The Code of Hammurabi contained provisions that can be interpreted as regulation of options embedded in commercial contracts (see Robert E. Whaley, 2006, *Derivatives: Markets, Valuation, and Risk Management* (Hoboken, NJ: Wiley), p 11.
- 4 <http://www.cftc.gov/IndustryOversight/TradingOrganizations/index.htm>.
- 5 Comex was established in 1933 through the combination of four other exchanges: the National Metal Exchange, the Rubber Exchange of New York, the National Raw Silk Exchange and the New York Hide Exchange.
- 6 The history of the CME is outlined in the recent 10-K company filing: "CME was founded in 1898 as a not-for-profit corporation. In 2000, CME demutualised and became a shareholder-owned corporation. As a consequence, we adopted a for-profit approach to our business, including strategic initiatives aimed at optimizing trading volume, efficiency and liquidity.[...]In 2007, CME Holdings merged with CBOT Holdings, Inc. and was renamed CME Group.[...]In 2008, [CME Group] acquired Credit Market Analysis Limited and its three subsidiaries (collectively, CMA)."
- 7 <http://www.cmegroup.com/trading/cme-direct/>.
- 8 "CME Group overview," UBS Global Financial Services Conference, May 8, 2012 (<http://investor.cmegroup.com/investor-relations/events.cfm>).
- 9 ECNs, electronic platforms for trading outside exchanges, are also known as alternative trading systems or alternative trading networks (ATS/ATN).
- 10 Victoria Woolley, 2002, "Step up to trade at the NYMEX," *Stocks & Commodities*, VI. (20)2, pp 38–42.
- 11 The employees of Nymex crawl on the floor of the pit wearing goggles, as the pit cards have sharp edges and may hurt their eyes.
- 12 The founder of ICE, Jeffrey Sprecher, acquired Continental Power Exchange, Inc, with the intention of creating a web-based platform for OTC energy commodity trading. In May 2000, IntercontinentalExchange (ICE) was established through the participation of a number of financial and merchant energy companies. Initial partners included BP, Deutsche Bank, Goldman Sachs, Morgan Stanley Dean Witter, Royal Dutch/Shell Group, SG Investment Banking, Totalfina Elf and Continental Power Exchange. In 2002, several leading US merchant companies entered into an agreement to purchase equity position in ICE (American Electric Power, Aquila Energy (a unit of UtiliCorp United), Duke Energy, El Paso Energy, Reliant Energy and Southern Company Energy Marketing (a unit of Southern Company)).
- 13 IntercontinentalExchange, 10-K for 2007.
- 14 ICE acquired European Climate Exchange (ECX) when it purchased in 2010 the Climate Exchange Plc. CFI stands for Carbon Financial Instruments.
- 15 In the US, the commodity exchanges own the clearinghouses. This arrangement is likely to change following the Dodd-Frank Act.
- 16 Intercontinental Exchange, Form 10-K, 2010, p 39 (<http://files.shareholder.com/downloads/ICE/1231523476x0xS1193125-11-28245/1174746/filing.pdf>).

- 17 This electronic platform supports trading of ICE energy futures and Canadian agricultural markets, in addition to the OTC markets. As of December 31, 2011, ICE supported 1,900 OTC participant firms and over 900 futures participants.
- 18 As explained in the ICE 10-K filing, block trades are executed in the voice broker market, typically over the telephone or through an instant messaging service, and then transmitted to ICE electronically for clearing.
- 19 The 2008 Farm Bill introduced the concept of the OTC contracts that serve a significant price discovery function. The CFTC identified 14 such contracts traded on ICE and the exchange was required to assume certain self-regulatory responsibilities (market monitoring, position limits).
- 20 Recent ICE decision to change OTC contracts to futures was taken too late to be covered in this book. Many statements of this paragraph may be incorrect in the light of this decision and the reader is encouraged to investigate this matter further.
- 21 [http://www.tmx.com/en/about\\_tsx/](http://www.tmx.com/en/about_tsx/).
- 22 NGX Overview, July 2010 (see the NGX website).
- 23 "Watt-Ex is an electronic trading system for electricity ancillary services, also known as operating reserves. Watt-Ex commenced operations in 2001 and was wholly acquired by NGX in November 2006," (see <http://www.ngx.com/wattex.html>).
- 24 76 Fed. Reg. 71,626 (November 18, 2011).
- 25 More information about the convoluted history of position limits can be found in the testimony by Dan M. Berkovitz, General Counsel of the CFTC, "Position Limits and the Hedge Exemption, Brief Legislative History," July 28, 2009 (see <http://www.cftc.gov/PressRoom/SpeechesTestimony/2009/berkovitzstatement072809>).
- 26 <http://www.cftc.gov/ucm/groups/public/@lrfederalregister/documents/file/2011-28809-1a.pdf>.
- 27 The significant price discovery concept was introduced by the Farm Bill of 2008.
- 28 H. R. 4173, p 349 ([http://www.cftc.gov/ucm/groups/public/@swaps/documents/file/hr4173\\_enrolledbill.pdf](http://www.cftc.gov/ucm/groups/public/@swaps/documents/file/hr4173_enrolledbill.pdf)) 7, U.S.C § 6(a)(1).
- 29 H. R. 4173, p 349 ([http://www.cftc.gov/ucm/groups/public/@swaps/documents/file/hr4173\\_enrolledbill.pdf](http://www.cftc.gov/ucm/groups/public/@swaps/documents/file/hr4173_enrolledbill.pdf)).
- 30 The spot month is defined in Regulation 151.3.
- 31 In other words, the position limits regime applies to all economically equivalent contracts.
- 32 The CFTC was created by Act of Congress in 1974. Its predecessor regulatory bodies included the Grain Futures Commission, which became the Commodity Exchange Commission in 1936.
- 33 For energy commodities, "[t]he spot month shall be the period of time commencing at the close of business of the third business day prior to the last day of trading in the underlying Core Referenced Futures Contract and terminating at the end of the delivery period for the [...] Referenced Contracts[...]" See 17 C.F.R § 151.3.
- 34 76 Fed. Reg. at 71,633.
- 35 In the case of Henry Hub natural gas, a trader will have to comply effectively with three rules: the spot month limits for physical delivery contracts, the spot month limit for natural gas cash-settled contracts (five times the physical limit) and the aggregate limit for both types of contracts (five times the physical contract limit).
- 36 Federal Register, 76(223), Friday, November 18, 2011, "Rules and regulations."
- 37 Federal Register, 76(223), Friday, November 18, 2011, "Rules and regulations" (17 C.F.R. § 151.4(b)(1)).
- 38 Federal Register, 76(223), Friday, November 18, 2011, "Rules and regulations" (17 C.F.R. § 151.5(b)(1)).
- 39 [http://www.nymex.com/ewd\\_margins.aspx](http://www.nymex.com/ewd_margins.aspx).
- 40 In the European Union, regulators are pushing for a similar reform and for the mandatory exchange trading and clearing of most derivatives.

- 41 The history of early clearinghouses is discussed in Peter Norman, 2011, *Risk Controllers: Central Counterparty Clearing in Globalised Financial Markets* (Chichester, England: Wiley).
- 42 ClearPort accepted for clearing transactions executed in the OTC markets as long as they were standardised and both counterparties agreed to take this route.
- 43 "For example, when ICE severed its ties with LCH.Clearnet in 2008 it managed to boost the amount of revenue from its benchmark Brent futures from US\$1.30 per contract to US\$1.53," according to Mr O'Shaughnessy's estimates. "That can be directly attributed to them bringing clearing in-house," he says. "Moreover, analysts see clearing as a growth industry in its own right. In the OTC markets in particular, the shift to clearing because of G20 reforms such as the US Dodd-Frank Act is likely to bring new business for exchanges that also have clearing." Jack Farchy and Jeremy Grant, 2012, "Prospect of LME's sale puts focus on clearing," *Financial Times*, February 22.
- 44 IntercontinentalExchange, Inc., Form 10-K For the Year Ended December 31, 2011, p 37.
- 45 Markets in Financial Instruments Directive (Directive 2004/39/EC), known as MiFID, came into force on November 1, 2007. MiFID II is the version of the original directive updated in the wake of the later financial crisis (Directive 2008/10/EC of the European Parliament and of the Council, March 11, 2008).
- 46 Peter Norman, *ibid*. A discussion of the conditions a clearing member has to meet can be found, for example, at <http://www.cmegroup.com/clearing/cme-clearing-overview/clearing-membership.html>.
- 47 <http://www.cmegroup.com/clearing/files/financialsafeguards.pdf>.
- 48 These cases are documented extensively in the book by Peter Norman, *ibid*.
- 49 <http://www.cmegroup.com/clearing/files/financialsafeguards.pdf>.
- 50 "Maintenance performance bond levels represent the minimum amount of protection against potential losses at which the clearing house will allow a clearing member to carry a position or portfolio. Initial performance bond reflects the minimum deposit a clearing member must obtain from a customer opening a new position. Should performance bonds on deposit at the customer level fall below the maintenance level, Exchange rules require that the account be re-margined at the required higher initial performance bond level." "CME Clearing requires 'gross' performance bonds for customer segregated positions in CME and NYMEX products." <http://www.cmegroup.com/education/events/pdfs/PerformanceBonds.pdf>
- 51 "CME Group's Audit Department routinely inspects the books and records of clearing members to ensure, among other things, their compliance with segregation requirements."
- 52 A detailed discussion of CME capital requirements is beyond the scope of the book. Every risk manager should review the relevant documents of CME/other exchanges they use for full information.
- 53 *Ibid*.
- 54 For example, as of December 30, 2011, CME Clearing held aggregate performance bonds of approximately US\$90 billion. Additional available resources that can be deployed in the event a defaulting firm has no resources left are discussed by the CME at <http://www.cmegroup.com/clearing/files/financialsafeguards.pdf>.
- 55 Richard Dale, 1998, *Risk Management in US Derivative Clearing Houses* (London, England: London Institute of International Banking, Finance & Development Law). The report cited by Richard Dale was prepared by the Committee on Payment and Settlement Systems of the central banks of the G10 Countries (May 1997).
- 56 During the bankruptcy of Lehman Brothers, the team from London Clearing House was not initially allowed on the premises of the bankrupt company, and the employees of Lehmans were not allowed to communicate with clearinghouse personnel. This was due to the confusion regarding the laws that applied in bankruptcy in this specific case. The system worked in the end, but this case demonstrates the risk that unexpected minor glitches can produce potentially catastrophic results. After all, a great kingdom was lost for the want of a nail.

- 57 In 2006, the system was licensed to 53 registered exchanges, clearing organisations, service bureaus and regulatory agencies.
- 58 The description of SPAN is based on the documents available from the CME: "CME SPAN, The Standard Portfolio Analysis of Risk," Chicago Mercantile Exchange 2006, "CME SPAN," undated presentation.
- 59 Paul H. Kupiec and Patricia White discuss extensively the historical and institutional ramifications of competing US margining systems in the working paper "Regulatory competition and the efficiency of alternative derivative margining systems," Federal Reserve Board, June 11, 1996.
- 60 The Federal Reserve Board was given margin authority under the Securities and Exchange Act of 1934. Excessively low margin requirements were seen as one of the roots of the 1929 stock market crash and were targeted by the New Deal architects as one of the critical systemic reforms.
- 61 Regulation T, issued by the Board of Governors of the Federal Reserve System, applies to extension of credit by the US securities dealers and brokers.
- 62 An additional margin of 10% is required for member customers and an additional margin of 35% for non-member customers.
- 63 Letter from Brian Regan, Nymex, vice president and counsel to the CFTC, dated February 6, 2004. <http://www.cftc.gov/files/tm/tmnymexfebruary62004letter.pdf>.
- 64 Maximum move is defined as 300% price scan range. The resulting gain or loss is then multiplied by 0.32.
- 65 Positive numbers indicate losses according to the conventions used in SPAN.
- 66 See, for example, Devraj Basu and Joëlle Miffre, 2009, "The performance of simple dynamic commodity strategies" ([http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1353059](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1353059)).
- 67 The description of COT can be found at: [http://www.cftc.gov/marketreports/commitmentsoftraders/cot\\_about.html](http://www.cftc.gov/marketreports/commitmentsoftraders/cot_about.html).
- 68 Current and historical reports can be found at: <http://www.cftc.gov/MarketReports/CommitmentsofTraders/index.htm>.
- 69 The information is available in 17 CFR 15.03
- 70 CFTC Regulation 1.3(z), 17 CFR 1.3(z) defines *bona fide hedging transactions and positions* – (1) *General definition*. Bona fide hedging transactions and positions shall mean transactions or positions in a contract for future delivery on any contract market, or in a commodity option, where such transactions or positions normally represent a substitute for transactions to be made or positions to be taken at a later time in a physical marketing channel, and where they are economically appropriate to the reduction of risks in the conduct and management of a commercial enterprise, and where they arise from:
- (i) The potential change in the value of assets which a person owns, produces, manufactures, processes, or merchandises or anticipates owning, producing, manufacturing, processing, or merchandising,
  - (ii) The potential change in the value of liabilities which a person owns or anticipates incurring, or
  - (iii) The potential change in the value of services which a person provides, purchases, or anticipates providing or purchasing.
- Notwithstanding the foregoing, no transactions or positions shall be classified as *bona fide hedging* unless their purpose is to offset price risks incidental to commercial cash or spot operations, and such positions are established and liquidated in an orderly manner in accordance with sound commercial practices and, for transactions or positions on contract markets subject to trading and position limits in effect pursuant to section 4a of the Act, unless the provisions of paragraphs (z) (2) and (3) of this section and §§1.47 and 1.48 of the regulations have been satisfied.
- 71 M. W. Masters and A. K. White, 2008, "The accidental Hunt Brothers: how institutional investors are driving up food and energy prices," (<http://accidentalhunbrothers.com/wp->

content/uploads/2008/09/accidental-hunt-brothers-080731.pdf); K. J. Singleton, 2011, "Investor flows and the 2008 boom/bust in oil prices," working paper, Graduate School of Business, Stanford University.

- 72 One explanation I can offer is that farmers usually hedge directly using futures exchanges; energy companies and metal producers rely to a greater extent on the OTC markets.
- 73 Form 40 is a five-page questionnaire that traders use for self-classification.
- 74 <http://www.cftc.gov/MarketReports/CommitmentsofTraders/DisaggregatedExplanatoryNotes/index.htm>.
- 75 This is how the CFTC is signalling the potential limitation of the data. "Commission staff reviews the reasonableness of a trader's classification for many of the largest traders in the markets based upon Form 40 disclosures and other information available to the Commission. As described above, the actual placement of a trader in a particular classification based upon their predominant business activity may involve some exercise of judgment on the part of Commission staff. Some traders being classified in the "swap dealers" category engage in some commercial activities in the physical commodity or have counterparties that do so. Likewise, some traders classified in the "producer/merchant/processor/user" category engage in some swaps activity. Moreover, it has always been true that the staff classifies traders not their trading activity. Staff will generally know, for example, that a trader is a "producer/merchant/processor/user" but we cannot know with certainty that all of that trader's activity is hedging. Staff is working on improvements to the Form 40 and other methodologies in order to improve the accuracy of the trader classifications. When large reporting or classification issues are found, an announcement is made and corrections are published as quickly as possible." (see <http://www.cftc.gov/MarketReports/CommitmentsofTraders/DisaggregatedExplanatoryNotes/index.htm>). It is worth noting that Enron was classified by the CFTC as a commercial entity in COT reports.
- 76 The information is collected under Part 17 of the CFTC regulations (see <http://www.cftc.gov/IndustryOversight/MarketSurveillance/LargeTraderReportingProgram/index.htm> for more details).
- 77 Izabella Kaminska, 2012, "The 'other reportables' oil mystery," *Financial Times*, May 14 (see <http://ftalphaville.ft.com/search?q=where+all+the+hedgers>).
- 78 Holbrook Working (1895–1985) was a professor of economics and statistics at Stanford University. He made significant contributions to the theory of hedging and futures prices.
- 79 Bahattin Büyüksahin and Michel A. Robe, 2009, "Commodity traders' positions and energy prices: Evidence from the recent boom–bust cycle," working paper, December 29; Bahattin Büyüksahin and Jeffrey H. Harris, 2011, "Do speculators drive crude oil futures prices?" *The Energy Journal*, 32(2).
- 80 Given how controversial this issue is, we shall use extensively quotations representing different positions taken in this debate.
- 81 Testimony of Elliot Chambers, Chesapeake Energy Corporation's corporate finance manager before the CFTC's "Hearing on energy position limits and hedge exemptions," August 5, 2009 (see [http://www.cftc.gov/ucm/groups/public/@newsroom/documents/file/hearing080509\\_chambers.pdf](http://www.cftc.gov/ucm/groups/public/@newsroom/documents/file/hearing080509_chambers.pdf)).
- 82 Testimony of Blythe Masters, managing director and head of the Global Commodities Group, JP Morgan CFTC, "Hearing on energy position limits and hedge exemptions," July 29, 2009 (see [http://www.cftc.gov/ucm/groups/public/@newsroom/documents/file/hearing072909\\_masters.pdf](http://www.cftc.gov/ucm/groups/public/@newsroom/documents/file/hearing072909_masters.pdf)).
- 83 The end users prevailed as far as the Dodd–Frank Act is concerned. As of April 2012, the fears of end users of derivatives were rekindled by new rules proposed by the Federal Deposit Insurance Corporation (FDIC), which would put commercial banks under obligation to demand (under certain circumstances) collateral from the buyers of hedge instruments.

- 84 Testimony of Thomas C. Deas, Jr., vice president and treasurer, FMC Corporation, Hearing before the Senate Committee on Banking, Housing, and Urban Affairs, "Building the new derivatives regulatory framework: Oversight of Title VII of the Dodd-Frank Act," April 12, 2011.
- 85 Ibid.
- 86 Of course, a producer has an option of hiring experienced traders or retaining consultants to improve terms on which they transact with hedge providers.
- 87 John E. Parsons and Antonio Mello, "The collateral boogeyman is back" (<http://bettingthe-business.com/2011/02/14/the-collateral-boogeyman-is-back>).
- 88 As one executive told me: "I have been tempted to set-up an exchange account because I do know that I am paying a significant amount of a wider bid-mid when I hedge with banks. I have not done it because of strong resistance from the rest of the C-suite because of having the cash tied-up." This push-back is understandable. We have seen many companies aggressively hedging with futures a very significant portion of production many years into the future and suffering a liquidity crisis as a result (in case of a significant shift in prices triggering margin calls).
- 89 Under SFAS 133, a company can match the results of a hedge (deemed effective) with the changes of the value of the hedged items. For a more detailed discussion (the rules vary for fair value, cashflow and currency hedges), see John D. Finnerty and Dwight Grant, 2006, "Testing hedge effectiveness under SFAS 133," *The CPA Journal*.
- 90 Under the Dodd-Frank Act, the end users of derivatives receive an *exception* from mandatory exchange trading and clearing. In the popular press, this is referred to as an exemption. An exception requires complying with certain documentation requirements that have not been entirely clarified at this point (spring 2012).
- 91 This observation is based on our conversations with a few CFOs and accountants.
- 92 As soon as we wrote this sentence, a news item flashed across the screen – as reported by Reuters, "the US Department of Justice has extended its derivatives investigation to include questions about London-based clearinghouse LCH.Clearnet [...]." As we have learned many times in life, there is a downside to being important (see <http://www.reuters.com/article/2012/06/08/usa-derivatives-probe-idUSL1E8H87VS20120608>).



## *Energy Markets: Market Participants and Regulatory Developments*

This chapter will offer a review of certain critical participants in the energy markets, such as exchanges, clearinghouses and brokers, as well as the users of their services. All these entities operate under a great deal of pressure arising from regulatory developments and technological progress in the area of information and communications networks. Another important group of participants are price-reporting agencies receiving increasing attention from the regulators, given how important they are to the functioning of the energy markets. We will start with a review of different classifications of market participants and provide a list of major categories of players. We will also discuss regulatory developments related to implementation of the Dodd–Frank Act (the Act) to the extent they affect operations of such market participants as exchanges, swap dealers and the end users of derivatives (the Act was also discussed in Chapter 6, in the context of its relationship with, and impact upon, energy market participants).<sup>1</sup>

The main provisions of the Act affecting the energy markets include the following.<sup>2</sup>

- ❑ Requirements that standardised derivatives be cleared. The definition of a standardised derivative is still open, but we can assume that a contract approved for clearing by a clearinghouse will have to be submitted for clearing. Commercial end users of derivatives were granted an exception following a general outcry by the companies using derivatives for hedging (see the section “To clear or not to clear” in Chapter 6).
- ❑ Swaps approved for clearing are to be traded on Swap Execution Facilities (SEFs) or designated contract markets;

- ❑ All swaps are to be reported to Swap Data Repositories (SDRs) and the volume and price information disseminated to the public as soon as it is practically possible.

This chapter is US-centric. There are many similarities between the US and European energy markets in terms of organisation and the characteristics of major market players. Many Asian countries are also rapidly catching up. The US remains, however, the leader, and solutions implemented in the US are likely to proliferate across the globe. This does not mean that we shall ignore markets outside the US – many specific issues related to the operations of the Asian and European markets will be covered in the chapters devoted to specific energy commodities.

### **MARKET PARTICIPANTS**

As mentioned in Chapter 1, the landscape of the energy markets is populated with many different participants, who differ in terms of their objectives, access to the physical markets, credit quality, the rules they operate under and the extent of regulatory oversight they are subject to, as well as their level of operational and financial sophistication. This creates a very rich and fascinating ecosystem and makes a career in this business a most interesting and intellectually stimulating experience. Some companies participate in many different activities, across many different physical commodity and geographical markets, and combine physical and financial trading, while others occupy narrow and specialised niches. The objective of this section is not to engage in a sterile exercise of building theory through classification. Our goal is rather to mention a number of dimensions along which energy market participants can be classified and produce a list of different activities in which they can engage. They can be generally categorised along these lines:

- ❑ physical commodity (natural gas, oil, coal);
- ❑ ownership: publicly owned companies (for example, Southern Company in the US), private companies (for example, the Koch Industries), government agencies (for example, Bonneville Power Authority in the US Northwest), municipal utilities (for example, Austin Energy in Texas), cooperative (for example, Oglethorpe Electric Membership Corporation in Georgia), etc;

- ❑ the extent of regulatory oversight – regulated versus unregulated entities, entities regulated by states versus federal government;
- ❑ location in the value chain – producers, servicing companies, midstream companies, marketers, end users;
- ❑ financial versus physical participants; and
- ❑ hedgers or speculators.

Some of the main categories of energy market participants are listed below. The list is by no means exhaustive and one has to remember that many entities cannot be easily classified as they may combine many types of operations across all the dimensions. For example, some integrated oil companies and big utilities operate specialised units engaging in proprietary trading and offering risk management instruments to the rest of the industry. Some financial institutions control physical assets, either directly or through contractual arrangements. An alien teleported onto a trading floor in Downtown Houston and then to Lower Manhattan could not see much difference between the two locations (except that kolaches would be served for breakfast in one place and bialys in the other).

- ❑ Fossil fuel producers, transporters and processors:
  - Major integrated oil and natural gas companies;
  - Independent oil and natural gas producers;
  - Upstream (exploration and production);
  - Midstream (transportation, storage); and
  - Downstream (refining, chemical plants).
  - National oil companies;
  - Coal producers;
  - Ethanol producers;
  - Oil and natural gas field servicing companies;
  - Natural gas gathering system operators and processing plants;
  - Natural gas and oil pipelines;
  - Shipping companies;
  - LNG liquefaction and regasification plants;
  - Natural gas and oil storage operators; and
  - Oil and coal terminal operators.
- ❑ Electricity producers and distributors:
  - Independent power producers;
  - Investor owned utilities;

- Municipal and cooperative utilities;
  - Government power agencies (for example, BPA and TVA); and
  - Retail electricity providers.
- Natural gas distribution companies.
- Financial institutions:
  - Investment and commercial banks;
  - Hedge funds;
  - Mutual funds;
  - Private equity funds; and
  - Fund managers, commodity pool operators.
- Brokers:
  - Voice brokers; and
  - Interdealer brokers.
- Exchanges.

These very different entities engaging in, or facilitating, energy trading have one thing in common. For all of them, the key to successful energy trading and marketing is a thorough understanding of the physical layer of the energy industry, energy-related financial market instruments and the regulatory regime. The command of these facts is what makes a difference between a mediocre and an outstanding trader, originator or risk manager. Different groups of market participants face different capital and regulatory constraints, have different objectives and come to the market at different times during the day and during the year. This information allows a trader to anticipate better the dynamics of market prices, although it does come at a price. Such awareness can only be acquired through years of experience, immersion in the industry and hard work. No university on the planet can teach these skills. The employees marketing energy commodities and offering hedge management instruments should understand their customers better than the customers understand their business, and succeed in the long term by offering superior service and impartial advice. This is the recommended strategy, however the author is far from being Pollyannaish. Predatory behaviour towards the customers can sometimes lead to riches before the person is ostracised by the industry, and some clever individuals will always stay one step ahead of the sheriff.

Energy trading organisations engage in several different types of activities:

- supporting the physical value chain of the industry;
- hedging;
- provision of financial instruments used by investors, hedgers, speculators; and
- speculation.

These different types of activities typically coexist in one organisation, and lines between different types of activities are usually very fuzzy. For example, a trading unit in an integrated oil company performs multiple functions. It helps to acquire crude oil for the company's refineries and to market refined products. A related important activity is asset optimisation: traders move crude oil in time and space and transform its quality through market transactions in order to facilitate the production of a basket of refined products with the highest market value. The same unit may take advantage of the company's strong balance sheet and financial engineering skills to offer risk management instruments to the end users and producers of different energy commodities. Finally, the traders can engage in speculation, taking advantage of their superior knowledge of the markets, access to information about the conditions of physical infrastructure and control of some critical assets.

One of the most important activities of any energy trading operation is hedging of its exposures or constructing/executing hedges on behalf of and for the benefit of its clients. Hedging decisions in the energy business are devilishly complicated and require a good understanding of many different layers of the industry (physical, financial, regulatory).

The decisions to hedge depend on a number of factors, including:

- the underlying business and financial strength of the company;
- current level of spot and forward prices;
- technological constraints embedded in the company's physical business;
- preferences of the investors (some investors want to acquire exposure to energy prices and do not want a company they invest in to hedge);

- industry structure (an airline will hedge differently depending on the extent of regulation of the markets in which it operates);
- the strategies of its competitors;
- the regulatory infrastructure and the extent to which regulators interfere with details of company's operations;
- accounting rules in place; and
- the overall macro conditions of the economy.

Sometimes, the best advice we can offer is to do nothing. In our career, the most difficult decisions were to advise a customer not to hedge. This can amount to exposing a customer to a potentially significant risk and disappointing an originator, who was usually a good friend and was seeing their bonus expectations reduced.

### **MARKET PARTICIPANTS: REGULATORY DEVELOPMENTS**

A discussion of the different classes of energy market participants would not be complete without discussing the legal status of major market participants – ie, what specific laws and regulations apply to different classes of market players. The energy business is heavily regulated and one has to monitor constantly the evolving legal framework and regulatory infrastructure. We are at a critical junction in the evolution of the regulatory framework of the US financial industry, with potentially very significant impacts on the providers of risk management instruments. The Dodd–Frank Wall Street Reform and Consumer Protection Act, signed into law by President Obama on July 21, 2010, created a number of new entities and institutional solutions with the consequences for our markets that cannot be fully predicted.

The Act applies to the financial markets, although its impact on physical layer of the industry is likely to be profound. Financial derivatives covered by the Act are referred to as “swaps.” The irony of the situation is that, at this point (late May 2012), swaps remain largely undefined (the CFTC and the SEC are mandated by the Act with development of an operational definition of a swap).

Under the Act there are three different types of participants in the swap markets:

- swap dealers;
- major swap participants; and
- other participants.

The highlights of proposed entity definition for swap dealers according to the Act include these criteria:

The term ‘swap dealer’ means any person who –

- (i) holds itself out as a dealer in swaps;
- (ii) makes a market in swaps;
- (iii) regularly enters into swaps with counterparties as an ordinary course of business for its own account; or
- (iv) engages in any activity causing the person to be commonly known in the trade as a dealer or market maker in swaps, provided however, in no event shall an insured depository institution be considered to be a swap dealer to the extent it offers to enter into a swap with a customer in connection with originating a loan with that customer.<sup>3</sup>

In addition, the CFTC formulated additional “distinguishing characteristics” in December 2010, including:<sup>4</sup>

- ❑ dealers tend to accommodate demand for swaps and security-based swaps from other parties;
- ❑ dealers are generally available to enter into swaps or security-based swaps to facilitate other parties’ interest in entering into those instruments;
- ❑ dealers tend not to request that other parties propose the terms of swaps or security-based swaps; rather, dealers tend to enter into those instruments on their own standard terms or on terms they arrange in response to other parties’ interest; and
- ❑ dealers tend to be able to arrange customised terms for swaps or security-based swaps upon request, or to create new types of swaps or security-based swaps at the dealer’s own initiative.

A major swap participant (MSP) has the following defining characteristics under the Act:

(A) IN GENERAL. – The term ‘major swap participant’ means any person who is not a swap dealer, and—

- (i) maintains a substantial position in swaps for any of the major swap categories as determined by the Commission, excluding –
  - (I) positions held for hedging or mitigating commercial risk; and
  - (II) positions maintained by any employee benefit plan (or any contract held by such a plan)[...]
- (ii) whose outstanding swaps create substantial counterparty exposure that could have serious adverse effects on the financial

- stability of the United States banking system or financial markets; or
- (iii) (I) is a financial entity that is highly leveraged relative to the amount of capital it holds and that is not subject to capital requirements established by an appropriate Federal banking agency; and
- (II) maintains a substantial position in outstanding swaps in any major swap category as determined by the Commission.

Being designated as a swap dealer or MSP would have a crucial impact on a company swept under the rule, exposing it to a number of potentially expensive requirements:

- mandatory clearing;
- capital and margin requirements;
- reporting and recordkeeping requirements;
- business conduct standards and operational requirements; and
- observing position limits.

By this point (late April 2012), it seems that some major energy companies (oil majors, big utilities) will be designated by the CFTC as swap dealers. One of the arguments the CFTC uses in support of its position is that some companies effectively self-classified themselves by joining ISDA as primary members. The definition of a primary member makes it very difficult to argue that one is engaging exclusively in hedging operations.

According to the Association's by-laws, every investment, merchant or commercial bank or other corporation, partnership or other business organisation that, directly or through an affiliate, as part of its business (whether for its own account or as agent), deals in derivatives shall be eligible for election to membership in the Association as a Primary Member, provided *that no person or entity participates in derivatives transactions solely for the purpose of risk hedging or asset or liability management* [emphasis added].<sup>5</sup>

On April 18, 2012, the CFTC announced pending publication of what was called "interim final rule" containing further details on swap dealers and MSPs. The rule provides for an additional extended 60-day comment period (hence the term interim). Some highlights of the Rule include interpretive guidance with respect to the language used in the Act ("hold out" and "commonly known"), as summarised below:<sup>6</sup>

- ❑ The determination of whether a person is a swap dealer should consider all relevant facts and circumstances, and focus on the activities of a person that are usual and normal in the person's course of business and identifiable as a swap dealing business; making a market in swaps is appropriately described as routinely standing ready to enter into swaps at the request or demand of a counterparty;
- ❑ a person making a one-way market in swaps may be a market maker, and exchange-executed swaps are relevant in the determination;
- ❑ examples of activities that are part of "a regular business," and therefore indicative of swap dealing, are entering into swaps to satisfy the business or risk management needs of the counterparty, maintaining a separate profit and loss statement for swap activity, or allocating staff and resources to dealer-type activities; and
- ❑ the SEC's dealer-trader distinction may be applied in identifying swap dealers.

In our view, this language (especially "maintaining a separate profit and loss statement for swap activity, or allocating staff and resources to dealer-type activities") would result in the classification of many energy companies as swap dealers. Some relief was offered through two additional provisions:<sup>7</sup>

- ❑ exclusion of certain hedging swaps from consideration in identifying swap dealers; and
- ❑ *de minimis* exemption from the definition of a swap dealer if "the aggregate gross notional amount of the swaps that the person enters into over the prior 12 months in connection with dealing activities must not exceed US\$3 billion."<sup>8</sup>

The second important clarification was related to the definition of "substantial position." The CFTC proposed two tests. The first is based on current uncollateralised exposure determined by marking the swap positions to market, with adjustments for the collateral and accounting for netting (to the extent it is allowed under the terms of applicable master netting agreements). The threshold adopted for this test is US\$1 billion. The second test is based on total exposure (current, as explained above, and future potential exposure), determined by:<sup>9</sup>

- ❑ multiplying the total notional principal amount of the person's swap positions by specified risk factor percentages (ranging from  $\frac{1}{2}\%$  to 15%) based on the type of swap and the duration of the position;

- discounting the amount of positions subject to master netting agreements by a factor ranging between zero and 60%, depending on the effects of the agreement; and
- if the swaps are cleared or subject to daily mark-to-market margining, further discounting the amount of the positions by 80%.

The definition of a substantial position excludes *bona fide* holdings of hedging or risk-reducing instruments but, as we would expect, not any positions held for “speculation, investing or trading.” A substantial counterparty exposure definition follows the lines of the substantial position definition, except that it includes all swaps positions (including those held for hedging).

The second important distinction applies to exchanges, and evolves around two basic legal constructs:

- designated contract market (DCM);
- exempt commercial markets (ECM);<sup>10</sup>

All futures contracts must be traded exclusively on an exchange that falls under the oversight of the CFTC and which approves DCMs.<sup>11</sup> A contract listed on a DCM must meet a number of requirements, and the exchange is obligated to enforce internally the CFTC rules and to publish daily information about trading volume, open interest and settlement prices. To obtain and maintain designation, a board of trade must comply with the following criteria:<sup>12</sup>

- general demonstration of adherence to designation criteria;
- prevention of market manipulation;
- fair and equitable trading;
- enforcement of rules on the trade execution facility;
- financial integrity of transactions;
- disciplinary procedures;
- public access to information on the contract market; and
- ability of the contract market to obtain information.

The reader can find the list of 18 core principles established by the CEA that DCMs have to comply with on the CFTC website.

The category of ECM was created by the Commodity Futures Modernization Act of 2000.<sup>13</sup> The CFMA added a new section 2(h) to the Commodity Exchange Act (CEA), dealing with bilateral contracts

between “eligible contract participants” (a new legal construct). Section 2(h)(3) exempted from CFTC oversight the “exempt commodities” traded directly (ie, not as brokered transactions) by “eligible commercial entities” (ECEs). The contracts falling under this section were energy commodities and metals.<sup>14</sup> This class of contract executed on an electronic trading facility<sup>15</sup> was removed from regulatory oversight, except for fraud and anti-manipulation provisions. Electronic exchanges were obligated to provide some limited information to the CFTC (and additional information on demand). To summarise, ECMs meet the following criteria:<sup>16</sup>

- trading platform: electronic trading facility;
- commodities traded: exempt commodities (metals, chemicals, energy products, pollution allowances, wood pulp, etc);
- intermediation: principal to principal trades only; and
- participants: eligible commercial entities.

Some provisions of the CFMA became known as the so-called “Enron loophole.” At the end of 2000, Enron introduced an electronic trading platform, EnronOnLine (EOL), under which it acted as a buyer to all sellers and a seller to all buyers (one-to-many platform). EOL was a phenomenal success in terms of the transaction volume it captured, the number of commodities trading and impact on the development of energy markets. Other energy companies and some financial institutions countered by developing competing platforms: one of them evolved into ICE.

In our view, the paradox of the EOL saga is that became a major contributing factor to Enron’s demise. Rapid volume growth, combined with price volatility, overwhelmed Enron’s balance sheet and the company’s ability to finance outstanding trading positions. Enron’s loophole, seen by many as custom-made legislation to benefit a specific market participant, is another proof that the road to hell is paved with good intentions and that friends may be often more dangerous than sworn enemies. Having said this, it is worth noting that EnronOnLine might not have needed this special treatment under the law, given its one-to-many design. However, the debate on this issue should be left to the lawyers.

The term “Enron loophole” is used typically in reference to Section 2(g) of the CFMA. This section removed from CFTC oversight any

“individually negotiated” transaction between “eligible contract participants” not executed on a trading facility. This exclusion was even more far reaching than the exclusion under section 2(h), although the emphasis was on direct negotiations. Our interpretation is that this section would not apply to EOL, but would cover complex bilateral transactions negotiated individually by Enron (or any Enron-type company) with commercial counterparties in the energy and metals space.

The Enron loophole was partially closed at the initiative of Senator Carl Levin, who had introduced in 2007 Senate Bill S. 2058. This bill was later attached to H.R. 6124, known as “The 2008 Farm Bill”.<sup>17</sup> President George W. Bush vetoed the bill, but the veto was overridden by both the House and Senate and, on June 18, 2008, the bill was enacted into law.<sup>18</sup>

The ECMs were subjected to nominal CFTC oversight, which was much more limited than the oversight of organised exchanges such as Nymex. One critical difference was that ECMs were not required to establish and enforce position limits. Many of the instruments traded on one of the most successful electronic trading platforms, ICE, allow market participants to acquire exposure to Nymex contracts through transactions known as *Nymex lookalikes* (Nymex swaps). These transactions are cash-settled bullet swaps, with a Nymex contract as the underlying. The growth of the market for lookalikes diluted the effectiveness of the position and accountability limits on Nymex.

As we mentioned in Chapter 6, The 2008 Farm Bill extended position limits to the contracts performing significant price discovery functions, traded on ECMs.

As mentioned before, the Dodd-Frank Act eliminated the category of ECMs, but, at the time of writing, they continue under temporary extensions, pending the promulgation of final rules.

### OTC VOICE BROKERS

OTC transactions are facilitated by voice brokers. For most people, the term “broker” invokes a vision of a messy, smoked-filled room, with a few men, smoking and with shirt sleeves rolled up, talking frantically on the phone (with some of them using two phones at the same time). In reality, modern brokerages are very sophisticated entities relying on advanced communications and information technology.

Technically, a broker is an entity that executes buy and sell order and acts as an intermediary facilitating and arranging bilateral transactions. The brokers do not take trading positions themselves and are paid through commissions and by investing excess funds deposited by customers with them (ie, they earn more interest on cash balances than they pay to their clients). A low interest rate environment eliminates for all practical reasons this source of revenues and sometime leads to excessively risky modifications of their business model. Another threat to their survival is the proliferation of electronic trading platforms that eliminate the need for their services.

It has been debated for the last few years whether voice brokers can survive the combined onslaught of technological revolution related to electronic trading and regulatory reforms favouring increased market transparency, eliminating the need for services of informed and well-connected intermediaries. The jury is still out, but our own conviction is that they are likely to survive, given their creativity and the customer's need for assistance in reducing search costs and arranging complex transactions.

A very important category in this industry is an interdealer broker, who specialises in facilitating transactions between major financial institutions, acting an intermediary. Such brokers combine old-fashioned voice intermediation with sophisticated electronic platforms. Such platforms use a number of methods for matching buyers and sellers. Two methods require special attention here. Under the interpretation of the Dodd-Frank Act as of spring of 2012, SEFs (the platforms for clearable derivative instruments) should rely either on the central limit order book (CLOB) or the request for proposal (RFP) method. Under the RFP approach, a potential buyer or seller submits a query to a number of dealers<sup>19</sup> and can act on the best offer. In the technical jargon, they may "hit the bid" (ie, sell at the highest price available to them), or "lift the offer" (ie, buy from the cheapest source). A customer cannot get a price within the bid–offer spread and can trade only with the dealer. The CLOB system is more transparent and anonymous. A customer can see the "stack" – ie, the magnitude and the price levels of bids and offers, allowing them to see the depth of the markets. A customer can trade with many dealers and with other customers.

Voice brokers face strong pressures related to competition arising from the more ubiquitous trading on electronic exchanges (ICE and

Globex in the US) and a wave of new regulatory requirements triggered by the Act. Currently (June 2012), final rules have not been promulgated yet by the CFTC, but the most important issues have been defined.

- *SEF definition.* Under the Act, derivative trading (with some exceptions) must migrate to exchanges or SEFs. The definition of a SEF has not been finalised, and it is not clear if voice brokers would qualify as SEFs.<sup>20</sup>
- *15-second rule.* The 15-second rule proposed by the CFTC requires that a broker exposes a potential transaction for 15 seconds to other market participants by displaying relevant information on a screen. This would give other market participants an opportunity to scrutinise the trade and, potentially, insert themselves into the deal or take it away. The delay requirement applies when one party is a customer and the other a trader or, in the case of crossing an order between two customers. Many brokers feel threatened by more agile competitors with more efficient and faster electronic platforms.

Given this new regulatory framework, which requires huge investments in IT and compliance infrastructure, it is not unreasonable to expect industry consolidation.<sup>21</sup>

### Futures Commission Merchants

FCMs are specialised brokers whose primary business consists in the execution of orders to buy or sell futures or options on futures. In addition, they collect and manage liquid assets collected from the customers to support such orders, primarily to manage margin calls and to guarantee derivative contracts. The profits of FCMs are derived from commissions and from the differential between interest paid to their customers on their excess funds held in the FCM accounts and interest received through reinvestment of these funds. The excess balances are maintained by FCM customers for convenience (to avoid frequent wire transfers) or because most FCMs require margins in excess of exchange requirements. There are about 65 FCMs in the US, holding over US\$150 billion in US customer collateral and nearly US\$40 billion in collateral related to transactions on foreign exchanges. As of March 2011, the total amount of customer funds held by the top 30 FCMs was more than US\$163 billion.

Customers' funds held by the FCMs should be, under the law, kept in segregated accounts. The CFTC Rule 1.20(a) is unambiguous in this respect:

All customer funds shall be separately accounted for and segregated as belonging to commodity or option customers. Such customers' funds when deposited with any bank, trust company, clearing organisation or another futures commission merchant shall be deposited under an account name which clearly identifies them as such and shows that they are segregated as required by the [Commodities Exchange] Act and this part ... No person ... that has received customer funds for a segregated account ... may hold, dispose of or use such funds as belonging to any person other than the option or commodity customers of the futures commission merchant which deposited such funds.<sup>22</sup>

Section 4d of the CEA and the CFTC Regulation 1.25 restrict permissible investments an FCM can make with customer funds. The general prudential standards require that all permitted investments be "consistent with the objectives of preserving principal and maintaining liquidity."<sup>23</sup>

FCMs are regulated by the CFTC. Additionally, the day-to-day oversight of FCMs is carried out by the designated self-regulatory organisations (DSRO). DSROs are responsible for periodic audits and sharing the information obtained in the course of the audits with the CFTC. As stated by Commissioner Jill Sommers:

Commission Regulation 1.10 requires FCMs to file monthly unaudited financial reports with the Commission and the DSRO. These reports include the FCM's segregation and net capital schedules, and any 'further material information as may be necessary to make the required statements and schedules not misleading.' [...] Commission Regulation 1.16 requires FCMs to file annual certified financial reports with the Commission and the DSRO.<sup>24</sup>

The business model of the FCM is currently under stress. Competition drove commission levels to very low levels, and the low interest rate environment all but eliminated another source of profits. This explains why some FCMs engaged in very risky strategies, leading in two well-known cases to the spectacular bankruptcies of REFCO and MF Global. In the second case, US\$1.6 billion of customers' money, supposedly held in segregated accounts, is still missing.

## PRICE REPORTING AGENCIES<sup>25</sup>

A documentary called *World without Oil*<sup>26</sup> illustrated the role of crude oil in our civilisation by taking to its logical limits the hypothetical situation of all the oil wells suddenly going dry. One could use a similar thought experiment and ask the question about how energy markets, and commodity markets in general, would function without PRAs. The obvious conclusion is that it is difficult to underestimate the role of these entities in the commodity markets, and especially in the energy markets. They are not only critical to the process of price discovery but also constitute one of the critical channels for dissemination of information about market transactions, regulatory and political developments, and the condition of the physical infrastructure. It is not an exaggeration to say that it is a vital component of the nervous system of the energy industry. In this section, we shall describe briefly the business model of the PRAs and provide brief descriptions of the more important entities. Additional information about specific techniques used to collect price information and calculate indexes is included in subsequent chapters in discussions of specific markets. This is by no means an exhaustive review; the objective is to alert the user to the importance of this poorly understood part of the market infrastructure and the need to monitor developments in this space.

### PRAs: The business model and services

PRAs provide a number of services that can be classified into several broad categories:

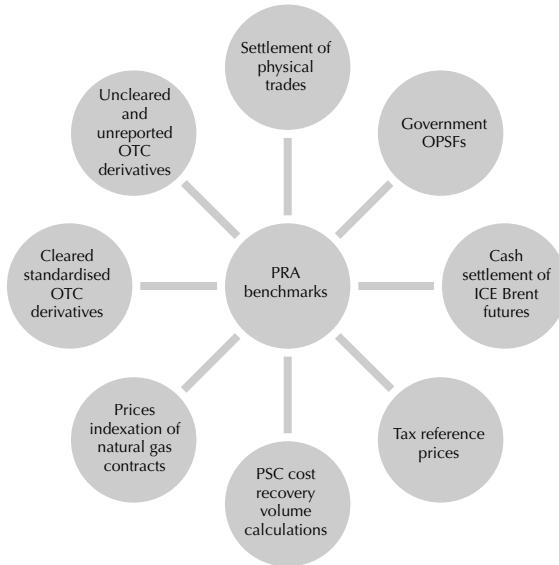
- ❑ collection of transaction data, calculation and dissemination of prices and price indexes for spot and derivative markets;
- ❑ assessments of forward price curves (as opposed to calculation of indexes from actual transaction data).
- ❑ collection of non-price data related to market trends, conditions of the physical infrastructure and other developments critical to the energy industry;
- ❑ dissemination of market information in real time (through Internet-based platforms) or through industry newsletters;
- ❑ development of databases and search engines; and
- ❑ additional non-core services, including:
  - training courses and industry conferences;

- consulting;
- risk management services; and
- transaction support (for example, confirmation services).

The impact of the PRAs on the energy business is obvious to anybody who has spent some time on the energy floor. A trader or analyst starts the day by reading several industry newsletters, reviewing index prices posted for their markets, following industry developments relying on the real-time services of their data providers. We still vividly remember the buses leaving for Houston from our town (located north of the Downtown) roughly between 5:30 and 6:00 am with row after row of passengers reading *Gas Daily* or *Megawatts Daily*. Industry conferences are also an occasion for bonding, socialising, creating pressure groups and information brokering.

The activities listed above vary in scope and importance from agency to agency, but are present in practically every entity covered in this section. The most profitable activity is the production and dissemination of price indexes. High profit margins in this activity can be explained by the way long-term contracts are structured in the energy industry. As will be explained in subsequent chapters, many contracts are based on floating prices that reset periodically to the published indexes. Many pipeline and local distribution company tariffs approved by the regulators refer to the price indexes as well. Once a sufficient critical mass of long-term contracts and tariffs accumulates, displacing the incumbents or moving to a different transaction regime becomes very difficult. It is crucial to remember that the importance of price indexes goes well beyond their use in pipeline tariffs and long-term commercial contracts. The indexes are used as benchmarks for the settlement of derivative contracts, as inputs in calculation of taxes and royalties collected by governments and private landowners, and the determination of official selling prices set by the governments controlling production in some countries. Figure 7.1 illustrates many potential uses of price indexes produced by the PRAs.

Having said that, the business model described above is likely to come under increasing pressure. The development of exchange-based trading combined with regulatory developments such as the Dodd–Frank Act<sup>27</sup> are likely to undermine the index production business that constitutes the anchor of the operations of PRAs. It is

**Figure 7.1** The uses of price indexes

Source: "Oil price reporting agencies," report by IEA, IEF, OPEC and IOSCO G20 Finance Ministers, October 2011.

Note: OPSF – Oil Price Stabilization Fund; PSC – production sharing contract

likely that the markets for commodity derivatives will become more transparent in the future. It is true that PRAs occupy a special market niche related to spot transactions (the Act applies to derivatives and not to physical transactions) but, as we shall see in subsequent chapters, spot prices are often based on the derivative markets and not vice versa. Increased transparency of the swap markets will be translated into increased transparency of spot prices. However, it remains to be seen if the business models of the PRAs will survive the test of time. The message of this section is: "Stay tuned and try to anticipate the consequences of whatever happens in your market."

These remarks emphasise the importance of understanding the issues related to price index construction, and of their advantages and shortcomings. We shall cover general topics in this section, with a more detailed review and examples in the more specific chapters. We urge traders and analysts to pay attention to even the smallest details related to the price indexes.

*Index construction<sup>28</sup>*

Price indexes allow the summarising of market trends in a single number that can be quickly absorbed by a trader and acted upon. The indexes may be calculated in a number of different ways.

- ❑ A specific statistical formula used for price calculations. Price indexes can be calculated as arithmetic averages or volume-weighted averages, although some more complicated formulas (such as, for example, an average of two different indexes) are possible.
- ❑ The exclusive use of outright prices of consummated transactions in calculations of indexes or reliance on additional information such as bids and offers, spreads or cross-calibration against other indexes in the absence of sufficient direct market data. In some case, price assessments are based on models and not on actual data.
- ❑ The level of discretion in elimination of transactions that do not meet the standards established by a PRA.
- ❑ The time window chosen for inclusion of transactions in index calculations. An example is provided by the use of the market-on-close (MOC) approach by Platts for certain indexes versus a reliance on transactions taking place during the entire day.
- ❑ The rules regarding completeness of transaction reported to the PRAs. In the US natural gas markets, at the insistence of the regulators, an entity reporting transactions to a PRA has an obligation to submit all the relevant deals. In the case of other markets and commodities, a potential for cherry-picking – ie, selective reporting – exists. Partial data submission could potentially influence the levels of reported indexes, benefitting some counterparties and hurting others. Examples of selective reporting include submitting purchases but not sales, including transactions without offsetting physical or financial hedges, such as contracts for difference in the case of crude oil (this term will be clarified in subsequent chapters).
- ❑ The rules for handling quality differences. For example, transactions in the oil markets included in a specific index may comprise deals related to different grades of crude. The question is whether price indexes should be based on raw prices or prices adjusted for quality differences.

There are many examples how a price index may obliterate some information about market trends and how different designs can influence the results. Suppose that prices were trending down or up during the day. The average can be the same for two days, but a trader would appreciate information about the trend when they start trading in the morning on the following day. Suppose that we are in a market where volume tends to be higher in the morning than in the afternoon (or vice versa). A volume-weighted index can mask the price trend or reverse it. Another example is a market with the same mid-level on two days and a wide bid–offer spread. Suppose that the market is subject to selling pressures on one day and buying pressure on the second day. The transactions happen at bid prices on one day and offer prices on the second day. A price index will show a price trend where there is none. These issues are not just academic curiosities but are a subject of heated discussion in the industry with a lot of prestige, influence and market share at stake. Should the price indexes be volume weighted or just plain leverages? Should prices be based on the entire day of trading or a certain time window? Should the atomic prices be reported along with the indexes? The examples of specific controversies will follow in the relevant chapters.

### *Manipulation*

Price indexes can be manipulated through submission of intentionally false information to the index publishers. This practice became endemic in the US natural gas markets (as discussed at length in the chapters on natural gas).<sup>29</sup> Under US law, this practice is illegal (although reporting to price index publishers is voluntary) and has resulted in a number of convictions. The sad reality is that for a long time the industry was cavalier about an issue that is at the heart of market efficiency and the viability of our way of life based on free and unobstructed markets:

Trust is an important lubricant of the social system. It is extremely efficient; it saves a lot of trouble to have a fair degree of reliance on other people's word [...] Trust and similar values, loyalty or truth-telling, are examples of what the economist would call 'externalities'. They are goods, they are commodities; they have real, practical economic value, they increase the efficiency of the system, enable you to produce more goods or more of whatever values you hold in high esteem.<sup>30</sup>

Another related issue is that a price index may be skewed without being intentionally manipulated. The set of the entities reporting to the newsletter is based on self-selection, making the index potentially non-representative. The Dodd-Frank Act and mandatory swap data reporting may change the field for derivatives, but the problems will persist with respect to the physical markets.

Another frequently raised criticism is related to differences in the construction of indexes across different PRAs and markets. As mentioned above, the approaches may vary from mechanical summary of the received quotes to fairly subjective procedures. In our view, a frequent reliance on common sense, experience and the good judgement of the staff of the PRAs, as well as flexibility, is inevitable given varying levels of maturity and liquidity of different energy markets. The extent of intervention by PRAs ranges from removing transactions from the reported data samples to coming up with assessments in the absence of any data. This latter practice became critical in the US natural gas markets when regulatory investigations in the early years of the first decade of the current century prompted many market participants to suspend data submissions to the PRAs in order to pre-empt potential legal liability. The CFTC and the Federal Energy Regulatory Commission (FERC) addressed the problem by offering safe harbour, a protection against legal liability in case of an honest, unintentional mistake. The safe harbour provision produced a significant increase in the number of reporting companies.

The discrepancies between prices reported by different PRAs relying on different data samples and approaches were investigated by the International Organization of Securities Commissions (IOSCO) and its partners:

Averaged over lengthy time periods, the differences among prices reported by different PRAs for the same crude oil grade is usually substantially less than US\$1.00/bbl. In the case of the key benchmark grade of "Dated Brent" this difference is about US\$0.01/bbl. In the case of other benchmarks, such as Dubai and Tapis, the differences over time can be more substantial. A significant component of this difference can be accounted for by the fact that PRAs use different delivery periods when assessing prices. After correcting for this important disparity, the differences in average prices reported by various PRAs remain less than US\$1.00/bbl for major crude streams other than Tapis. On any given day, however, the differences in prices reported by different PRAs for the same crude oil grade can be substantial, even when allowance is made for the different delivery periods considered.<sup>31</sup>

Given leverage inherent in the derivative contracts, even small differences can be magnified into big P&L impacts. This explains why the PRAs receive from time to time complaints from the users of price indexes.

### **Regulatory scrutiny of PRAs**

From time to time, the role PRAs play in the industry attracts the attention of regulators, and this creates the opportunity for market participants to obtain more insights into their operations.<sup>32</sup> Increased interest of governments and regulators in the operations of PRAs typically coincides with periods of increased industry turbulence and price volatility. The most recent period of elevated scrutiny followed the financial crisis of 2008–09, and the unprecedented levels crude oil prices reached in the summer of 2008. It is difficult to find a better proof of the importance of PRAs than being on the agenda of a G8 or G20 meeting. We could say that this is good news and bad news at the same time. A few key developments related to initiatives by G8 and G20 countries are reviewed below.

The Task Force on Commodity Futures Markets was created in September 2008 following the G8 finance ministers meeting in Osaka, Japan, in June 2008, during which concerns were raised regarding the high volatility of energy and food prices. With a mandate to investigate this issue, the task force was formed by the technical committee of IOSCO and was co-chaired by the US CFTC and the UK's FSA. The final report, published in March 2009,<sup>33</sup> made recommendations that included the following:

- ❑ It is crucial to improve the transparency of both market fundamentals (supply, demand, inventories, transport capacities, etc) and physical commodity market transactions. Improvement is necessary both in the availability and the quality of information on all the major physical markets. Regarding oil markets specifically, the International Energy Forum (IEF), the International Energy Agency (IEA) and OPEC should pursue their efforts to improve the reporting by their members with regard to the completeness and timeliness of physical oil data.
- ❑ Financial regulators should determine whether any physical commodity reference prices used in their markets are reliable. If not, such reference prices may facilitate manipulation.<sup>34</sup>

The same concerns were reiterated by leaders of the G20 group during their Pittsburgh 2009 and Seoul 2010 meetings. Following

these meetings, the task force was directed to orient its focus primarily to the oil markets. Some of the related documents are recommended to the reader for closer scrutiny.<sup>35</sup>

This topic received also close attention from European regulators and financial industry think tanks:

In addition to increasing the availability and accuracy of data available on commodity markets, many regulators consider that research should be conducted, for example for energy markets on:

- The transaction flow of the physical and derivative markets: in order to understand the detail, quantity, value and proportion of transactions in the OTC and listed derivatives markets and to define the appropriate approach for data collection
- The process of price formation and the sources of prices volatility: value and accuracy of prices generated by price reporting agencies, value of existing price benchmarks, stability of spread differentials used in conjunction with price benchmarks, convergence issues between the physical and derivatives markets, impact of listed derivatives pricing on OTC contracts and vice-versa, sources of price volatility.<sup>36</sup>

It is easy to understand the interest of governments and regulators in the operations of the PRAs. The agencies play a critical role in the commodity markets at times of huge stress on the financial system and political turbulence, both of which can largely be fed by increasing commodity prices. The PRAs have to navigate volatile, constantly evolving markets dominated by powerful market participants with conflicting interests. It is not surprising that they find themselves in the cross-hairs from time to time. What is surprising is that they still succeed in doing a good job, publishing price indexes, rain or shine, in markets with varying levels of liquidity and evolving at a very high rate.

The reviews of PRAs by IOSCO and other cooperating international institutions are focused around two basic issues:

- ❑ Governance of PRAs, including whether or not the PRAs have appropriate safeguards and procedures and whether or not independent oversight is needed; and
- ❑ Whether the PRA assessment techniques in physical and derivative markets are adequately transparent and the impacts that such assessment and data collection have on the markets.<sup>37</sup>

An interim report by IOSCO, following a preliminary report of March 2012, was presented to the G20 group meeting in Los Cabos,

Mexico, in June 2012 (the report was expected later that year). As reported in the *Wall Street Journal*,<sup>38</sup> “It contained fewer recommendations and dropped some of the toughest, including the idea of a regulator. It retained the prospect of requiring that all trades be reported, as well as requirements for better bookkeeping for all price-reporting agencies.” Our view is that eventually the regulators in the US and UK will implement a version of regulation light, following the precedent established in the US by the FERC and the CFTC after 2001 with respect to the natural gas markets. Natural gas price indexes are quoted in the pipeline tariffs, which have to be approved by the FERC. The FERC and the CFTC used this leverage to push through certain requirements, such as the reporting of all transactions if a company chooses to report at all. Many price indexes are used for settlements of derivative contracts and the regulators can refuse to list a contract on an exchange unless the underlying price index satisfies certain conditions.

The regulatory scrutiny of PRAs has had a number of positive outcomes. The reports produced by IOSCO in collaboration with international organisations such as IEA, IEF and OPEC helped to educate the industry about the importance and operations of PRAs. Public and civilised debate between informed parties is always the best way to address potential issues. The PRAs responded by developing the “Price Reporting Code for Independent Price Reporting Organizations” (the Code), which is likely to be adopted by the industry. The Code contains a number of provisions that should help to reinforce public confidence in the price discovery process. The provisions of the Code range from fairly general principles (maintenance of robust governance, arrangements for elimination of conflicts of interest, publication of price assessment methods (already widely practiced) and dealing fairly and consistently with all market participants), to more specific procedures related to the prompt and fair handling of complaints. The Code contains a requirement for the signatories to publish an Attestation of Compliance (AOC) or an Explanation of Material Non-compliance with the proposed standards, audited by an internationally recognised external audit firm or an independent internal audit group within the corporate organisation of an independent price reporting organisation (IPRO).

## PRAs: A review

### *Platts*

Platts is of critical importance to a number of regional markets, with uncontested leadership in US energy commodities. It started as a company called National Petroleum News, formed by the Cleveland, Ohio, journalist Warren Cumming Platt. In 1923, the first newsletter, *Platts Oilgram*, was published. In 1953, Platts was acquired by The McGraw-Hill Companies, and now provides a wide array of services to the energy industry, including more than 8,500 daily price assessments, used as benchmarks across the energy industry, based on data compiled in 25 offices, and publishes more than 60 newsletters and reports. The information about construction of price indices and the scope of market coverage can be found at <http://www.platts.com/MethodologyAndReferencesHome>.

### *Argus*

Argus was founded in 1970 by publisher and oil industry journalist James Ashley Nasmyth, as a newsletter with a fairly limited distribution. The breakthrough came in 1979 when, as described in the *Times of London*, he launched a daily oil market report, known initially as the *Argus Telex*. This report has grown over time into a diversified media and market information company.<sup>39</sup>

The information about construction of price indices and the scope of market coverage can be found at <http://www.argusmedia.com/Methodology-and-Reference>.

### *ICIS Heren*

Heren Energy was started in 1993 by Patrick Heren, as a provider of price indexes for the UK wholesale natural gas markets. Over time, the company expanded into additional markets, including LNG, emissions, coal, chemicals and oil. Heren has the strongest presence in the European gas and power markets. In 2008, Heren Energy was acquired by ICIS, part of Reed Business Information. Following the acquisition, Heren became referred to as ICIS Heren.

### *Canadian Enerdata*

Canadian Enerdata was established in 1984 and combines information collection, distribution and analysis services. Publications include the *Canadian Gas Price Reporter* (monthly and daily), the

*Canadian Natural Gas Storage Survey* and *The Natural Gas Lookout*. The indexes are used as benchmarks for contracts traded on the Calgary-based NGX, ICE and in the OTC markets. The company is of critical importance to the Canadian markets.

### *OPIS*

OPIS stands for Oil Price Information Service, a company started in 1977 as an *Oil Express Newsletter*. In 1980, the company commenced coverage of the rack market<sup>40</sup> and currently collects more than 70,000 rack prices each day. The market coverage was extended in 1981 to spot price assessments for all refined products, and now covers (among others) markets in spot gasoline, diesel and jet fuel prices. In 1999, retail gasoline market coverage was started. The information collected under this programme is used by AAA, Microsoft (MSN Autos), Garmin and MapQuest.

### *Intelligence Press*

Intelligence Press (IP) is a publisher of the *Natural Gas Intelligence (NGI)* newsletter providing coverage for the US natural gas market. The first issue of *NGI* was published in 1981. Over time, IP has expanded to offer coverage of markets outside the US, including Canada and Mexico, as well as Central and South America, Europe and Asia. As reported on its website, IP's other publications include: *NGI's Weekly Gas Price Index*; *NGI's Bidweek Survey* (both started in 1988); *NGI's Daily Gas Price Index* (1993); and *Power Market Today* (2001).

## **CONCLUSIONS**

The very rich ecosystem of the energy industry not only makes this business fascinating: it also creates a very fertile ground for traders who design trading strategies around behavioural biases and the idiosyncrasies of different market participants. This is true of micro-strategies for trading natural gas at a specific hub or along a specific pipeline, and of macro-strategies developed to make bets on oil price, taking into consideration developments in major economies. Understanding what makes different market players act and how their collective actions add up to certain markets outcomes is a necessary condition of successful trading. It is not, however, a sufficient condition. We have to also understand the physical layer of the

industry. Human emotions rise and fall; the laws of science remain unchanged. In the next sections of this book, we shall continue with a review of the two critical areas of critical importance to energy trading: the physical and institutional layers of the industry, starting with the natural gas industry.

- 1 This chapter should not be construed as an extensive and complete review of the Act. This is impossible for a number of reasons. Implementation of the provisions of the Act by the CFTC and the SEC is not complete and the specific rules are still being developed and are often challenged in court. This task is also beyond our pay-cheque level: it would take a regulatory lawyer to provide such a review.
- 2 This is not an exhaustive review of the Act. A summary of the energy derivative provisions of the Act can be found at Paula S. Greenman, Ann A. Hawkins, John W. Osborn and Philip McBride Johnson, 2010, "Energy derivatives under the Dodd-Frank Act," Skadden, Arps, Slate, Meagher & Flom LLP, July 22.
- 3 R. 4173, <http://www.sec.gov/about/laws/wallstreetreform-cpa.pdf>. The full text of the Dodd-Frank Act is also available from the CFTC website ([http://www.cftc.gov/ucm/groups/public/@swaps/documents/file/hr4173\\_enrolledbill.pdf](http://www.cftc.gov/ucm/groups/public/@swaps/documents/file/hr4173_enrolledbill.pdf)).
- 4 CFTC, Release No. 34-63452; File No. S7-39-10
- 5 <http://www2.isda.org/about-isda/join-isda/primary-members>.
- 6 [http://www.cftc.gov/ucm/groups/public/@newsroom/documents/file/msp\\_ecp\\_fact-sheet\\_final.pdf](http://www.cftc.gov/ucm/groups/public/@newsroom/documents/file/msp_ecp_fact-sheet_final.pdf).
- 7 Ibid.
- 8 We shall have to wait for more information to interpret this provision. It seems to conflate the concept of "notional amount," which is a stock concept and is measured at a specific point in time, with a flow, which is measured over a period of time. The *de minimis* threshold will be phased-in over a three-year period. Over this time period the limit will be US\$8 billion.
- 9 Ibid.
- 10 Another exempt market includes exempt boards of trade, trading so-called "excluded commodities". According to the CFTC, "Section 1a(13) defines an "excluded commodity" to mean among other things an interest rate, exchange rate, currency, credit risk or measure, debt instrument, measure of inflation, or other macroeconomic index or measure. The term "security" for purposes of this provision also includes any group or index thereof or any interest in, or instrument based on the value of, any security or group or index of securities." (<http://www.cftc.gov/IndustryOversight/TradingOrganizations/EBOTs/ebot>).
- 11 The CFTC website defines an ECM as "boards of trade (or exchanges) that operate under the regulatory oversight of the CFTC, pursuant to Section 5 of the Commodity Exchange Act (CEA), 7 USC 7. DCMs are most like traditional futures exchanges, which may allow access to their facilities by all types of traders, including retail customers."
- 12 <http://www.cftc.gov/IndustryOversight/TradingOrganizations/DCMs/index.htm>.
- 13 The Dodd-Frank Act eliminates this category of market participants but they continue to operate, at the time of writing, under temporary extension. Given these conditions of regulatory uncertainty the mention of ECMs is warranted.
- 14 After the demise of Enron, ICE became the primary beneficiary of the CFMA.
- 15 A "facility or system in which multiple participants have the ability to execute or trade agreements, contracts, or transactions by accepting bids and offers made by other participants that are open to multiple participants in the facility or system." (see <http://www.sec.gov/news/press/2011/2011-35.htm>).
- 16 [http://www.cftc.gov/stellent/groups/public/@newsroom/documents/file/derivatives-markets\\_091207.pdf](http://www.cftc.gov/stellent/groups/public/@newsroom/documents/file/derivatives-markets_091207.pdf).

- 17 The Food, Conservation, and Energy Act of 2008.
- 18 In our view, the loophole was only partially closed.
- 19 Currently proposed rules require a minimum of five dealers; see Aline van Duyn, 2010, "Derivatives trading brought into public view," *Financial Times*, December 16.
- 20 H.R. 2586, Swap Execution Facility (SEF) Clarification Act (To refine the definition of swap execution facility in the provisions regulating swap markets added by title VII of the Dodd–Frank Wall Street Reform and Consumer Protection Act), sponsored by representatives Scott Garrett (R-NJ), and Carolyn Maloney (D-NY) is working its way through Congress. A reader should pay attention.
- 21 Alexander Osipovich, 2012, "Endangered species? Can the boutique energy brokerage survive Dodd–Frank?" *Energy Risk*, May 3.
- 22 <http://ecfr.gpoaccess.gov/cgi/t/text{text=idx?c=ecfr&sid=f5e80071c661ed12fddd7ce349a8563b&rgn=div8&view=text&nnode=17:1.0.1.1.0.4.15&idno=17>.
- 23 <http://www.gpo.gov/fdsys/pkg/FR-2011-12-19/html/2011-31689.htm>.
- 24 Testimony of Jill E. Sommers, Commissioner of the CFTC, before the United States Senate Committee on Agriculture, Nutrition and Forestry, December 13, 2011.
- 25 PRAs are also referred to as Independent Price Reporting Organisations (IPROs).
- 26 <http://www.youtube.com/watch?v=2htGadwrhcl&feature=related>.
- 27 As explained elsewhere, the Dodd–Frank Act requires reporting swap transaction data to the swap data repositories and dissemination of price information as close to real time as practical.
- 28 A white paper by ICIS Heren, 2010, "Price determination in European gas markets," London, May, contains many additional observations about different approaches to index construction, which inspired some of the comments here.
- 29 The natural gas market was most effectively investigated from the point of view of index generation. This topic will be covered in Chapter 24.
- 30 Kenneth J. Arrow, 1974, *The Limits of Organization* (New York, NY: Norton), p 23.
- 31 "Oil price reporting agencies," report by IEA, IEF, OPEC and IOSCO to G20 Finance Ministers, October, 2011.
- 32 For example, the abuses of natural gas price reporting to the index publishers in the 1990s resulted in an in-depth review of industry practices by the FERC and the CFTC. This topic will be covered in the chapters on natural gas.
- 33 <http://www.iosco.org/library/pubdocs/pdf/IOSCOPD285.pdf>.
- 34 "Task Force on Commodity Futures Markets Report to the Financial Stability Board," April 2011, p 7 (see <http://www.nefactioncenter.com/PDF/iosc.pdf>).
- 35 "Oil price reporting agencies," report by IEA, IEF, OPEC and IOSCO to G20 Finance Ministers, October, 2011 (see <http://www.iosco.org/library/pubdocs/pdf/IOSCOPD364.pdf>); OICU-IOSCO, 2012, "Functioning and oversight of oil price reporting agencies," consultation report, March (see <http://www.iosco.org/library/pubdocs/pdf/IOSCOPD375.pdf>); Task Force on Commodity Futures Markets, 2010, report to the G20 Technical Committee of IOSCO, OR 08/10, November.
- 36 "Improving the Transparency and Stability of Commodity Markets" (see [http://www.eurofi.net/pdf/2011/G20/Eurofi\\_Commodities\\_main.pdf](http://www.eurofi.net/pdf/2011/G20/Eurofi_Commodities_main.pdf)).
- 37 Remarks by Commissioner Scott D. O'Malia, CFTC, Eighth Annual Energy Trading Conference, Bauer College Global Energy Management Institute, University of Houston, March 23, 2012.
- 38 Carolyn Cui, 2012, "Regulators back off tougher curbs on oil," *The Wall Street Journal*, June 21.
- 39 <http://www.timesonline.co.uk/tol/comment/obituaries/article4827314.ece>.
- 40 "Price at which the majors and independent refineries sell branded or unbranded gasoline to jobber/wholesalers. It is related to the commodity spot price, but adjusted for transportation, overhead, and profit." (<http://www.petromarket.net/glossary/rack-price>).

Section 3

## **Natural Gas**



## *Natural Gas: Upstream*

This chapter covers the production and processing of natural gas prior to injection into the long-haul pipeline systems. Natural gas consumed at the burner tip consists primarily of methane with some additives and trace quantities of other substances. Natural gas at the wellhead is a mixture of hydrocarbons, other gases, water and other substances. In both cases, the same term, natural gas, is used in the industry. Industrial processes that convert gas extracted from the ground into gas consumed by residential, commercial and industrial users (through separation of methane from other hydrocarbons) will be explored in the section on natural gas processing plants, an important but relatively poorly understood segment of the industry. In the US, the importance of gas processing plants and the natural gas liquids they produce is increasing for the reasons that will be explained in detail in the next chapter.

We also provide information about the levels of production and consumption of natural gas in the US and worldwide, and consider the limitations of the available data. We conclude the chapter with a brief introduction to modelling the production flows of natural gas over time (decline curves), for individual wells and a portfolio of wells. This is an important topic for anybody who involved in decisions regarding the purchase/sale of gas-producing properties or in the purchase /sale of future volumetric flows (see the section on the volumetric production payments in Chapter 11). The controversies regarding the potential of shale natural gas resources evolve around estimates of the parameters of decline curves, and it helps to have some understanding of the modelling challenges in this area.

Some of what we cover in this chapter may seem unnecessary to some readers. However, in our experience, understanding a little bit

about the technology of natural gas never hurts. We can recall a successful transaction involving the purchase of a natural gas field, followed by a small celebratory party. The lead originator on the team asked after a few beers the question that was always on his mind: "How do you find those huge underground caverns filled with gas?" As the eyes started to roll around the table, we struggled to come up with a polite answer. "The easiest way is to make such caverns in the first place and then finding them is relatively easy unless one is forgetful."

## ORIGINS AND CHEMICAL PROPERTIES OF NATURAL GAS

Most natural gas consumed around the world was created along with crude oil during the geological period known as the Phanerozoic era, 5.3 to 570 million years ago.<sup>1</sup> According to the organic theory<sup>2</sup> of natural gas and oil formation, organic matter is deposited at the bottom of lakes, rivers, coastal waters along with sediments such as rock and sand, before changing into kerogen, the insoluble fraction of organic matter. Over time, the remains of plants and animals mixed with non-organic sediments evolve under conditions of high pressure and temperature, into source rock containing natural gas and/or oil. The composition of hydrocarbons found in the underground reservoirs depends both on the characteristics of the organic matter and the geological conditions under which it was accumulated. Organic matter comes from lakes, river deltas and marine basins. Natural gas comes primarily from land plants, with kerogen characterised by low hydrogen to carbon ratios and high oxygen to carbon ratios. Marine plankton and algal organic matter (with medium hydrogen to carbon ratios and low oxygen to carbon ratios) generates both natural gas and oil. Natural gas produced along with oil is known as associated gas. Marine and lacustrinal, mainly algal, organic matter (with high hydrogen/carbon ratios and low oxygen/carbon ratios) results in oil.

Organic matter, buried under layers of rocks, evolves into hydrocarbons under conditions of high temperature and pressure. The thermal gradient is equal on average to 4.7°F/100 metres. The process of transformation of kerogen into oil, called catagenesis, happens between temperatures of 140°–302°F, at depths ranging from two to six kilometres. Above this temperature level (up to 482°C), the process known as metagenesis leads to the development

of dry gas (methane) that can be found down to 10 kilometres. At a more shallow depth of between one and two kilometres, the process known as diagenesis leads to the development of methane mixed with other gases, such CO<sub>2</sub> and H<sub>2</sub>S.

Both oil and gas reservoirs are made of porous and permeable rocks into which the hydrocarbons migrate from the source rock. The migration process stops when oil and natural gas reach the surface or when a layer of impermeable rock is reached and the hydrocarbons are trapped under the ground. Geometric arrangements of rocks that permit significant accumulation of oil and/or natural gas are known as traps. The distinction between source rocks and reservoir rocks is critical to understanding the technology of natural gas production. We shall revisit it one more time in the next chapter on unconventional natural gas.

This brief discussion explains why natural gas found in nature is mixed with many other hydrocarbons and other trace gases. Table 8.1 shows a typical composition of natural gas produced at the well-head. One has to recognise, however, that the composition of natural gas may vary significantly from location to location. Technically, natural gas delivered to consumers is methane (CH<sub>4</sub>), with some other additives introduced for safety (methane is odourless and requires other gases to alert the end users when a leak occurs), blending (for example, nitrogen may be added to make natural gas suitable for pipeline transportation) and energy level enhancement (for example, ethane may be left in the gas stream when the heat content is especially valuable).

As mentioned, natural gas produced at the wellhead is a mixture of different gaseous hydrocarbons. In some reservoirs, natural gas is

**Table 8.1** Typical natural gas composition at the wellhead

Methane	CH <sub>4</sub>	70–90%
Ethane	C <sub>2</sub> H <sub>6</sub>	
Propane	C <sub>3</sub> H <sub>8</sub>	
Butane	C <sub>4</sub> H <sub>10</sub>	
Carbon dioxide	CO <sub>2</sub>	0–8%
Oxygen	O <sub>2</sub>	0–0.2%
Nitrogen	N <sub>2</sub>	0–5%
Hydrogen sulphide	H <sub>2</sub> S	0–5%
Rare gases	A, He, Ne, Xe	Trace

Source: AGA, <http://www.naturalgas.org/overview/background.asp>

dissolved in oil and is called associated gas, or wet gas. The term wet gas is also used to describe natural gas that contains a high percentage of heavier hydrocarbons (natural gas liquids). The term dry gas is used to describe natural gas after processing – ie, the removal of hydrocarbons other than methane, or natural gas that comes from geological formations where it consists almost exclusively of methane. Natural gas dominated by methane as the main component is also called lean gas.

Other terms used in the context of production of natural gas are:

- ❑ sour gas: natural gas containing relatively high quantities of hydrogen sulphide (as opposed to sweet gas);
- ❑ residue gas: natural gas after processing and removal of other hydrocarbons; and
- ❑ casing-head gas: natural gas that was dissolved in oil and was separated at the wellhead.

In addition to natural gas produced from the underground reservoirs, natural gas may be produced through industrial processes from coal or through fermentation of organic matter (manure or residential waste). In the first case, natural gas is called coal gas, town gas or syngas. In the second case, natural gas is called landfill gas.

Conventional natural gas is one of the terms whose precise meaning evolves over time. Non-conventional gas is natural gas that could not be, or was not, produced recently for technological or economic reasons. Breakthroughs in technology or changes in market conditions made the exploitation of these resources feasible. Currently, the US natural gas industry defines non-conventional natural gas as:<sup>3</sup>

- ❑ natural gas from shale formations;
- ❑ coal-bed methane; and
- ❑ tight sands natural gas.

Subsequent sections of this chapter will provide more information about the production trends and consumption patterns of natural gas in the US. Chapter 9 provides more information about US non-conventional natural gas supplies.

## GATHERING AND PROCESSING NATURAL GAS

### Gathering systems

The complexity of composition of natural gas explains why it has to pass through a number of processes before it can be injected into the interstate and intrastate pipeline system, and before it can reach the final user. The definitions of pipeline quality gas can vary from installation to installation and are enumerated in the natural gas tariffs. Most common conditions are related to the following characteristics:<sup>4</sup>

- ❑ heat content – natural gas should be within a specific Btu content range (1,035 Btu per cubic feet, +/–50 Btu);
- ❑ a specified hydrocarbon point dew<sup>5</sup> temperature level – below which any vapourised gas liquid in the mix will tend to condense at pipeline pressure, causing operational problem and potentially corrosion;
- ❑ only trace amounts of elements such as hydrogen sulphide, carbon dioxide, nitrogen, water vapour and oxygen are allowed; and
- ❑ removal of particulate solids and liquid water that could be detrimental to the pipeline or its ancillary operating equipment.

Producing wells are often dispersed over a large area and natural gas is collected through a network of low diameter and relatively low-pressure pipelines, known as a gathering system. A gathering system is typically connected to a processing plant, although in some cases dry gas can be injected directly into the pipeline system after decontamination. This is accomplished by installing so-called skid-mount plants, small-scale units located near the wellheads.

Associated gas has to be separated from oil at the wellhead in a special unit called a separator (this is required for roughly 25% of US-produced natural gas). Other processes taking place before natural gas can be safely injected into the pipelines include:

- ❑ in-field compression;
- ❑ solid particle removal;
- ❑ natural gas heating;
- ❑ natural gas blending;
- ❑ removal of water (dehydration);

- removal of elements such as helium, CO<sub>2</sub>, H<sub>2</sub>S and other rare gases; and
- removal of natural gas liquids (accomplished primarily in the natural gas processing plants).

The gathering systems often deploy heaters that prevent the temperature of natural gas from falling below the level at which hydrates are formed. Hydrates are crystalline solids that may block valves and pipes. There are many good reasons to remove water from the gas stream. Water can form ice that can choke valves and other components of the pipelines. Water combined with carbon dioxide and hydrogen sulphide can result in corrosion and the thinning of pipelines. Condensation of water in the pipelines can increase inlet pressure and cause a pressure drop downstream. Scrubbers are used to remove impurities such as sand from the pipelines. Another service offered by the gathering systems is known as in-field compression. As the reservoirs age, the internal pressure drops causing not only a drop in natural gas flow rates but also making it difficult to inject gas into the high-pressure pipelines. This problem is addressed through a gradual increase of pressure in the gathering system pipes, using small compressors that can be installed in small plants that can also serve as field-level processing facilities for decontamination, dewatering and some natural gas liquids removal. In-field compression is an important process as it allows increased production of natural gas from ageing reservoirs with falling internal pressure that could otherwise not overcome normal pipeline pressure. The gathering systems serve not only as conduits between the wellheads and the long-distance pipelines and distribution networks, but also as a buffer between the systems of different pressure levels.

Natural gas blending is an example of another process used to modify gas composition. Blending may be, for example, a process of mixing gas containing unacceptable levels of other gases (such as CO<sub>2</sub>, H<sub>2</sub>S) with natural gas of better quality. This solution is sometimes used during outages of the natural gas processing plants, due to maintenance or other causes, such as natural catastrophes. Blending was used on a large scale following hurricanes Katrina and Rita, which flooded many gas processing plants. Blending allowed routing natural gas around the inoperable plants and injecting it into

the pipeline system. An alternative to blending is treatment, consisting of a continuous circulation of gas over a chemical called amine that has a chemical affinity for CO<sub>2</sub> and H<sub>2</sub>S, and acts as an absorber. The impurities can be removed later by heating the amine.

The operations of natural gas processing plants can include some or all of the processes described above, in case they were not carried out upstream, but their main function is to remove the natural gas liquids from the natural gas stream. The (somewhat confusing and unintuitive) terms used frequently in the industry in the context of natural gas processing are *recovery* and *rejection*. Rejection means that a given hydrocarbon is not removed (separated) from the natural gas stream produced at the wellhead as a separate product, and is rather sold as a component of the gas stream available at the tailgates of the processing plants. Recovery is the opposite of rejection.

### Natural gas liquids

Natural gas liquids (NGLs) represent a poorly understood but increasingly important niche market with some unique features. There are many reasons why analysing natural gas liquids markets pose difficulties.<sup>6</sup>

- ❑ Natural gas liquids have multiple uses that make forecasting demand and data collection quite challenging and time consuming. In order to understand the demand (and price fluctuations), one has to understand the dynamics and economic fortunes of multiple industries using NGLs as raw materials, with unique features, complex technological processes and exposed to foreign competition. The prices of NGLs fluctuate depending on the values of many competing uses and the values of substitute products.
- ❑ The hydrocarbons known collectively as NGLs have multiple links to other parts of the energy complex, related to:
  - some natural gas liquids are produced in the refineries (and are known as liquefied petroleum gases or LPG) in addition to the NGL volumes related to natural gas production; and
  - some gas liquids are used in technological processes outside the natural gas complex (see the section on gasoline blending in the chapters on oil and refined products).

- ❑ There are many technological processes used for the extraction of NGLs from the natural gas stream, and their mix is constantly evolving as new facilities are added and old facilities are retired.
- ❑ The technologies for transportation and storage of NGLs are relatively poorly understood outside the NGL business.
- ❑ There are significant technological flexibilities (as will be explained later) in the extraction of NGLs that can be taken advantage of under different market conditions.
- ❑ The market for NGLs is much less transparent than the market for natural gas.
- ❑ Demand for some NGLs (such as propane) is seasonal, and driven by certain agricultural uses and weather conditions.
- ❑ The price dynamics of different NGLs follow different patterns. Some NGLs prices are closely correlated to natural gas prices (for example, ethane), while other liquids follow closely the prices of crude oil.
- ❑ The price correlations are unstable and fluctuate over time.

The bad news is that this is only a partial list of the challenges an analyst faces in this market.

The basic natural gas liquids are:

- ❑ Ethane ( $C_2H_6$ ) is a critical feedstock used in the chemical industry as an input to the production of many plastics – primarily ethylene, and also acetic acid and vinyl chloride. Ethane can be used as a refrigerant in cryogenic refrigeration systems. It is often left in the natural gas stream, for two primary reasons. One reason is that natural gas may be produced in the area where demand for ethane is low or where there is no infrastructure to move it to the chemical plants. At the time of writing this has been a serious problem in the Marcellus basin in the Northeastern US. The industry is working actively to address this challenge by developing new transportation options, storage facilities and considering the construction of ethylene crackers in the area. The second reason may be dictated by the relationship of market prices and the existing contractual arrangements between a plant and its customers (as explained in Chapter 11).
- ❑ Propane ( $C_3H_8$ ) is often called a liquefied petroleum gas,<sup>7</sup> as it is also a product of crude oil processing and may be transported as

a liquid after compression. Propane is used as fuel, both for residential and industrial applications. Residential use revolves around cooking or special-purpose heating (propane is called a cooking gas in many parts of the world). Some special uses of propane include camping equipment in the US, and even hot air balloons. Propane is an important replacement for wood or coal in many developing countries.<sup>8</sup> Industrial uses of propane include heat generation for many different processes, including laundries and agricultural crop drying and burning in the US. The latter use explains why the US propane prices display a seasonal peak in October, in addition to another winter peak associated with heating. Propane is used increasingly as a vehicle fuel (mostly for specialised vehicles such as forklifts), and is known outside the US as autogas. Many other uses, include applications as a feedstock in oil refining (see the chapters on oil and refined products), a feedstock in the manufacturing of plastics and chemicals, such as propyl alcohol, fuels for flame throwers and even special effects in movies (all these great fires and explosions we enjoy during a night out).

The industrial demand for propane is concentrated in the Gulf Coast area (PADD III). PADD stands for Petroleum Administration Defense District. Residential (space heating) and agricultural demands are concentrated in PADD I and PADD II.

- Normal butane ( $\text{CH}_3\text{CH}_2\text{CH}_2\text{CH}_3$ ) is also called n-butane. It is used as a very important input for seasonal gasoline blending, a process that is crucial to any transportation system (to be discussed at length in the chapters on oil). Other uses include as a fuel, either alone or in a mixture with propane (in this case, butane may be called LPG; the terms used worldwide lack standardisation). It can be also used as a feedstock for the manufacture of ethylene and butadiene, a key ingredient of synthetic rubber, or as a feedstock for steam cracking in refining (see the chapter on oil and refined products). Auxiliary uses include as fuel in cigarette lighters and propellant for aerosol sprays.
- Isobutane ( $\text{C}_4\text{H}_{10}$ ) is used by refiners to enhance the octane content of motor gasoline (see the chapters on oil), and also as a gas for refrigeration systems and propellant for aerosol sprays.

- ❑ Pentane ( $C_5H_{12}$ ), a main component of natural gasoline, is a very low-density liquid hydrocarbon, sometimes called natural gas condensate or simply condensate. Another frequently used term is natural gasoline. Natural gasoline is principally used as a motor gasoline blendstock or petrochemical feedstock. Other isomers<sup>9</sup> of pentane include iso-pentane and neo-pentane.

Given the molecular structure of natural gas liquids, these hydrocarbons are often referred to as:

- ❑ methane – C1;
- ❑ ethane – C2;
- ❑ propane – C3;
- ❑ isobutane and butane – C4; and
- ❑ pentanes – C5+.

The most important industry relying on NGLs as feedstock are the olefin plants (called sometimes ethylene plants, as this is the most important compound in the olefin group). Olefins (also discussed in the chapters on oil and refined products) include such chemicals as ethylene, propylene, butylene, butadiene and benzene. The olefin plants use as raw materials ethane, propane, butane, naphtha and gas oil, with the ability to switch from one feedstock to another, varying from plant to plant. For example, the plants cannot use butane from refineries because it is contaminated with fluoride. The decision of what feedstock to use depends on their relative yields and prices. Relative yields are shown in Table 8.2.

**Table 8.2** Olefin plant yields (pound per pound)

Outputs	Ethane	Propane	Butane	Naphtha	Gas Oil
Ethylene	0.80	0.40	0.36	0.23	0.18
Propylene	0.03	0.18	0.20	0.13	0.14
Butylene	0.02	0.02	0.05	0.15	0.06
Butadiene	0.01	0.01	0.03	0.04	0.04
Fuel Gas	0.13	0.38	0.31	0.26	0.18
Gasoline	0.01	0.01	0.05	0.18	0.18
Gas Oil	–	–	–	0.01	0.12
Pitch	–	–	–	–	0.10

Source: Donald L. Burdick, William L. Leffler, "Petrochemicals in Nontechnical Language," Penwell Publishing Company, Tulsa, Oklahoma, 1990

The infrastructure for the production and distribution of NGLs includes processing plants (covered later), pipelines, storage facilities and rolling stock for transportation (specialised railway cars, bobtails (local delivery trucks), etc). The pipelines connect major hubs and storage/distribution centres, including Conway: Kansas, Mont Belvieu: Texas, Medford: Oklahoma, Hattiesburg: Mississippi, Chicago: Illinois and Sarnia: Ontario in Canada. The Conway market is a cluster of several hubs, including Conway, Bushton: KS, McPherson, KS and Medford: OK.

### **Processing plants**

Natural gas processing plants usually perform two critical functions:

- separation of methane and heavier hydrocarbons; and
- NGLs fractionation.

As reported by BentekEnergy, there are several technologies used in the gas plants:

There are basically three types of natural gas plant technologies. (1) The oldest plants use lean oil absorption, which means that oil is used to soak up the liquids from the gas, and then the liquids are boiled out. This was the way gas processing was done 50 years ago, and most of these plants are either out of commission or they have been upgraded. The yield out of these plants is about 70% of butanes and 90% of pentanes (natural gasoline) in the gas stream. Propane recoveries are low, and the yield of ethane is effectively zero. (2) One step up is the refrigeration plant, in which the gas is chilled down to approximately minus 30 degrees. This process extracts essentially all of the pentanes and butanes, 70% of the propane, and 40% of the ethane. (3) The primary technology used for most new plants for the past 30 years is cryogenic, in which the gas stream is chilled to about -120°F. This process is much more energy efficient, and recovers almost all pentanes, butanes and propane and 65–95% of the ethane.<sup>10</sup>

Output from natural gas processing plants is referred to in the industry jargon as raw mix. Fractionation allows the separation of different types of NGLs into marketable products that are used as an important feedstock to many chemical processes. NGLs are then distributed through a pipeline system to petrochemical plants, refineries, propane distributors and other end users. The liquids pipelines transport usually what is referred to as an E/P mix

---

(ethane/propane) which is typically 80% ethane and 20% propane by volume. For a number of practical reasons listed below, propane represents no more than 25% of the mix.

- ❑ “Inadvertent rejection of propane. Depending on a plant’s capability, some propane usually gets rejected along with the ethane. That’s a bad thing with propane prices above US\$1.20/gal.
- ❑ Quality spec of tailgate gas. When ethane is rejected, the Btu content of the residue stream increases. Depending on the gas quality, rejection can result in violation of natural gas pipeline specifications.
- ❑ When plant co-owners and/or multiple processors are involved in a plant, it can be difficult to get consensus on rejection strategies, particularly in the middle of a month. In recent years, this has become less of an issue due to more efficient contracting.”<sup>11</sup>

High NGL content may be sometimes a mixed blessing for natural gas producers. The development of the new producing regions relying on shale natural gas requires not only massive investments in the development of local infrastructure (gathering systems, gas liquids plants, additional pipeline capacity) but also finding markets for NGLs. The Marcellus basin, described in the chapter on non-conventional natural gas, is a good example of the dilemma facing some producers.<sup>12</sup> Natural gas produced in the western part of Pennsylvania is characterised by high Btu levels and NGL content.<sup>13</sup> The producers are able to find buyers in the North East for all NGLs, with the exception of ethane. Propane can be used locally for crop drying and space heating. Butanes and natural gasoline can be used by local refineries and the chemical industry. Ethane, however, represents a major challenge as practically all US ethylene plants are located in the Gulf Coast area. The choices the industry faces include:

- ❑ shipment of ethane south to the US Gulf Coast using a number of alternative modes of transportation (pipelines, barges);
- ❑ shipment of ethane north to Ontario (the Sarnia area);
- ❑ blending of ethane with dry natural gas produced in the area (primarily from the wells located in Eastern Pennsylvania); or
- ❑ shutting down the wells as solutions for the disposal of ethane are developed.

The industry is currently considering a number of solutions, described in detail in the report referenced in footnote 10. This example is a good illustration of the complexities of modern energy business and the need to keep track of many moving parts in analysing market trends.

The profitability of NGLs operation may be captured through a number of indicators, including the most popular fractionation spread (frac spread). The calculations of crack spread are shown in Table 8.3. The calculations require specifying the basket proportions of different NGLs, and it is obvious that the percentages used will vary from region to region and may not apply to a specific producer. The conversions translate the prices of NGLs stated by conventions in cents (dollars) per gallon into dollars per MMBtu. The conversion factors vary slightly between different sources. In the example below, the prices are reduced by seven cents per gallon to account for transportation and fractionation (T&F) costs. In the calculations, the conversion factor is applied to the reduced price. In other words, we are applying a conversion factor to the T&F cost. An alternative calculation would treat this cost as a lump-sum adjustment to the fractionation spread. The fractionation spread, as with most indicators, is based on certain conventions, and debating which convention is better makes little sense.<sup>14</sup> We have to understand different conventions and use the one that seems most reasonable. The next step is multiplying the basket weights by the prices of NGLs (in US\$/MMBtu). The spread is calculated by subtracting the price of natural gas (US\$4.62/MMBtu in this case) from the basket price of NGLs US\$14.83/MMBtu.

**Table 8.3** Fractionation spread calculations

	NGL basket (%)	NGL prices (US\$/gallon)	T&F charges	Conversion (MMBtu/ gallon)	US\$/ MMBtu	Weighted US\$/ MMBTtu
Ethane	40.00	0.75	0.07	0.06633333	10.25	4.10
Propane	29.00	1.522	0.07	0.09166667	15.84	4.59
Normal butane	7.00	1.818	0.07	0.10364286	16.87	1.18
Iso butane	10.00	1.951	0.07	0.09957143	18.89	1.89
Natural gasoline	14.00	2.464	0.07	0.10945238	21.87	3.06
Total	100.00					14.83
Henry Hub Spread						4.62
						10.21

Source: BentekEnergy

The most recent summary information about processing plants is available from EIA for 2009.<sup>15</sup> This information is already obsolete, given the growth of natural gas production in new producing regions such as Marcellus. In 2009, there were 493 processing plants operating in 49 states, with a combined processing capacity of 77.5 billion cubic feet (Bcf) per day. Four states (Texas, Louisiana, Alaska and Wyoming) accounted for 71% of overall capacity.

#### *Processing plants: Summary*

To summarise, natural gas processing plants perform two critical functions in the natural gas supply chain:

- ❑ the removal of valuable hydrocarbons from the natural gas stream – these hydrocarbons, known collectively as NGLs, are inputs to many different production processes in the chemical industry and, by decree of the gods of the markets, became critical to the financial results of natural gas production from shale formations in the US (next chapter will cover this in more detail); and
- ❑ the conditioning of natural gas for injections into the pipeline system.

A comment is required at this point. In the US and the UK, it is a conventional wisdom in the natural gas industry that molecules are perfectly fungible once delivered into the pipeline grid. This is true for two reasons.

- ❑ Historically, natural gas has come mostly from similar formations and had similar characteristics.
- ❑ To the extent that gas was coming from different sources, the industry succeeded in homogenising it through pre-processing and processing before injection into the pipeline grid. Processing guarantees minimum methane amount and maximum levels of ethane, propane, butane, pentanes and heavier hydrocarbons, as well as nitrogen, oxygen and carbon dioxide.

This is not true, however, of other parts of the world where natural gas may flow from multiple locations (think Europe and natural gas coming through pipelines from Norway, Russia and Algeria). An increase in LNG flows complicates this picture even further, with the

potential for cargos arriving at regasification terminals with gas of varying quality.<sup>16</sup>

The industry is addressing this problem by incorporating into contracts inter-changeability parameters for handling gas from multiple supply sources. Globally, one can identify three major regions with respect to the calorific value of natural gas.<sup>17</sup>

- ❑ Asia (Japan, South Korea, Taiwan): heating value in excess of 1,090 Btu/scf (standard cubic foot);<sup>18</sup>
- ❑ Continental Europe: 990–1,160 Btu/scf; and
- ❑ UK, US: generally less than 1,065 Btu/scf.

Gas interchangeability provisions in the supply contracts are often stated in terms of the so-called Wobbe Index (or number), defined as:

$$WI = LHV \times \sqrt{SG}$$

where *LHV* stands for lower heating value, and *SG* for specific gravity. Specific gravity is calculated by dividing the density of a substance by the density of a reference substance (usually water). See the chapters on oil for more information. *LHV* is the amount of energy released during combustion less the energy required to vapourise the water present in a given substance.

### NATURAL GAS – UNITS OF MEASUREMENT

The obvious challenge for anybody engaged in the cross-border trading of natural gas is multiplicity of units and the difficulty of measuring actual flows of natural gas through the system. In the US (and, to some extent, the UK), a dual system of units is used, based either on heat content or on volume. The prices are stated in dollars per MMBtu and volumes in cubic feet. The reason for two parallel accounting systems is obvious: for transportation and storage, it makes practical sense to rely on volume. We trade natural gas for its energy content and the use of volume would not be practical, given that under different pressures the same volume would correspond to different energy amounts. An even more important reason to use energy content in trading is that this is what end users care about. However, shippers and pipelines care about volumes. In practice, we rely on a rule of thumb, according to which one cubic foot is equivalent to 1,000 Btus. In practice, this conversion is just an

approximation, given that the calorific content of natural gas produced in the US may vary from location to location and from year to year. For example, the average (unweighted) heat content of natural gas in the US calculated for 50 states + DC varied between 1025 and 1029 Btus per cubic foot (for dry natural gas delivered to consumers between 2003 and 2009).

Internationally, differences at the wellhead, before processing, are even more significant. We should always keep this in mind and distinguish between gas prior and post processing and from different locations. As mentioned above, prices are stated in dollars per million Btus (MMBtu). M stands for one thousand (as M is a Roman numeral) in this context (and MM for 1,000 times 1,000).<sup>19</sup> One should keep in mind that most countries use the “k” symbol (kilo) for thousand and “m” (mega) for one million. This corresponds to the prefixes used in the SI (International System of Units). We offer a few prefixes for reference in Table 8.4, as we shall encounter some of them in the energy landscape.

Daily consumption of natural gas or pipeline flows are reported in billions of cubic feet (Bcf) per day. Billion in the US means 1,000 time one million (or  $10^9$ ). One should remember that 10 raised to power 9 in some European countries is referred to as “milliard.” Achieving a useful definition of these units is not a trifling matter. After the Prussian–French war of 1870, France was obligated to pay reparations of 5 billion gold francs. The interpretation of the word “billion” mattered. Annual consumption/production of natural gas is measured in the US in trillions of cubic feet. Trillion stands for  $10^{12}$ . Another useful term is therm, or a 100,000 Btus. One MMBtu is equal to 10 therms (or a dekatherm).

Many countries use a different system of units than MMBtu to report the energy content of natural gas. Canadians use joules and

**Table 8.4** SI prefixes

---

$10^{12}$	tera	T
$10^9$	giga	G
$10^6$	mega	M
$10^3$	kilo	k
$10^2$	hecto	h
$10^1$	deka	da

---

Source: <http://physics.nist.gov/cuu/Units/prefixes.html>

**Table 8.5** Energy equivalent conversions

	Million Btu	Giga joules	TOE (metric tons of oil equivalent)	TCE (metric tons of coal equivalent)
<b>Million Btu</b>	1.00000	0.94782	39.68320	27.77824
<b>Giga joules</b>	1.05506	1.00000	41.86800	29.30760
<b>TOE (metric tons of oil equivalent)</b>	0.02520	0.02388	1.00000	0.70000
<b>TCE (metric tons of coal equivalent)</b>	0.03600	0.03412	1.42857	1.00000

Source: U.S. Energy Information Administration

many Asian and European countries have switched to reporting natural gas production and consumption in kilowatts. A few energy equivalent conversions are shown in Table 8.5.

One watt is approximately 3.41214 Btu/hour.<sup>20</sup> By extension, one Btu is 0.293071 watthours.<sup>21</sup> This conversion can be used if one has to translate the European statistics of natural gas consumption and production into the system of units used in the US. Anybody complaining that this is a headache should show some empathy with the plight of British consumers of natural gas:

So how do we convert hundreds of cubic feet of gas to kilowatt-hours? For this example take the current reading of your gas meter and deduct the previous meter reading from it. If your previous figure was an estimate this estimated reading may actually be higher, in which case deduct the previous reading from the current reading. This does not affect the arithmetic it just means either you have more to pay, or your energy company owes you something back.

Either way the difference between the numbers is net hundreds cubic feet of gas. Take this result and multiply it by 2.83, this converts hundreds of cubic feet to cubic metres. Multiply this result by a correction factor of 1.022640 and then by the calorific value shown on your last gas bill. Finally divide the result by 3.6 to give kWh.<sup>22</sup>

The bad news is that our problems are only just beginning. Natural gas is increasingly distributed in compressed form (compressed natural gas, or CNG) as a transportation fuel.<sup>23</sup> CNG may be measured using mass, energy or gasoline equivalent units. The *NIST Handbook* defines the gasoline gallon equivalent (GGE) as 5.660 pounds of natural gas. One GGE is equal to 114,118.8 Btus.

LNG in the international markets is traded in tons or cubic metres. The conversion used for LNG are as follows:<sup>24</sup>

1 ton LNG	= 48,700 cubic feet;
	= 1,379.03 cubic metres of natural gas;
	= 48.7 MMBtu.
1 cubic metre of LNG	= 20,631 cubic feet;
	= 5842 cubic metres of natural gas;
	= 20.63 MMBtu.

In Europe and in Asia, natural gas is traded using the unit of cubic metres. One cubic metre of natural gas is equal to about 35.3146667 cubic feet and about 35,314.6667 Btu. The latter is based on the assumption that one cubic foot is equal to 1,000 Btus. As we know, this is an approximation that ignores differences in the quality of natural gas.

Visit <http://www.natgas.info/html/natgasunitsconversion.html> for additional unit conversion information. Be aware that the conversions may vary from source to source and country to country (as we can see comparing the information above and Table 8.6). The conversions depend often on the assumptions of the ambient conditions under which measurements are carried out and the units as defined. The best practice is to:

- always rely on the government-established standards; and
- specify the units and their definitions in the contract.

Finally, Table 8.6 provides a summary of unit conversion for natural gas.

## NATURAL GAS CONSUMPTION AND PRODUCTION DATA

In this section, we will start with a summary of the statistical data for natural gas production and consumption trends and patterns in the US, followed by a discussion of the data sources available for the US and its shortcomings. It is important for any fundamental analyst covering the natural gas business to understand the highly imperfect nature of the weekly and monthly statistics the industry relies on.

### US production and consumption trends

Natural gas is an important US source of energy and one of the critical feedstocks to many industrial processes, in addition to industrial uses as a source of heat. The overall consumption of natural gas in the US in 2011 was equal to 24,267,483 MMcf, translating roughly

**Table 8.6** Unit conversion for natural gas

<p>1 tcf (trillion cubic feet) = <math>10^{12}</math> cubic feet is equal to:</p> <p>28.32 bcm 1,000 trillion (<math>10^{15}</math>) Btu ~1 Quad (<math>10^{15}</math>) Btu 172,4138 MM boe 1,054.615 PJ 20.53 MM ton (million metric ton) LNG</p>	<p>1 tcm m<sup>3</sup> (trillion cubic metres) is equal to:</p> <p>35,310.7345 bcf 35,314.67 trillion (<math>10^{12}</math>) Btu 35.31 Quad (<math>10^{15}</math>) Btu 6,088.0577 MM boe 37,239.2302 PJ 724,9294 MM ton (million metric tons) LNG</p>
<p>1 ton LNG is equal to</p> <p>48,709.2060 cubic feet 1,379.4447 cubic metres 8.3981 boe 48.7092 MMBtu</p>	<p>Other energy equivalents</p> <p>1 BTU = 0.252 kcal = 1.0546 kJ 1 kWh = 860 kcal = 3,600 kJ = 3,412 Btu 1 kWh = 3.41 cubic feet = 0.0966 cubic metres = 0.00059 boe 1 bbl crude oil = 5.8 MMBtu 1 bbl diesel = 5.825 MMBtu 1 bbl gasoline = 5.25 MMBtu 1 bbl fuel oil = 6.29 MMBtu</p>
<p>1,000 Btu energy equivalent to:</p> <p>1 cubic feet 0.0283 cubic metres 0.000172414 boe 1,054.6150 kJ</p>	<p>1 boe energy equivalent to</p> <p>5,800 cubic feet 164.256 cubic metres 5.8 MMBtu 6,116,767 kJ</p>

Source: [www.natgas.info](http://www.natgas.info)

Note: The information is available for most conversions above as on-line calculators

Assumes 1,000 cf = 1 MMBtu

boe – barrels of oil equivalent

into 24 Tcf (trillion cubic feet). The EIA reports consumption of natural gas by major sectors, and the breakdown in 2011 and 2001 is shown in Table 8.7.

Lease and plant fuel is related to the use of natural gas as fuel by the processing plants and other production-related uses such as fuel consumed in drilling operations, heaters, dehydrators and field compressors (as explained in detail earlier in this chapter). The pipeline and distribution category corresponds to the use of natural gas primarily as a fuel for compressors moving gas along the pipeline grid.

**Table 8.7** US natural gas consumption by sector (2011 and 2001)

Sector	2011 Consumption (MMcf)	2011 (%)	2001 (%)
Lease and plant fuel consumption	1407122	5.80	5.08
Pipeline and distribution use	682692	2.81	2.83
Residential consumption	4459030	18.37	20.66
Natural gas deliveries to commercial consumers in the US	3039210	12.52	13.35
Industrial consumption	6768516	27.89	33.37
Vehicle fuel consumption	32940	0.14	0.07
Deliveries to electric power consumers	7877973	32.46	24.63
Total	24267483	100.00	100.00

Source: U.S. Energy Information Administration

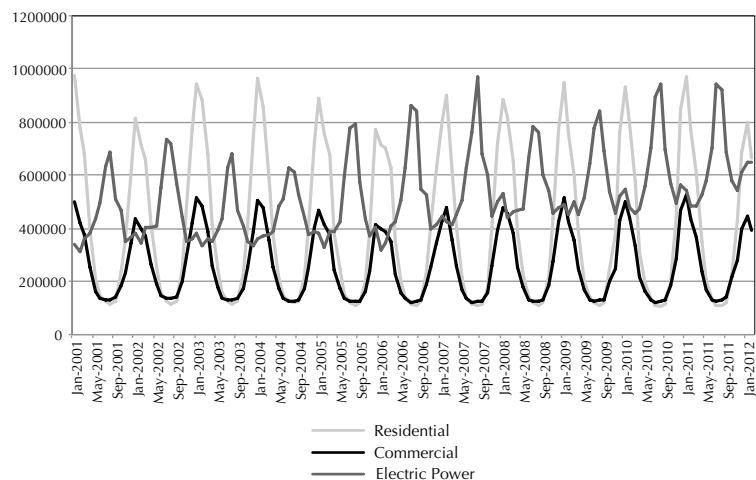
Notes: March through February of the next year

It is important to recognise that natural gas consumption, both at the national and sectoral level, is characterised by pronounced seasonality and variation over time, reflecting changes in the structure of the US economy. Figure 8.1 illustrates fluctuations of consumption of natural gas by three sectors with both significant usage levels and seasonality (commercial, industrial and electric generation).

Table 8.7 shows an obvious tendency towards an erosion of the share of industrial load and an increase in the share of natural gas used for power generation. As one can see, the shift happened primarily between industrial and power generation uses, at the expense of industrial consumption. This shift reflected the shutdown of many chemical plants in the US and shifting of industrial production to offshore locations,<sup>25</sup> where the prices of natural gas and other feedstocks are lower, and which are more competitive in other respects compared to the US. The low prices of natural gas make it more competitive with respect to other fuels used in power generation (such as coal). The drop in industrial use of natural gas reflects also changes in the reporting of natural gas used in industrial power plants (behind the fence generation).

As seen in Figure 8.2, the *aggregate* consumption of natural gas is highly cyclical, with the peaks occurring in the winter months and troughs taking place between April and October. Another, smaller seasonal peak in the summer months reflects the growing role of

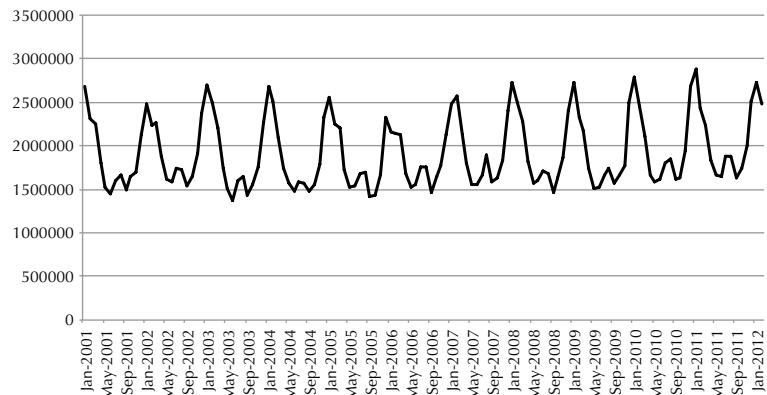
**Figure 8.1** US natural gas consumption by sector, monthly data (2001–2/2012, MMcf)



Source: U.S. Energy Information Administration

natural gas as fuel for power generation. This general seasonal pattern is a result of a combination of seasonal patterns of different sectors. Residential, commercial and industrial sectors peak in the winter (for the industrial sector, the picture is somewhat more complicated, as will be explained below). Power generation-related load peaks in the summer. A very strong seasonal component in natural gas consumption, related to heating load in the winter and the air conditioning load in the summer (natural gas is used as a fuel for power generation), makes temperature a very obvious choice as an explanatory variable in the regression models that can be used to produce short-term demand forecasts and weather-normalised forecasts (forecasts of consumption based on average (normal) temperatures). This is a sound approach that produces typically very reliable forecasts, although a few words of caution should be offered. First, there is a difference in the nature of industrial demand in a few states, such as California, Louisiana and Texas, and the rest of the US. Figure 8.3 shows industrial consumption of natural gas in these three states, while Figure 8.4 shows industrial consumption in the remaining states. One can see that, in the case of the three states, the seasonal component is not very pronounced. This can be explained

**Figure 8.2** US natural gas total consumption, monthly data (2001–2/2012, MMcf)



Source: U.S. Energy Information Administration

by the fact that the industrial consumption of natural gas in these states is driven primarily by its use as a feedstock or fuel (for example, in aluminium smelters) and is dependent more on the general level of activity for the chemical and metallurgical industry. In the remaining states, industries use natural gas primarily for space heating, so temperatures serve as useful explanatory variables.

The data related to the US natural gas production data can be summarised as follows, following the conventions used by the EIA:

Natural gas gross withdrawals from gas wells  
Natural gas gross withdrawals from oil wells

---

US natural gas gross withdrawals  
Less:  
Natural gas repressuring<sup>26</sup>  
Nonhydrocarbon gases removed from natural gas  
US natural gas vented and flared

---

Natural gas marketed production  
Less:  
US natural gas extraction loss

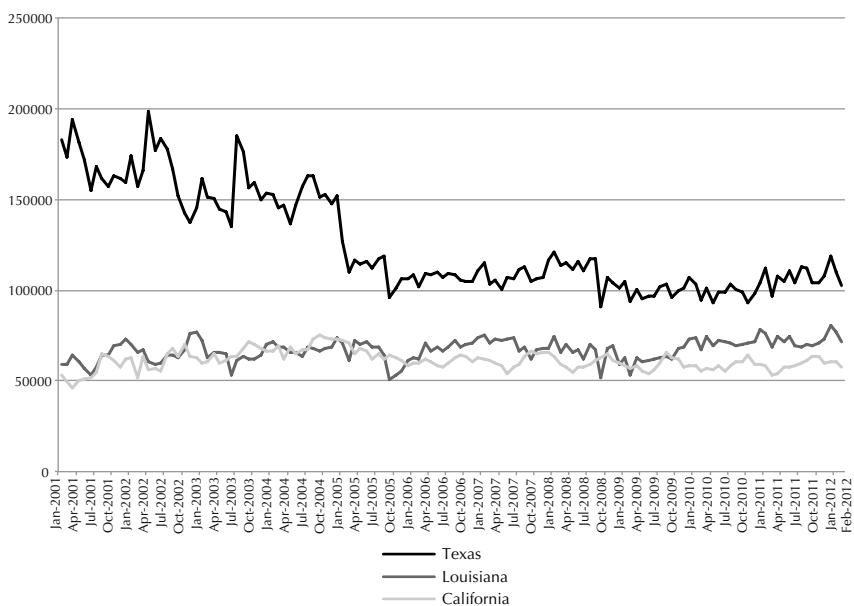
---

Dry<sup>27</sup> natural gas production

---

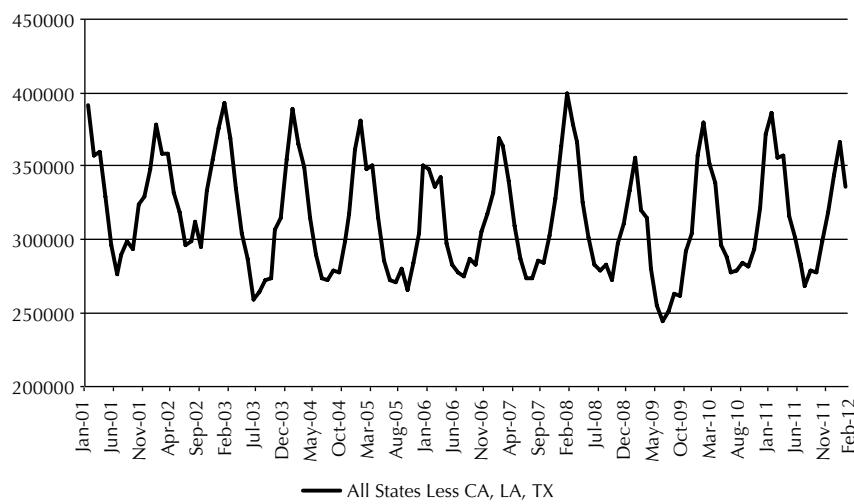
Figure 8.5 shows dry natural gas production by month, which requires two comments. First, one can see sharp drops in production

**Figure 8.3** US natural gas industrial consumption, monthly data (CA, LA, TX) (2001–2012, MMcf)



Source: U.S. Energy Information Administration

**Figure 8.4** US natural gas industrial consumption, monthly data (without CA, LA, TX) (2001–2012, MMcf)



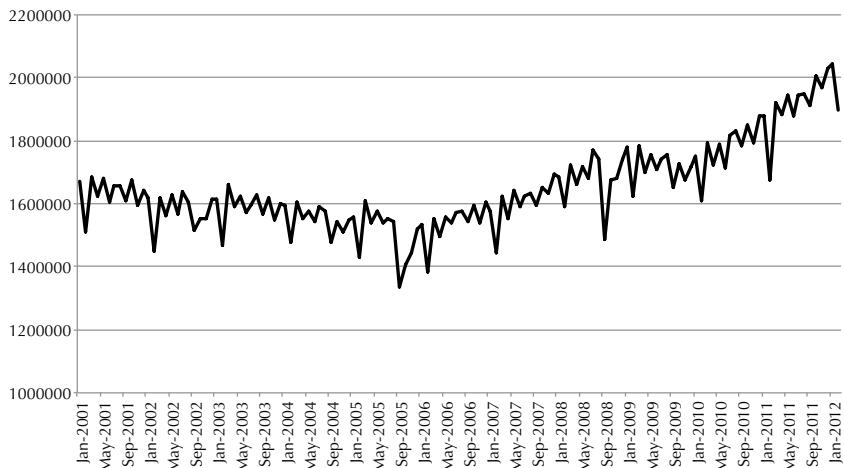
Source: U.S. Energy Information Administration

from time to time. These drops were caused by hurricanes (such as Rita and Katrina) temporarily shutting down production of natural gas in the Gulf of Mexico. Second, one can see an increase in the production of natural gas beginning around 2006, the result of the shale revolution covered in the next chapter. Third, Figure 8.5, like the figures illustrating changes in consumption, does not make adjustments for different number of days in a month, and sometimes fluctuations of the flows from month to month are artificially amplified. Figure 8.6 shows the evolution of natural gas production per day.

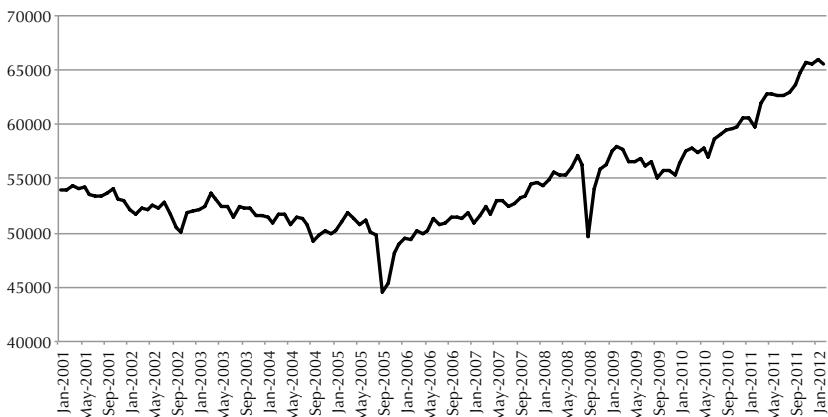
A quick glance at Figures 8.2, 8.5 and 8.6 demonstrates an obvious dilemma for the US natural gas industry. Production tends to be rather stable from month to month, but the consumption is highly seasonal. This requires creation of a buffer between production and consumption flows. This buffer is provided by the natural gas storage, an exceptionally important component of the industry infrastructure.

Figure 8.7 shows the composition of the sources of US natural gas. This figure illustrates the evolving role of imports of natural gas in supplying the US market. We can see again a significant increase in the domestic output of dry gas (naturally dry gas or gas after

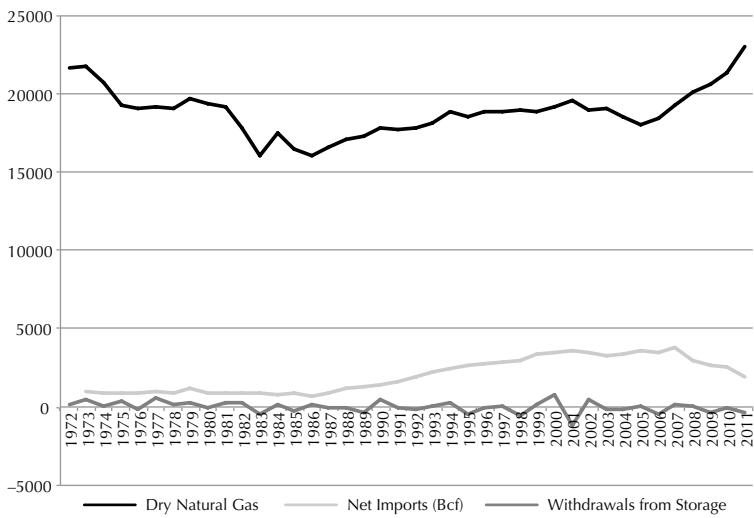
**Figure 8.5** US dry natural gas production by month (2001–2012, MMcf)



Source: U.S. Energy Information Administration

**Figure 8.6** US daily natural gas production (2001–2012, MMcf)

Source: U.S. Energy Information Administration

**Figure 8.7** US sources of natural gas by year, Bcf (1980–2008)

Source: U.S. Energy Information Administration

removal of natural gas liquids) and a decreasing reliance on imports in the last few years. The net contribution of gas from storage is small on an annual basis, but oscillates between injections and withdrawals during the course of a year.

### World production of natural gas

The principal source of data regarding production of natural gas worldwide used by the industry is the BP "Statistical Review of World Energy". Figure 8.8 shows production of natural gas (daily, Bcf) by major geographical regions. The total daily output for the world in 2010 was equal to 309 Bcf. About two thirds of the output was accounted for by a relatively small number of producing countries, a group that included the US, the Russian Federation, Canada, Iran, Qatar, Norway, China, Saudi Arabia, Indonesia and Algeria (see Figure 8.9).

A sobering realisation is that, out of 10 countries representing two thirds of the world output of natural gas, several producers have either net deficits (domestic consumption exceeds consumption) or have relatively small or stagnating volume available for exports. Other countries with significant excess production volumes over consumption in 2010 (but smaller overall production levels than the 10 countries listed in Figure 8.9 include Trinidad and Tobago (2 Bcf/day), Turkmenistan (1.9 Bcf/day), Uzbekistan (1.3 Bcf/day), Egypt (1.6 Bcf/day), Australia (1.9 Bcf/day) and Malaysia (3.0 Bcf/day).<sup>28</sup>

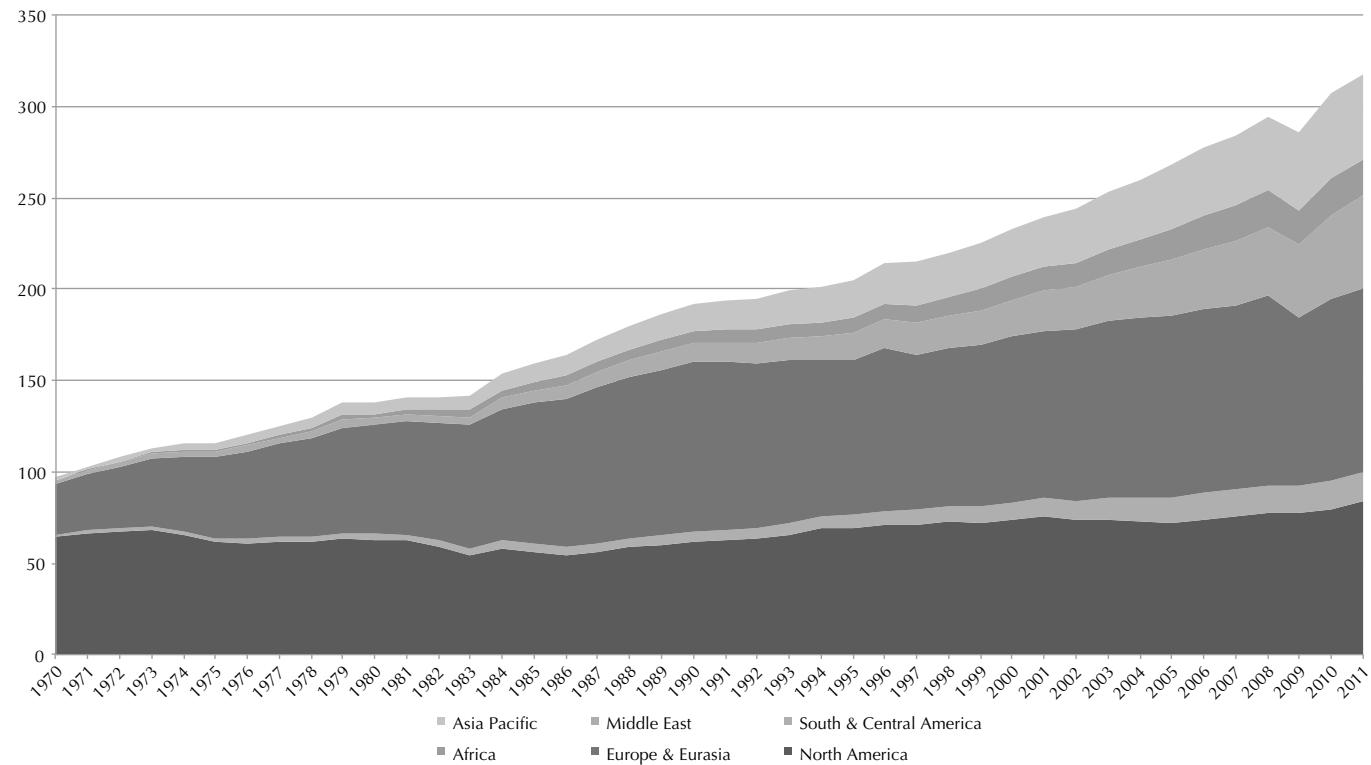
### Data sources and data quality issues

Any user of massive volumes of statistics should spend some time reviewing the statistical procedures used in data generation. The available data for the production and consumption of natural gas are produced through a complex process that combines collection of raw data and estimates based on partial samples. The statistical methods and procedures used by government agencies introduce into the process unavoidable bias that has to be understood and corrected by an analyst.

The current methods for collection of the natural gas production data have been implemented by the EIA in April 2010.<sup>29</sup> As in any statistical process, two critical steps are sampling and estimation, the process of extending the sample information to the entire universe of examined data.

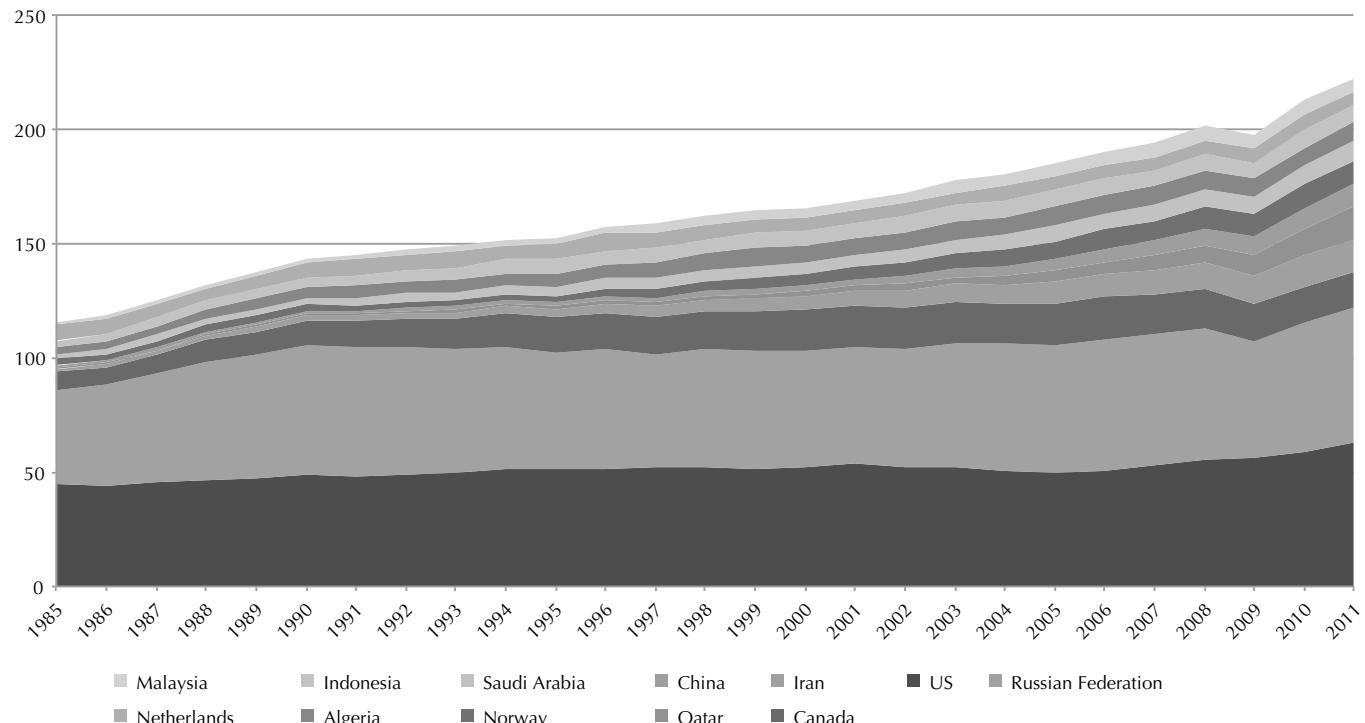
The EIA collects natural gas production data on a monthly basis using the form EIA-914, with a sample of companies that is refreshed every month. Figure 8.11 illustrates the procedures for collecting the data.<sup>30</sup> Refreshing the sample is required for a number of reasons,

**Figure 8.8** Production of natural gas by region (Bcf/day)



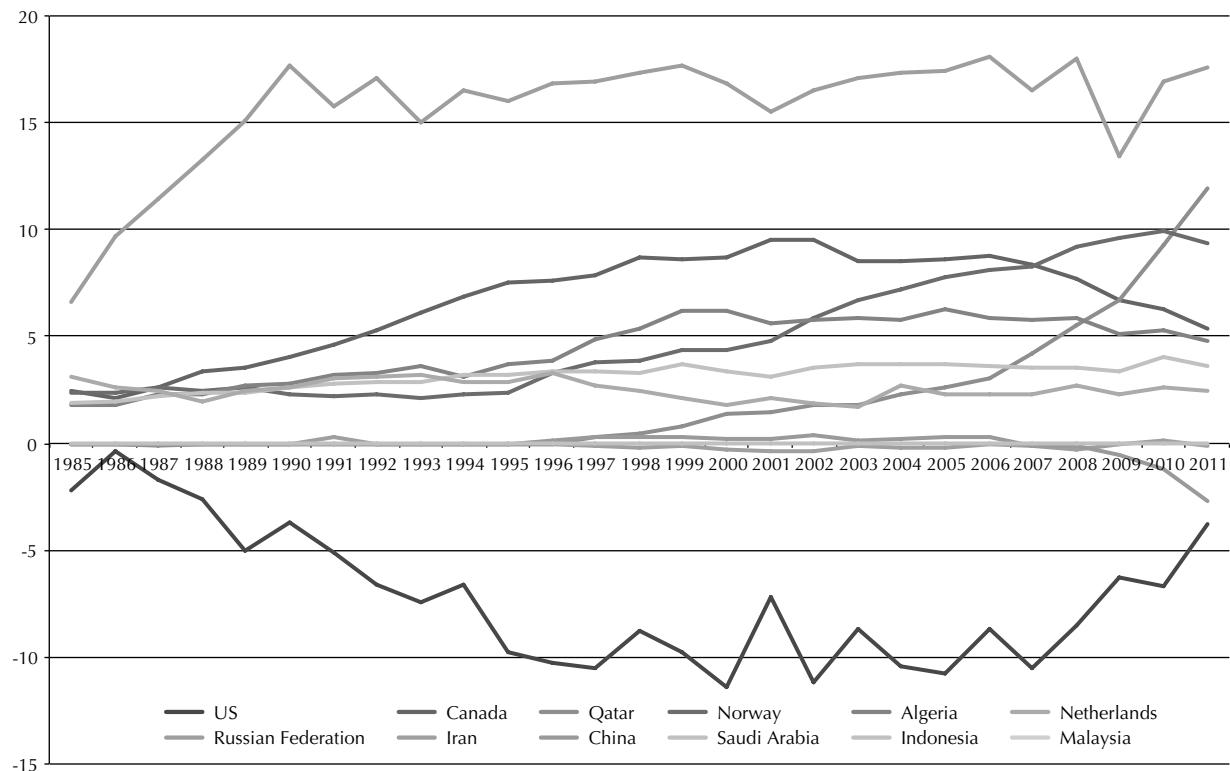
Source: BP, "Statistical Review of World Energy 2011"

Figure 8.9 Production of natural gas, 10 countries (Bcf/day)



Source: BP, "Statistical Review of World Energy 2011"

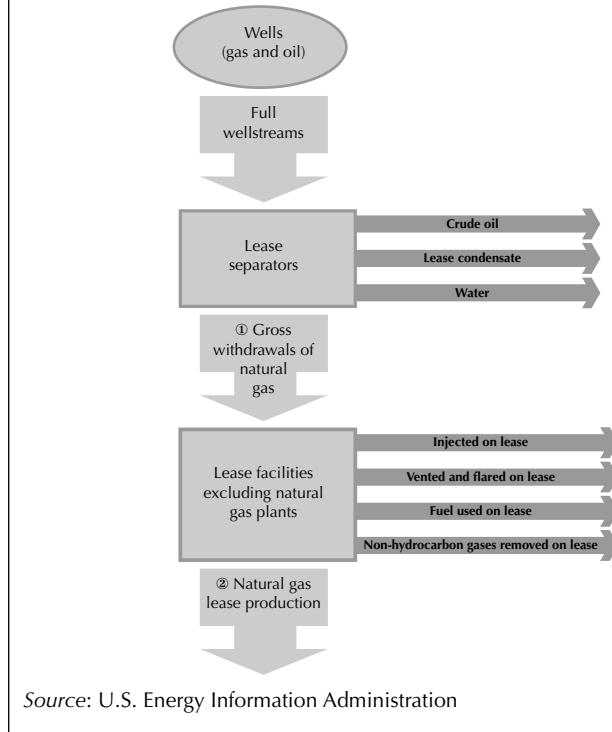
**Figure 8.10** Net production of natural gas (production less consumption), 10 countries (Bcf/day)



Source: BP, "Statistical Review of World Energy 2011"

including mergers and acquisitions, as well as due to fluctuations in production levels. The sampling process uses a cut-off level eliminating small operators, and a respondent falling below the minimum threshold for a period of time is removed from the survey. Both the sampling and the estimation processes are supported by the HPDI<sup>31</sup> database. This database contains information acquired from state regulatory agencies at a well or lease level on a monthly basis. Information for certain states is missing or inadequate (Illinois, Indiana, Kentucky, Pennsylvania and Tennessee), and is supplemented with data collected from the EIA-23 data survey (Annual Survey of Domestic Oil and Gas Reserves). The sampling process uses a threshold level of 20 MMcf/day, except for Oklahoma, where the threshold is 10 MMcf/day – a lower threshold for Oklahoma reflects a lower level of industry concentration in this state, compared with the rest of the US. Companies falling below the

**Figure 8.11** EIA-914 process for reporting natural gas production



threshold for a period of time are removed from the sample, and the companies exceeding the threshold may be added. The procedure for removal and addition from and to the sample gives some discretion to the company administering this process.

Estimation is based on the simple ratio (SR) method for the states listed in the "Monthly Gross Natural Gas Production Report." This method is based on application of the ratio of total production to the current sample's production, calculated with a time lag, to the current reported sample volume. The time lag means that the ratio for a December sample may be calculated with a lag of six months, using total production and the production corresponding to the current sample for the month of June of the same calendar year. The lags vary by state and range from six to 18 months. The use of lagged ratios is necessary because the current months' HPDI data is not complete. The lagged ratios for a given state are averaged over a six-month period. The procedure is summarised through the following equations:

$$AvgSR_{i-L} = \frac{\sum_{i-L}^{i-L-5} \frac{TP_i}{SP_i}}{6} \quad (8.1)$$

$$TPest_i = S_i \times AvgSR_{i-L} \quad (8.2)$$

where  $AvgSR$  denotes a simple ratio, six-month average

$TP$	Total production from HPDI
$SP$	Sample production from HPDI
$L$	Lag time in months
$TPest$	Total production estimate for the current month
$S_i$	Sampled production for the current month
$i$	Current or estimation month

For other states, the ratios of EIA-895 survey (Annual Quantity and Value of Natural Gas Production Report) to the EIA-914 survey annual volumes for the previous calendar year is used to inflate the current month EIA-914 sample volumes to produce the estimate of the current production.

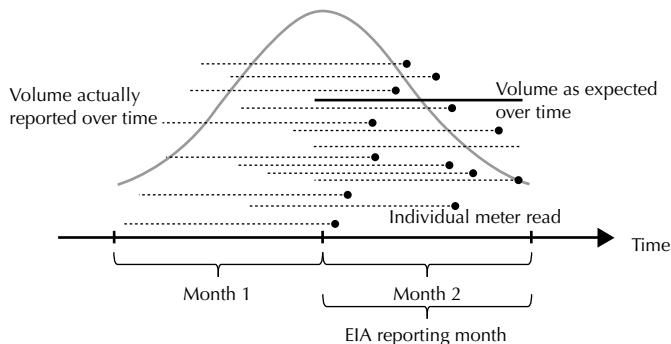
The most frequently used EIA consumption data are reported by month, state and several different categories of end users, such as:

- residential;
- commercial;
- industrial;
- natural gas vehicle fuel; and
- electric power consumers.

The data is available with a lag of a few months and is subject to *ex post* revisions. A caution is required also with respect to EIA consumption data. The commercial and residential consumption data is based on information reported by local distribution companies (LDCs), based on the billing months that do not coincide with the calendar months. This problem is illustrated in Figure 8.12 from a presentation by the EIA. The discrepancy between true monthly and billing cycles can be successfully cured using econometric techniques developed by a number of quantitative research groups in the industry. The author worked with BentekEnergy on the implementation of such algorithms. One obvious regularity is that EIA-reported volumes are understated in the periods of rising monthly consumption and overstated under conditions of falling volumes. Even if the average annual consumption changes slowly, the seasonality of demand for natural gas will induce significant month-to-month distortions between true and reported numbers.

The deliveries of natural gas to consumers have a number of well-known defects that have to be corrected using statistical techniques, or at least recognised in the analytical process.

**Figure 8.12** EIA natural gas consumption data collection



Source: Steve Harvey, 2009, "US Natural Gas Balances. Issues in Monthly Consumption Data," Benposium, June 4, Houston.

Given the shortcomings of the official production and consumption data discussed above, the industry is searching for alternative sources of information. One viable alternative is the use of pipeline nominations data discussed in Chapter 10. These data have a number of obvious advantages over EIA-reported natural gas consumption and production numbers.

The pipeline flow data are available daily and represent a very detailed and precise snapshot of the level and flows of natural gas in the US, including flows across the borders. It is important to recognise that actual flows may be slightly different than nominations, but the discrepancies are very small and will be likely awash, given positive and negative deviations. The EIA data are available with a time lag of several months and are subject to a number of updates and corrections.

The nomination data are often meter-specific, and many meters can be associated in turn with certain energy-related assets, such as power plants, natural gas storage facilities and industrial plants or gathering systems and gas processing plants. In many cases, such meters serve unique customers. The association of meters with energy assets allows the creation of real-time (daily) samples of demand by different user classes, at levels of granularity exceeding the highly aggregated reporting categories used by the EIA. This information is, however, not complete as certain meters may be supporting a large number of end users, and it is impossible to allocate the flows precisely to specific classes of customers.

The inability to allocate pipeline flows to different categories of end users can be addressed in a number of ways. We can argue that, for most trading desks, only aggregate consumption of natural gas at the national and regional levels matters, and this is certainly true. We can also argue that the allocated flows represent a sufficient sample to detect early trends in the structure of consumption of natural gas (ie, changes in the level of consumption by different users) and this again is all that matters. An alternative approach used by most analysts is to develop a statistical model relating consumption data by end-user class based on pipeline flows to the EIA-reported numbers. A statistical model based on historical data allows us to translate currently available flow sample data to the EIA-equivalent aggregates. This approach depends critically on the quality of the mapping from pipeline meters to the energy assets, available from the data vendors or developed in-house.

---

The industry is making a transition from the traditional reliance on EIA data to more extensive use of pipeline flow data. The importance of this trend can hardly be exaggerated. The availability of high-granularity and real-time data will allow the industry to develop a new class of quantitative forecasting models. The trading organisations that will be first to exploit this opportunity will capture a considerable competitive advantage.

### DECLINE CURVES AND DEPLETION ANALYSIS

It seems reasonable to discuss the decline curves and tools for modelling the depletion of reservoirs in this chapter. These rather simple mathematical models are used across the energy industry whenever we have to deal with an exhaustible resource and seek to project future volumetric flows from historical data. Decline curve analysis is used extensively in developing cashflow projections for individual oil and natural gas wells, and for larger properties containing multiple wells. Cashflow projections are critical to the valuation of producing properties and to hedging decisions. The debate regarding shale gas (see Chapter 9) profitability (or lack of it) evolves around the calibration of the equations discussed in this section to empirical data, equations that have been the bread and butter of professional geologists but ignored by everybody else. These equations are very important and our recommendation to the finance professionals is not to skip this section – although, as we know, the devil will tolerate an aspersion with holy water better than an MBA will a mathematical formula.

The concept of decline curves was introduced by J. J. Arps in 1945, and survived the test of time due to its simplicity and mathematical tractability.<sup>32</sup> Denoting the initial production rate by  $r_0$ , one can postulate that the production rate at time  $t$ ,  $q(t)$ , is given by one of the three equations:

$$\text{Exponential : } q(t) = r_0 \exp\{-\lambda(t - t_0)\} \quad (8.3)$$

$$\text{Hyperbolic : } q(t) = r_0 \left\{1 + \lambda\beta(t - t_0)\right\}^{-1/\beta} \quad (8.4)$$

The harmonic equation is the hyperbolic equation with  $\beta$  equal to 1. The parameter  $\lambda$  in equations 8.3 and 8.4 is the decline rate. Coefficient  $\beta$  in equation 8.4 is the shape parameter. Starting  $t_0$  time is usually taken as equal to 0. The cumulative production is given by:

$$Q(t) = \int_{t_0}^t q(u)du \quad (8.5)$$

The  $Q(t)$  equations are given by:

$$Q(t) = \frac{r_0}{\lambda} \left\{ e^{-\lambda t_0} - e^{-\lambda(t-t_0)} \right\} \quad (8.6)$$

$$Q(t) = \frac{r_0}{\lambda(\beta-1)} \left[ (1 + \lambda\beta(t-t_0))^{1-\frac{1}{\beta}} - 1 \right] \quad (8.7)$$

for the exponential and hyperbolic equations, respectively.

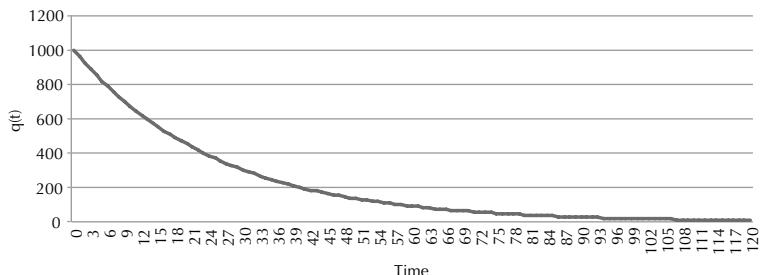
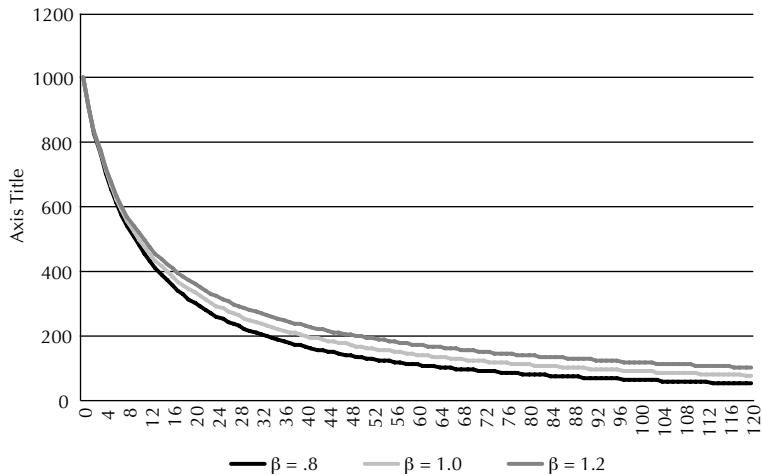
If the cut-off point for production (ie, the point at which the volumetric flows do not cover the costs of operating a well) is determined, we can use the formulas to calculate the shutdown time. Figure 8.13 shows the plot of an exponential decline curve. It is important to emphasise that even small differences in assumptions may have significant consequences for the economic outcomes of different projects. Figure 8.14 shows the plot of hyperbolic decline curves for different values of the  $\beta$  coefficient.

A paper published by a group of Schlumberger engineers (see Table 8.8 below for reference) and available from the Society of Petroleum Engineers (SPE) contains estimates of decline curves coefficients for a number of basins, rock formations and well types. We include the reported values to give a reader an idea of the values one can encounter in practice. The values vary from location to location and study to study. There is a potential for some estimates to be affected by the aggregation process, and the results should be treated with an extreme caution.

A few useful definitions follow. These terms are used often in practice (with some differences between different sources) and it makes sense to offer a set of internally consistent definitions. The level of decline rates is a very important and controversial topic, especially in the oil industry.<sup>33</sup> The initial reserves are denoted by  $R_0$ . The remaining reserves (what remains in the ground) are given by  $R_r = R_0 - Q_t$ . The depletion level is defined as the ratio of cumulative production and initial reserves at time  $t$ :

$$D_t = \frac{Q_t}{R_0} \quad (8.8)$$

with the depletion rate at time  $t$  given by:

**Figure 8.13** Exponential decline curve ( $\lambda = 0.04$ )**Figure 8.14** Hyperbolic decline curves ( $\lambda = 0.1$ ,  $\beta = 0.8, 1.0, 1.2$ )

Note: Time measured in months

$$d_t = \frac{q_t}{R_0} \quad (8.9)$$

The depletion rate that applies to the remaining reserves is given by:

$$d_{\delta t} = \frac{q_t}{R_r} = \frac{q_t}{R_0 - Qt} \quad (8.10)$$

This rate is equal to the inverse of reserve-to-production ratio used often in practice to characterise a given producing property.

For exponential decline curves, the depletion rate is equal to the decline rate. In the hyperbolic case:

**Table 8.8** Decline curve estimation results

Case	Reservoir	Well type	$\beta$	$\lambda$
Barnett	Shale	Horizontal	1.5933	0.0089
Fayetteville	Shale	Horizontal	0.6377	0.0325
Woodford	Shale	Horizontal	0.8436	0.0227
Haynesville	Shale	Horizontal	1.1852	0.0632
Eagle Ford	Shale	Horizontal	1.694	0.0826
Cotton Valley	Tight gas sandstone	Horizontal	0.7259	0.0248
Cleveland	Tight gas sandstone	Horizontal	1	0.0149
Cotton Valley (1980)	Tight gas sandstone	Vertical	1.2778	0.0021
Cleveland (1980)	Tight gas sandstone	Vertical	2.3483	0.0022
Cotton Valley (>2005)	Tight gas sandstone	Vertical	1	0.0175
Barnett (1980s)	Shale	Vertical	1.9366	0.0046

Source: Jason Baihly, Raphael Altman, Raj Malpani and Fang Luo, 2011, "Study assesses shale decline rates," *The American Oil and Gas Reporter*, May.

$$d_{\hat{a}t} = \frac{\lambda(1-\beta)}{1+\lambda\beta(t-t_0)} \quad (8.11)$$

Most "how to" books recommend using semi-log graph paper for plotting the production data. The reason for doing this is obvious. The general form of an exponential function is given as:

$$y = ae^{rt} \quad (8.12)$$

where  $r$  denotes the rate of growth<sup>34</sup> (constant for the exponential function) and  $t$  and – time. Parameter  $a$  is a constant. Taking natural log of both sides, we obtain

$$\log(y) = \log(a) + rt \quad (8.13)$$

The graph of the log of variable  $y$  versus time is a straight line. Plotting the logs of production volumes against time, we shall get a straight line if the decline curves are in fact exponential. The production levels at times  $t_1$  and  $t_2$  allow to come up with an estimate of the growth rate  $r$ :

$$\log(y_1) = \log(a) + rt_1 \quad (8.14)$$

$$\log(y_2) = \log(a) + rt_2 \quad (8.15)$$

$$r(t_2 - t_1) = \log(y_2) - \log(y_1) \quad (8.16)$$

$$r = \frac{\log(y_2) - \log(y_1)}{t_2 - t_1} \quad (8.17)$$

$$r = \frac{\log(y_2 / y_1)}{t_2 - t_1} \quad (8.18)$$

Decline curve analysis is a viable approach after certain minimum number of data points becomes available. In practice, it is difficult to resist the temptation to extrapolate production trends relying on a limited set of historical data before one can determine what statistical model fits a specific situation. Many questionable deals have happened due to buyers failing to predict high decline rates or not asking for the relevant information. In practice, decline curves are often estimates for datasets containing many wells (well-by-well analysis is time intensive and requires a lot of effort to generalise the results). This approach is not without dangers, as has been pointed out by Art Berman.<sup>35</sup> Throwing a large number of wells with different completion dates into a common pot requires aligning their individual production data according to an artificial timeline. For example, the data for the well completed on June 15, 2010, and the well completed on July 15, 2011, are shifted to the same starting point of the timeline. The dangers of such approach should be obvious to any statistician. The wells drilled at different point in time may not be directly comparable as rig operators acquire more experience and become more efficient.<sup>36</sup> Efficiency gains in different natural gas shale formations have been well documented.<sup>37</sup> The data may be distorted due to well-stimulation efforts and from shutdowns resulting from pipeline constraints.<sup>38</sup> The most important factor is the survivorship effect: the underperforming wells are abandoned early, and this may create an illusion of slower production declines over time. Art Berman recommends well-by-well analysis or analysis based on data corresponding to different vintages of wells (for example, wells completed during the year 2010).

An alternative to decline curve analysis, which requires a reasonably long history or production flow, is a technique that allows for developing initial estimates of oil/gas initially in place (OIIP/OGIP). It should be emphasised that this approach has to be supplemented with an assumption of the recovery factor and does not provide

information about the time profile of volumetric flows. OIIP can be calculated as follows:<sup>39</sup>

$$N = \frac{7758\phi(1-S_w)hA}{B_0} \quad (8.19)$$

where  $h$  denotes the formation thickness (measured in feet);

- $A$  drainage area (in acres)
- $\phi$  porosity (in percentages)
- $S_w$  formation water saturation
- $B_O$  oil formation volume factor.<sup>40</sup>

For example, for a productive zone of 15 feet, with water saturation of 35% and an oil formation factor of 1.215, OIIP is about 747,000 barrels for a property of 80 acres.

The corresponding calculation for GIIP is given by:

$$N = 43560\phi(1-S_w)hAB_g \quad (8.20)$$

$B_g$  stands for the gas formation volume factor. One has to remember that oil shrinks or expands when brought to the surface. The recovery factor will vary from formation to formation and will be different for oil and gas. The rule of thumb recovery factor for oil is about 30%, but one should be very careful in coming up with an estimate, and a qualified geologist should be consulted. A detailed discussion of this issue is beyond the scope of the book and the author's abilities.

## CONCLUSIONS

This chapter covered the chemical and physical properties of natural gas and the steps necessary to prepare this fuel for injection into the pipeline system. Natural gas liquids are a critical link in the value chain of the US natural gas industry and have recently become very important to its overall profitability. This topic will be covered in the next chapter. Natural gas consumption is characterised by pronounced seasonality, which explains the need for storage, serving as a buffer between relatively stable production flows and fluctuating demand. Storage facilities and the importance of inventory reports will be discussed in Chapter 10. Natural gas tends to be produced at locations removed from population and industry centres, which is why an extensive and capital-intensive transporta-

tion infrastructure is required. Pipelines and LNG, two competing modes of transportation for natural gas, will also be covered in Chapter 10.

- 1 The discussion of the origins of natural gas and oil is based on the PhD dissertation by Fredrik Robelius, 2007, "Giant oil fields: The highway to oil", Uppsala Universitet (available at <http://publications.uu.se/abstract.xsql?dbid=7625>).
- 2 According to abiotic theory, natural gas is not a finite resource but is constantly produced inside the earth's crust under conditions of high pressure and temperatures. As it migrates to the surface, it becomes contaminated with organic matter, leading to the formulation of incorrect theories about its origins.
- 3 According to the EIA, "'unconventional gas' refers to natural gas extracted from coalbeds (coalbed methane) and from low-permeability sandstone and shale formations (respectively, tight sands and gas shales)" (see [http://www.eia.gov/oiaf/analysispaper/unconventional\\_gas.html](http://www.eia.gov/oiaf/analysispaper/unconventional_gas.html)).
- 4 EIA, 2006, "Natural gas processing: The crucial link between natural gas production and its transportation to market," Office of Oil and Gas, January ([http://www.eia.doe.gov/pub/oil\\_gas/natural\\_gas/feature\\_articles/2006/ngprocess/ngprocess.pdf](http://www.eia.doe.gov/pub/oil_gas/natural_gas/feature_articles/2006/ngprocess/ngprocess.pdf)).
- 5 "The hydrocarbon dew point, or HCDP, is the temperature and pressure at which heavy components of the stream condense and begin to form liquids. It is not uncommon in some parts of the country for ambient temperatures to cool a natural gas stream down to its hydrocarbon dew point and cause condensation within the pipeline." See Darin L. George, Andy M. Barajas and Russell C. Burkey, 2005, "The need for accurate hydrocarbon dew point determination," *Pipeline & Gas Journal*, September.
- 6 One can find timely updates on the dynamics of the NGL prices on the following blog: <http://www.rbnenergy.com/>.
- 7 Propane is mixed in this case with butane and other hydrocarbons.
- 8 The author, who grew up in Poland, remembers the use of propane-filled steel cylinders, delivered by specialised trucks, in his mother's kitchen. By the way, outside the US, propane is mixed with butane, and this means that US standard equipment should not be used elsewhere and vice versa. The percentage of butane may be as high as 75%.
- 9 Isomers are the molecular structures which have atoms bonded together in different orders.
- 10 BentekEnergy, 2010, "Market alert: The rich get richer, shale gas and NGLs," December 30.
- 11 <http://www.rbnenergy.com/The-Decline-and-Fall-of-Conway-Ethane-Part-II>.
- 12 See BentekEnergy.
- 13 The NGL content reaches 1,350 Btu. See, <http://www.pikepa.org/Planning/Gas/Scott%20Blauvelt%20Marcellus%20Shale%20Natural%20Gas.pdf>.
- 14 The calculations in the text correspond to the conventions used by BentekEnergy. There are many different conventions used for the calculation of fractionation spread. In Canada, for example, frac spread is calculated typically for the C3+ basket (gas liquids less ethane). A good review of economics of gas processing plants can be found at <http://www.rbnenergy.com/>.
- 15 More information is available from EIA (see EIA, 2006, "Natural gas processing: The crucial link between natural gas production and its transportation to market," Office of Oil and Gas, January ([http://www.eia.doe.gov/pub/oil\\_gas/natural\\_gas/feature\\_articles/2006/ngprocess/ngprocess.pdf](http://www.eia.doe.gov/pub/oil_gas/natural_gas/feature_articles/2006/ngprocess/ngprocess.pdf)); EIA, "Natural gas processing plants in the United States: 2010 Update" ([http://www.eia.gov/pub/oil\\_gas/natural\\_gas/feature\\_articles/2010/ngpps\\_2009/overview.cfm](http://www.eia.gov/pub/oil_gas/natural_gas/feature_articles/2010/ngpps_2009/overview.cfm))).
- 16 For example, some of LNG arriving at Cove Point (MD) in the past was injected after regassification directly into the local distribution network. The high liquid content of gas was

- causing damage to rubber seals at customers' installations, with obvious consequences.
- Varying natural gas quality may also cause serious damage to turbines in power plants.
- 17 David Wood and Saeid Mokhatab, 2007, "Natural gas interchangeability in focus as source of LNG widen," *LNG Journal*, February, pp 14–18.
- 18 Scf is based on 60° Fahrenheit, one atmosphere pressure at sea level.
- 19 A person inclined to make your life difficult would point out that MM in the Roman mathematical notation stands for 2,000, but nobody will have their MBA revoked for forgetting this.
- 20 [http://en.wikipedia.org/wiki/British\\_thermal\\_unit](http://en.wikipedia.org/wiki/British_thermal_unit).
- 21 Power and energy are often confused. Power is the rate at which work is performed or energy is consumed. Watt is a unit of power (one joule per second).
- 22 <http://www.ukenergy.co.uk/pages/gas-kwh.html>.
- 23 "How natural gas is measured" (see <http://www.tulsagastech.com/measure.html>); also *NIST Handbook 44* Appendix D (<http://ts.nist.gov/WeightsAndMeasures/Publications/H44-09.cfm>). The standards are defined by the National Conference of Weights & Measurements (NCWM).
- 24 <http://www.natgas.info/html/natgasunitsconversion.html>.
- 25 The shale gas revolution and low prices of natural gas are likely to bring some industrial production back to the US, after it had been lost to off-shoring.
- 26 Repressuring is the production technique consisting in injecting natural gas or another inert gas back into the oil well in order to increase the rate of flow of oil by decreasing its viscosity and density.
- 27 Dry natural gas is effectively methane, or what remains from wellhead gas after the liquefiable hydrocarbons and any significant amounts of non-hydrocarbon gases have been removed.
- 28 BP, "Statistical review of world energy 2011."
- 29 EIA-914 Monthly Gas Production Report Methodology Current as of April 2010 ([http://www.eia.doe.gov/oil\\_gas/natural\\_gas/data\\_publications/eia914/eia914meth.pdf](http://www.eia.doe.gov/oil_gas/natural_gas/data_publications/eia914/eia914meth.pdf)).
- 30 Form EIA-914, Monthly Natural Gas Production Report Instructions.
- 31 HPDI is a software and information services company for the energy industry owned by DrillingInfo. HPDI provides well-level information based on the data reported to states.
- 32 J. J. Arps, 1945, "Analysis of decline curves," *Transactions of AIME*, **160**, pp 228–47; see also Kewen Li and Roland N. Horne, 2003, "A decline curve analysis model based on fluid flow mechanisms," SPE 83470," for an explanation of how the Arps' curves are related to the solution of a fluid flow equation. This section is based mostly on Mikael Höök, 2009, "Depletion and decline curve analysis in crude oil production," Uppsala University, May. See also [www.petroprojects.com](http://www.petroprojects.com).
- 33 There is an important distinction between decline rates and depletion rates, which is often ignored and leads to embarrassing mistakes. As explained by Euan Mearns: "The term decline rate refers to the percentage annual reduction in the rate of production (in barrels/day) from an individual field or a group of fields. [...] The term depletion rate refers to the percentage of recoverable resources (in barrels) in a field or region that are being produced each year." The definition of recoverable resources may vary from source to source. See, <http://www.theoildrum.com/node/9327#more>.
- 34 In the case of production, the rate of growth is negative. It is, therefore, natural to use the term *decline rate*.
- 35 Arthur E. Berman and Lynn Pittinger, 2011, "U.S. shale gas: Less abundance, higher cost" (<http://www.theoildrum.com/node/8212#more>).
- 36 This is referred to in economics as "learning by doing," productivity improvements accomplished through repetitive execution of the same tasks.
- 37 See, for example, the investor presentations by Southwestern Energy.
- 38 Pipeline constraints may force well shutdowns in areas of rapidly growing production with midstream infrastructure lagging behind.

- 39** M. A. Mian, 2002, *Project Economics and Decision Analysis* (Vol 1) (Tulsa, OK: PennWell Corp).
- 40** The units of  $B_O$  are defined as (RB/STB): reservoir barrels per stock tank barrel and account for hydrocarbon expansion when moving from high pressure/temperature conditions to the normal conditions prevailing at the surface (see <http://pubs.usgs.gov/of/1998/ofr-98-0034/PA.pdf>). Stock tank barrel refers to oil after the dissolved gas was removed or escaped.

## *Non-conventional Natural Gas*

A few short years ago, the US energy industry was criticised for making an expensive and dangerous bet on an energy source likely to dry up in the near future. “The coming shortage of natural gas in the United States and Canada, compounded by global oil peak and decline, will try the energy and economic systems of both countries to their limits. It will plunge first the US, then Canada, into a carbon chasm, a hydrocarbon hole, from which they will be hard put to emerge unscathed.”<sup>1</sup> The US was seen as a country with a growing dependence on LNG imports required to fill a growing wedge between domestic production and demand. However, this has all changed with amazing speed.

One of the messages we try to convey to students is that they are witnessing the transformation of the energy industry, and especially the oil and natural gas business, through a number of technological innovations compressed into a relatively short period of time. These innovations include new seismic research techniques, the use of powerful computers to optimise the operations of physical assets, the growth of the financial markets and the development of financial instruments enhancing industry’s access to the capital markets and ability to control risks. Unquestionably, the most important innovation is technology that allows accessing reserves of oil and natural gas trapped in the shale rock formations, which were considered to be outside our reach not so long ago. The new technology involves a combination of techniques, such as horizontal drilling and hydraulic fracturing, that have been known for a long time but have become enhanced and reached a critical level of efficiency in around 2007. This was proof of the ingenuity of US engineers working for the energy industry and of the risk-taking ability of management. At the same time, it has become the focus of multiple controversies, ranging

from its environmental impact to its profitability. Any professional in the energy industry should be aware of the public dispute surrounding shale natural gas and oil: the problems do not go away when they are ignored. This book will not offer final answers to the many questions hanging over the industry but, hopefully, will help to summarise the arguments of both supporters and sceptics.

The unquestionable success of the shale revolution from a technological point of view does not mean that any investment in this area will be automatically successful. After all, the revolutions are known to devour, like Saturn, their own children,<sup>2</sup> as any beheaded Jacobin would confirm if they still had the chance to speak. Riding a wave of new transformational innovation creates a few winners but also many failures, as shown by the history of some railways and innumerable dot com companies. Hopefully, this chapter will point readers to a number of useful sources, helping to take advantage of this great historical opportunity and to avoid potential perils.

## **CLASSIFICATION OF NON-CONVENTIONAL NATURAL GAS DEPOSITS**

The definitions of non-conventional natural gas change over time in the same way as the definitions of a non-conventional lifestyle. What was non-conventional a few years ago may be routine practice today, as in the case of our mores. Non-conventional sources of natural gas are defined by contrasting them against the reserves that were historically first to be accessed, due to location, production cost, proximity to the markets, environmental impact and other considerations. As the producing fields in the US and other locations matured, the industry moved to the reserves representing more technological challenges. The progress in modelling underground oil and gas reservoirs and in drilling techniques allowed the access of reserves that would have been either uneconomical to exploit or technically beyond the reach of the industry in the past.

A widely used definition of conventional crude oil and natural gas is available from the EIA.<sup>3</sup> It is “crude oil and natural gas that is produced by a well drilled into a geologic formation in which the reservoir and fluid characteristics permit the oil and natural gas to readily flow to the wellbore.” By logical completion, all other types of crude and natural gas can be classified as unconventional. According to this definition, the reasons for low or non-existent flow rates may

vary. The openings in the rocks may be too small for the molecules to push through and migrate to the wellbore or crude oil may have high viscosity preventing it from flowing, and making transportation and processing difficult.

There are three types of non-conventional natural gas according to industry definitions:

- tight gas;
- coal-bed methane; and
- shale gas.

Tight gas refers to natural gas trapped in rocks characterised by low levels of permeability and porosity. Natural gas trapped in such rocks (called tight rocks) cannot easily migrate to the wellhead and flows at a very low rate, if at all. The most important characteristics that distinguish between traditional natural gas reservoirs and tight sands are porosity and permeability. Porosity determines how much gas and oil can be stored in a rock formation; permeability defines how easily the hydrocarbons can flow towards the wall of the well. Total porosity is percentage of the void space divided by the volume of the rock formation.<sup>4</sup> Effective permeability depends on the extent to which the void spaces in a rock are connected. Porosity between 5–10% is characterised as poor; porosity above 20% is considered to be excellent. The permeability of a formation is described by the Darcy law of fluid dynamics:<sup>5</sup>

$$Q = 0.001127 \times \kappa \times \frac{A}{\mu} \times \frac{\partial P}{\partial L} \quad (9.1)$$

where  $Q$  stands for the fluid rate (barrels per day)

- |                           |   |
|---------------------------|---|
| $\kappa$                  | – permeability  |
| $A$                       | – cross-sectional area of flow (square feet)          |
| $\mu$                     | – viscosity of the fluid                              |
| $\partial P / \partial L$ | – pressure gradient (pounds per square inch per feet) |

Permeability is measured in milidarcies,<sup>6</sup> viscosity in centipoises.<sup>7</sup> The number 0.001127 is a conversion factor. Permeability in the range of 100–1,000 is considered excellent, 10–100 is considered good. It is important to recognise that high porosity does not guarantee high permeability, and vice versa.<sup>8</sup> Both porosity and

permeability may vary significantly over a relatively small area. Potential investors should remember that drilling close to a very productive well is no guarantee of success.

The industry developed several techniques to enhance flows from such formations, primarily through rock fracturing and acidisation. The most promising tight gas plays in the US are:<sup>9</sup>

- the Bossier/Cotton Valley/Vernon fields in East Texas and Northern Louisiana;
- in the Rocky Mountain region, the Green River Basin, Piceance Basin and the Unita Basin;
- the West Texas Canyon Sands; and
- the Clinton–Medina development in Ohio.

Coalbed methane is found either in coal seams or in the surrounding rocks. In many mining operations, methane is a safety threat as it seeps into the underground mines and may cause serious explosions. Exploitation of coalbed methane is highly controversial, given its ecological impact, related to the high quantities of contaminated water being delivered to the surface as a byproduct.

## SHALE NATURAL GAS

The recent revolution in the natural gas industry in the US is associated with the development of techniques for the extraction of natural gas from shale formations. Devonian shales are a very fine-grained sedimentary rock, forming parallel layers that trap natural gas in-between. Here is a more technical definition of shale, from Schlumberger:<sup>10</sup>

A fine-grained, fissile, detrital<sup>11</sup> sedimentary rock formed by consolidation of clay- and silt-sized particles into thin, relatively impermeable layers. It is the most abundant sedimentary rock. Shale can include relatively large amounts of organic material compared with other rock types and thus has potential to become a rich hydrocarbon source rock, even though typical shale contains just 1% organic matter. Its typical fine grain size and lack of permeability, a consequence of the alignment of its platy or flaky grains, allow shale to form a good cap rock for hydrocarbon traps.

The difference between shale and tight natural gas is that shale formations serve as both source and reservoir rocks. In the case of other formations, natural gas migrates from the source rock to the

reservoir rock, where it is trapped by a layer of impermeable rock, which prevents gas from escaping to the atmosphere. Shale formations do not need such a cap: natural gas does not flow anyway. The most concise description of the difference between shale and conventional natural gas has been offered by Jan de Jager:

Generally speaking you could say that unconventionalals are easy to find but hard to produce. When it comes to conventionals, it's the other way round: difficult to find but easy to produce.<sup>12</sup>

The formation of shale natural gas required first the burying of rocks rich in organic matter deep enough to produce sufficient heat and pressure for chemical processes to take place, followed by the upward migration of such rocks. According to de Jager:

It's precisely this geological sequence of deep burial followed by upward pressure that's occurred in large parts of the United States. The substrata contains shales that are millions of years old, stretching across the territories of several states, such as the Marcellus Shale in the eastern US. Around 10 million years ago these layers were pushed upwards once more and they now lie some 1.5 kilometres below the earth's surface, an ideal depth for gas production. Here in the Netherlands it remains to be seen whether the shales with sufficient organic matter have been buried deep enough or been hot enough in the past to actually result in gas generation. What's more, even if they are, we don't yet know if they can be developed economically.<sup>13</sup>

If Jan de Jager's concerns apply to formations outside the Netherlands, this would be bad news for European countries hoping to reduce their dependence on imports of natural gas, often from potentially unstable or unreliable suppliers. Developments in Poland in 2012 related to shale natural gas seem to corroborate de Jager's opinion.<sup>14</sup> This is also another good reason to appreciate the good fortune of the US.

Another important property of shale formations is that they are highly anisotropic – ie, their properties vary considerably across space, even if one takes small steps from a given location. The word anisotropic should be written in flaming letters on the wall of any industry executive or analyst. The practical consequences are that two wells drilled on adjacent lots may have very different levels of productivity and, in a given play, one should concentrate exploration and production effort in the best locations.<sup>15</sup> When multiple wells are drilled from the same pad, identifying the optimal

directions in order to access the “sweet spots” becomes quite important.

A critical difference between conventional and shale natural gas wells is a high initial decline rate of output, reaching in some cases 85–90%. Decline curves covered in the previous chapter are analytical tools used to model volumetric flows from a well or a portfolio of wells. High initial decline rates can be captured through proper calibration of the parameters of these curves. This property of flows from shale wells has important implications for the dynamics of natural gas prices and for the capital expenditures of the industry. We shall revisit this topic shortly.

The extraction of natural gas from shales is quite expensive, and ultimate recovery levels are quite low (around 10%). Most of the natural gas containing Devonian shale in the US is located around the Appalachian Basin. The Barnett Shale system is a giant gas field in Texas (primarily in Wise, Denton and Currant counties), which was the first shale play to be developed by the industry. The average gas in place in the Barnett Shale is 160 Bcf of gas per square mile. The development of this field explains the significant shifts of natural gas production in Texas.

The shales are one of the most promising formations for the future exploration and production of natural gas. The most important US basins are listed in Table 9.1. This list is extracted from a report commissioned by EIA from INTEK, which produced an assessment of onshore Lower 48 states technically recoverable shale gas and shale oil resources. Table 9.2 contains assessments of shale potential worldwide.<sup>16</sup> It should be kept in mind that similar estimates are arriving at a fast and furious pace from different sources, and that this is a work in progress – as technology and market prices evolve, it changes the definition of what can be produced. The reader should use all these estimates with caution, as assumptions and methods vary from study to study. Downward revisions of the Polish shale resources and similar revisions in the US are a reason to proceed with caution. The study summarised in Table 9.2 contains projections for Poland that were eventually reduced 10 times a few months after they were produced. Even at this reduced level, Poland has still a lot of natural gas, although it is not clear if it can be profitably produced.

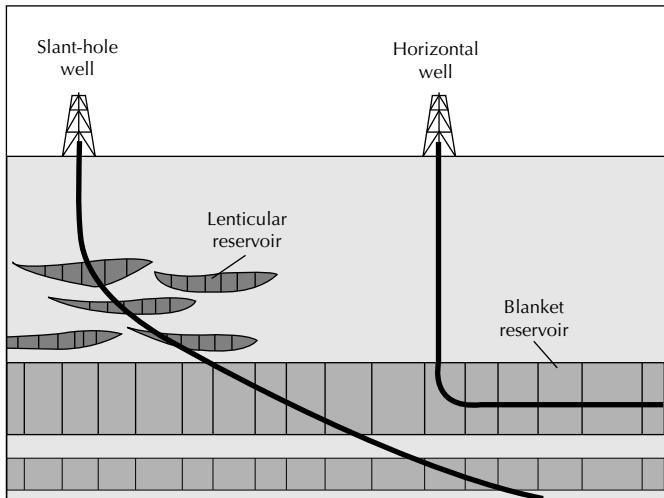
Production of natural gas from shale formations has been made possible through improvements in drilling technology, especially

horizontal and slant drilling, that allow one well to access pockets of natural gas and oil that are located at a distance from the well, as opposed to being located directly under the drilling rig (see Figure 9.1).

Even in the late 1990s, the shale formations were considered by the industry to be outside the scope of economically viable exploitation. Many industry analysts came to the conclusion that the US made a potentially expensive and erroneous bet on natural gas as a fuel used for electricity generation, given the expectation of limited domestic and Canadian supplies, and the potential difficulty of attracting LNG flows into our markets.<sup>17</sup> A decade later, shale natural was changing the dynamics of the US market, creating the challenge of finding additional uses for this resource, such as new modes of transportation (compressed natural gas cars), new technologies (gas-to-liquids)<sup>18</sup> or even exports of natural gas to markets outside North America (through the construction of LNG liquefaction plants). What are the main factors explaining this dramatic change to market conditions?

The first factor is the ubiquitous presence of shale formations in the US; the second, very rapid technological progress that made exploitation of these resources economically viable.

**Figure 9.1** Horizontal vs. vertical drilling



Source: United States Geological Survey (USGS)

**Table 9.1** Natural gas shale basins in the US

Play	Technically Recoverable Resource Gas (Tcf)	Area (sq. miles)		Average EUR Gas (Bcf/well)
		Leased	Unleased	
Marcellus	410.34	10,622	84,271	1.18
Big Sandy	7.4	8,675	1,994	0.33
Low Thermal Maturity	13.53	45,844		0.30
Greater Siltstone	8.46	22,914		0.19
New Albany	10.95	1,600	41,900	1.10
Antrim	19.93	12,000		0.28
Cincinnati Arch	1.44	–		0.12
<b>Total Northeast</b>	<b>472.05</b>	<b>101,655</b>	<b>128,272</b>	<b>0.74</b>
Haynesville	74.71	3,574	5,426	3.57
Eagle Ford	20.81	1,090		5.00
Floyd–Neal & Conasauga	4.37	2,429		0.90
<b>Total Gulf Coast</b>	<b>99.99</b>	<b>7,093</b>	<b>5,426</b>	<b>2.99</b>
Fayetteville	31.96	9,000		2.07
Woodford	22.21	4,700		2.98
Cana Woodford	5.72	688		5.20
<b>Total Mid-Continent</b>	<b>59.88</b>	<b>14,388</b>		<b>2.45</b>
Barnett	43.38	4,075	2,383	1.42
Barnett Woodford	32.15	2,691		3.07
<b>Total Southwest</b>	<b>75.52</b>	<b>6,766</b>	<b>2,383</b>	<b>1.85</b>
Hilliard–Baxter–Mancos	3.77	16,416		0.18
Lewis	11.63	7,506		1.30
Williston–Shallow Niobraran	6.61	–		0.45
Mancos	21.02	6,589		1.00
<b>Total Rocky Mountains</b>	<b>43.03</b>	<b>30,511</b>		<b>0.69</b>
<b>Total Lower 48 US</b>	<b>750.38</b>	<b>160,413</b>	<b>136,081</b>	<b>1.02</b>

Source: U.S. Energy Information Administration (July 2008)

As we said, the Barnett Shale Basin was the first to be explored to a significant extent, contributing to a significant jump in the production of natural gas in Texas.<sup>19</sup> Development of Barnett was followed by hectic exploration and development activity in other locations, primarily in Haynesville, Fayetteville, and Marcellus. Another area of very hectic drilling activity at this time (early 2012) was the Eagle Ford area of Texas. Shale formations align well with the existing pipeline grid in the US and are often located close to big consuming centres (for example, the Marcellus Basin in the North East).

**Table 9.2** Estimated shale gas technically recoverable resources for select basins in 32 countries, compared to existing reported reserves, production and consumption during 2009

	2009 Natural Gas Market (trillion cubic feet, dry gas basis)				
	Production	Consumption	Imports (Exports)	Proved Natural Gas Reserves (trillion cubic feet)	Technically Recoverable Shale Gas Resources (trillion cubic feet)
<b>Europe</b>					
France	0.03	1.73	98%	0.2	180
Germany	0.51	3.27	84%	6.2	8
Netherlands	2.79	1.72	-62%	49.0	17
Norway	3.65	0.16	-2156%	72.0	83
U.K.	2.09	3.11	33%	9.0	20
Denmark	0.30	0.16	-91%	2.1	23
Sweden	—	0.04	100%		41
Poland	0.21	0.58	64%	5.8	187
Turkey	0.03	1.24	98%	0.2	15
Ukraine	0.72	1.56	54%	39.0	42
Lithuania	—	0.1	100%		4
Others	0.48	0.95	50%	2.7	19
<b>North America</b>					
United States	20.60	22.8	10%	272.5	862
Canada	5.63	3.01	-87%	62.0	388
Mexico	1.77	2.15	18%	12.0	681
<b>Asia</b>					
China	2.93	3.08	5%	107.0	1,275
India	1.43	1.87	24%	37.9	63
Pakistan	1.36	1.36	—	29.7	51
<b>Australia</b>					
Australia	1.67	1.09	-52%	110.0	396
<b>Africa</b>					
South Africa	0.07	0.19	63%	—	485
Libya	0.56	0.21	-165%	54.7	290
Tunisia	0.13	0.17	26%	2.3	18
Algeria	2.88	1.02	-183%	159.0	231
Morocco	0.00	0.02	90%	0.0	11
Western Sahara	—	—	—	—	7
Mauritania	—	—	—	1.0	0
<b>South America</b>					
Venezuela	0.65	0.71	9%	178.9	11
Colombia	0.37	0.31	-21%	4.0	19
Argentina	1.46	1.52	4%	13.4	774
Brazil	0.36	0.66	45%	12.9	226
Chile	0.05	0.1	52%	3.5	64
Uruguay	—	0	100%		21
Paraguay	—	—	—	—	62
Bolivia	0.45	0.1	-346%	26.5	48
<b>Total of Above Areas</b>	<b>53.10</b>	<b>55</b>	<b>-3%</b>	<b>1,274.0</b>	<b>6,622</b>
<b>Total World</b>	<b>106.50</b>	<b>106.7</b>	<b>0%</b>	<b>6,609.0</b>	

Source: U.S. Energy Information Administration (April 2011)

### Technological breakthrough

The technological progress that made shale formation viable is related to a combination of horizontal drilling and the hydraulic fracturing of rock formations. Both techniques have been known and used for a long time (horizontal drilling since the 1930s and hydraulic fracturing since the 1950s), but have been greatly improved and refined for the shale plays. The first horizontal well was drilled near Texon, Texas, in 1929.<sup>20</sup> Hydraulic fracturing began in 1947 in the Hugoton gas field in Grant County, Kansas, with a small-scale liquid injection to open a well clogged by mud. The first patent was issued in 1949 to the Halliburton Oil Well Cementing Company, which then performed the first two commercial fracturing treatments in Oklahoma and Texas.<sup>21</sup>

Horizontal drilling is a two-step process of drilling a vertical well first and then expanding from a certain depth in possibly multiple lateral directions. The main advantage of horizontal drilling is that it produces a bigger exposure to rock formations. A horizontally drilled well can access natural gas located under urban areas (for example, by drilling a vertical well in a public space, such as a park or an airport, and then expanding horizontal wells under residential areas). The critical aspect of horizontal drilling, however, is the presence of natural or man-made vertical fractures in rock formations that provide a conduit through which natural gas may flow into the wellbore. A horizontal well will cross more such fractures than a vertical well through a given layer of shale (see Figure 9.1). Natural vertical fractures are, for example, a characteristic of the Marcellus Shale.

Another advantage of horizontal drilling is the reduced footprint for a given output of natural gas. For example, the development of one square mile field could require sixteen vertical wells located on separate well pads. But the same field could now be accessed with six to eight, and even as little as four, horizontal wells. This translates not only into a lower immediate footprint for the pad, but also a smaller area of access roads, pipelines, processing equipment and a reduced impact on wildlife habitat and risk of adverse ecological impact.<sup>22</sup>

Hydraulic fracturing consists in creating artificial fractures in rock formations or expanding the naturally occurring ones.<sup>23</sup> It starts with the firing of explosive charges into rocks, followed by the injection

into the well of large volumes of liquids which consist primarily of water with a proppant (finely ground sand or ceramic particles) and about 2% by volume of chemical additives. The pressurised liquids create fractures in the rock formations that are prevented from closing under the weight of the layers of rocks above the well once the liquids are removed. This is accomplished by pushing the propellant into the openings created in the rock by the fracturing liquids. Hydraulic fracturing is carried out segment by segment of the well (these segments, which are temporarily isolated from the rest of the well, are called stages).

This process is fairly complicated and may include up to eighteen steps,<sup>24</sup> starting with acid treatment of the drilled well hole in order to clean fractures in the rock that may be clogged with debris and mud from the drilling operation. In the next step, the hole is filled with the slickwater, a mixture of water and polymers that are added in order to reduce friction.<sup>25</sup> In the next steps, other polymers are used, designed to keep the particles of proppant suspended in the liquid and preventing them from settling out at the bottom of the well.

### **Shale gas resources**

US natural gas shale resources are very abundant and are likely to cover many years of future consumption.<sup>26</sup> One of the most widely quoted sources over the last few years are the studies of the Potential Gas Committee (PGC). However, the PGC combines proved recoverable reserves estimated by the EIA (and established by drilling) with the estimates of “potential resources” to derive “future gas supply.” As explained on the Committee website:

The Potential Gas Committee reports its assessments of potential resources in three categories of decreasing geological certainty:

1. Probable resources (discovered but unconfirmed resources associated with known fields and field extensions; undiscovered resources in new pools in both productive and nonproductive areas);
2. Possible resources (undiscovered resources associated with new field/pool discoveries in known productive formations in productive areas); and
3. Speculative resources (undiscovered resources associated with new field/pool discoveries in nonproductive areas).

For each category, a “minimum,” “most likely” and “maximum” volume is assessed (with no implied numerical quartiles or other

quantification of certainty, only the assessor's best judgment). The three respective "most likely" values are summed arithmetically to yield the total "most likely" resource at the province, area and national levels.<sup>27</sup>

The report, released on April 27, 2011 (for year-end of 2010), put the total US resource base at 1,898 tcf.<sup>28</sup>

One way of putting natural gas resources available to the US in perspective is to start with the PGC resource estimate (1,898 tcf) and assume that the US consumption of natural gas will grow at 3% per annum (and assuming no imports for simplicity). In this case, US domestic supplies will last for about forty years. Even using a very optimistic PGC assessment of available gas in place, the statement of about one hundred years of US supplies – repeated over and over in the media – seems unreasonable. As physics professor Albert Allen Bartlett stated once "The greatest shortcoming of the human race is our inability to understand the exponential function." Having had to check a countless number of home works that required using exponential function, we have to agree.

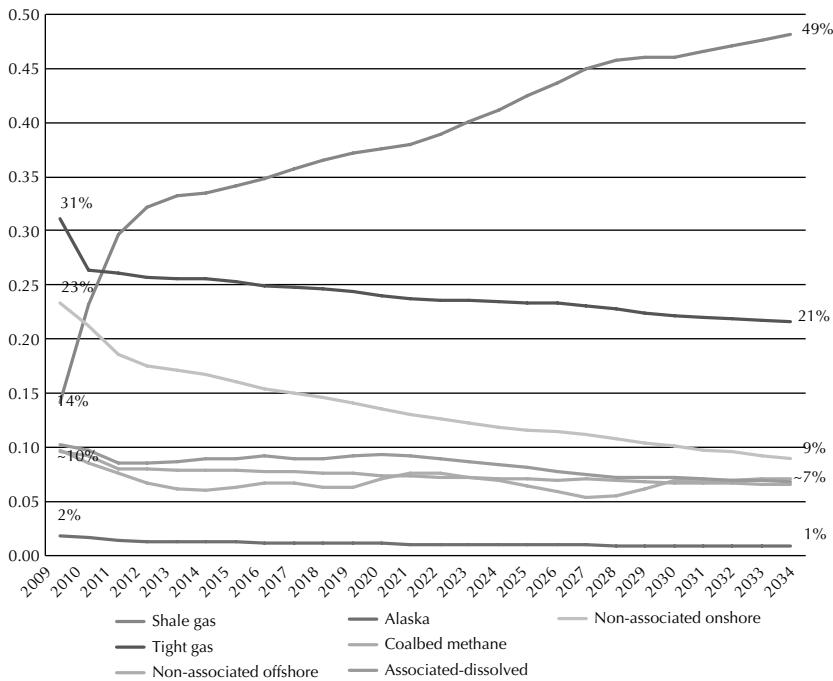
As explained, the EIA data is not directly comparable to PCG numbers, and is much less speculative. The history of shale gas projections, irrespective of the source, is characterised by significant swings as more information becomes available. As always, any technological revolution goes through a stage of euphoria, until cooler heads prevail.

The US EIA has increased dramatically the estimates of shale natural gas resources in its "2011 Annual Energy Outlook":

The technically recoverable unproved shale gas resource is 827 trillion cubic feet (as of January 1, 2009) in the AEO2011 Reference case, 480 trillion cubic feet larger than in the *Annual Energy Outlook 2010* (AEO2010) Reference case, reflecting additional information that has become available with more drilling activity in new and existing shale plays. The larger resource leads to about double the shale gas production and over 20 percent higher total lower 48 natural gas production in 2035, with lower natural gas prices, than was projected in the AEO2010 Reference case.<sup>29</sup>

The EIA Annual Energy Outlook 2012 Early Release Overview contained a much lower overall assessment of shale gas potential, which was quite a disappointment to the industry.

In the AEO2012 Reference case, the estimated unproved technically recoverable resource (TRR) of shale gas for the United States is 482

**Figure 9.2** US dry natural gas production (EIA projections)

Source: U.S. Energy Information Administration (2012)

trillion cubic feet, substantially below the estimate of 827 trillion cubic feet in AEO2011. The decline largely reflects a decrease in the estimate for the Marcellus shale, from 410 trillion cubic feet to 141 trillion cubic feet. [...] Drilling in the Marcellus accelerated rapidly in 2010 and 2011, so that there is far more information available today than a year ago. Indeed, the daily rate of Marcellus production doubled during 2011 alone.

This downward revision does not change the overall EIA forecast of growing US reliance on non-conventional natural gas (with shale gas accounting for 49% of dry gas production in 2035), as shown in Figure 9.2.

One should recognise that the estimates of shale gas resources and reserves (the two should not be conflated)<sup>30</sup> will change as more production and exploratory wells are drilled and as technology progresses, increasing overall recovery factors.<sup>31</sup> As explained in a paper by Leonardo Maugeri,

the huge differences in permeability, porosity, and thickness of a shale/tight oil formation require many more exploration wells be drilled in different areas of the field before making it possible to have an idea of the effective recoverability rate from the whole formation. The rapid output increase and decline of shale/tight oil producing wells further complicates matters, which makes shale/tight oil operations a ‘drilling-intensive’ activity.[...] For these reasons, it is impossible to make any reasonable evaluation of the future production from a shale/tight oil formation based on the analysis of a few wells data and such limited activity.<sup>32</sup>

In view of the above, it is not surprising that the estimates vary and are sometimes quite controversial. The industry tends to be very optimistic, with a few sceptics recognising the overall potential of shale gas but expressing a view that the ultimate recovery will fall short of excessively optimistic predictions.<sup>33</sup>

### **SHALE GAS CONTROVERSY**

The natural gas shale revolution is not free from controversy, and one can expect that debate to intensify and become even more acrimonious in the future.<sup>34</sup> There are three main types of objections raised by the critics of hydraulic fracturing. One is related to the environmental impact of hydraulic fracturing, the second to the estimates of reserves and costs of production of shale natural gas. We shall review these issues in detail below. The third group of arguments is related to the general claim that natural gas is a bridge fuel, characterised by lower (as compared to coal) emissions of CO<sub>2</sub> – which will buy time as the world makes the transition to a less carbon-intensive future. This last topic is addressed in detail later in the chapter.

#### **Environmental impact**

The concerns related to the process of hydraulic fracturing revolve around the quantity of water that is used per well and the potential environmental impact on water supplies. It is estimated that a single well may require as much as three to five million gallons of water (although some wells may require up to eight million gallons). In the parts of the US where water supplies are limited, this is not a trivial amount. This concern may be addressed by recycling water by piping or trucking it from one drilling site to another. The concerns related to potential pollution are much more serious. Fracturing liquids contain many chemicals, in addition to polymers, and their

composition is treated as a commercial secret (despite many lists circulating on the Internet). The liquids removed from the well after the fracturing process is finished are stored in tanks or retention ponds that may spill or leak, and then seep into ground water or lakes and rivers. In addition to man-made chemicals, the liquids may become contaminated inside the well with certain contaminants – including heavy metals or naturally occurring radioactive materials (NORMs), such as radon. A bigger concern is related to the risk that the wall of a well hole may be perforated and the fracturing liquids may contaminate the aquifers.

The industry has come under growing pressure to disclose the composition of the drilling liquids. Several states introduced regulations requiring mandatory reporting of the chemicals contained in the fluids.<sup>35</sup> Some companies (Range Resources, Chevron Corp, BP) started voluntary disclosure using the website FracFocus.org, a joint project of the Ground Water Protection Council and the Interstate Oil and Gas Compact Commission.<sup>36</sup> The measures taken by several states were generally supported by the industry. In our view, the disclosure was long overdue, given how emotional this issue has become and the relative ease with which it is possible to obtain information about the chemicals in the fracking liquids, as long as one can show some ingenuity.<sup>37</sup>

The water-related issues are complicated, as is the case in other energy industry activities, by overlapping state and federal jurisdictions. Several different acts passed by US Congress apply to the unconventional natural gas productions, including:

- the Clean Water Act (CWA) – regulates surface discharges of water;
- the Safe Drinking Water Act (SDWA) – regulates the discharge of water that may affect underground sources of drinking water; and
- the Clean Air Act (CAA) – limits emissions from stationary sources such as compressor engines and sets ambient air quality standards.

In addition, the National Environmental Policy Act (NEPA) mandates the assessment of environmental impacts before exploration and production permits are granted on federal lands.

A very controversial amendment of the SDWA passed by

Congress in 2005 exempted certain oil- and gas-related activities, including hydraulic fracturing (except with diesel), and deferred them to the control of individual states.<sup>38</sup> Oil and gas facilities were also exempted from the requirements related to reporting fracturing fluids composition to the Toxics Release Inventory under the Emergency Planning and Community Right to Know Act.<sup>39</sup>

The growing scale of production of natural gas from shale formations led to growing grass-roots opposition to hydraulic fracturing, which had a significant impact on the political process. There are several new laws under consideration in the US Congress, including the Fracturing Responsibility and Awareness of Chemicals Act of 2009, to repeal the exemption for hydraulic fracturing in the SDWA. It is not obvious if this act will be passed by Congress following the 2010 elections.<sup>40</sup>

Preliminary results of a study carried out at the University of Texas, released in 2011, found no direct link between hydraulic fracturing and ground water contamination.<sup>41</sup> The study does not exclude the potential for contamination due to mishandling of flow-back water stored in containment ponds or due to poor casing or cement jobs. The issue is what technology can be used to reduce the probability and severity of potential incidents. Hopefully, the full study will help to make the debate less emotional and anchor it in scientific and engineering fact.

Unfortunately, it may be a long wait. An independent review of the preparation of the report cited above, published at the end of 2012, led to the retirement of the chief researcher, and the resignation of the director, of the Energy Institute at the University of Texas. For more information see, Andrew Revkin, "Damning Review of Gas Study Prompts a Shakeup at the University of Texas," The New York Times, December 6, 2012."

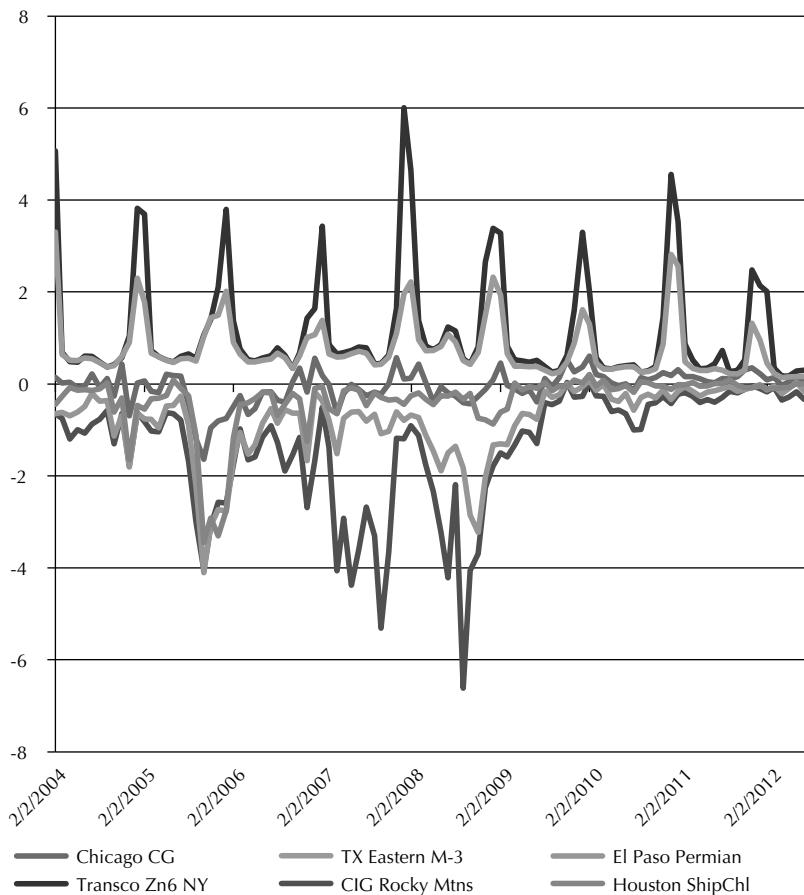
### **Prices and production costs**

At the time of writing, conditions in the US natural gas markets can be described using two terms: excess supply and depressed prices. The impact on the industry goes beyond these two fairly simplistic statements.

*Price dynamics.* An increase in the production of natural gas combined with a depressed economy resulted not only in the fall of

prices but also in the compression of prices across space and the suppression of volatility. Price compression is reflected in shrinking absolute price differentials between different market centres and Henry Hub, the market hub used as a central reference point for US natural gas. The shrinking basis (as such differentials are called) is the result of more ubiquitous production, with many fields located close to the population and to industrial centres (the Marcellus field is a good example) and the construction of new pipelines liberating previously stranded gas. Figure 9.3 provides an illustration of

**Figure 9.3** US natural gas basis evolution over time, selected locations



Source: Platts, © 2012 by Platts, a division of The McGraw-Hill Companies, Inc.

shrinking basis differentials across multiple locations in the US. A general availability of gas has suppressed the volatility that historically attracted traders to this market. The outcome is that a once-vibrant market turned into a Rodney Dangerfield version of the commodity business: it gets a lot of attention but no respect. We can observe many hedge funds specialising in trading natural gas heading for the exits, and some new hedge funds with the same proposed profile failing to find willing investors. We shall revisit this issue in Chapter 11, where we express an opinion that this is a temporary development. Natural gas traders should not think about retirement yet.

*Asset prices.* Low volatility and compressed basis levels spilled over into the pricing of some critical assets and contracts. We can see the depressed prices of natural gas storage facilities and the erosion of the value of long-term transportation contracts. The valuations of storage and pipeline assets depend on basis levels and price volatility, and are ultimately linked to the risk of potential scarcity of natural gas at a given location. One interesting property of natural gas wells in the shale formations is that they may serve as *de facto* operational on-line storage. In many instances, shutting down a shale well does not have the same adverse consequences as the shutdown of a conventional well (one should remember, however, that not all shale rock formations were not created equal). For a conventional well, future volumetric flows may be impaired; in the case of shale wells, flows may even be enhanced. The impact on natural gas pipelines is equally significant. Some pipelines are bursting at the seams, struggling with excessive nominations (see Chapter 10 on pipeline operations) due to surging natural gas output. Some pipelines are running at a fraction of capacity, as increased production is changing spatial patterns of natural gas flows.<sup>42</sup>

*Cashflow and balance-sheet stress.* Many exploration and production companies are highly leveraged and face at the same time high capital expenditures (for reasons to be explained later in this chapter). In the low price environment, the solution is sale of valuable assets or joint ventures with foreign and domestic equity investors.

### The reasons for depressed prices

A person with a sense of history and no financial exposure to the natural gas markets can simply shrug their shoulders and say nihil novi sub sole. The history of commodity markets is replete with events of overproduction and depressed prices. One of the first models a quantitative economist learns is the cobweb model,<sup>43</sup> which explains price fluctuations between the two extremes of high and low prices, through supply curves based on lagged prices and demand curves dependent on current prices. The period of depressed natural gas prices has, however, some interesting unique underlying factors. These factors are well understood in the industry and will be summarised briefly here.

- The industry has hedged aggressively, taking advantage of high levels of forward prices, supported to a large extent by the inflows of investment funds into natural gas markets. As hedges roll off and the forward prices fall, this factor will be less important going forward.
- Many wells produce natural gas rich in NGLs or associated gas (natural gas dissolved in oil). High prices of oil and NGL-related revenues change the economics of such wells, and simple breakeven analysis does not capture this part of the revenue stream. The industry is migrating drilling rigs to locations where natural gas is wet (ie, it is characterised by a high content of natural gas liquids) or is co-produced with crude oil (associated gas). Roughly speaking, this is a belt extending through the US from Eagle Ford in Texas to Bakken in North Dakota and parts of Marcellus. A contribution of NGLs, condensate and crude oil can make the production of natural gas profitable for many companies, as long as valuable byproducts are extracted as well. The irony of this situation is that NGLs, for a long time a byproduct of natural gas, provide a critical subsidy for what used to be the principal commodity. The question remains if the increase in production of natural gas will eventually translate into a significant drop of prices of NGLs. The prices of NGLs have been on a downward trajectory (at the time of writing), with prices being particularly depressed at the Conway Hub (as compared with Mont Belvieu). More ethylene cracker capacity will come eventually online in the US, but it will have to compete for markets with

petrochemical plants developed in the Persian Gulf and elsewhere.

- ❑ Initial exploration costs and lease costs were covered by investors (many foreign) interested in acquiring an exposure to the shale business and obtaining access to hydraulic fracturing technology. An injection of equity into the industry happened through the direct acquisition of reserves or joint ventures.
- ❑ Many companies treat initial exploration and lease acquisition costs as sunk costs (irrespectively of the kindness of strangers, who covered these expenses in some cases) and make current operational decisions based on the marginal cost of production. In business and economics, “bygone is bygone” – historical costs do not matter.<sup>44</sup> One of the consequences is that the industry’s day-to-day operational decisions are driven by what is called in the industry “halfway breakeven cost” (or “half-cycle breakeven cost”). This concept is important and will be covered in detail shortly. A related issue is the determination to continue producing natural gas as long as prices exceed variable costs in order to stay in business, hoping that the prices will eventually go up.
- ❑ Some exploration and production (E&P) companies may be obligated to drill by a number of different contractual arrangements. Some companies promised to contribute volumes to newly constructed pipelines or made commitments under volumetric production payments (see Chapter 11) or under joint venture agreements. In some cases, lease terms require drilling by a certain deadline to pre-empt lease expiration. Some companies drill just one well in such cases to preserve the lease. Such drilling is referred to in the industry as “held by production.”
- ❑ In many instances, lease acquisition happened under conditions reminiscent of the gold rush – with some companies rushing to capture the best spots and overpaying in the process. As mentioned, the prospect of losing a valuable lease forces drilling decisions even before the necessary infrastructure (such as gathering systems and pipeline connections) is developed. The outcome is a huge inventory of wells ready to produce that will eventually come on-line. As reported by BentekEnergy in the spring of 2012:

Northeast Pennsylvania dry gas production can grow 31% over the next 16 months, even with zero rigs operating Even if all rigs

were taken out of the Northern Marcellus and current drilling activity ceased, regional dry gas production can grow 1.3 Bcf/d over the next 16 months. Bentek's Northeast Observer reveals that due to the large non-producing well inventory in the Northern Marcellus, coupled with high IP rates in the region, production in Northeast Pennsylvania can still grow from approximately 4.1 Bcf/d today to 5.4 Bcf/d by September 2013, a 31% increase that results exclusively from working off the existing backlog of 1,000 non-producing wells.<sup>45</sup>

- ❑ Some E&P companies perceive themselves as takeover targets and operate with an objective to maximise reserves. Drilling is the obvious strategy to increase the value of a company, given that valuations are often based on the reserves the acquiring company can obtain through acquisition.
- ❑ Major oil and natural gas companies are increasingly moving into shale formations after having initially left this field to a number of highly entrepreneurial companies. Big companies tend to take a longer view and are less sensitive to current market conditions, given their overall financial strength. Some big integrated companies do not mind the possibility of low-priced natural gas (as long as this is not a permanent condition) because such market conditions create attractive merger and acquisition opportunities.
- ❑ Exploration efforts related to shale natural gas help to boost the reserves of E&P companies. Many producers report reserves in terms of barrels of oil equivalent and, given the difficulty of replacing oil reserves, drilling for gas may be considered the next-best solution. The drive to increase reserves is a powerful motivating force behind many mergers and acquisitions, even if natural gas is not the real McCoy (compared to oil).
- ❑ Exploration of shale natural gas is increasingly similar to an industrial assembly line process, with commitments (for example, leases of drilling rigs) that are difficult to cancel on a short notice.

### **Half-cycle breakeven cost**

This section explains the calculations for the half-cycle breakeven cost of producing natural gas from shale formations.<sup>46</sup> The example is based on the following assumptions:

	Assumption	Units
Initial production	5,000	Mcf/day
Royalty interest	15	% of gross revenue
Operating cost	1.35	US\$/Mcf
Production taxes		
Discount rate	10	%
EUR	3.46	Bcf
Drilling and completion cost	4,000,000	US\$
Annual decline rate year 1	0.7	
Annual decline rate year 2	0.3	
Annual decline rate year 3	0.17	
Annual decline rate year 4	0.15	
Annual decline rate year 5	0.13	
Annual decline rate year 6	0.11	
Annual decline rate year 7	0.09	

We show calculations for a few months in the text that follows. The approach we take is based on calculating the average daily production for the first months of successive years using assumed decline rates and then linear interpolation to calculate monthly production for the interim month. Assuming that the well in this example starts operating in January of year 1 at the daily flow level 5,000 Mcf a day, the production rate for January of year 2, would be calculated as  $5,000 \times (1 - 0.7) = 1500$ . It is assumed that daily production does not change during a month. Daily production for a given month is calculated through linear interpolation between daily production levels for January of the current year and January of the next year. This means that, during the first year, daily production drops by  $(5000 - 1500)/12 = 291.666$  from month to month.

Monthly production is obtained by multiplying daily production by number of days in a month. Multiplication of monthly production by the breakeven price (which remains to be determined) gives gross cashflow for a month. Gross cashflow is reduced by royalties (we assume the rate of 15%), and well operating costs that are assumed to be constant and equal to US\$1.85 per 1,000 cubic feet. For convenience, we ignore taxes (it could be a well in Pennsylvania, where there is no severance tax). Monthly net cashflows are discounted back to the current level using a rate of 10% pa. The Excel Solver is used to find the price of natural gas, which will yield the NPV of future cashflows (after subtraction of the well drilling cost) equal to zero. In this case, the price is equal to US\$3.43 per 1,000 cubic feet.

In the spreadsheet, the well is shut down when we reach the target

**Table 9.3** Half-Cycle Breakeven Cost Calculation

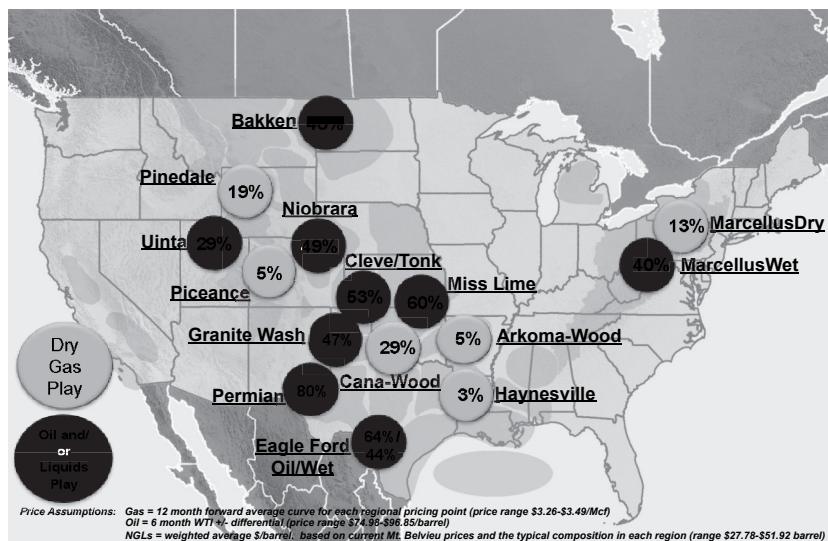
#	Year	Month	Days	Daily production	Production	Cumulative production	Cum prod (Bcf)	Gross revenue (US\$)
1	1	1	31	5000	155000	155000	0.155	531,828
2		2	28	4708.333333	131833.333	286833.3333	0.286833333	452,340
3		3	31	4416.666667	136916.667	423750	0.42375	469,782
4		4	30	4125	123750	547500	0.5475	424,605
5		5	31	3833.333333	118833.333	666333.3333	0.666333333	407,735
6		6	30	3541.666667	106250	772583.3333	0.772583333	364,560
7		7	31	3250	100750	873333.3333	0.873333333	345,688
8		8	31	2958.333333	91708.3333	965041.6667	0.965041667	314,665
9		9	30	2666.666667	80000	1045041.667	1.045041667	274,492
Gross revenue (US\$)				Revenue net of royalties (US\$)	Revenue net of royalties and operating cost (US\$)	Discounted net revenue (US\$)		
531,828				452,054	242,804	240,797		
452,340				384,489	206,514	203,115		
469,782				399,314	214,477	209,203		
424,605				360,914	193,852	187,522		
407,735				346,575	186,150	178,584		
364,560				309,876	166,438	158,354		
345,688				293,835	157,823	148,916		
314,665				267,465	143,659	134,431		
274,492				233,318	125,318	116,299		
252,618				214,726	115,332	106,147		
214,447				182,280	97,905	89,363		
190,572				161,986	87,005	78,758		

estimated ultimate recovery (EUR) of 3.22 Bcf. In practice, a decision will be made sometime in the future to decommission the well when operating expenses exceed net revenues. Such a decision-making process is not modelled in this very simplified example, included in the book to come up with an order of magnitude of a breakeven price. A more realistic example would include additional cashflow streams from selling NGLs and condensate (and possibly crude oil), which may change significantly the financial outcomes for those fortunate enough to produce associated gas or wet gas. The magnitude of such additional flows will vary depending on the well locations. As explained above, the industry is looking to relocate rigs to locations in liquids-rich basins to reduce the impact of low prices of natural gas on the bottom line.<sup>47</sup>

The calculations of half-cycle breakeven costs are controversial, as they do not capture the full cost of production. The omitted items include the initial lease cost, which may be significant for the firms that lacked discipline in the process of negotiating leases or succumbed to irrational exuberance and overpaid the landowners. According to Art Berman,<sup>48</sup> other omitted costs include interest expense and general and administrative (G&A) (overhead), dry hole cost, and plug and abandon (P&A) expense. This is certainly true, but we still believe that half-cycle cost is a useful number. These costs are important because they determine at the margin the behaviour of many natural gas producers. The initial outlays are treated as the sunk costs and do not influence current decisions. This explains why the prices of natural gas may remain depressed for a period of time. Rational behaviour at an individual level often adds up to a sub-optimal outcome at the collective level.

BentekEnergy, a company which pioneered the calculation of break even costs, captures economic results of drilling through annualised internal rates of return for various basins. Figure 9.4 illustrates the differences between dry and wet, as well as, associated gas wells economic outcomes.

But what are the all-inclusive costs of production for shale natural gas formations? According to Art Berman, the full costs are between US\$6 and US\$8/MMBtu for most E&P companies. Figure 9.5 shows realised prices (i.e. prices adjusted for additional cash flows from hedges) for a number of US natural gas producers (not identified by name) vs. estimated 5-year breakeven cost of production.

**Figure 9.4** Internal rate of return by basin

Source: BentekEnergy

A recent study by a group of Schlumberger experts contains a study of decline curves for a number of natural gas producing basins combined with estimates of breakeven prices.<sup>49</sup> According to the authors, “for wells analyzed in core plays area in 2008 and 2009, only

**Table 9.4** Breakeven price for a number of shale basins

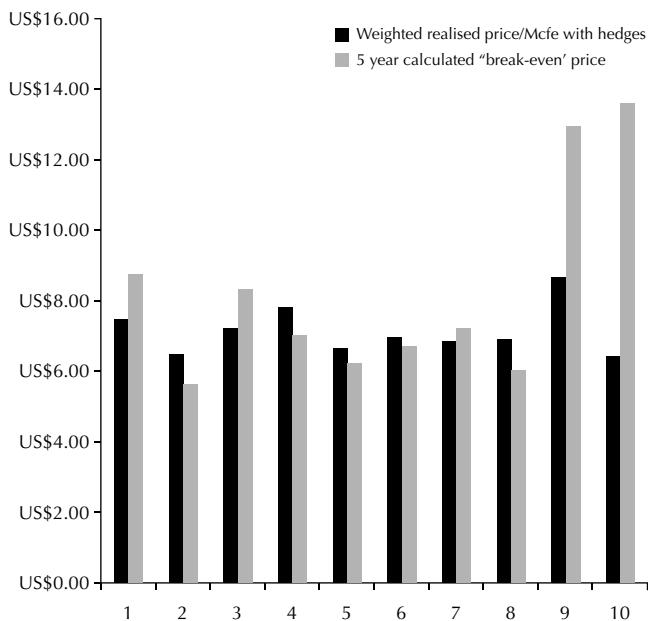
Basin	DOFP	EUR (Bcf)	Breakeven price (US\$/MMBtu)
Barnett	2008	2.895	3.7
Barnett	2009	2.867	3.74
Fayetteville	2008	2.463	3.65
Fayetteville	2009	3.401	3.2
Woodford	2008	2.544	7.35
Woodford	2009	3.389	6.22
Haynesville	2008	4.579	6.95
Haynesville	2009	6.092	6.1
Eagle Ford	2009	3.793	6.24

Source: Jason Baihly, Raphael Altman, Raj Malpani and Fang Luo, 2011, “Study assesses shale decline rates,” *The American Oil and Gas Reporter*, May

Note:

DOPF – Date of First Production

EUR – Estimated Ultimate Recovery

**Figure 9.5** Selected Company 5-Year Imputed Production Costs/Mcfe

Source: Labyrinth Consulting Services

wells in the Barnett and Fayetteville were deemed to be profitable under spot gas prices.” One should point out that natural gas prices were at much higher levels when this article was written. As of June 2012, the Nymex prompt month natural gas prices have hovered around US\$2.50/MMBtu, and cash prices at many locations were under US\$2/MMBtu during the spring of 2012. The breakeven prices reported in the article are shown below in Table 9.4.

All the factors discussed above created toxic pricing conditions for the US natural gas industry. Historically, similar circumstances (an industry or a segment of an industry operating at a loss and fixated on survival by making production decisions to generate cash to cover variable cost) have led to significant price jumps. We would not be surprised if this happened as well in the US. In the meantime, we can expect that many E&P companies active in the shale basins will struggle. The further consolidation of the industry cannot be excluded.

## Industry countermeasures

Low prices of natural gas, persisting for the reasons outlined above triggered a number of countermeasures by E&P companies. One trend is the optimisation of drilling activities and the relentless pressure to improve fracturing technology. Our experience is that the ingenuity of scientists working in the US energy industry should not be bet against. Examples include:

- ❑ drilling multiple wells from the same pad using rigs placed on rails and moving them a just a few yards before a new well is started;
- ❑ extending the lateral length of a well;
- ❑ accessing the most productive sections of a given lease from the same pad, given an improving understanding of the geology of shale resources; and
- ❑ reducing the number of days required to complete a well.

All these techniques resulted in a significant improvement in US drilling rig productivity, a development that was captured and can be monitored through the BentekEnergy rig productivity index. Higher rig productivity explains why many natural gas production forecasts were too conservative. Industry analysts often used the historical time series of rigs in operation to predict production flows, and missed the possibility of higher output combined with a lower number of rigs drilling for natural gas.

One example of the measures taken by E&P companies to counter the effect of low natural gas prices can be found in the presentations of Chesapeake Energy, one of the early pioneers of shale natural gas. The company has announced that it will concentrate on liquid-rich locations and will reduce its drilling programme until prices reach US\$6/MMBtu. This strategy can be summarised as follows:<sup>50</sup>

- ❑ “Reduce drilling of natural gas wells except for those required to HBP<sup>51</sup> leasehold or to use a drilling carry provided by a joint venture partner until such time as natural gas prices rise above US\$6.00 per Mcf;
- ❑ Lease and develop substantial new liquids-rich plays in which the company can acquire very large leasehold positions of 250,000–750,000 net acres;

- ❑ Within one year of acquisition, sell a minority interest in a new play, recovering all or virtually all of the cost to acquire the leasehold in the play and to fund approximately a significant portion of Chesapeake's future drilling costs in the play;
- ❑ Accelerate drilling of liquids-rich plays until year-end 2012 when the company's drilling capital expenditures are balanced approximately 50/50 between natural gas plays and liquids-rich plays;
- ❑ Continue adding proved reserves, net of monetisations and divestitures, of approximately 2.5–3.0 tcf (415–500 million barrels of oil equivalent, mmboe) annually; and
- ❑ Accomplish these goals without the issuance of additional equity and with a reduction of debt levels such that the company becomes investment grade within the next few years."

The costs of production between US\$6 and US\$8 will eventually translate into higher prices of natural gas when the temporary countervailing factors cease to operate.

### **Geopolitical repercussions of shale natural gas**

It remains to be seen whether worldwide natural gas production from the shale formations will follow the US trajectory. One reason why this may not happen could be geology (see the statement by Professor Jan de Jager quoted at the beginning of this chapter). Putting that aside, one has to recognise the US has the advantage of unique institutional and business conditions supporting fast growth of shale gas production that may not be present in other countries.

- ❑ *Mineral rights.* In most countries, mineral rights (below a certain depth) are retained by governments. This may help in obtaining a licence to drill but does not create incentives for landowners to cooperate with E&P companies. In many countries (for example, Poland), the density of population is high and farms tend to be small, complicating the task of construction of access roads, obtaining water rights and minimising community impact.
- ❑ *Business culture.* Most countries with huge shale potential do not have the same business culture, characterised by the presence of innovative, entrepreneurial companies that were at the forefront

of shale developments in the US. It will also take time to train local labour, as many potential shale gas producers have to build this industry from scratch.

- ❑ There is no existing midstream infrastructure (gathering systems, processing plants, pipelines), and this means that significant capital expenditures are required. Potential investors will be reluctant to make commitments to the industry unless the economic potential of shale gas is proven with extensive drilling.

In case even mildly optimistic predictions regarding shale potential outside the US are correct, the potential impact on the world natural gas markets and the geopolitical consequences will be significant. Countries and regions with high dependency on imports will improve energy security and see improvements in their balance of trade. Pipeline and LNG flows of natural gas will be rearranged and price levels will change, putting some huge infrastructure investments at risk. One can already detect growing nervousness in some countries that are big natural gas exporters. Even a casual review of the daily press shows growing concerns of destabilising technological change in this industry.<sup>52</sup> Full analysis of this topic would require a separate volume. All we can do is to urge energy traders and analysts to stay tuned.

### **HOW GREEN IS GREEN?<sup>53</sup>**

Natural gas is often promoted as a bridge fuel, the best of the fossil fuels in terms of environmental impact as compared to coal and oil. The advantage natural gas has over coal is obvious. Burning coal produces CO<sub>2</sub>, the main greenhouse gas, SO<sub>x</sub> and NO<sub>x</sub>,<sup>54</sup> ash, and traces of mercury, arsenic and selenium. Burning natural gas produces NO<sub>x</sub> and CO<sub>2</sub>, but the emissions of CO<sub>2</sub> per unit of energy produced are much lower than in the case of coal. The EIA provides the comparison of CO<sub>2</sub> emissions in Table 9.5 on its website.

**Table 9.5** Pounds CO<sub>2</sub> per MMBtu

Anthracite	227.400
Bituminous	205.300
Subbituminous	212.700
Lignite	215.400
Methane	115.258

*Source:* U.S. Energy Information Administration

**Table 9.6** Emission comparison between EPA and industry data

Source category	EPA		API/ANGA		Impact on source category emissions
	Metric tons of CH4	% of EPA emissions total	Metric tons of CH4	% of revised emissions total	
Gas wells liquids unloading	4,501,465*	51%	637,766	14%	-86%
Unconventional well re-fracture rates	712,605*	8%	197,311	4%	-72%
Other production sector emissions**	3,585,600	41%	3,585,600	81%	
Total production sector emissions	8,799,670		4,420,677		-50%

Source: Terri Shires and Miriam Lev-On, 2012, "Characterising pivotal sources of methane emissions from unconventional natural gas production. Summary and analysis of API and ANGA survey responses," URS Corporation and The LEVON Group, June 1.

Methane produces 115.258 pounds of CO<sub>2</sub> per MMBtu, compared with 205.3 pounds in the case of bituminous coal. One additional factor has to be considered here. The thermal efficiency of a combined cycle power plant reaches 60% for the most efficient units; the thermal efficiency of a coal plant can be assumed to be about 35%.<sup>55</sup> Lower emissions per unit of energy and higher efficiency translate into a very significant advantage of natural gas over coal.

These comparisons stop short, however, of a consideration of additional factors. Methane is a much more potent GHG than CO<sub>2</sub>, with actual translation coefficients being still a subject of debate among scientists. Estimates range between the multiples of 21 and 35, but may go as far as 105.<sup>56</sup> This means that potential leaks of natural gas from pipelines, distribution networks and drilling rigs can quickly offset the greenhouse advantages of methane over coal.

The issue of reducing methane emissions came up on the eve of the Cancun Conference on climate change. Two academics pointed out that a relatively easy way of slowing down global warming, before full international consensus can be reached, is lowering the emissions of "three short-lived gases – methane, some hydro fluorocarbons and lower atmospheric ozone – and dark soot particles."<sup>57</sup> The authors pointed out that reduction of methane emissions can be accomplished through a number of methods:

Take methane, for example, which is 25 times more powerful than carbon dioxide in causing warming. It is emitted by coal mines, landfills, rice paddies and livestock. And because it is the main ingredient in natural gas, it leaks from many older natural-gas pipelines. With relatively minor changes – for example, replacing old gas pipelines, better managing the water used in rice cultivation (so that less of the rice rots) and collecting the methane emitted by landfills – it would be possible to lower methane emissions by 40 percent. Since saved methane is a valuable fuel, some of this effort could pay for itself.<sup>58</sup>

These comments point to the importance of systematic comparative studies on the impact of different competing fossil fuels on the environment. The author is not competent to evaluate scientific evidence presented in a number of different academic studies on this topic; however, it is quite obvious that the industry should pay close attention to the debate on this subject, as it will affect not only public perception but also public policy. There are several critical issues which we can identify.

The potency of methane compared to other GHGs has two dimensions. Methane is much more harmful than CO<sub>2</sub>,<sup>59</sup> but it remains in the atmosphere for a shorter period of time. It is very important to identify the time period over which the environmental footprints of CO<sub>2</sub> and other fossil fuels are compared.

There is a significant difference between conventional and shale natural gas in terms of contributions to greenhouse effect. Two recent developments in this area require a closer look. The first is a paper from scientists at Cornell University,<sup>60</sup> which identified three conduits for greenhouse emissions.

- ❑ *Direct emissions.* Emissions of CO<sub>2</sub> related from using natural gas as a fuel for electricity generation, transportation and space heating.
- ❑ *Indirect emissions.* Emissions related to the consumption of other fossil fuels in production of natural gas (for example, gasoline used in the transportation of equipment for drilling natural gas wells).
- ❑ *Fugitive emissions* of methane in different stages of production and distribution. These emissions are much more significant in the case of shale gas compared to conventional natural gas, and occur primarily because of the following.

- Significant volumes of methane gas are released when liquids used for fracking come back to the surface (flow-back). Additional fugitive emissions are related to the drill-out process. Conventional natural gas production does not create these problems. Flow-back corresponds to the period after completion of the well but before the start of commercial production. Fracking liquids removed from a well contain natural gas in quantities that may not be sufficient for sales.
- Routine venting and equipment leaks. Post-completion methane emissions are most likely of the same order of magnitude for conventional and non-conventional natural gas.
- Losses of methane during natural gas processing.
- Losses during transportations, storage and distribution.

According to the Cornell paper, the combination of volumes of direct, indirect and fugitive emissions, in conjunction with the chemical and physical properties of methane, does have an impact.

Methane contributes substantially to the greenhouse gas footprint of shale gas on shorter time scales, dominating it on a 20-year time horizon. The footprint for shale gas is greater than that for conventional gas or oil when viewed on any time horizon, but particularly so over 20 years. Compared to coal, the footprint of shale gas is at least 20% greater and perhaps more than twice as great on the 20-year horizon and is comparable when compared over 100 years.

The second related development was the revision by the Environmental Protection Agency (EPA) of the procedures used to calculate greenhouse emissions for the production sector of natural gas systems.<sup>61</sup> The revisions included changes to calculations of emissions from existing sources (as well as liquids unloading, condensate storage tanks and centrifugal compressor seals), and from two additional sources (unconventional gas well completions and work-overs). The cumulative result of these changes was an increase of 204% of estimated GHG emissions.

Several dissenting voices require closer attention. A paper by IHS CERA<sup>62</sup> pointed out that the EPA's conclusions were based on a small sample covering around 8,800 wells, or ~2% of the wells covered in the EPA's national GHG inventory. According to the paper, EPA methane emission factors were based on just four data

points, with some of the information related to methane captured for sale. If the equipment for methane capture is not installed, it does not mean that that all natural gas is vented. It may be flared, producing CO<sub>2</sub>, which is not as potent as methane from the perspective of the greenhouse effect. The paper also contained comments on the paper by Howarth *et al*, who based some of their conclusions on the CERA article “US industry highlights, February–March, 2009.” The bone of contention is the use of data about methane captured as the basis for conclusions about methane released.

An alternative study is based on a survey conducted under the auspices of American Petroleum Association (API) and the American National Gas Association (ANGA).<sup>63</sup> Collected data included information on gas well types, gas well venting/flaring from completions, work-overs and liquids unloading, and the use of centrifugal compressor and pneumatic controllers. The survey covered 91,028 out of the 355,082 gas wells in the EPA inventory. The main conclusion was that:

Survey results in two source categories – liquids unloading and unconventional gas well re-fracture rates – [being] substantially lower [than] EPA’s estimated emissions from natural gas production and shift Natural Gas Systems from the largest contributor of methane emissions to the second largest (behind Enteric Fermentation, which is a consequence of bovine digestion).

Table 9.6 from the cited study compares the results of the EPA and API/ANGA study.

Several scientists from Cornell University dispute the conclusions of Robert W. Howarth and his co-authors. In a paper published in 2012, they concluded that “using more reasonable leakage rates and bases of comparison, shale gas has a GHG footprint that is half and perhaps a third that of coal.”<sup>64</sup>

The papers cited in this section are in the public domain and the reader is advised to read them. Our recommendation for any trader of natural gas is to follow this discussion. The controversy is unlikely to go away and the likely outcome over time will be a decision by the industry to invest more aggressively in methane capture technology, with some impact (difficult to quantify at this point) on the cost of production.

It is reasonable to expect that the natural gas industry will at some point come under increased pressures to reduce emissions, not only

in the transportation of natural gas but also in drilling. This is evidenced, for example, by a number of reports related to air quality in areas where natural gas is produced from shale rock formations, close to big urban centres such as Barnett Shale, where an increase in benzene concentrations has been detected.<sup>65</sup>

Of course, additional analysis may further reduce the perceived superiority of natural gas over other fuels. For example, one could include in the analysis emissions in the production process of different competing fuels. Production of natural gas from shale formations requires the construction of new access roads and many trips by trucks delivering water, fracking liquids and disposing of wastewater. The very high decline rates that are typical of wells drilled in shale formations cause more frequent trips by delivery trucks and higher levels of drilling activity, compared to conventional production. These factors indicate that more studies are required before one can arrive at final conclusions.

Finding that natural gas may not be superior to coal from the point of view of greenhouse emission, or that its advantages are smaller than previously assumed, does not mean that one should contemplate curtailing production of natural gas, especially that produced from shale formations. Natural gas offers many advantages from the point of view of sulphur and mercury content and, most importantly, US energy security. The environmental risks of hydraulic fracking cannot be denied. The large-scale industrial processes of the extractive industry are usually a nasty and dangerous business, but it does not mean that we can give up as a society the benefits we derive from the exploitation of mineral resources. A rational approach consists of recognising the potential problems and finding the most efficient ways of mitigating the risks. Unfortunately, professionals working in the natural gas industry are often excessively defensive and deny *any* risk related to shale natural gas production.<sup>66</sup> Critics, on the other hand, operate as if our society had an inexhaustible source of perfectly safe and abundant energy. This contradiction was recognised more than half a century ago by George Orwell when he was dispatched on a fact-finding mission to investigate conditions in the UK coal industry. As much as he was shocked by the high rates of unemployment and dangerous working conditions, he understood the necessity of producing coal:

Our civilization [...] is founded on coal, more completely than one realizes until one stops to think about it. [...] Practically everything we do, from eating an ice to crossing the Atlantic, and from baking a loaf to writing a novel, involves the use of coal, directly or indirectly. [...] Whatever may be happening on the surface, the hacking and shoveling have got to continue without a pause or at any rate without pausing for more than a few weeks at the most.<sup>67</sup>

The same observations remain true today, with respect to coal and other sources of energy. The business of exploring and producing energy commodities has to go on if we are to preserve the ability to address the ills of our civilisation. We have no choice but to walk a tightrope between necessity and the dangers that no responsible person should recklessly dismiss as non-existent.

## CONCLUSIONS

Developments related to shale natural gas are a critical issue from the point of view of our energy future and, by extension, that of energy trading. Many future market developments, ranging from the levels of natural gas prices in the US, to international trade flows of natural gas (in the form of LNG), to international pricing regimes for LNG and pipeline natural gas delivered into Europe and Asia, depend on what happens with respect to shale gas in the US. The market impact will be huge and will be felt in the power markets (as gas generation capacity expands), coal markets, prices of liquid fuels (due to the expansion of fleets of cars running on compressed natural gas) and prices of natural gas liquids (and, by extension, prices of plastics). An abundance of natural gas can reverse the trend towards the transfer of some US industries abroad. Daily monitoring of developments in this area is probably the number one task for energy traders and analysts.

<sup>1</sup> Julian Darley, 2004, *High Noon for Natural Gas: The New Energy Crisis* (White River Junction, Vermont: Chelsea Green Publishing). This is not intended as a criticism of the otherwise excellent book by Mr. Darley. The author does not think anybody could have predicted the shale boom in 2004, it is rather intended as a warning against making any type of dramatic prediction. Let's not forget that Dante (*Inferno*, Canto XX) placed the fortune tellers and diviners (who were trying to predict the future when alive) in the 8th circle, with their heads twisted backward, condemned to look behind for eternity.

<sup>2</sup> Attributed to Danton, though more likely pronounced by Pierre Vergniaud.

<sup>3</sup> <http://205.254.135.7/tools/glossary/index.cfm?id=C>.

<sup>4</sup> Frederik Robelius, 2007, "Giant oil fields – The highway to oil. Giant oil fields and their importance for future oil production," *Acta Universitatis Upsaliensis*.

<sup>5</sup> *Ibid*, p 23.

- 6 "A porous substance has a permeability of 1 darcy if, in 1 second, 1 cubic centimetre of a gas or liquid with a viscosity of 1 centipoise will flow through a section 1-centimetre thick with a cross section of 1 square centimetre, when the difference between the pressures on the two sides of the section is 1 atmosphere. Sandstone typically has a permeability of a few darcys." (see <http://www.sizes.com/units/darcy.htm>).
- 7 Darcy's law describes the flow of fluid through a porous medium.
- 8 For example, pumice has porosity of about 90% but the pores are isolated, making it a very poor reservoir rock.
- 9 <http://www.theoildrum.com/story/2006/3/8/222920/5485?nocomments>.
- 10 <http://www.glossary.oilfield.slb.com/Display.cfm?Term=shale>.
- 11 Detritus means disintegrated, eroded matter or accumulated material (debris).
- 12 Annemieke van Roekel, 2012, "Unconventionals are easy to find but hard to get," *European Energy Review*, January 16 (see <http://www.europeanenergyreview.eu/site/pagina.php?id=3462>).
- 13 Ibid.
- 14 ExxonMobil discontinued their exploration effort in Poland after drilling two non-commercial wells (see, Jan Cienski, 2012, "Poland shale: Exxon exit," *Financial Times*, June 18). The Polish Geological Institute has reduced the estimates of Polish natural gas shale resources by 90% (compared to the EIA numbers). Still, even with lower estimates, Poland has resources covering 24 years of consumption. The cost of production is still an open issue. See, Andrew Kureth, 2012, "The real disappointment about shale gas," *Warsaw Business Journal*, March 12.
- 15 "There was a time you all were told that any of the 17 counties in the Barnett Shale play would be just as good as any other county," McClendon said. "We found out there are about two or two and a half counties where you really want to be." *Bloomberg News* October 14, 2009.
- 16 EIA, 2011, "World shale gas resources: An initial assessment of 14 regions outside the United States," April. This study was commissioned from Advanced Resources International.
- 17 See, for example, the following interview with Andrew Weissman: "I think we are heading towards an enormous problem," Weissman said. "More often than not in the next decade we will have severe shortages." (<http://www.silverbearcafe.com/private/dryup.html>). We were equally convinced of a potential long-term supply problem.
- 18 Gas-to-liquids (GTL) is an application that allows one to transform natural gas into gasoline or diesel. This technology is described in the chapter on non-conventional oil.
- 19 In 2009, the Barnett Shale play in Texas produced 6% of all natural gas produced in the lower 48 States ("Modern shale gas development in the United States: a primer," prepared for the US Department of Energy Office of Fossil Energy and National Energy Technology Laboratory by the Ground Water Protection Council and ALL Consulting. Referred to subsequently as Primer ES-1).
- 20 [http://www.spe.org/jpt/print/archives/2006/02/JPT2006\\_02\\_tech\\_bits.pdf](http://www.spe.org/jpt/print/archives/2006/02/JPT2006_02_tech_bits.pdf).
- 21 <http://www.kgs.ku.edu/Publications/PIC/PIC32r1.pdf>.
- 22 "Modern shale gas development in the United States: A primer," op. cit.
- 23 See, for example, [http://www.chk.com/Media/Educational-Library/Fact-Sheets/Corporate/Hydraulic\\_Fracturing\\_Fact\\_Sheet.pdf](http://www.chk.com/Media/Educational-Library/Fact-Sheets/Corporate/Hydraulic_Fracturing_Fact_Sheet.pdf). The website <http://www.hydraulicfracturing.com/Pages/information.aspx> also contains a lot of technical information about hydraulic fracturing.
- 24 [www.theoildrum.com](http://www.theoildrum.com), "Making holes and cracks around oil and gas wells," Posted by Heading Out, November 22, 2009.
- 25 Such polymers are used also in liquids developed for crowd control (the banana water) as they make ground too slippery to stand on.
- 26 See, for example, <http://www.mines.edu/Potential-Gas-Committee-reports-unprecedented-increase-in-magnitude-of-U.S.-natural-gas-resource-base>.

- 27 <http://potentialgas.net/what-we-do-2>.
- 28 "But the jump pales in comparison to the giant 36 percent leap the committee reported two years ago, mostly gas from emerging shale gas formations, such as the Marcellus in the Northeast and the Haynesville in the South. That report cemented the arrival of the shale gas booms, incessantly declared a "game-changer" in the energy world." Mike Soraghan, 2011, "Potential US natural gas supplies have jumped 3%, industry experts say," *The New York Times*, April 27.
- 29 [http://www.eia.gov/forecasts/aeo/pdf/0383er\(2011\).pdf](http://www.eia.gov/forecasts/aeo/pdf/0383er(2011).pdf).
- 30 The concept of reserves is discussed in the section on oil markets.
- 31 The ratio of effectively producing reserves to total gas in place.
- 32 Leonardo Maugeri, 2012, "Oil: The next revolution. The unprecedeted upsurge of oil production capacity and what it means for the world," Belfer Center for Science and International Affairs, John F. Kennedy School of Government, June.
- 33 One can follow the discussion on [www.theoldrum.com](http://www.theoldrum.com) (see also <http://petroleumtruthreport.blogspot.com/2012/06/arthur-berman-publications.html>).
- 34 See, for example, Ian Urbina, 2011, "Drilling down: insiders sound an alarm amid a natural gas rush," *New York Times*, June 26. The website <http://www.energyindepth.org/> is a good portal to information supporting different aspects of hydraulic fracturing. One of the best-known sceptics of shale gas is Houston-based geologist Art Berman. His position is summarised best in his address to the ASPO 2012 Conference in Vienna, "Learning from US shale gas experience" (see <http://www.youtube.com/watch?v=1386Jt17myY>).
- 35 "On [June 24, 2011] Texas Gov. Rick Perry signed into law a bill that will require companies to make public the chemicals they use on every hydraulic fracturing job in the state. While a handful of other states have passed similar measures, Texas's law is significant because oil and gas drilling is a key industry in the state and the industry vocally supported the measure." Ben Casselman, 2011, "'Fracking' disclosure to rise: gas drillers begin supporting laws requiring them to list chemicals they use," *Wall Street Journal*, June 20. Companies may request withholding information about certain chemicals from the public if they believe that it represents a trade secret.
- 36 <http://fracfocus.org>.
- 37 For example, fracking liquids exported to other countries have to be accompanied by detailed listing of their components, as required by foreign laws.
- 38 See, for example, "The Halliburton loophole," *New York Times*, November 2, 2009. "Among the many dubious provisions in the 2005 energy bill was one dubbed the Halliburton loophole, which was inserted at the behest of – you guessed it – then-Vice President Dick Cheney, a former chief executive of Halliburton."
- 39 "Toxic chemical release reporting: Community right-to-know". *Code of Federal Regulations*. Title 40. Pt 372.23. 2007.
- 40 On June 30, 2011, France became the first country to prohibit hydraulic fracturing. See, "Gaz de schiste : le Parlement interdit l'utilisation de la fracturation hydraulique," *Lemonde.Fr* avec AFP, June 30, 2011. See also "Gaz de schiste, le rêve polonais," *Le Monde*, June 1, 2011, for a different approach by some other European countries. Bulgaria joined France in banning hydraulic fracturing (in 2012). Vermont took the same path in the US.
- 41 [http://energy.utexas.edu/index.php?option=com\\_content&view=article&id=145:early-results-from-hydraulic-fracturing-study-show-no-direct-link-to-groundwater-contamination&catid=34:press-releases&Itemid=54](http://energy.utexas.edu/index.php?option=com_content&view=article&id=145:early-results-from-hydraulic-fracturing-study-show-no-direct-link-to-groundwater-contamination&catid=34:press-releases&Itemid=54).
- 42 The Rocky Express (REXX) pipeline is a dramatic example of changes in the natural gas flows. "Kinder Morgan Energy Partners LP is now considering reversing its Rockies Express Pipeline, acknowledging that with plenty of natural gas along the East Coast, the pipeline may better serve customers in the West." *Pipeline & Gas Journal*, August 2011, 238(8) (<http://www.pipelineandgasjournal.com/rockies-express-pipeline-may-reverse-flow-move-shale-gas>).

- 43 B. P. Pashigian, 2008, "Cobweb theorem," in Steven N. Durlauf and Lawrence E. Blume (Eds), *The New Palgrave Dictionary of Economics* (2e) (Basingstoke, England: Macmillan).
- 44 The phrase, "bygone is bygone" in economics is attributed to the British economist William Stanley Jevons. "The fact is that labour once spent has no influence upon the future value of an article: it is gone and lost forever. In commerce bygones are forever bygones; and we are standing clear at each moment, judging the value of things with a view to future utility."
- 45 <http://www.bentekenergy.com/InTheNews.aspx#Article7192>.
- 46 The spreadsheet, developed by the author, follows the template used in the BentekEnergy publication "The Beast in the East™ Marcellus Shale Poised to Disrupt U.S. Gas Market," BentekEnergy, March 19, 2010.
- 47 Rig count information is available at no cost from Baker Hughes ([http://investor.shareholder.com/bhi/rig\\_counts/rc\\_index.cfm](http://investor.shareholder.com/bhi/rig_counts/rc_index.cfm)).
- 48 Arthur E. Berman, 2010, "Shale gas – abundance or mirage?" Labyrinth Consulting Services, Montréal, Québec, December 2.
- 49 Jason Bainly, Raphael Altman, Raj Malpani and Fang Luo, "Shale gas production decline trend comparison over time and basins," SPE technical paper 135555. A summary of the paper can be found in Jason Bainly, Raphael Altman, Raj Malpani and Fang Luo, 2011, "Study assesses shale decline rates," *The American Oil and Gas Reporter*, May. The quote is from the summary paper.
- 50 <http://seekingalpha.com/article/218878-chesapeake-energy-making-all-the-right-moves>.
- 51 Held by production – ie, the case of drilling required to keep the lease.
- 52 The Russian press tended to be dismissive of shale gas potential describing this development as a temporary fad. This is changing as exemplified by a recent article in *Novaya Gazeta*, "Труба тревоги нашей. Америка готовит сланцевую революцию, и это может окончательно подорвать устои российской экономики," ("The pipeline of our panic. America prepares shale revolution and this may finally undermine the foundations of the Russian economy,") May 4, 2012.
- 53 This subsection was inspired by the post in *The Oil Drum* by Chris Vernon, June 24, 2010, (see <http://europe.theoildrum.com/node/6638#more>), who pointed out the need to go beyond simplistic analysis of the environmental benefits of natural gas exclusively in terms of CO<sub>2</sub> emissions in power.
- 54 SOx and NOx (pronounced as "socks" and "nocks") denote different sulphur and nitrogen oxides (such as, for example, SO<sub>2</sub>).
- 55 This issue is discussed in more detail in Chapter 20 on the electricity markets.
- 56 Drew T. Shindell, Greg Faluvegi, Dorothy M. Koch, Gavin A. Schmidt, Nadine Unger and Susanne E. Bauer, 2009, "Improved attribution of climate forcing to emissions," *Science*, October 30, 326(5953), pp 716–18.
- 57 Veerabhadran Ramanathan and David G. Victor, 2010, "To fight climate change, clear the air," *New York Times*, November 27.
- 58 Ibid.
- 59 The International Panel on Climate Change used a conversion factor of 21 back in 1992. Recent studies point to much higher levels of potency of methane compared to CO<sub>2</sub>.
- 60 Robert W. Howarth, Renee Santoro, and Anthony Ingraffea, 2011, "Methane and the greenhouse-gas footprint of natural gas from shale formations," working paper, Cornell University.
- 61 United States Environmental Protection Agency (EPA), "Inventory of greenhouse gas emissions and sinks: 1990–2009," Washington DC, April, 2011(b) (<http://www.epa.gov/climatechange/Downloads/ghgemissions/US-GHG-Inventory-2012-Main-Text.pdf>).
- 62 Mary Lashley Barcella, Samantha Gross and Surya Rajan, "Mismeasuring methane: Estimating greenhouse gas emissions from upstream natural gas development," IHS CERA private report.

- 63 Terri Shires, Miriam Lev-On, 2012, "Characterizing pivotal sources of methane emissions from unconventional natural gas production. Summary and Analysis of API and ANGA Survey Responses," URS Corporation and The LEVON Group, June 1.
- 64 Lawrence M. Cathles III, Larry Brown, Milton Taam and Andrew Hunter, 2012, "A commentary on 'The greenhouse-gas footprint of natural gas in shale formations,'" by R.W. Howarth, R. Santoro and Anthony Ingraffea, *Climatic Change* (DOI: 10.1007/s10584-011-0333-0), January 3. Response to their article can be found in Robert W. Howarth, Renee Santoro and Anthony Ingraffea, "Venting and leaking of methane from shale gas development: response to Cathles *et al.*," *Climatic Change* (DOI 10.1007/s10584-012-0401-0).
- 65 "Nearly one-fourth of the sites monitored in North Texas' Barnett Shale natural-gas region had levels of cancer-causing benzene in the air that could raise health concerns, state regulators said Wednesday. They emphasised, however, that gas companies have fixed the worst emission problems and are working on less-serious sites where the state still wants benzene levels to come down." Randy Lee Loftis, "High benzene levels found on Barnett Shale," *The Dallas Morning News*, Thursday, January 28, 2010 (see <http://www.dallasnews.com/sharedcontent/dws/news/localnews/stories/DN-shale.ART.State.Edition2.4bf8062.html>).
- 66 We have heard many statements that the risks of hydraulic fracturing are exactly zero. We have reason to believe that these risks can be reduced to a very low level but it does not mean that they go away. Humans will be sloppy, overworked, sometimes poorly trained and will make mistakes.
- 67 George Orwell, 1937, *The Road to Wigan Pier* (London, England: V. Gollancz), Chapter 2.



## *Natural Gas Transportation and Storage*

This chapter will discuss the transportation and storage of natural gas – ie, the *midstream* segment – a critical link between the producers and consumers of natural gas. The details covered here may seem technical at first glance, but they are critical to both trading natural gas in the short run and to understanding long-term location price differentials. Some of the data used by the traders to make decisions are related to storage and pipeline grid operations. The Weekly US storage report available from the EIA can be compared, in its impact on the gas market, to the impact a monthly non-farm payroll report has on the financial markets. The reason why storage reports have been such an important source of information is that they provide in condensed form information about the balance between supply and demand. Weekly storage data has to be adjusted for the short-term impact of weather conditions, and this creates a great opportunity for fundamental analysts who compete for the title of the best storage forecaster. Several news organisations run weekly scorecards and tabulations ranking different forecasters.<sup>1</sup>

An understanding of pipeline operations is very important from the point of view of trading strategies built around geographical price difference (basis in the US), pipeline constraints or pipeline expansion removing constraints, short-term disruptions related to outages and operational flow orders. An important development over the last few years has been the emergence of several companies providing information about nominations on natural gas on the pipeline system. The reason why this is important is that nominations approximate very closely realised actual pipeline flows. The monitoring of pipeline flows of natural gas helps to create a snapshot of supply and demand practically in real time. Ignoring these and other data sources discussed in this chapter should be done only at

your own risk. The mechanics of the nomination process may appear to be a somewhat boring topic to most new entrants to the industry aspiring to be high-flying and dashing traders. They are correct in labelling such details as the best alternative to sleeping pills, but this activity can have a profound impact on the short-term physical natural gas markets and should be well understood.

Natural gas may be transported to end users in a number of ways, characterised to a varying extent by the transformation of methane.<sup>2</sup>

- ❑ Pipeline transportation is the most straightforward way of delivering natural gas to consumers.
- ❑ Natural gas can be converted into liquid by lowering its temperature to about  $-160^{\circ}\text{C}$  ( $-260^{\circ}\text{F}$ ); this topic will be covered in the section on LNG.
- ❑ Natural gas can be compressed and transported in special tankers (this technology is known as compressed natural gas).
- ❑ Natural gas can be transformed through chemical processes into products that are liquid in the ambient temperatures, such as methanol or gasoline (this promising technology is known as gas-to-liquids).
- ❑ Natural gas can be used as fuel, transformed into electricity and transmitted over the power grid (natural gas by wire).

We will start with a review of technical and regulatory issues related to the US pipeline system.

## NATURAL GAS: TRANSPORTATION

### US pipeline grid

After processing at the gas plants, natural gas can be delivered to the long-distance natural gas pipelines (trunk lines) that transport the molecules to local distribution companies or directly to large industrial and commercial customers. Natural gas pipelines in the US constitute an efficient and reliable network that is more than 306,000 miles long. Its smooth functioning is supported through the telemetric system of data collection, known as supervisory control and data acquisition (SCADA). The control rooms receive continuous datastreams with information about flow volumes, pressures and gas temperatures. Compressor stations used to manage pipeline pressure are controlled from the operation rooms, with personnel on

duty around the clock, and often thousands of miles away from the assets they supervise and dispatch. The US integrated natural pipeline grid is one of those technological marvels that the public is generally unaware of.

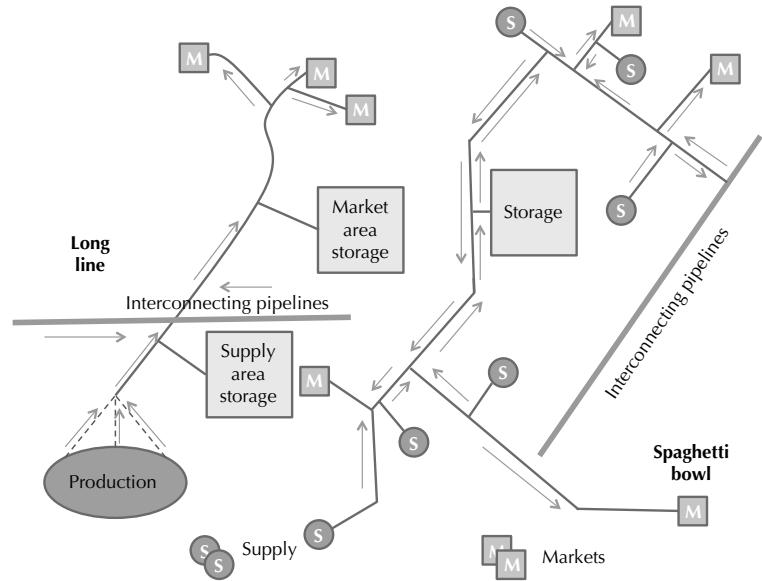
The US natural gas pipeline network is composed of large diameter steel pipelines operating under high pressures, ranging typically from 600 to 1,200 psi. Transmission pipes are usually 24–42 inches in diameter and made of steel 0.25" to 0.75" thick.<sup>3</sup> Smaller branches extending from a larger pipeline and connecting to large customers or groups of customers are known as laterals, and are made from smaller diameter pipes (6–16").

The US natural gas pipelines connect production regions with regions characterised by a concentration of population and industry. The density of pipelines in the sparsely populated mountain west of the country is much lower than in the eastern part of the country. Historically, the pipelines were running from the south to the north, with only a few important east-west pipes. This has changed with the construction of several "horizontal" pipelines, such as the Rockies Express (Rex). One of the challenges the industry is facing is adjusting the pipeline grid to new sources of supply from the shale formations.

The pipelines follow two basic designs, known as trunk lines and grids (See Figure 10.1). A trunk line (also known as "gun barrel" pipeline) is characterised by the concentration of receipt points at one end, and of delivery points at the other end. Such pipelines usually connect a large producing area with an important, big size market. A grid is a system of multiple laterals that serve a large number of regional markets. Such a design is called sometimes a Spaghetti Bowl (a reticulated pipeline or a "web-like" pipeline).

Natural gas is propelled through the pipeline system by compressors located approximately every 50 to 100 miles, typically rated from 1,000–15,000 horsepower (1–13 MW). The compressors operate by burning part of the natural gas flowing through the pipe, although a new design of compressors using electric motors has been successfully tested. The compressors, based on a centrifugal or reciprocating design, propel natural gas along the pipeline by increasing the local pressure and pushing natural gas towards the segments of the pipeline where the pressure is lower.

Pipeline capacity can be increased in two different ways. One

**Figure 10.1** Different pipeline designs

Source: Based on a drawing from J. Richard Moore, "The US Natural Gas Industry: an Overview"

solution is to increase pipeline diameter or to use the so-called *looped design*, by locating several parallel pipes along the same route, corresponding to the right-of-way. Additional pipes may be used as temporary storage space available for accepting overflow production during periods of low demand. Pipelines often store additional gas by temporarily increasing the pressure, in anticipation of periods of high demand, corresponding to very hot or very cold weather. The ability to adjust the volume of gas in a pipeline, known in the industry jargon as *linepack*, creates a buffer between production and demand, supplementing permanent storage facilities. Another way of increasing capacity is the addition of compressor stations or installation of more powerful compressors. An increase in pressure reduces the space natural gas occupies, enabling the transporter to ship more Btu per cubic foot.

### Pipeline transportation contracts

There are two basic types of contracts for the pipeline transportation of natural gas:

- ❑ firm service, which can be denied only under conditions of *force majeure* (ie, acts of God, natural disasters and, in many cases, equipment failure); and
- ❑ an interruptible service offers a lower level of reliability (offset by a lower cost). It is not guaranteed and can be cancelled at short notice, or without notice at all, if pipeline capacity is required to serve higher priority customers (ie, firm-service subscribers).

In the case of a firm contract, a shipper pays the so-called demand (or reservation) charge, which is comparable to an option premium and gives the right, but not the obligation, to use the service, and a commodity charge that is related to the actual volume of natural gas that flows. In the case of interruptible service, only a commodity charge is paid by the shipper. The logic behind this tariff design is that reservation charges related to the firm contracts should pay for the capital costs of building the facility, and the commodity charges should cover the operating costs. We shall encounter similar arguments in the case of other tariffs, especially in the electricity markets.<sup>4</sup>

The simple distinction between firm and interruptible transportation contracts has to be further qualified, given the variety of practical arrangements used in practice. As in the case of a restaurant manager described in the novel *Master and Margarita* by Mikhail Bulgakov, who offered salmon with many degrees of “freshness,” there are many shades of firmness and interruptibility.<sup>5</sup> The rules under which such a shipper may be interrupted, or *bumped* in the industry jargon, are somewhat complex and constantly evolving. Other types of contracts that may be encountered in the tariffs are shown below.<sup>6</sup>

- ❑ *No-notice*. This service can be defined as a premium or platinum firm transportation contract. The shipper can receive gas (both nominated and un-nominated) on a daily basis up to the maximum contract level, without incurring daily balancing and scheduling penalties. No-notice service, as expensive as it is, is becoming increasingly important, given the growing importance of gas-fired power generation units, which are often dispatched on a very short notice (within minutes) and cannot predict

natural gas needs for the next day with a high degree of reliability.

- ❑ *Other primary firm.* This type of service has a somewhat lower degree of firmness than gold-plated firm service. A shipper has a reasonable expectation of an uninterrupted service.
- ❑ *Secondary firm.* Many contracts specify secondary delivery/receipt points in addition to the primary points. The rights associated with such delivery points have typically a lower priority service than primary delivery points, but higher priority than interruptible transportation.
- ❑ *Authorised overrun (AO).* The AO service can be acquired in conjunction with firm or interruptible service and provides a licence to exceed daily (or weekly, monthly volumes) up to a limit and at a price.

In the case of interstate pipelines, transportation cost is based on a pipeline-specific tariff approved by FERC. Interruptible rates (as well as the transportation rates in the case of capacity release) are negotiated. The example in Table 10.1 shows the rates for firm and interruptible service contained in the Rex tariff, effective April 1, 2008. The reservation fees are monthly charges for pipeline space, whereas the commodity rates are in dollars per dekatherm scheduled daily.<sup>7</sup>

The fuel, in addition to the lost and unaccounted for (L&U) charges for Rex, are available on a separate sheet. L&U is a percentage of the total daily scheduled volume to compensate the pipeline for losses and measurement inaccuracies. Typically, the item is less than 1%.

As explained, the charges paid by shippers of natural gas are broken into two tiers. Fixed charges (also called demand charges or reservation fees) are paid irrespective of whether the pipeline capacity is actually used. These charges can be compared to the payment for renting a car, which is due irrespective of the number of miles actually driven (and have to be paid even if the car is not driven at all). In other words, they can be compared to an option premium, giving the option holder the right, but not the obligation, to use the service. Variable costs (also called commodity charges) are paid only if the service is actually used and depend on the level of throughput.

**Table 10.1** Rockies Express Pipeline

FERC gas tariff alternate first revised sheet No 5

Second revised volume No 1 superseding

Original sheet No 5

## CURRENTLY EFFECTIVE RATES 1/

Rate schedule FTS (firm transportation service) 3/

<b>Receipt zone 2</b>	<b>Delivery zone 2</b>	<b>Maximum reservation (US\$)</b>	<b>Maximum commodity (US\$)</b>	<b>Minimum commodity (US\$)</b>
Zone 1	Zone 1	7.0720	0.0030	0.0030
	Zone 2	23.5546	0.0049	0.0049
	Zone 3	50.2197	0.0074	0.0074
Zone 2	Zone 1	23.5546	0.0049	0.0049
	Zone 2	16.4826	0.0019	0.0019
	Zone 3	43.1477	0.0044	0.0044
Zone 3	Zone 1	50.2197	0.0074	0.0074
	Zone 2	43.1477	0.0044	0.0044
	Zone 3	26.6651	0.0025	0.0025

Rate schedule ITS (interruptible transportation service)

Zone 1	Zone 1	0.2355	0.0030
	Zone 2	0.7793	0.0049
	Zone 3	1.6585	0.0074
Zone 1	Zone 1	0.7793	0.0049
	Zone 2	0.5438	0.0019
	Zone 3	1.4230	0.0044
Zone 1	Zone 1	1.6585	0.0074
	Zone 2	1.4230	0.0044
	Zone 3	0.8792	0.0025

An example below shows how transportation is priced based on a specific tariff. Transportation tariff for Rex lists demand charges for transportation of natural gas between different pipeline zones. For example,<sup>8</sup> the charges applied to the shipment of natural gas between Zone 1 and Zone 3 listed in the tariff are given as monthly rates of US\$50.2197/MMBtu (or daily rates of 1.6511). The relationship between these two rates can be explained as follows:

$$50.2197 / 30.41667 = 1.651059^9$$

The monthly rate of 50.2197 divided by 30.41667 (an average number of days in a month) translates into a daily rate of 1.6511 (after rounding).

**Table 10.2** Rockies Express Pipeline (fuel and L&U)

FERC gas tariff fifth revised sheet No 6  
 Second revised volume No 1 superseding  
 Fourth revised sheet No 6

## CURRENTLY EFFECTIVE RATES

ITS/FTS fuel and lost and unaccounted for reimbursement percentages

Receipt zone 2	Delivery zone 2	Maximum Percentage	Minimum Percentage
Zone 1	Zone 1	0.69	0.10
Zone 2		2.42	0.10
Zone 3		3.59	0.10
Zone 2	Zone 2	1.52	0.10
Zone 3		2.69	0.10
Zone 1	Zone 3	0.96	0.10

The cost of reserving space for 10,000 MMBtus/day over an average month is equal to:

$$10,000 \text{ MMBtu} * 50.2197 \text{ US\$/MMBtu} = \text{US\$}502,197/\text{month}$$

The variable cost of transportation has a number of components. The static components are independent of the current level of natural gas prices. The commodity charge for the transportation of natural gas from Zone 1 to Zone 3 of Rex is given as US\$0.0074/MMBtu. The annual charge adjustment (ACA) component is a surcharge assessed by the FERC under Section 154.38 (d) (6) of their regulations to permit interstate pipeline companies to recover from the shippers' total annual charges. The ACA depends on the monthly pipeline volume. In our case this charge is US\$0.0019/MMBtu.

Compressor fuel and L&U charges cover the cost of fuel used to maintain pipeline pressure and lost and unaccounted for natural gas. This cost is specified in terms of the percentage (3.59% in the case used for illustration) of natural gas being shipped, and actual payments will vary with the market price of natural gas. Assuming the price of natural gas to be US\$4/MMBtu, this charge will be equal to  $0.0359 \times 4 = \text{US\$}0.1436/\text{MMBtu}$ , or  $4/(1-0.0359) = 0.1489$ . These calculations demonstrate two different ways used by the industry to calculate the cost of compressor fuel. One way to think about the cost calculation is to assume that the shipper will nominate, for each unit

of gas being shipped, an additional volume, calculated as  $1/(1 - \text{compressor fuel percentage}) - 1$ . In our example, the nominated volume is  $1/(1-0.0359) = 1.03724$ . The extra volume, multiplied by the price of natural gas equal to US\$4/MMBtu, translated into US\$0.1489/MMBtu ( $0.03724 \times 4$ ). Under an alternative calculation, the shipper receives, for each nominated unit, a slightly smaller volume (ie,  $1 - \text{compressor fuel percentage}$ ). Another method to handle fuel and L&U is for the shipper to schedule 10,359 Dth from the producer and deliver 10,000 Dth to market.

Total variable cost per MMBtu in our case is equal to US\$0.1582/MMBtu ( $0.0074 + 0.0019 + 0.1489$ ). Total loaded cost for natural gas shipment between Zone 1 and Zone 3 of Rex is equal to US\$1.8903/MMBtu ( $0.1582 + 1.6511$ ). This last number includes both the reservation and commodity charges. The reservation charges represent 91.3% of the total loaded cost, and the variable charges represent 8.7%. The breakdown between the fixed and variable charges varies from pipeline to pipeline, with pipelines shifting most of the cost to the fixed (SFV) charges (as the overriding objective is faster amortisation of the capital cost). SFV stands for the straight fixed-variable rate design approach. Under SFV, a pipeline recovers all of its fixed costs through a demand charge and all of its variable costs through a commodity charge. The older pipelines, with fully amortised capital investment costs, rely on tariffs with a higher share of variable cost (ie, MFV). Under MFV (the modified fixed-variable method), a portion of the pipeline's fixed costs (return on equity and related income taxes) is included in the commodity charge, not the demand charge.

Many modelling efforts related to the valuation of transportation contracts rely on the use of option pricing theory, and specifically on the application of spread option technology. This is not necessarily a wrong approach (as will be discussed later in the book), but caution should be used in interpreting market information. First of all, precious little natural gas would flow in the system if decisions were made based on the Gas Daily prices (these prices are explained in Chapter 11) and the full tariffs. Most time differentials between daily prices at two locations seldom spike above the corresponding full (maximum) tariff.<sup>10</sup> According to the FERC data:

Focusing on fluctuations in the market value of transportation service as shown by basis differentials between Louisiana and

Chicago and between the Permian Basin and the California border, respectively, these figures show that for most of the year, the value of transportation service is less than the maximum transportation rate of Natural Gas Pipeline Company of America and El Paso Natural Gas Company, respectively. [...] [D]uring brief peak-demand periods the value of transportation service is measurably greater than the maximum transportation rate.<sup>11</sup>

One may ask the question why the shippers are at all willing to pay the maximum tariff rate. A somewhat homely analogy may help to explain it. The working hours on a trading floor are usually very long and we used to pop into a local supermarket on our way home to shop for groceries. Late shoppers would often be offered perishable items (such as rotisserie chicken) for a nominal price. Such bargains are very attractive, but it is not the most desirable way to feed a family. The supply is not guaranteed and other shoppers may snatch it first if it materialises at all. Many natural gas buyers, like average shoppers, are willing to pay extra for guaranteed and immediate supply. The end users who have other options (for example, local gas storage access or dual burner capacity and propane or resid tanks) may take the risk of relying on interruptible contracts and lower the overall cost of acquisition of natural gas. The local distribution companies, expected to provide reliable uninterrupted supply, do not have this luxury. Another advantage of holding firm capacity is that the holder can resell/release unused capacity at a price higher than the maximum tariff regulated rate for pipelines. Of course, a lot of pipeline capacity is sold at prices lower than the maximum tariff, either under interruptible contracts or through capacity release.<sup>12</sup>

Once a shipper enters into an agreement with a pipeline and commits to paying reservation charges, their day-to-day decisions are based on variable costs only. This reflects one of the basic economic principles – that a sunk cost (ie, a historical or incurred cost) should be ignored in making forward-looking decisions.

### The nomination process

Nominations are procedures carried out in order to exercise the rights acquired under the transportation and supply and storage contracts. Put simply, it is a request to transport natural gas or reserve capacity on a pipeline, under the general terms of a negotiated transportation agreement. In the case of gas supply contracts,

the nomination is requested by the purchaser and submitted to the seller – known as a title transfer nomination. The transportation transfer nomination comes from the shipper and is made with the transporter. Custody transfer nominations happen when gas is transferred along the pipeline from one party to another (for example, from one transporter to another).

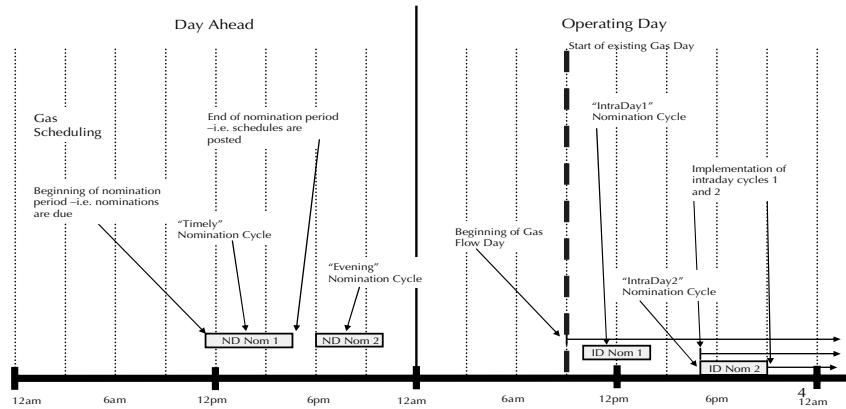
The nomination process has been standardised over time under the auspices of the Gas Industry Standards Board (GISB), which changed its name in 2002 to the North American Energy Standards Board (NAESB).<sup>13</sup> Uniformity of procedures is important to lower the costs and avoid mistakes. A template nomination document includes certain basic data required to schedule the gas flows and reconcile later the differences between the scheduled and actually flown volumes (imbalances).<sup>14</sup> The nomination schedules followed by most firms in the industry are summarised in Figure 10.2 and Table 10.3.<sup>15</sup>

The nomination process can be broken up into several stages. During the “timely” cycle, occurring during the day ahead, with the nomination deadline of 11.30 CST and results being posted by 4.30 CST, shippers get guaranteed service as long as the scheduled amounts do not exceed contract quantities.<sup>16</sup> The capacity that remains unscheduled can be offered to a customer with lower priority. During the gas day, capacity is scheduled only if previously scheduled firm shippers do not use capacity.<sup>17</sup> Late nominations include the evening window (6.00 deadline, CST)<sup>18</sup> and two intra-

**Table 10.3** Pipeline nomination cycles

Before the Gas Day	During the Gas Day
<b>Timely</b> <ul style="list-style-type: none"> <li>• 11.30 – Nominations due</li> <li>• 12.00–15.30 – Confirmation Done</li> <li>• 16.30 – Scheduled Quantities Posted</li> <li>• 9.00 (Next Day) – Gas Flow Begins</li> </ul>	<b>Intraday 1 (ID1)</b> <ul style="list-style-type: none"> <li>• 10.00 – Nominations due</li> <li>• 10.30–13.00 – Confirmations done</li> <li>• 14.00 – Scheduled Quantities Posted</li> <li>• 17.00 – Gas Flow Begins</li> </ul>
<b>Evening</b> <ul style="list-style-type: none"> <li>• 18.00 – Nomination Due</li> <li>• 18.30–21.00 – Confirmations Done</li> <li>• 22.00 – Scheduled Quantities Posted</li> <li>• 9.00 (Next Day) – Gas Flow Begins</li> </ul>	<b>Intraday 2 (ID2)</b> <ul style="list-style-type: none"> <li>• 17.00 – Nomination Due</li> <li>• 17.30–20.00 – Confirmations Done</li> <li>• 21.00 – Scheduled Quantities Posted</li> <li>• 21.00 – Gas Flow Begins</li> </ul>

Source: Carolina Gas Transmission

**Figure 10.2** Gas nomination process

Source: This figure is based on the "NAESB Final Report on the Efforts of the Gas-Electric Interdependency Committee," Docket No. RM05-28-000, February 24, 2006, page 14

day nominations on the gas day. Meeting the 6.00 deadline is a condition of gas flow starting at the beginning of gas day (9.00 CST). Intraday nomination deadlines are, respectively, 10.00 and 17.00 CST. During the second intraday cycle, "bumping" – ie, service cancellation to uninterruptible customers – is not allowed under NAESB standards. Many pipelines may use a more complicated nomination schedule.

The nomination process is a critical process in moving natural gas between the entities controlling the supply (ie, the producers and marketers) and end users. The nomination cycle starts with a shipper submitting a request for service that specifies the receipt points at which natural gas will be delivered to the pipeline and the delivery points at which natural gas will be taken out of the pipeline. Typically, the receipt points are upstream from the delivery points, given the physical direction of the flow of natural gas. A special service, called backhaul (and, sometimes, paper flow) delivers natural gas to a delivery point upstream of the receipt point. This happens not through reversing the physical flow (although some pipelines have this ability) but through displacement. This service is offered under a special rate schedule or service, which recognises that no compressor fuel is used in the process. The nomination request specifies in addition the date range, volumes, scheduling priority, contract at both the receipts/delivery locations, and a trans-

portation contract under which the service is requested. The request can usually be submitted by telephone, fax or electronically.

As mentioned, a pipeline compares nominations to operational capacity and adjusts the volumes, if necessary. This process follows the terms specified in the tariff of the pipeline in question, with the lowest priority contracts (discounted interruptible contracts) being cut first, based on the rate. The interruptible volumes, secondary firm volumes and firm volumes follow, with allocations done on a *pro rata* basis for each category.

The nominations follow several cycles, called timely, evening, ID1 and ID2, as described in Table 10.3.<sup>19</sup>

In compliance with the regulations adopted in Order No. 637, interstate pipelines post daily information on the Internet about scheduled natural gas volumes for most of the continental US. Under FERC Order 720, the intrastate pipelines will also have to post the same information; in addition, the interstate pipelines will have to post some additional information (ie, no-notice service). However, Order 720 is being contested in the court system and the future of this rule is uncertain at the time of writing. The nomination information posted by the pipelines is one of the most important sources of fundamental data (as will be explained later in the chapter).

During periods of capacity deficiency, the pipelines curtail the volumes of gas flowing to the interruptible shippers using the convention of elapsed prorated scheduled quantity (EPSQ), which distributes curtailments according to the elapsed time of the gas flow. For example, if gas has been already flowing for a third of the gas day, the reduction could be as much as 2/3 of the daily volume. If the flow started later during the day, the cut would be smaller.

The demand for natural gas is highly cyclical, with both high frequency and seasonal fluctuations driven by the weather conditions. Most end users, especially power producers and local distribution companies, cannot exactly predict their daily and hourly loads. They manage their needs by diverging from the volumes of natural gas to which they are entitled under the contracts with the pipelines. The differences between daily and hourly takes and contractual volumes (imbalances) are settled through a process called balancing, which takes place under the conditions specified in the tariffs.

In the early days of the natural gas business, prior to deregulation

and the emergence of trading operations, balancing would typically happen on a monthly basis. A shipper who exceeded the contracted volumes by a margin allowed under the tariffs and negotiated agreements would have to return the excess volume to the pipeline in a number of ways:

- ❑ over/under delivering the volumes over the next time period;
- ❑ buying natural gas in the market and returning it to the pipeline; and
- ❑ finding another shipper with an opposite imbalance and negotiating a swap.

This system functioned reasonably well prior to deregulation, when the relationships between the pipelines and their customers were based on long-term relationships and cooperation rather than competition. The arrival of energy trading entities changed the rules of the game. Many players took advantage of the system and developed a business model based on extracting natural gas from the pipelines under the conditions of scarcity and high prices, and returning the excess volumes when prices dropped. The trading operations effectively had free options that they were exercising very aggressively. This led to the switch to more frequent balancing, reconciling differences within weeks or days.

Daily balancing does not prevent very large, temporary imbalances that occur when the shippers exceed their contractual volumes (typically during the periods of extreme weather) or leave gas in the pipelines (when the weather is mild). In both cases, the pipeline pressure may deviate from safe levels and undermine pipeline integrity. The first line of defence for a pipeline in such situations is to fall back on natural gas from the storage facilities used as a buffer to maintain pipeline integrity. Many pipelines either invest in their own storage capacity or acquire storage space in order to protect themselves against imbalances. If the pipeline gas inventory is significantly out of step with demand, the pipeline can declare in such case what is known as an operational flow order (OFO) notice – which imposes, in addition to strict requirements to balance gas volumes daily within tolerance bands, heavy penalties for short-term imbalances. The penalties may be quite severe (in some cases, exceeding US\$50/MMBtu), leading to a hectic race to acquire or sell natural gas

at certain locations, in order to meet pipeline requirements. Penalties are typically defined as the market price of natural gas plus non-compliance charges.

An example of a Pacific Gas and Electric (PG&E) OFO settlement agreement, which the California Public Utilities Commission (CPUC) approved on February 17, 2000, can be found at: [http://www.pge.com/pipeline/library/regulatory/ofo\\_update\\_20000401.shtml](http://www.pge.com/pipeline/library/regulatory/ofo_update_20000401.shtml). An OFO will only be called if inventory is forecast to fall outside the established pipeline inventory limits (the pipeline inventory limits are 600 MMcf apart). The OFO typically states:

- the OFO stage;
- the system inventory level (high or low);
- the noncompliance charge; and
- the tolerance band (percent of allowable variance between scheduled volume and actual volume).

The penalties in the case of PG&E and tolerance bands are shown in Table 10.4.<sup>20</sup>

As one can see, the penalties may be quite severe in the case of Stage 3 or 4 OFO. This explains why an OFO may lead to a spike in prices (to the downside or, more often, to the upside). A physical trader will be willing to pay a very high price (or accept a very low price) to avoid penalties. An efficient trading operation would have procedures in place to monitor conditions of the pipelines at the locations at which it has positions because failing to anticipate an OFO or missing an announcement even by a few minutes may be very costly. Some firms offer a service under which they continuously monitor the pipeline websites and notify the clients by email when an OFO is declared.<sup>21</sup>

### **Pipeline regulation**

From the point of view of regulatory oversight, pipelines may be classified as interstate pipelines or intrastate pipelines. One straightforward definition of an intrastate pipeline is based on its path: a facility completely enclosed within the borders of one state and not crossing the state line. Intrastate pipelines are not subject to the oversight of the Commission (ie, the FERC). A special category of intrastate pipelines are the so-called Hinshaw Pipelines, which

"operate within a single state, but can receive gas from outside their state without becoming subject to the Commission's NGA jurisdiction."<sup>22</sup>

*Regulation by FERC of interstate natural gas pipelines*

Under the Natural Gas Act of 1938 (NGA), the transportation of natural gas in interstate commerce is subject to FERC regulation. For FERC's regulatory purposes, a "transportation" service includes a storage service. The authority of the FERC to regulate to regulate the services rendered in natural gas pipeline transportation in interstate commerce includes:

- certification and construction of new facilities;
- extension or abandonment of services and facilities;
- maintenance of accounts and records;
- acquisition and disposition of facilities;
- initiation and discontinuation of services; and
- various other matters.

Under the NGA, the rates for service on interstate facilities must be just and reasonable and not unduly discriminatory. The pipelines are operated pursuant to tariffs which set forth terms and conditions of service. These tariffs must be filed with and approved by the FERC pursuant to its regulations and orders. The tariff rates may be lowered by the FERC, on its own initiative, or as a result of challenges to the rates by third parties if they are found unlawful and the FERC could require refunds of amounts collected under such unlawful rates. The rates are derived based on a cost-of-service methodology.<sup>23</sup>

The NGA gives the FERC a lot of discretion in the rate-making

**Table 10.4** Example of OFO penalties

Tolerance band as % of usage	Noncompliance charge US\$/Dth
Stage 1	up to ±25%
Stage 2	up to ±20%
Stage 3	up to ±15%
Stage 4	up to ±5%

Source: [http://www.pge.com/pipeline/library/regulatory/ofo\\_update\\_20000401.shtml](http://www.pge.com/pipeline/library/regulatory/ofo_update_20000401.shtml)

process. The cost-of-service approach, known also as a revenue requirement, calls for the identification and determination of different cost items that should be covered by the revenues that the pipeline is allowed to earn. The next step is the assignment of the cost to individual customers through the rate-design process. The guiding principles of this process are as follows:<sup>24</sup>

- ❑ *Functionalisation.* The customers pay only for the services, or pipeline functions, they use.
- ❑ *Cost classification.* Cost classification is based on the fundamental distinction developed in microeconomics between fixed and variable cost. The fixed cost does not vary with the level of output or service and, in the case of pipeline, includes depreciation, interest payment, rent for leased office buildings and other facilities. The variable cost changes depending on the output level and includes for example compressor fuel.
- ❑ *Cost allocation.* The costs are allocated to different classes of customers making them responsible for recovery of the revenue requirements for which they are responsible.
- ❑ *Rate design.* Order 636 adopted the concept of the SFV rate design. Under this design, the fixed costs are covered through a monthly reservation charge. The approach is not mandatory but the FERC made it clear that this a recommended approach. The FERC reversed in Order 636 its previous adoption in 1983 of the MFV method, which “was devised to help pipelines sell gas by moving all fixed costs except for return on equity and related taxes to the reservation charge.”<sup>25</sup>

A cost of service includes the following components:

- ❑ operation and maintenance expenses;
- ❑ depreciation, depletion and amortisation;
- ❑ taxes; and
- ❑ return on investment.

Return on investment is calculated by multiplying the rate base by the allowed rate of return.<sup>26</sup> The rate base includes:

- ❑ utility plant less accumulated depreciation, depletion and amortisation;

- working capital;
- allowable unamortised development and research costs; and
- less: accumulated deferred income taxes;

The permitted rate of return determination is, as one can expect, more art than science. A number of different methods are used, including:

- discounted cashflow;
- capital asset pricing model;
- a risk-premium analysis; and
- a comparable earnings analysis.

The entire approach to rate making can be summarised through the following identity:<sup>27</sup>

$$\text{Rates} \times \text{expected volumes} = \text{Reasonable expenses} + (\text{Rate base} \times \text{Permitted rate of return})$$

The construction of new pipelines often results in the discrete and sudden realignment of prices, as previously stranded gas can be taken to more lucrative markets, and existing markets are flooded with cheap supply from alternative sources. The price changes can be quite dramatic. The expansion of the Rex pipeline was a very positive development for producers in the Rockies, who obtained access to new markets and saw a significant increase in prices relative to the Henry Hub. Producers of natural gas in the Gulf of Mexico and Midcontinent saw this development as a modern version of Sherman's march through Georgia, as their volumes were pushed from the markets back into the supply areas. This explains why monitoring pipeline projects and analysing their potential market impact is critical to any trading operation. This is equally important to equity stock analysts. New pipelines result not only in more gas-on-gas competition but also render parts of the existing physical infrastructure of the industry obsolete in an economic sense. Some older pipelines may become empty and storage facilities may be idled, with big potential consequences to the stock prices of midstream companies. It is not unusual to see a pick-up in the swap and option activity before significant additions to pipeline grid. More agile traders are trying to take advantage of uninformed

counterparties, who tend to read only the sports pages in the newspapers.

A critical step in any pipeline project is obtaining a certificate of public convenience and necessity from the FERC pursuant to Section 7c of the NGA and a number of FERC policy statements. The FERC seeks to determine whether a new project is in the public interest and, if this is the case, how it should be paid for (on an incremental or a rolled-in basis). Policy statements of 1995 and 1999 specifically endorsed reliance on the “open season” process:

The Commission has a two-step process for determining whether the market finds an expansion project economically viable. The first step, which occurs prior to the certificate application, is for the pipeline to conduct an open season in which existing customers are given an opportunity to permanently relinquish their capacity. This first step ensures that a pipeline will not expand capacity if the demand for that capacity can be filled by existing shippers relinquishing their capacity. [...] The second step is that the expansion shippers must be willing to purchase capacity at a rate that pays the full costs of the project, without subsidy from existing shippers through rolled-in pricing.<sup>28</sup>

A transparent process for new pipeline approvals helps fundamental analysts to monitor developments on this front and anticipate price realignments.

#### *Intrastate pipeline regulation*

Intrastate natural gas pipeline operations generally are not subject to rate regulation by the FERC, but they are subject to regulation by various state agencies (typically, public utilities or public services commissions) in states where they operate.<sup>29</sup> One exception to this general rule is the use of intrastate pipeline systems to transport natural gas in interstate commerce. The rates, terms and conditions of such transportation services are subject to FERC jurisdiction under Section 311 of the Natural Gas Policy Act (NGPA), which regulates, among other things, the provision of transportation services by an intrastate natural gas pipeline on behalf of a local distribution company or an interstate natural gas pipeline. Section 311 of the NGPA allows an interstate pipeline company to sell or transport gas on behalf of any intrastate pipeline or local distribution company without prior FERC approval. Under Section 311, rates charged for transportation must be fair and equitable, and amounts

collected in excess of fair and equitable rates are subject to refund with interest.

### **MONITORING NATURAL GAS PIPELINE FLOWS<sup>30</sup>**

Energy trading is an information-intensive activity requiring the continuous processing of huge volumes of data, and physical players have historically had a significant competitive advantage over purely financial traders. Their involvement in handling commodity flows has provided these players with unique insights into the market impact of changes in physical infrastructure and levels of demand and supply at different locations. Over time, however, improvements in access to market data and the information technologies used to process this data have eroded this competitive edge, opening new channels of information flows and allowing for the instant acquisition and assessment of vast volumes of data. This trend is especially visible in the North American natural gas business, in which daily availability of immense datasets related to natural gas flows in the pipeline grid is transforming the market for physical natural gas. Due to the tight interaction between physical natural gas and other related markets – including financial, gas, power and other energy commodities – this development is exerting an influence over the entire energy complex.

#### **Physical pipeline flows**

The flow of natural gas occurs through the complex interaction of many factors; three of the most important will now be discussed.

##### *Seasonality*

The consumption of natural gas is characterised by seasonal patterns, with consumption peaking in winter, dropping in spring and then increasing to a smaller seasonal peak in summer, due to demand for air-conditioning. Production flows, however, tend to be distributed uniformly over time. The production of natural gas is concentrated in certain parts of the US and Canada, often far removed from the residential, commercial and industrial demand primarily located near large population centres. This results in an uneven distribution of flows over the pipeline system, with certain pipelines having excess capacity during some periods of the year and running full during periods of high load in the markets they are

serving. Both production and demand are influenced by weather-related shocks, which, in the case of demand, interact with seasonal patterns. Demand shocks are due to periods of unseasonably warm or cold weather, or disruptions in other parts of the energy complex. For example, a nuclear power plant outage often creates additional demand for natural gas related to the generation of replacement power. Shocks to production result from events that disrupt the normal operations of the natural gas industry infrastructure. Hurricanes are the most obvious examples of such shocks, but there are also many more pervasive and persistent events causing localised problems, which are less well known to a casual observer of the industry – such as mechanical failures or wellhead freeze-offs. These can occur due to severe cold weather in production fields, and can temporarily reduce gas flows during periods of highest demand.

#### *Operational constraints*

The shifting distribution of natural gas flows in the pipeline grid and the operational characteristics of the physical pipelines themselves often result in constraints, which can lead to price spikes at certain locations, both to the upside and the downside. The price spikes are often mirror images of each other. A constraint in the pipeline grid may suppress prices in the producing region and create shortages and higher prices in the market region.<sup>31</sup> Successful traders, armed with reliable weather forecasts, information about current pipeline operational capabilities and up-to-date information covering the levels of production and demand, are able to anticipate constraints in the pipeline grid and profit from them.

#### *Changes to infrastructure*

In the longer run, changes of natural gas flows in the pipeline grid result from changes to physical infrastructure, such as the addition of new power and industrial plants. The biggest shocks result from discrete additions or eliminations of pipeline system capacity – both can change the balance between supply and demand in the production regions and the market regions that a new pipeline serves.

One example of such a fundamental shift followed the construction of the Alliance pipeline from Alberta to the Chicago area. The project dramatically changed the basis prices at Alberta's AECO trading hub and at Chicago Citygate. Another example is the Rockies

Express pipeline from the prolific Rocky Mountains producing region to the high-demand regions of the eastern US. The ability to anticipate the changing patterns of natural gas flows in the pipeline grid is critical to the success of an energy trading operation.

The most important structural shift reshaping the US natural gas industry is the changing spatial production profile. Increased production from shale formations, distributed across the US territory, interacts with evolving pipeline grid and results in new patterns of natural gas flows between different regions. Reliance on historical experience may lead to incorrect trading decisions and flawed forecasts: the industry needs granular and up-to-date information about production and demand, and the pipeline flow data is the best window into the rapidly evolving landscape of the industry.

#### *Natural gas storage forecasts*

Another critical insight one can derive through monitoring natural gas flows is the ability to better forecast changes to the levels of natural gas inventories. Consider the following implications of two important characteristics of the US–Canadian natural gas grid. First, the US–Canadian gas grid comes very close to being a closed system (ignoring, for the moment, LNG flows and flows through the Mexican border). In a closed system, production is equal to consumption plus the change of the inventory level of natural gas in storage. By monitoring the injections of natural gas at receipt points and withdrawals at delivery points, up-to-date information about changes to production and consumption levels can be developed. A point can be a physical location (for example, a natural gas processing facility or an electric power plant). Alternatively, measurement points can be meters or compressor locations on gas mainline transportation systems that record receipt or delivery information. Measurement points can also comprise energy storage locations or portions of a transmission system (for example, segments of a natural gas pipeline). Additionally, certain receipt/delivery points can be associated, uniquely and exclusively, with natural gas storage facilities. By observing the flows at these points, it is possible to assess injections and withdrawals into and out of storage and thus estimate the level of storage inventories.

Second, using flow data from specific receipt/delivery points, it is possible to estimate the impact of changes in pipeline line fill.

Pipelines can absorb, and later release, excess flows of natural gas during periods of short-term downward spikes in demand. The pipelines can accomplish this by adjusting pressure, within the tolerance limits allowed by safety standards and regulations. During periods of upward spikes in demand, pipelines can provide additional gas supply by reducing pressure. On many occasions, natural gas pipelines, in anticipation of lower or higher demand, decrease or increase gas pressure, respectively, to accommodate fluctuating gas flows. In pipeline parlance, these activities are said to be adjusting pipeline line fill (or line pack). The line fill or line pack can be considered an extension of the natural gas storage facilities, a natural buffer between the producers and consumers of natural gas. The changes in the line fill often complicate the task of predicting storage levels. Even if one can produce a very precise forecast of production and demand, the difference between the two can be split between the storage injections/withdrawals (a reported number) and changes in line fill.

### **Data sources**

A few years ago, information on natural gas flows and capacities was generally unavailable or available only at a considerable expense. The emergence of companies that specialise in the collection and dissemination of natural gas flow data in the US and Canadian pipeline grid has brought about a fundamental change in the North American natural gas market.

The gap in the information about natural gas flows in the US was eliminated by regulatory and technological developments. On the regulatory side, the FERC initiated the process of deregulating natural gas prices and transportation capacity in the late 1980s. The FERC also instituted certain rules designed to provide greater market transparency to all industry participants. These rules were expanded over time, culminating in an extensive set of procedures issued on February 9, 2000, in Order No. 637. Among these rules is a requirement based on Standard 4.3.6, circulated by the former Gas Industry Standards Board (now the North American Energy Standards Board), requiring all interstate pipelines in the US to post “operationally available pipeline capacity” on their company websites several times each day.

In practical terms, this requirement commits interstate pipelines to

release each day information regarding the volume of all natural gas that they receive, transport, store and deliver within each delivery period, prior to the day of flow, with updates during the day of flow. Early postings indicate what is scheduled to flow in the day ahead. The later, or intra-day, postings show the final schedule and, in most instances, accurately represent actual physical flows. Thus, raw pipeline data on gas flows is theoretically available for near real-time analysis of supply, transportation, storage and demand. However, these datasets are posted independently on each pipeline's website in a variety of formats and structures. The presentation is generally designed for operational purposes and provides little or none of the aggregated information necessary to make the information usable for analytic purposes. The result is that raw pipeline gas flow and capacity data is extremely complex and difficult to interpret. Because the data structure from one pipeline to the next is so different, this "raw" natural gas flow and capacity data has proven to be of limited use to decision-makers for a number of years. However, earlier in the 2000s, the usefulness of this data was improved significantly by companies providing services for the collection and analysis of pipeline data. This information is provided by a number of data vendors, including LCI Energy Insight (LCI), Genscape and BENTEK Energy.<sup>32</sup> These companies use proprietary software to extract the information about pipeline delivery and receipt volume data that all interstate pipeline and storage companies are required to post on their websites. They use technology that automatically gathers, validates, standardises and warehouses the data, subsequently delivering the date in a variety of formats.

Some of these data providers enhance the value of this information through the association of different delivery and receipt points with metadata that specifies the location, function, operator and other important information about the point. For example, each receipt or delivery point may be identified by the type of natural gas infrastructure associated with each point, such as a gathering system, natural gas processing plant, natural gas storage facility, a power plant or a local distribution system. Several of these data providers also offer analytical and decision-support systems, which provide capabilities to transform these huge volumes of data into useful information resources about trends within the natural gas market. For example, receipts from gathering systems, gas plants

and other production points can be organised in a manner that allows analysts to compute regional gas production volume. Flows through pipeline hubs (multiple pipeline interconnections), constrained pipeline points and segments, and other gas flow locations, are used by analysts to identify bottlenecks and market imbalances. Pipeline deliveries can be identified in a standardised manner such that analysts can aggregate total demand at specific utility systems.

Until 2010, the framework developed by these companies to monitor natural gas flows did not provide coverage for certain pipelines that are not in the FERC jurisdiction – for example, intrastate pipelines. Such pipelines were initially not obligated to post natural gas flow and capacity data on their websites, leaving some very important regions with incomplete coverage. This is especially true of Texas, Oklahoma and Louisiana, where many pipelines of great importance to the industry and of considerable capacity are wholly contained within a state and do not cross any state borders. In 2006, natural gas volumes delivered to consumers from Texas, Oklahoma and Louisiana were 23.8% of the total gas deliveries within the US, while industrial and electric customers in these states consumed 33.9% of total US industrial and power generation demand. One example in Texas is the Oasis pipeline serving the Houston Ship Channel and Katy locations, which are critically important to the US manufacturing industry, given the high concentration of chemical plants in the area surrounding Houston.

The gap in the information about natural gas flows in the US was eliminated for a short period of time due to an initiative taken by the FERC. On April 19, 2007, they issued a notice of proposed rule-making (NOPR) that put on the table an extension of interstate reporting rules to intrastate pipelines. Based on authority granted to the FERC in the Energy Policy Act of 2005, these new rules became binding on October 1, 2010, under FERC Order No. 720. The intrastate pipelines successfully contested the order in the court system. The US Court of Appeals for the Fifth Circuit vacated FERC's Order Nos. 720 and 720-A (the Texas Pipeline Association and the Railroad Commission were the petitioners). The Fifth Circuit held that Order Nos. 720 and 720-A exceeded the scope of the FERC's authority under the NGA. The FERC argued the NGA and Section 23 of Energy Policy Act (EPA) of 2005 authorised

the Commission to require that intrastate pipelines post daily nominations.

For the flow data in the case of BENTEK Energy (the company the author is most familiar with), the data is collected twice a day (evening cycle and I2 cycle) and delivered in the morning between 5 and 6 am Mountain Time. The number of records for the interstate data is around a magnitude of 25,000 records. This means that a client receives about 50,000 records a day from two cycles. Many energy trading organisations invest heavily in data processing, concentrating on the visualisation of the data (to allow traders to identify quickly any emerging patterns). This is important, as the data arrives when physical natural gas trading gets under way. Other research efforts are related to building regional models of natural gas flows in order to predict the price impact of changes in local consumption levels (driven, in the short term, primarily by weather) on prices. Other uses of the flow data include:

- ❑ the monitoring of production and demand trends in real time;
- ❑ improved quality of natural gas demand and storage level forecasts;
- ❑ forecasting of pipeline flows and pipeline capacity utilisation rates;
- ❑ optimisation of transportation, storage and supply contracts by the end users of natural gas (LDCs and big industrial/commercial users); and
- ❑ in the future, availability of the pipeline flow information will facilitate the algorithmic trading of physical natural gas.

Trading organisations with limited resources rely on daily summary reports available from many different sources specialising in the analysis of flow data. Some sample reports can be found, for example, on the website of BENTEK Energy.<sup>33</sup>

### **LIQUEFIED NATURAL GAS**

LNG is made by cooling methane to about -260°F (about -160°C). This results in the condensation of gas into liquid form under normal atmospheric pressure. This process shrinks the volume of natural gas to about 1/600 of the original level in gaseous form, making transportation over long distances economically viable. LNG is key to the

development of the world's natural gas markets. From the point of view of the producing nations with reserves removed from the consumption centres, LNG is one of several technologies available to deliver the commodity to end users. There are other reasons why LNG is an important component of the natural gas industry. On the supply side, one persistent aspect of available supply is the existence of significant volumes of stranded gas. This term refers to natural gas produced along with oil that is flared because there is no local demand. This term is used also with respect to natural gas reserves that cannot be used locally.

An inability to export natural gas from some locations is related to the limitations of the distance over which natural gas can be shipped profitably by pipeline. The estimates differ and depend on the parameters of the pipeline (for example, its diameter) and the character of the terrain over which the pipeline would have to be built. The breakeven points between pipelines and LNG tankers range between 1,500 and 3,000 miles.

Geopolitical issues are another important driver behind the trend towards increasing reliance on LNG. Historically, many western European countries have been dependent on imports of natural gas from Algeria and Russia, transported over long-range pipelines. The need to diversify supply sources and reduce the risk of the unexpected curtailment of delivered volumes has led many European countries to a greater reliance on LNG. In the last few years, disputes between Gazprom and Ukraine (and occasionally Belarus) were occurring with an almost clockwork regularity every winter, exposing many importing countries in western Europe to serious gas shortages. About 25% of western Europe natural gas consumption comes from Russia, with about 80% of the volumes crossing Ukrainian territory in the past. However, this is due to change with the construction of two undersea pipelines bypassing Ukraine and other countries. The Nord Stream twin pipeline runs from Vyborg, Russia, to Lubmin, Germany (Line 1 was completed by March 2012). South Stream, under the Black Sea, is under consideration, with construction likely to start (as of the time of writing) at the end of 2012.

From the point of view of consuming nations with an ageing reserve base or no domestic reserves, LNG often is seen as a critical component of the supply portfolio. In the case of the US, the growth

of the LNG industry was driven by the expectations of imminent shortages of natural gas in the US, with domestic demand outstripping the North American supplies. Forecasts available from the US government and industry experts inevitably contained a graph of a widening wedge between US production and consumption of natural gas, with LNG imports seen as a solution to future shortfall. Close to 30 applications were filed with the FERC for permits to build new regasification facilities, with the total number in the spring of 2012 reaching 12 US installations with a total capacity of about 15 Bcf/day. Parallel developments in countries producing natural gas consisted in the expansion of liquefaction plants intended for exports to the US, and LNG tankers built with the same objective in mind.

The reality turned out to be very different. The expansion of natural gas production from non-conventional sources invalidated dramatically the predictions of future US dependence on imports of LNG. As a matter of fact, the US currently indirectly exports LNG through displacement. LNG produced in facilities built in the expectation of US demand now flows to other markets, primarily in Europe and the Far East (South Korea, Japan and Taiwan). The existing regasification terminals in the US are greatly under-utilised and may be converted into re-export or even liquefaction terminals.

The LNG market is key to the shape of the future natural gas industry. The optimists see LNG as a crucial link between different regional markets, with common pricing structures emerging in the North Atlantic and the Far East, leading to the emergence of global, integrated natural gas market. The pessimists point to the periodic discrepancies between liquefaction capacity and regasification capacity. What may be looming, they warn, are periods of shortages and destructive competition between consuming nations for the limited output, alternating with periods of excessive liquefaction capacity and depressed prices.

### **LNG technology and markets**

LNG supply chain consists of three main links:

- liquefaction facilities;
- transport by LNG tanker; and
- regasification.

Further components of the system include the producing field and, depending on the location of the field, a pipeline to the shore where a liquefaction facility and loading terminal are built. In addition, natural gas may have to be treated to satisfy the requirements of the liquefaction process in the producing area and/or the processes at the regasification terminal required to meet the pipeline gas quality specifications.

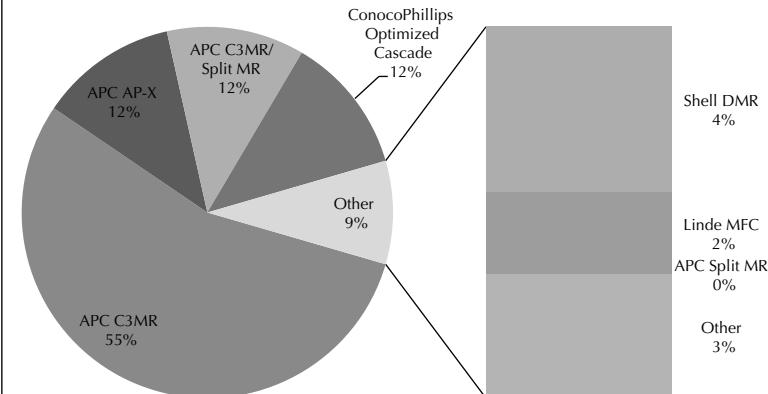
Liquefaction is a process of reducing the temperature of natural gas to the point at which it remains liquid at normal atmospheric pressure (about  $-161^{\circ}\text{C}$ ). The process of liquefaction may be carried out under pressure and, once gas is liquefied, its temperature has to be further lowered as normal pressure is being restored. A unit for the liquefaction of natural gas is called a train. Trains range in capacity from three to eight million tons per year. The trend is towards developing bigger trains in order to exploit economies of scale.

There are several liquefaction technologies used worldwide, with different versions of Air Product technology used in about 80% of trains, followed by ConocoPhilips Optimized Cascade. LNG liquefaction capacity by technology type is showed in Figure 10.3.<sup>34</sup>

By the end of 2010, total liquefaction capacity had reached 270.9 MMtpa (millions ton per annum), with 94 liquefaction trains in operation. The breakdown of capacity by country is shown in Table 10.5 and Figure 10.4. The largest increment of liquefaction capacity between 2006 and 2010 took place in Qatar. In coming years, capacity expansion will happen primarily in Australia, given the discoveries of significant natural gas resources in the north of the continent (for example, the Gorgon project).<sup>35</sup>

LNG carriers can carry loads typically around 125,000 m<sup>3</sup>. As in the case of trains, the trend is towards building bigger carriers with some tankers under construction reaching the size of 215,000 m<sup>3</sup>. By the end of 2010, the world's LNG fleet consisted of 360 ships with a combined capacity of 53 million m<sup>3</sup>.<sup>36</sup>

The tankers are constructed either using integrated-tank technology, with forces exerted by the cargo transmitted to the ship's inner hull by a metallic membrane,<sup>37</sup> or based on a number of self-supporting spherical tanks, a technology originated by Norwegian firm Moss Rosenberg. The heat loss that happens during transportation yields natural gas known as boil-off, which is used for the

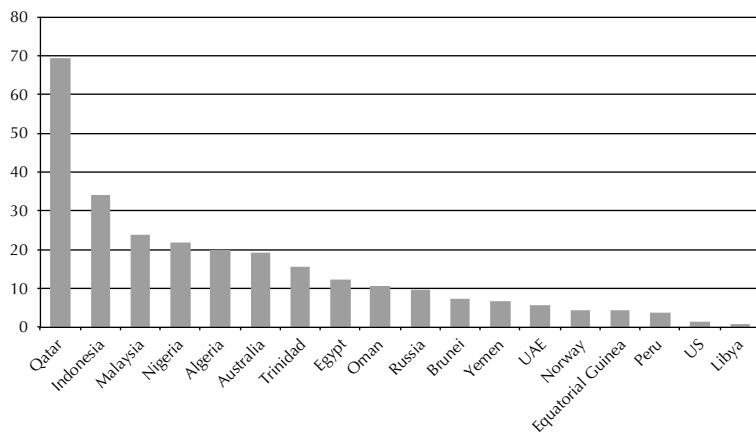
**Figure 10.3** Liquefaction capacity by technology (2010)

Source: International Gas Union (<http://www.igu.org/igu-publications-2010/IGU%20World%20LNG%20Report%202010.pdf>)

tanker's propulsion system, amounts to about 0.2% of the LNG cargo per day. New technological developments reduced the boil-off to 0.1% of the cargo per day.<sup>38</sup>

A very important factor increasing LNG market flexibility will be the evolution of transportation and regasification technology. The capacity of tankers tends to increase over time and fuel efficiency improves. This trend is exemplified by the introduction of a new class of tankers, the Q-max and Q-flex ships. Q-max is a category of LNG membrane tankers that were introduced in 2007, with a capacity of between five and six Bcf. The letter "Q" stands for Qatar and "max" signals that these tankers are the biggest ships able to dock in Qatar. The Q-max ships have a re-liquefaction system that reduces natural gas losses due to boil-off. This improves energy efficiency and reduces methane emissions (ie, the emission of a potent green house gas (GHG)). The Q-flex tankers range from 210 to 216,000 m<sup>3</sup> in capacity (smaller than Q-max tankers reaching 266,000 cubic metres) and use slow diesel engines (unlike traditional LNG tankers burning boil-off), that help to improve energy efficiency. The use of larger, more efficient tankers reduces the unit cost of transporting LNG.

Two other LNG-related developments require additional attention. One is the increasing proliferation of flexible storage and

**Figure 10.4** Liquefaction capacity by country, MMtpa (2010)

Source: International Gas Union, <http://www.igu.org/igu-publications-2010/IGU%20World%20LNG%20Report%202010.pdf>

**Table 10.5** LNG liquefaction by country (MMtpa, 2010)

Country	MMtpa
Qatar	69.2
Indonesia	34.1
Malaysia	23.9
Nigeria	21.9
Algeria	19.9
Australia	19.3
Trinidad	15.5
Egypt	12.2
Oman	10.8
Russia	9.6
Brunei	7.2
Yemen	6.7
UAE	5.8
Norway	4.5
Equatorial Guinea	4.5
Peru	3.7
US <sup>39</sup>	1.5
Libya	0.7
Total capacity	271.0

Source: International Gas Union, <http://www.igu.org/igu-publications-2010/IGU%20World%20LNG%20Report%202010.pdf>

regasification units (FSRUs). Energy Bridge is a technology developed initially by El Paso that places the regasification plant on the tanker. As they explain, "Energy Bridge is the proprietary offshore liquefied natural gas (LNG) regasification and delivery system developed by Excelerate Energy. [...]Energy Bridge Regasification Vessels, or EBRVs, are purpose built LNG tankers that incorporate onboard equipment for the vaporisation of LNG and delivery of high pressure natural gas." This solution addresses two critical issues related to LNG. In many potential destinations there is a strong grassroots opposition to the construction of LNG plants, given safety concerns.<sup>40</sup> In many locations, there are no regasification plants and construction of the offshore collection system is a cheaper, time-efficient alternative.

Shell's floating LNG (FLNG) facility<sup>41</sup> places the liquefaction plant on the tanker. This technology will allow the production of LNG by accessing natural gas produced in the offshore fields, pre-empting the need to develop an underwater pipeline and expensive onshore infrastructure.

Regasification is carried out in vertical tubes or coils submerged in heated water or exposed to water running outside the tube. Vapourised gas may be further treated, especially if it originates from reservoirs rich in heavier hydrocarbons that have to be removed from the stream before it enters the pipeline system. A regasification plant may be connected to a local storage facility serving as a buffer between the unit and the pipeline system.

As of the end of 2010, world regasification capacity had reached 572 MMtpa in 83 plants (10 of them floating units).<sup>42</sup> Most of the capacity is concentrated in Japan (31.5%) and the US (19%), followed by Korea (14.6%), Spain (6.8%) and the UK (5.8%). The US capacity is, of course, massively underutilised, a good illustration of the risk of committing capital in this industry. Regasification capacity is expected to increase by 110 MMtpa by 2015.

### The costs of the LNG supply chain

The entire integrated LNG supply chain is characterised by high capital intensity and high initial cost. This explains why LNG projects are being undertaken by some of the biggest energy companies with the strongest balance sheets, and are often operated through joint ventures with other partners in order to spread the

risks and obtain access to sources of supply and distribution infrastructure.

There are several studies in the public domain that have produced estimates of the unit cost of bringing one MMBtu of natural gas to a specific market. In the case of the US market, many estimates revolve around the magic number of US\$3.50 per MMBtu, creating an impression that as long as the domestic prices at Henry Hub exceed this threshold, flows of LNG will materialise. Such estimates should be taken with a pinch of salt, not because they represent a superficial, back-of-the-envelope number, but because there are many other factors to consider for which it is difficult to obtain reliable forecasts.

First, such estimates are ageing very quickly, given the dynamic nature of the industry. Over the last few years we have seen a race between growing costs, due to an increase in prices of certain materials (especially high-quality steel and cement), and technological improvements to the design of tankers and liquefaction facilities. Larger liquefaction trains and carriers reduce the unit cost. The other side of the coin is that the gas field behind the liquefaction facility has to exceed a certain size to support a large plant.

Second, the cost estimates are often based on certain generic assumptions – such as the cost of natural gas or the definition of the liquefaction capacity. A frequent assumption is that, at many locations, the cost of natural gas is zero or negative because it is flared and, in addition, valuable gas liquids can be removed from the stream and monetised. This may be true, but natural gas has to be delivered to the plant and this requires the construction of a gathering system. Many liquefaction plants are located in places that lack a basic infrastructure and the cost of developing access roads, power grid and ports has to be considered. The basic infrastructure may be subsidised or shared by many projects and the related costs may be excluded in the final count, but cannot be completely ignored.

Finally, the cost of the entire chain cannot be estimated by ignoring the distance between the point of supply and the destination of LNG tankers. A longer distance translates into higher fuel usage, requiring more tankers to shuttle between the two points, creating more potential for scheduling and logistical problems.

The best (although a relatively old) estimate of the cost of a new LNG chain is available in a report by Simmons and Company (see

Table 10.6), which has the unit cost of LNG for a relatively large liquefaction plant at 7.5 million metric tons of LNG, 1 Bcf/d regasification unit and a 6,500 nautical miles distance. The breakeven cost is calculated assuming an 8% return on different operations, except for the upstream, where higher risk of exploration and production requires a higher return (15% was assumed).

Estimates from Wood McKenzie Ltd (as of the time of writing) reflect the escalation of costs over the last ten years. The cost of a five MMtpa integrated LNG project is estimated to be between 11 and 16 billion USD (four to six billion upstream, four to six billion for the liquefaction plant, one billion for transportation and two to three billion for regasification).

## NATURAL GAS STORAGE

### Natural gas storage facilities

Natural gas storage facilities represent an important buffer between supply and demand. Production takes place at roughly equal rates through the year, but demand tends to be highly cyclical. The excess supply is injected during summer into storage and withdrawn during the winter months. Understanding the operations and technology of natural gas storage facilities is important for anybody trading natural gas. EIA storage reports, covered in detail below, offer a quick way of assessing supply and demand trends, and predicting the future direction of prices. Many trading desks develop trading strategies around their real or perceived superior ability to predict storage levels and to front run other trading operations ahead of weekly storage reports released at 9.30 am Central on Thursdays. Storage facilities can be looked at as portfolios of calendar spread options, and some companies use mathematical algorithms to trade

**Table 10.6** Cost of the LNG supply chain

	Capital US\$MM	Variable US\$/MMBtu	Cost US\$/MMBtu	Required US\$MM/year	Return US\$/MMBtu	Breakeven US\$/MMBtu
Upstream	1,644	0.30	56	0.15	247	0.68
Liquefaction	1,950	0.18	91	0.25	156	0.43
Shipping	1,572	0.22	60	0.16	126	0.34
Regasification	877	0.08	55	0.15	70	0.19
<b>Total</b>	<b>6,042</b>	<b>0.77</b>	<b>261</b>	<b>0.71</b>	<b>598</b>	<b>3.13</b>

Source: Simmons & Co

around their storage positions. Some traders see storage as a way of establishing speculative positions and wait for a hurricane or other types of extreme weather. This is usually an expensive way to speculate (injections and withdrawal costs eat into profits), and works only if conditions develop which place a sufficiently high premium on physical gas. This strategy may be recommended at locations where there is no liquid forward market and where spot prices are characterised by frequent spikes.

Two basic types of storage are cryogenic storage and underground storage. Cryogenic storage is used for peak shaving, the delivery of natural gas to satisfy local needs during the conditions of excessive stress, and when demand spikes or interruptions of supply occur due to accidents and/or pipeline outages. This type of storage relies on relatively small LNG tanks using either double wall or membrane technology. Some storage tanks may be buried underground for safety reasons.

Underground storage uses three basic technologies that have in common the fact that they all take advantage of specific geological formations but vary in terms of capacity and the rates at which natural gas can be injected and withdrawn. All underground storage facilities require cushion gas, which remains permanently in the facility and is treated as part of the investment outlay. The function of this gas is to maintain internal storage facility pressure to allow for the removal of working gas that is cycled in and out.

Depleted oil and natural gas fields have rock formations characterised by high porosity and permeability. Their geological characteristics are well understood because they have contained oil and/or natural gas in the past and can be quickly converted into storage facilities. They are typically operated as seasonal storage, with the injection season extending typically from March/April into October/November. They are emptied during the winter months between December and February, although the withdrawal season may start in November and extend into March.

Aquifer storage uses water-holding rock formations that can be reconditioned for natural gas storage. The cost of developing aquifers is higher than the cost of developing depleted reservoirs, and they are used only in the areas where there are no better alternatives. They also require more cushion gas, increasing the initial cost.

Salt dome storage facilities are created as underground caverns by dissolving salt formations capped by impermeable rock. The salt dome storage is characterised by high deliverability – natural gas can be injected and removed quickly.

Depleted reservoirs and aquifers are used primarily to satisfy demand during winter and, given their characteristics, it usually takes the entire spring and summer season to fill them. They may sometimes be used to satisfy summer peak load during periods of unusually hot weather when more natural gas is used as fuel in power plants, or during supply outages caused primarily by hurricanes. This can be done in two ways. The first line of defence is to divert natural gas earmarked for injections to market, extending the injection schedule into the autumn. In the case of serious market disruptions, natural gas may be withdrawn from storage and taken to the market. The operators of such storage facilities often talk about a “hurricane option” associated with natural gas storage. The energy trading floors are the only place in the US where people sometimes pray for hurricanes.<sup>43</sup>

### EIA storage reports

Natural gas storage inventory statistics are among the most important data points followed by the energy traders. Historically, this information was collected by the American Gas Association, but the task has now passed on to the EIA, a unit of the Department of Energy. The data is collected through the survey of big storage operators. “The [...] sample of 63 respondents is estimated to account for more than 90 percent of the average reported working gas in storage. [...] Respondents provide estimates for working gas in storage as of 9 am Friday each week. The estimates are released on Thursday between 10:30 and 10:40 am (Eastern Time) on EIA’s website, except for certain weeks that include Federal holidays.”<sup>44</sup>

The example of the screen containing EIA storage statistics is shown in Table 10.8. The release of the data coincides often with frantic trading and increased volatility, as traders who predicted the number correctly take profits and those who were wrong cut their losses. It is important to emphasise that the EIA-reported data show the level of inventories, not the level of injections and withdrawals. This distinction is important because the industry participants often refer to this report as injection/withdrawal data. In reality, the EIA

data allow calculation of implied injections (or withdrawals), by calculating the differences between two subsequent absolute levels of natural gas in storage. Such levels may, however, be affected by the reclassification of inventories (from cushion to working gas) and by other accounting changes or corrections of previous errors. This is an additional complication facing any fundamental analyst in this business.

Table 10.7 summarises the characteristics of different types of storage, while Table 10.8 contains the information about the existing US storage capacity by region.

The weekly report based on sample information collected through Form EIA-912, "Weekly Underground Natural Gas Storage Report" is supplemented by information collected on Form EIA-191, "Monthly Underground Gas Storage Report." This latter form is used to collect the census, not sample, data, which by definition will be more reliable. Figure 10.5 shows the difference between the two sets of data and, as one can see, the quality of weekly reports has been systematically improving, with some minor problems developing in 2007 and again in 2011. The monthly data does not receive the same level of attention as the weekly reports. This is true of all economic data. The first release, which often contains a large error, moves the market. Few analysts pay attention to subsequent revisions.

The storage number is important because it provides indirect information about current consumption levels and about how much natural gas will be available for the heating season or how much remains in storage. Low injections (relative to historical numbers and expectations) in the summer (or even a withdrawal) provide an indication that demand is high relative to production and that the industry may enter the heating season with a low inventory buffer.

**Table 10.7** Natural gas storage facilities – technical characteristics

Type	Cushion/ working gas	Injection period (days)	Withdrawal period (days)
Aquifer	50–80%	200–250	100–150
Depleted reservoirs	50%	200–250	100–150
Salt caverns	20–30%	20–40	10–20

Source: Current State of and Issues Concerning Underground Natural Gas Storage, FERC, Docket No. AD04-11-000.<sup>45</sup>

**Table 10.8** Weekly natural gas storage report

Released: January 10, 2008 at 10.30 am (Eastern time) for the week ending January 4, 2008.

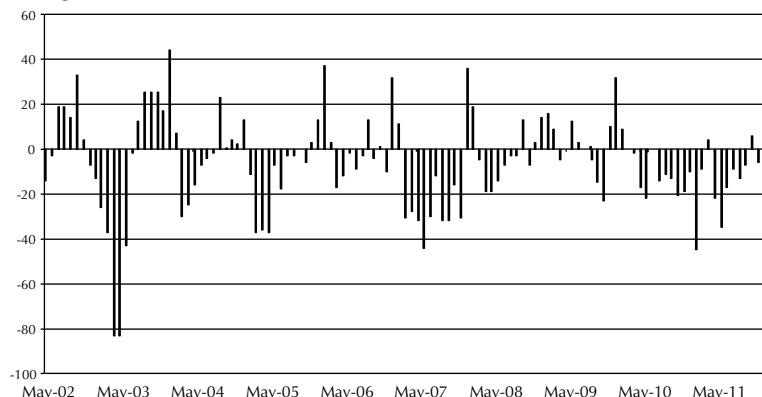
Next release: January 17, 2008

Working gas in underground storage, Lower 48

Region	Stocks in Bcf			Historical comparisons			
	01/04/08	12/28/07	Change	Year Ago (01/04/07)	Five-year (2003–07) average		
East	1,511	1,604	-93	1,716	-11.9	1,504	0.5
West	375	395	-20	382	-1.8	350	7.1
Producing	864	922	-58	934	-7.5	775	11.5
Total	2,750	2,921	-171	3,032	-9.3	2,628	4.6

Source: U.S. Energy Information Administration<sup>46</sup>

The market typically reacts to such developments by sending prices higher, both for the short term and for the winter months. High levels of natural gas in inventory close to the end of winter send a strong bearish signal. This is because many natural gas storage facilities have to be emptied towards the end of winter for technological or accounting reasons. The technological requirements are related to the potential loss of structural integrity or migration of natural gas away

**Figure 10.5** Differences between the weekly and monthly natural gas storage data

Source: U.S. Energy Information Administration

**Table 10.9** Summary of underground natural gas storage, by region and reservoir type, close of 2007

Region	Depleted Gas/Oil Fields			Aquifer Storage			Salt Cavern Storage			Total		
	Sites	Working Gas Capacity (Bcf)	Daily Withdrawal Capability (MMcf)	Sites	Working Gas Capacity (Bcf)	Daily Withdrawal Capability (MMcf)	Sites	Working Gas Capacity (Bcf)	Daily Withdrawal Capability (MMcf)	Sites	Working Gas Capacity (Bcf)	Daily Withdrawal Capability (MMcf)
Central	40	466	4,639	8	90	1,585	1	1	0	49	557	6,224
Midwest	88	927	22,549	31	271	5,890	2	2	85	121	1,200	28,524
Northeast	107	791	14,733	0	0	0	3	5	470	110	796	15,203
Southeast	27	149	3,160	3	7	68	4	39	3,622	34	195	6,850
Southwest	45	914	13,677	0	0	0	21	126	9,526	66	1,040	23,203
Western	19	281	7,372	1	22	850	0	0	0	20	303	8,222
Total US	326	3,528	66,130	43	390	8,393	31	173	13,703	400	4,091	88,226
(Marginal sites) <sup>1</sup>	(34)	(73)	(695)	(2)	(4)	(108)	(2)	(5)	(0)	(38)	(82)	(803)

<sup>1</sup> Marginal sites: very little or no activity reported during the 2007 calendar year. Marginal sites included in State/Regional total.

Note: Bcf – Billion cubic feet. MMcf = Million cubic feet.

Source: Energy Information Administration, Gas Transportation Information System, December 2008.<sup>47</sup>

from the facility. High levels of remaining inventories late in February or March may result in the natural gas market being flooded with unusually high (relative to demand) supply. A similar mechanism operates in late September or October if the natural gas storage facilities are nearly full. The producers may be forced to sell natural gas at depressed prices as the storage operators leave the market.

There are two warnings regarding this seemingly easy strategy. First, the full US storage capacity is an estimate based on technical studies and current operational practices. The operators may take risks and inject natural gas beyond rated capacity levels. Construction of new storage facilities and a demonstrable ability to squeeze more gas into existing fields has increased available capacity beyond 4,000 Bcf. EIA estimates are shown in Table 10.10. One should always remember that the capacity limit is an elastic rubber band and not a concrete wall, and be careful when making statements about running out of storage capacity when the summer inventories exceed 4,000 Bcf.

Second, the mechanisms described above are well understood, but many inexperienced energy traders tend to jump on the same band-wagon and make bets based on natural gas inventory levels. This can quite often create an overbought or oversold market, vulnerable to manipulation. In the commodity markets, it is necessary to always evaluate the fundamental trading strategy through the lens of existing trading positions. When almost everybody marches to the same tune the trading positions may become dangerously skewed and vulnerable to bear or bull raids.

The importance of natural gas inventory data resulted in the development of a cottage industry of forecasters who produce predictions of injections or withdrawals. There are two basic techniques that can be used to generate these forecasts. The first is based on the prediction of demand for natural gas over the reporting week (Friday morning to Friday morning). The US natural gas market is a contained system in which consumption is equal to production, with an adjustment for cross-border flows and changes in inventory level. The basic identity is given by:

$$\begin{aligned} & \text{Production} + \text{Imports} - \text{Exports} + \text{Storage withdrawals} \\ & - \text{Storage injections} = \text{Consumption} \end{aligned} \tag{10.1}$$

Imports and exports flow through a number of points and can be monitored through a few websites that provide up-to-date information. Production is fairly stable and can be guessed as a model implied flow (an explanation to follow). Natural gas consumption can be estimated using an econometric model that ties natural gas consumption to certain explanatory variables, related primarily to weather. Most models use actual temperatures or temperatures transformed into heating and cooling degree days.<sup>48</sup> Consumption (realised demand) data is available, with a considerable time lag, from the EIA, by state and by consumer class.<sup>49</sup> The regression model explaining consumption is run at the aggregate US level, at regional level or state by state and by end-user class. The model can be calibrated by using the estimated demand, in conjunction with reported storage data, to calculate the implied production. The implied production is calculated from equation 10.1, under the assumption that all the other items are known or are estimated correctly. After a few weeks of collecting implied production numbers, one has a fairly good guess of production levels and is ready to produce injection/withdrawal forecasts. This approach relies critically on the assumption that production changes very slowly from week to week. This is true in general, but such an approach may break down during periods of supply disruptions – for example, due to hurricanes.

This method, used across the industry, has some obvious shortcomings. First, EIA consumption data is fairly aged and usage levels may evolve quickly. Demand destruction due to high prices or slowing macroeconomic activity, especially in the case of industrial and commercial loads, can be very rapid. Second, some types of demand are not weather-sensitive. For example, industrial demand in states with a high concentration of chemical industries (Texas, Louisiana) that use natural gas as feedstock or fuel is not weather-dependent. Third, this approach misses additional storage capacity associated with line pack, the volume of natural gas in the pipeline system. The pipelines often take advantage of the ability to adjust internal pressure in order to store natural gas short term in anticipation of higher demand.

An alternative approach is based on monitoring the natural gas flows through specific meters associated with natural gas storage facilities. One can obtain very precise information about injections and withdrawals by monitoring nominations on the meters

associated with these facilities and use econometric techniques to extrapolate the injections/withdrawals for the monitored facilities to the rest of the industry. This approach has two shortcomings. First, many storage facilities are attached to intrastate pipelines that are not reporting daily nominations (with a few exceptions). The second, more serious, problem arises because many storage facilities are located behind city gates (ie, inside the local distribution company service area) and it is difficult to obtain information about the distribution of natural gas volumes flowing to different customers.

### **Storage contracts**

Storage contracts represent a careful compromise between meeting customers' needs and the necessity to maintain the integrity of the facility and its operational constraints. A contract giving a customer significant discretion with respect to the level and timing of injection/withdrawal decisions could create congestion, result in an inability to meet contractual obligations on a high demand day and lead to market disruptions and costly litigation.

Storage contracts include a number of standard terms that restrict a customer's freedom of action, such as:

- maximum storage quantity (MSQ);
- maximum daily injection quantity (MDIQ);
- maximum daily withdrawal quantity (MDWQ); and
- maximum daily transportation quantity (MDTQ).

The load-following contracts that require immediate response to the customer's need (for example, to support a peaking power generation unit) require additional contractual provisions, including:

- hourly injection quantity (HIQ);
- hourly withdrawal quantity (HWQ); and
- hourly transportation quantity (HTQ).

The contracts or tariffs specify the deadlines for nominations (typically by 10.30 am CST on the day preceding gas flow or an hour before the deadline for intraday nominations). The flexibility of a storage customer is also limited by the procedures of the pipeline connected to a given storage facility. The storage services provided,

as with the pipeline services, can be classified as firm or interruptible, within the limits defined by the maximum storage, injection, withdrawal quantities, as well as the constraints related to the load-following provisions.

The fees that are charged under a storage agreement include:

- reservation fee (applied to the maximum storage quantity);
- injection fee;
- withdrawal fee;
- fuel charge;
- hub services fee;
- parking services fees;
- loaning services fee; and
- authorised overrun charge.

#### *Storage rates*

The storage facilities in the FERC jurisdiction charge either cost-based rates (under Section 7(c) of NGA) or market rates under Section 311 of NGPA.

The cost-based method typically uses the equitable approach under which the fixed costs (such as capital costs, interest on debt, taxes) are recovered through rates:

- 50% based on storage deliverability; and
- 50% based on storage capacity.

while variable costs are recovered through injection and withdrawal charges.

Most low-deliverability storage facilities offer services that are cost-based. According to the FERC Staff Report,<sup>50</sup> a two-cycle per year facility costs US\$5–6 million per Bcf of working gas capacity. This compares with US\$10–12 million for a salt dome cavern with 6–12 cycles per year. The median cost-of-service per one Dth of working natural gas per year was estimated at US\$0.64 (based on the 20 tariffs on file). The corresponding number for salt caverns is close to US\$2.93 per Mcf. This is based on the following assumptions: a 5 Bcf salt dome facility, 13% targeted return-on-equity, 8% cost of debt, 100% debt to equity ratio, 34% tax rate, 3% state *ad valorem* tax rate, annual cost of service of US\$14.53 million, a 20-year life and a 10-year

useful life for the purpose of tax calculations. The cost-based rates allow the seasonal storage facilities to compete in the marketplace against more flexible competitors.

High-deliverability storage facilities are often operated by marketers and other entities that have a high-risk appetite, and are willing to take price risk. However, they require market-based rates, offering returns closer to 20% pa. Such rates can be granted by the FERC on condition the storage operator can demonstrate that they have no market power or that market forces will mitigate their market power. To demonstrate a lack of market power, the applicants must first identify the relevant product and geographic markets for the proposed storage services. The FERC's rule relies on a very general definition of the relevant product market and includes close substitutes for gas storage services.<sup>51</sup>

The analysis provided by applicants must identify the alternatives available to potential customers of the proposed storage services. The potential for exercise of market power is measured by analysing the size and concentration of the market. The market power screen is based on the Herfindahl–Hirschman Index (HHI). The threshold level for the HHI adopted by the FERC is equal to 1,800. The HHI is calculated by squaring the percentage market share of different participants in the relevant market (as described above), and calculating the sum of the squares. The maximum value of the index is

**Table 10.10** Estimates of natural gas storage capacity and historical maximum storage volumes (as of April 2010 and April 2011)

	Working gas design capacity						Demonstrated peak working gas Capacity as percentage of working gas design capacity	Maximum end-of-month working gas inventories <sup>3</sup>	
	Demonstrated peak working gas capacity	Including inactive fields	Active fields only (excluding inactive fields)	Apr-10	Apr-11	Apr-10			
Region	Apr-10	Apr-11	Apr-10	Apr-10	Apr-11	Apr-10	Apr-11	Apr-10	Apr-11
East	2,196	2,205	2,281	2,276	2,288	96%	96%	2,097	2,097
Producing	1,297	1,340	1,383	1,366	1,406	94%	95%	1,224	1,263
West	556	558	688 R	688	694	81%	80%	524	524
Lower 48	4,049	4,103	4,353 R	4,331	4,388	93%	94%	3,833	3,847

Source: U.S. Energy Information Administration

Notes: Billion cubic feet, unless otherwise stated

Totals may not equal sum of components because of independent rounding.

R = revised.

10,000 (corresponding to one firm representing 100% of the market), and the minimum is 0% (corresponding to an infinite number of equally small firms). Sometimes, only 50 firms with the biggest market shares are considered, but there is no theoretical justification for this approach. The FERC considers the index of 1,800 to be a factor contributing to a decision rather than a firm limit. It pointed out, however, that:

“The 1,800 HHI level is not a bright-line test below which an applicant would automatically qualify for market-based rates, or above which an applicant would be excluded from market-based rates. Rather, the Commission uses the 1,800 HHI level as an indicator of the level of scrutiny to be given to the applicant. As explained in the Policy Statement, if the HHI is above 1,800 the Commission will give the applicant closer scrutiny because the index indicates that the market is more concentrated and the applicant may have significant market power. Conversely, an HHI below 1,800 would result in less scrutiny of the applicant’s potential to exercise significant market power because it would indicate that the market is less concentrated.”<sup>52</sup>

Other pertinent factors include the barriers to entry in a defined market region and services provided by non-storage facilities that represent alternatives available to potential storage customers, such as local gas supply, LNG supply, financial instruments (virtual storage – ie, a derivative instrument which replicates the economic outcome realised through actual control of physical storage) and pipeline capacity which effectively competes with narrowly defined storage facilities. The barriers to entry may be created through multiple factors, including strong competitive position of the entrenched incumbents and geological characteristics of the area.<sup>53</sup> Additionally, an applicant must demonstrate that it would not be able to maintain a price increase of 10% over competitive levels for an extended period of time. This is a reasonable requirement, but proving or disproving it *ex ante* may be difficult in practice.

## CONCLUSIONS

This chapter covered the segment of the natural gas industry referred to usually as midstream. This segment is as important to trading natural gas as it is uninspiring. Those who understand that trading is about hard work and collecting and processing huge volumes of fundamental data treat this part of the natural gas value chain as an

important source of information, a window into an important area of the energy industry. The next chapter will examine market transactions and strategies that can be structured around the physical processes discussed in this and the two previous chapters.

- 1 For example, a weekly summary of storage forecasts and commentary is available from John Sodergreen at the Energy Metro Desk.
- 2 A. Rojey and C Jaffret (Eds), 1997, *Natural Gas: Production, Processing, Transport (technical edition)* (Paris, France: Institut Francais du Pétrole Publications).
- 3 See: Bob Shively and John Ferrare, 2007, *Understanding Today's Natural Gas Business* (Energydynamics).
- 4 This design of the transportation tariff was one of the requirements of Order 636. This straight fixed-variable rate design has the impact of increasing the costs of reserving capacity but at the same time lowering the variable cost of transportation. Shippers whose peak-period needs for capacity are very high compared with their average needs are particularly affected by this change.
- 5 Ken Costello, 2006, "Efforts to harmonize gas pipeline operations with the demands of the electricity sector," *Electricity Journal*, December, (19)10, pp 7–27.
- 6 *Op. cit.* Ken Costello, 2006.
- 7 The tariff was obtained from the pipeline website.
- 8 The example is borrowed from the presentation by BENTEK Energy at Benposium, June 9, 2010.
- 9 Recall that 1 Dth = 1 MMBtu.
- 10 Historical data supporting this claim can be found in the FERC Order 712 (<http://www.ferc.gov/whats-new/comm-meet/2008/061908/G-4.pdf>). In this Order, the FERC removed, on a permanent basis, the rate ceiling on capacity release transactions of one year or less.
- 11 *Ibid*, p 36.
- 12 The capacity release market was established under Order 636. This market allows shippers to resell unused firm transportation market transportation capacity.
- 13 This change was dictated by the growing integration of the power generation and natural gas industries, given the growing importance of natural gas fired power plants.
- 14 "Nomination through allocations," BCI Training Systems, Houston, Texas, 1995.
- 15 "NAESB Final Report on the Efforts of the Gas-Electric Interdependency Committee," FERC, Docket No. RM05-28-000.
- 16 When an interruptible service is available, it is not unusual for marketers to over-nominate at a given meter to capture a higher percentage of the gas that is ultimately scheduled. This is because, in case of over-nomination, the pipeline allocates a percentage to each entity nominating gas based on the original nominations.
- 17 Ken Costello, p 10.
- 18 CST stands for Central Standard Time. The timezone, which includes Chicago and Houston, is one hour behind New York.
- 19 *Ibid*. Most of BENTEK's Pipe2Pipe reports are based on the ID2 cycle.
- 20 As above.
- 21 For example, such a service is offered by BENTEK Energy.
- 22 <http://www.troutmansandersenergyreport.com/2010/05/ferc-announces-new-reporting-rules-for-intrastate-and-hinshaw-pipelines/>.
- 23 The discussion is based on James H. McGrew, 2003, "FERC," Federal Energy Regulatory Commission, American Bar Association, Basic Practice Series.
- 24 These principles are explained in detail in the Order 636 (<http://www.ferc.gov/legal/maj-ord-reg/land-docs/restruct.asp>).

- 25 As above. The Seaboard method recovered 50% of fixed costs in the commodity charge. The United method used the percentage rate of 75%. The names of the methods refer to specific rate cases.
- 26 James H. McGrew, op. cit., p 93.
- 27 Steven P. Schneider, 1997, "Natural gas pipeline regulation and its impact on value," The International Association of Assessing Officers Public Utility Seminar, Nashville, TN, March 20–22.
- 28 88 FERC 61,227; Docket PL99-3-000, Certification of New Interstate Natural Gas Pipeline Facilities, Statement of Policy (issued September 15, 1999).
- 29 In Texas, the regulatory body for pipelines is the Texas Railroad Commission (TRRC).
- 30 This panel is based on the following article: Rusty Braziel and Vincent Kaminski, 2007, "Going with the flow," *Energy Risk*, August.
- 31 Such disruptions have happened in the US western markets (the Rockies), with natural gas prices collapsing temporarily due to maintenance-related outages on area pipelines. This was the case for natural gas becoming stranded in the producing region, without sufficient transportation capacity to take it to market. Some marketers who had anticipated this situation entered into swaps with producers that allowed them to take advantage of this market disruption (they were receiving fixed price and paying floating).
- 32 The author served as a consultant for BENTEK Energy and Russell Braziel was an employee of the company.
- 33 <http://www.bentekenergy.com>.
- 34 A detailed discussion of LNG technology is beyond the scope of this book. See: Valerie Rivera, Ayema Aduku and Oluwaseun Harris, 2008, "Evaluation of LNG technologies," April 30 ([http://www.lngpedia.com/wp-content/uploads/Evaluation\\_of\\_LNG\\_Production\\_Technologies\\_-\\_Univ\\_of\\_Oklahoma.pdf](http://www.lngpedia.com/wp-content/uploads/Evaluation_of_LNG_Production_Technologies_-_Univ_of_Oklahoma.pdf), for more information).
- 35 <http://www.chevronaustralia.com/ourbusinesses/gorgon.aspx>.
- 36 International Gas Union (<http://www.igu.org/igu-publications-2010/IGU%20World%20LNG%20Report%202010.pdf>).
- 37 A. Rojey and C Jaffret (Eds), 1997, *Natural Gas: Production, Processing, Transport (technical edition)* (Paris, France: Institut Francais du Pétrole Publications).
- 38 As above.
- 39 The Kenai LNG plant in Alaska was decommissioned in 2011. Construction of Cheniere Energy Partners Sabine Pass, LA, plant was due to start in the first part of 2012.
- 40 The author is unable to assess issues of the safety of regassification process. We have read papers ranging from predictions of catastrophic impacts to zero danger.
- 41 [http://www.shell.com/home/content/aboutshell/our\\_strategy/major\\_projects\\_2/prelude\\_flng/revolution\\_natural\\_gas\\_production/](http://www.shell.com/home/content/aboutshell/our_strategy/major_projects_2/prelude_flng/revolution_natural_gas_production/).
- 42 One of them is a gravity-based system located near Rovigo in the Veneto region of Italy (see <http://www.pipelineandgasjournal.com/world%E2%80%99s-first-offshore-gravity-based-lng-terminal-near-venice>).
- 43 The value of this option has been reduced through expansion of onshore shale natural gas production. The consequences of hurricanes for the US natural gas market may be still significant through combined infrastructure and demand impact.
- 44 The details of the sampling algorithm used by the EIA to produce weekly storage numbers can be found at <http://ir.eia.gov/ngs/methodology.html>.
- 45 Available at <https://www.ferc.gov/EventCalendar/Files/20041020081349-final-gs-report.pdf>.
- 46 [http://www.eia.doe.gov/oil\\_gas/natural\\_gas/ngs/ngs.html](http://www.eia.doe.gov/oil_gas/natural_gas/ngs/ngs.html).
- 47 [http://www.eia.doe.gov/pub/oil\\_gas/natural\\_gas/analysis\\_publications/ngpipeline/und\\_rgrnd\\_storage.html#overview](http://www.eia.doe.gov/pub/oil_gas/natural_gas/analysis_publications/ngpipeline/und_rgrnd_storage.html#overview).
- 48 The heating and cooling degree days are calculated by using a threshold temperature level (typically 65°F), corresponding to weather conditions seen as comfortable. If the average daily temperature is equal to 73°F, it translates into eight cooling degree days and zero

heating degree days. If the average daily temperature is equal to 40°F, it translates into 25 heating degree days and zero cooling degree days.

- 49 [http://www.eia.gov/dnav/ng/ng\\_cons\\_sum\\_dcu\\_nus\\_m.htm](http://www.eia.gov/dnav/ng/ng_cons_sum_dcu_nus_m.htm).
- 50 "Current State of and Issues Concerning Underground Natural Gas Storage," Federal Energy Regulatory Commission Staff Report, September 30, 2004, Docket No. AD04-11-000. This report remains the most current review of the storage business in the US (at the time of writing).
- 51 "In general, the relevant geographic is the geographic area containing those suppliers that can affect any attempt by the applicant to exercise market power." Order 678.
- 52 FERC, "Rate Regulation of Certain Natural Gas Storage Facilities," June 19, 2006, (Docket Nos. RM05-23-000, AD04-11-000; Order No. 678).
- 53 For example, New England has little geologic potential for the development of underground storage facilities.

## *US Natural Gas Markets*

Voltaire is quoted as saying that should God not exist, it would be necessary to invent him. The same may be said about the US natural gas market from the point of view of an aggressive trader or a speculator. There are many structural features of this market which make it an optimal platform for trading by smart, well-informed and agile traders, but also from time to time by people willing to abuse the system. It is a market whose microstructure can be exploited, but it can also ruin an imprudent person.

This is how we opened the chapter when we started writing the book in 2007. Unfortunately, it is no longer true as far as speculators are concerned – although there is some good news from the point of view of consumers. The prices of natural gas in the US are depressed, volatility is suppressed and locational price differences have been eliminated or significantly compressed. It is difficult to find consenting adults willing to invest in start-up natural gas-oriented hedge funds, as some of our friends have found out. Trading desks in financial institutions are scaling down. The factors behind this rapid change of fortunes are the abundance of natural gas from shale formations and expansion of the pipeline network. If our opinion that natural gas prices in the US are just temporarily depressed through a combination of market forces (as discussed in Chapter 9) is correct, this market is likely to make a comeback and volatility is likely to reassert itself. A growing dependence on shale formations means remaining on a treadmill, with high decline rates dictating continued intensive drilling effort. Relatively small variations in prices may either slow down or accelerate drilling, creating preconditions for a much bigger price shock. This market, in our view, will remain a great opportunity for patient speculators (if such creatures exist), and it is still too early for greying industry veterans

to turn off their computers and switch to playing golf and sipping bourbon.

The natural gas market in the US is very mature and sophisticated, combining standardised physical and financial transactions, and complex structured transactions. We will start by discussing daily and monthly OTC physical transactions and corresponding financial transactions, primarily swaps and basis swaps. We will continue with a look at exchange-based contracts (Nymex and ICE), followed by an exploration of structured transactions: volumetric production payments (VPP) and municipal prepayes.

## SPOT, FORWARD AND FUTURES MARKETS

### US natural gas market: The essential characteristics (supply and demand curves)

The US natural gas markets have many unique features resulting from the historical development and cross-pollination between the energy and financial markets. Some of the most important features are listed below.

- ❑ Natural gas is traded actively at many locations, referred to as market hubs, the intersections of multiple intra and interstate pipelines, and interfaces of the pipelines with local distribution systems.
- ❑ The natural gas markets are based both on physical (ie, involving physical delivery) and purely financial (cash-settled) transactions. The settlements of financial transactions are based on publicly available price benchmarks of natural gas in physical and/or other financial markets.
- ❑ In the past, physical markets were organised around monthly transactions for base load gas, delivered at roughly equal quantities over the course of a calendar month. The prices of base load gas at specific locations are referred to as monthly indexes. Monthly markets are still very important but the importance of daily markets is increasing.
- ❑ The development of the natural gas industry led to the emergence of short-term markets for gas flowing over the next day. Increasingly, we can see frequent transactions for natural gas traded intra-day, with pipelines offering multiple nomination cycles (up to four per day).

- ❑ Another unique feature of the North American natural gas market is the use of a central pricing benchmark, the natural gas futures contract with delivery at Henry Hub, Erath, Louisiana. Natural gas at other locations can be priced by adding the basis to the Nymex futures price – the basis represents the differential between the index (ie, locational price) and the Nymex contract settlement price of the expiring contract.
- ❑ The locational basis in the US and Canadian natural gas markets is traded, for most locations, as a standalone contract. This creates a unique triangular relationship between three types of prices: Nymex futures contract price, locational index and basis. One can replicate an index position by entering simultaneously into a Nymex position and a basis position. One can enter directly into an index transaction and a Nymex transaction, implicitly assuming a basis position.
- ❑ The US is becoming increasingly dependent on electricity generated from gas-fired power plants. Interactions between natural gas and electricity and transactions referencing prices of these two commodities create unique price dynamics, which favour traders present in both markets and well familiar with the fundamentals of both industries.

As in the case of any market, its underlying properties can be explained by the characteristics of the supply and demand curve. These properties reflect the technology of production and distribution of natural gas, and also patterns observed in the behaviour of buyers and sellers, their varying objectives, access to capital and the extent they are subject to regulatory oversight. Many interesting characteristics of this market result from the interactions of the buyers and sellers with diametrically different objectives. At one extreme, hedge funds are driven by a pure profit motive and are condemned to very aggressive and high-risk strategies (as their investors would expect). At the other extreme, regulated utilities follow the “prudent man” principle, and operate to minimise the risk of shortages and to guarantee long-term security of supplies achieved at a reasonable cost. In between these two extremes, we can see market participants with multiple objectives, following varying and evolving strategies, creating a fascinating market ecosystem.

*Demand curve*

The demand curve is relatively inelastic,<sup>1</sup> for a number of reasons. The residential and commercial users are largely insulated from short-term fluctuations in intra-month prices through the existing tariffs and through the rigidity of the billing process. Seasonality of demand may mask the impact of prices on the bills received by households and commercial users, and delay the adjustment process for most consumers. For many industrial users, the contribution of natural gas may be a relatively small percentage of the overall cost and an expense that is simply unavoidable. For some industrial buyers and power producers, the change in price levels of natural gas relative to the prices of other feed stocks and fuels may result in sudden adjustments, ranging from input substitution to complete shutdowns (as in the case of ammonia plants, for example). Such sudden actions may result in amplified changes to price levels.

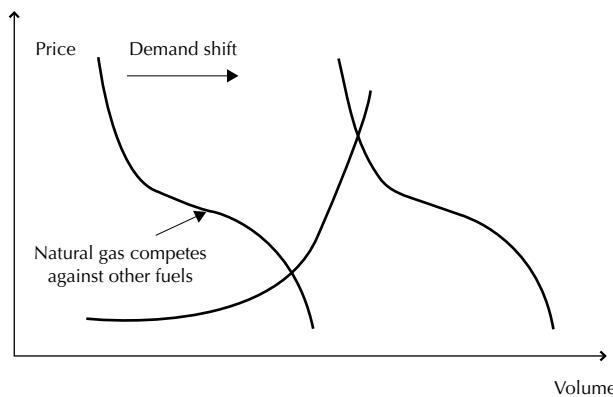
The most important factor explaining low demand elasticity is the behaviour of buyers responsible for the acquisition of natural gas by local distribution companies. Their behaviour is dictated by the overriding principle of reliability and continuity of supply under winter conditions. These axioms influence both decisions related to natural gas storage injections as well as gas purchase contracting. The trading community developed an intuitive understanding of this behaviour and translated it into two simple empirical rules.

- ❑ Always go short into winter. This rule is based on the notion that utility buyers will always overpay for protection and the winter will not (for most of the time) be as cold as the industry fears. This strategy backfires occasionally, especially when it is overdone.
- ❑ Sell out-of-the-money winter and hurricane season calls, buy at-the-money calls. This strategy is based on the notion that buyers will overpay for insurance against tail events offered by the purchase of options with high strike price. A long position in the calls benefits from short-term vibrations of prices around current levels, often induced by excessive trading. Sometimes, financial traders, who do not understand the microstructure of this market are seduced by the instances of price spikes and go long volatility through purchase of out-of-the money options – just to be crushed when the market subsides.

Another unique aspect of natural gas trading is a heavy dependence on weather forecasts in the case of the short-term physical markets. Heat and cold waves at certain regions of the US result in reconfiguration of natural gas flows in the US, with prices adjusting in order to release gas at some locations and attract gas to other locations, subject to available capacity on the pipeline grid. The best physical traders are often former meteorologists or those who constantly interact and cooperate closely with weather forecasters. Several weather forecasting firms specialise in producing forecasts for energy traders.

The demand curve also reflects the fuel-on-fuel competition arising from the ability of substitution of one fuel for another. This is reflected in the flat segment of the demand curve, characterised by high price elasticity and high cross-price elasticity (elasticity with respect to the price of substitute fuel). Historically, this potential existed with respect to natural gas and residual fuel oil in electricity generation, although the scope of this substitution seems to be rather limited. One can detect intensifying natural gas versus coal competition in certain regions of the US, reflecting the falling prices of natural gas. The potential for fuel substitution is captured in a relatively flat section of the demand curve that is characterised by much higher price elasticity than other segments of the curve. Fuel substitution happens primarily in power generation in two different ways. Some generation units have dual burner capacity and can switch from one fuel (for example, natural gas) to a competing fuel (for example, residual fuel oil). The crossover from one fuel to another depends on the relative prices of the two fuels and the thermal efficiency of both units. In some cases, a switch happens regardless of the price ratios and is driven by supply constraints in winter in the natural gas markets. This type of switching, called seasonal, was particularly true of New England in the 1990s. An alternative to switching fuels in the case of the same generation unit is to expand production in one set of units (for example, natural gas units) and shut down coal-fired units. This type of substitution has became an important factor in the natural gas and electricity markets in the US following a drop in the natural gas prices due to the expansion of output from the shale formations.

**Figure 11.1** Natural gas market: A stylised representation of demand and supply curves.



### *Supply curve*

The supply curve in the short-run is characterised by high convexity. This technical term describes a well-known fact that for a relatively wide range of prices the supply is quite responsive – ie, the supply curve is close to being horizontal, with small price increases being sufficient to induce additional volumes, either transported from other locations, drawn from storage or coming from expanded production. At some point, rigidity of the existing infrastructure restricts additional flows, and the supply curve becomes very inelastic (as shown by the vertical section of the curve).

These characteristics of natural gas market can be illustrated with a well-known graph (see Figure 11.1) summarising the properties of natural supply and demand curves. This graph explains why prices may be characterised by big jumps even in conjunction with relatively small and temporary changes in fundamentals. This basic model has been extensively used in the literature to explain the properties of the natural gas market. The same basic approach has been expanded by capturing the potential for competition of natural gas against multiple fuels in the demand curve.<sup>2</sup>

### **A balancing market**

Another important feature of a short-term physical natural gas market is that it is effectively a balancing market, a market to dispose of surpluses and cover deficits in the short run. This is a general char-

acteristic of many energy markets, a co-existence of long-term and short-term contracts, with the latter serving primarily to smooth fluctuations in supply and demand in the short term. Most market participants manage their long-term needs through base load contracts and enter the short-term markets to make marginal adjustments to their positions. In these marginal transactions, the prices may be, by general standards, unreasonably high or unreasonably low. This may happen for a number of reasons.

- ❑ Many market participants are relatively price insensitive and are motivated by reliability concerns or operational constraints. As explained above, they may transfer the price risk to another party (for example, rate payers);
- ❑ Sometimes, the marginal adjustments to volumetric positions represent a small percentage of the total needs and paying even a high price (or receiving a low price) has a small impact on overall financial results. Given the market microstructure, a small volume of current market transactions may have a significant impact on a large number of floating price contracts and contracts based on formulaic prices – ie, prices referencing other market indexes or variables. This tail-wags-dog property of many commodity markets, very pronounced in the case of natural gas, can contribute significantly to high price volatility.
- ❑ In case a pipeline imposes an OFO due to operational constraints, the imbalance penalties may reach very high levels (tens of dollars per MMBtu). In such cases, the market prices may spike, as even unusually high market price may be preferable to a settlement with a pipeline.

### **Trader's perspective**

Another factor contributing to price volatility is a tendency that manifests itself with amazing regularity: most speculative traders tend to make the same directional bet based on analysis of the same fundamentals. The traders watch the same fundamentals and socialise in the same bars. A position that is perfectly justified from the point of view of a single trader may collectively result in a highly imbalanced market, vulnerable to sudden shocks. The speculative positions often represent the other side of positions treated by some counterparties as hedges and assigned to existing primary

exposures. Under FAS 133, a hedge has to be assigned to the existing position and documented at inception of the transactions. Such positions are unlikely to be unwound ahead of schedule. A sudden, often minor, shift in market fundamentals creates selling or buying pressure, may produce a run-away market. The speculators rush to modify their positions but hit a brick wall of other market participants who are unwilling to trade. The best example of a position that is characterised by exceptionally high volatility is the March/April spread (H/J spread). The prices of these two contracts are highly sensitive to weather conditions and the storage situation. A small change in the market fundamentals, combined with big bets, can result in explosive market changes.

Price shocks can be magnified through a number of special characteristics of the market microstructure. For one, some traders can use certain features of the trading systems – such as stop-loss orders. A stop-loss order “is an order placed with a broker to buy or sell once the stock [or a financial instrument in general] reaches a certain price. A stop-loss is designed to limit an investor’s loss on a security position.”<sup>3</sup> The stop-loss practice occurs in commodities as well. Such orders represent a tool designed to avoid a significant loss resulting from a sudden change in market price, especially when an investor is unable to monitor market trends in real time. Many traders can correctly guess the levels at which large numbers of stop-loss orders are posted and can make an effort to push the market prices through the levels at which a large number of orders are placed, creating a chain reaction of automatic transactions, that further influence market prices, amplifying the result of the initial transactions. Such a practice is often described as running the stops, or gunning for stops.

This mechanism is well known to the market, as evidenced by opinions expressed by many experienced industry insiders.

Jeff Nichols, a precious metals consultant, says a well-capitalised floor trader can occasionally pull off such a move in small, thinly traded markets, particularly if other traders sense what is going on and join in. But in larger markets, such as currencies, it takes a lot of money to gun for stops successfully because the player has to be able to buy or sell contracts in significant quantities. Well-known commodities trader Richard J. Dennis says he tries to anticipate where technical traders have placed their stops and gauge the effect that activation of the stops will have on prices. “If you look at charts, you can make a reasonable guess about where the stops are,” Mr.

Dennis says, adding that he uses this information to avoid those areas. "They're a little bit like land mines going off, and you don't want to walk into the mine field."<sup>4</sup>

In markets based on open-outcry systems, some traders can infer this information by observing action in the pit, catching a glimpse of the order cards of other traders or by talking to other traders. In many cases, the traders can correctly guess that other market participants will place stop-loss orders around the levels implied by technical analysis. Technical analysis attempts to identify patterns in the historical price series that can be used to predict the future course of the market. Most methods of technical analysis rely on patterns discerned in historical volumes and prices through visual inspection, but sometimes more advanced quantitative and statistical tools are used. Many technical analysis systems define the so-called support or resistance level, the price levels at which the prices are likely to fluctuate for a period of time, with sharp movement to the upside or the downside once these levels are penetrated. Similarly, such self-reinforcing price movements can be produced through option hedging strategies.

As explained in Chapter 4, an option is a contract that gives the right, but not an obligation, to buy (sell) a certain commodity at a given price by (or on) a certain day. In this market, some traders (known as market makers) sell options to buy or sell NG futures contracts, and hedge those options positions by also buying or selling the underlying futures contracts (or other instruments, whatever the underlying happens to be). The purchases of the underlying instruments are determined by calculating the option delta (hence, delta hedging). When a burst of trading activity succeeds in moving the price, it starts a process that may feed on itself. The option market makers have to buy or sell the underlying contract just to rebalance the hedge positions, and this in turn leads to a cascade of further price movements and additional hedging-related trades in the same direction. This, again, is an example of market participants becoming accidental and unintentional collaborators of market manipulators who can start an avalanche by throwing a small snowball.

### Monthly and daily transactions

The natural gas markets in the US were historically organised around monthly transactions for base load gas delivered ratably over the course of a calendar month. The transactions were negotiated during a period of five business days at the end of the calendar month prior to the delivery month, known as the bid week. For a long time, the volumetric adjustments over the period of the month would be accomplished by overdrawing or underdrawing natural gas from a pipeline on a daily basis as the flows were usually balanced over the monthly time period. These days, the volumetric adjustments happen through intra-month transactions, for the next day delivery or for the balance of the month. The monthly market developed around the market hubs, usually located at the intersection of multiple pipelines or at the points of interface between the pipelines and local distribution systems (citygate hubs).

Price discovery for monthly bid week transactions was traditionally based on surveys conducted by the industry newsletters that compile transaction information from market participants and calculate “index prices,” computed as the volume-weighted averages of individual transaction prices. The transactions used to calculate the indexes are the fixed-price physical deals and physical basis transactions (as will be explained below). The market for indexes is dominated by Platts, followed by Argus and Natural Gas Intelligence. The monthly indexes are published in the industry newsletters on the first business day of the month.

An example will help to explain the mechanics of this market. One of the most important market hubs in the US natural gas industry is the Houston Ship Channel (HSC), a location with access to many intra and interstate pipelines that supply one of the biggest clusters of chemical plants in the world, and move natural gas from production areas to other parts of the US. Transactions for base load natural gas at HSC and other hubs happen over a period of a few days, around the Nymex futures contract expiration dates in the bid week.<sup>5</sup> Bid week transactions can take place OTC or on electronic trading platforms such as ICE.

Information about transactions, such as volumes and agreed prices, is collected by the industry newsletters, which use the individual transaction data to calculate quantity-weighted price indexes. The indexes are published on the first business day of a calendar month. The difference between the index price and the final settle-

ment price of the Nymex expiring contract for the same calendar delivery month is known as the realised basis. There are 12 specific basis numbers calculated during the course of the year for a given location for which a published index is available. Basis may be either positive or negative, depending on the relationship between the prices of natural gas at a given place and the Nymex contract price. As a rule of thumb, locations east and north of Henry Hub are associated with positive basis, and locations to the west are associated with negative basis. Most energy traders use price indexes calculated and published by the Platts' newsletter "Inside FERC Gas Market Report." The price index published by this newsletter is referred to as an IFERC index.

The monthly price indexes for natural gas are calculated in two different ways, using fixed-price transactions or physical basis transactions. Fixed-price transactions are very straightforward: for example, the buyer commits to pay US\$4.5/MMBtu for base load natural gas delivered over the next month at a given location. Physical basis is negotiated as a fixed difference in price, or price differential, to Nymex before the final settlement price of the Nymex is known (such as "the Nymex final settlement price plus three cents"); once the Nymex price settles on the contract expiration day, the negotiated differential is added to the Nymex price, and the total determines the price of natural gas at the specific location. In other words, after the Nymex contracts settles, the transaction mutates from floating to fixed, and the monthly price crystallises.

Physical basis transactions are very important in certain natural gas hubs. The best way to describe their significance is to go to the source, Platts, as this company compiles most of the price data used in the physical natural gas markets. The *Platts North American Natural Gas Methodology and Specifications Guide* (October 2012) states:

For the monthly bidweek price survey, bidweek is defined as the last five business days of each month. For each day of bidweek, report all fixed-price physical deals negotiated that day for delivery throughout the next month. Also report all physical basis deals in which the basis value is negotiated on one of the first three days of bidweek and the price is set by the final closing value of the near-month NYMEX futures contract plus or minus the negotiated basis. Platts' current policy is to use physical basis deals for points east of the Rocky Mountains, except in the Permian Basin region at Waha, El Paso Natural Gas Co., Permian Basin and Transwestern Pipeline Co., Permian Basin.

As specified in its methodology, Platts uses physical basis deals at price locations east of the Rocky Mountains, except for El Paso Natural Gas and Transwestern Pipeline in the Permian Basin and the Waha market centre. Physical basis deals are used at 33 of 41 delivered-to-pipeline locations and at 14 of 22 market centre locations in the monthly survey. Further information regarding physical basis transactions can be found in Figure 11.2. The different shades indicate the percentages of bid week transactions at index points across the country that are physical basis transactions. This graphic illustrates the importance of physical basis transactions to the formation of monthly prices at different trading hubs through the US, for which the monthly indexes are published.

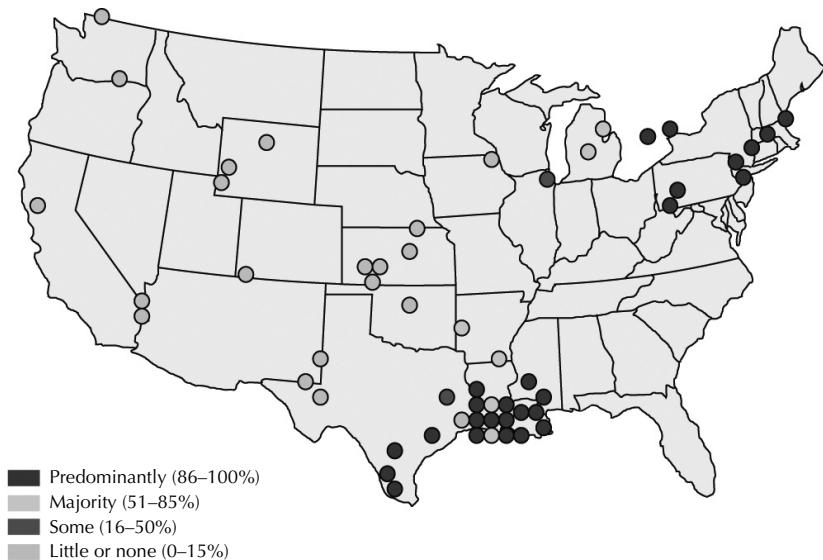
In February 2005, Platts began publishing in “Inside FERC Gas Market Report,” a table breaking out the physical basis deals used to compile monthly indexes at points where basis trading is most prevalent. That information clearly has been useful to those interested in the natural gas market; FERC’s 2006 report included an analysis of the use of physical basis transactions in Platts monthly gas indexes. The practice of detailed reporting of the physical basis transactions is described as follows:<sup>6</sup>

To make its reporting of basis deal more accessible to all subscribers, beginning in February Platts is including the separate basis tables for delivered-to-pipeline and market center price locations in all of the formats in which it reports its monthly gas prices – in the first-of-the-month supplement to *Platts Inside FERC’s Gas Market Report*, in the biweekly *IFGMR*, in the monthly *Gas Daily Price Guide* and on *Natural Gas Alert*, our electronic news service. [...] In addition, the basis table now lists each price location for which physical basis deals are accepted even if no basis deals were received at that point for the given month.

### The importance of index prices

The importance of index prices for the natural gas industry is difficult to overstate. Many market participants use the monthly prices for index transactions that are effectively floating price deals, with natural gas delivered over long time periods at prices determined during the bid week and calculated by the index publishers. Such transactions are typically priced at index, adjusted by the “premium,” a differential contained typically in the range of (-2, 2) cents. In many markets, the producers receive prices equal to an agreed percentage of a specified index. Index prices are also used as

**Figure 11.2** Use of physical basis in natural gas indexes at major trading points (2006)



Source: FERC

in financial swaps and basis swaps (explained below). There are many different justifications for a premium to an index in a floating-price transaction:

- ❑ the premium may be charged to cover transaction costs and the profit margin of a marketer, and is the equivalent of a bid–offer spread;
- ❑ the producers often want to maximise market share and are willing to sell gas at a small discount to the index in order to attract customers; and
- ❑ sometimes natural gas is delivered not at the hub location but at a meter at some distance from the market, and the premium captures the physical delivery cost differential (extra transportation cost).

Insights into the role of price indexes in the US natural gas markets can be obtained from the data collected by the FERC on Form 552, “Annual Report of Natural Gas Transactions.” The submissions on

this form, following FERC Order 704 December 26, 2007, apply to market participants that sold and purchased 2.2 million MMBtu or more of natural gas annually. This is approximately the amount of gas used by a 90 MW peaker power plant running every day for nine hours. It is obvious that this form does not capture all transactions, but it still provides a good snapshot for the bulk of the trades. In 2009, the volume of reported physical gas market transactions was equal to about 56 Tcf (compared to the annual production of about 22 Tcf). This reflects the fact that the same molecules are usually traded multiple times before reaching the end user. The pricing structures used are shown in Figure 11.3. As one can see, 70% of the underlying volumes were priced based on the published indexes; fixed-price deals represented just 22% of the volumes.

The volumes of index transactions and financial swaps often exceed by a very large factor the volume of transactions executed at a given location during the bid week. This creates obvious temptation for many market players to influence published prices, and by the end of the 1990s market manipulation of published price indexes had become an endemic phenomenon. Chapter 24 reviews the different methods of index price manipulation and discusses the measures taken by the regulators to protect the integrity of the markets.

Daily physical transactions take place every morning for the next day delivery<sup>7</sup> at many trading hubs. The daily transaction prices are reported to such publications as *Gas Daily* or *Argus Natural Gas Americas* publication and become available on the day of the transaction, when the newsletters are distributed electronically to the subscribers, usually between 5 and 6 pm Central Time. The publication uses fixed-price transactions in calculation of the volume-weighted indexes. It is important to remember that the reported prices may be identified by the transaction date or flow date. Ignoring this obvious distinction can lead in practice to many costly snafus.

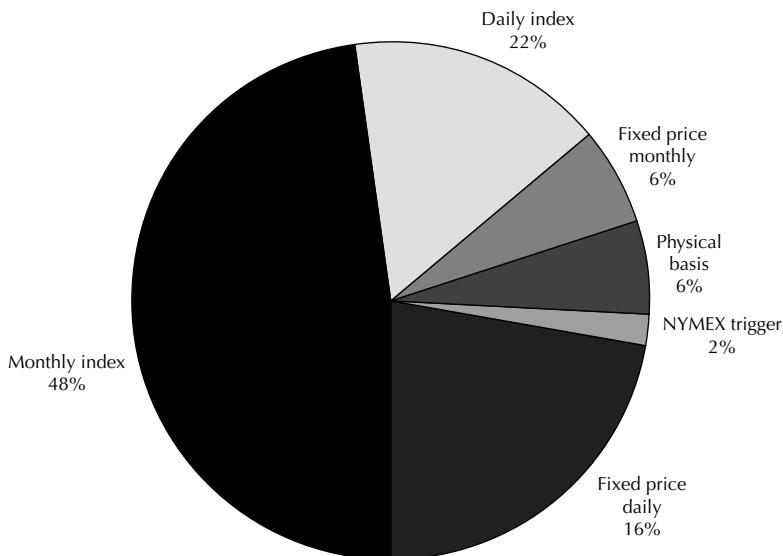
### Natural gas swaps

The concept of a swap was reviewed in general terms in Chapter 4. At the cost of some repetition, the basic terms will be revisited. In the natural gas market, swaps are used for both purely financial, cash-settled transactions and for pricing physical deals that require delivery of the underlying commodity. As mentioned in Chapter 4, a swap, in principle, denotes an exchange of two streams of cashflows,

one based on a fixed rate or price, the second based on variable (floating) rate or price. A swap defined in terms of two floating prices is called a basis swap. In the case of a physical swap, there are two flows: one monetary, the other physical. The swap pricing principles still apply, however, as the recipient of natural gas can take the commodity to the market and receive the current – ie, floating – market price.

Under the conventions established in the financial markets, the payer of the fixed price in a fixed-for-floating swap is known as the swap buyer. The actual cashflows are calculated by multiplying the respective prices by the agreed notional or actual volume. In pricing a swap, one should pay attention to the settlement dates (ie, the dates when a floating price for a given period becomes known) and the cashflows dates. These dates are not usually identical. By industry convention, the physical transactions settle 25 days after the end of the month in which physical gas flew. This convention goes back to the days when the gas meters were mechanical and read by pipeline employees driving from location to location and mailing or faxing

**Figure 11.3** US natural gas physical transactions (2009)



Source: <http://www.ferc.gov/market-oversight/reports-analyses/overview/gas-transactions2009.pdf>

the logs with the flow numbers. It took a long time to collect the information and reconcile the numbers. Financial swaps typically settle within three to five business days.

One of the special features of the commodity markets is that in many markets one can directly trade the spreads between two different prices. In other words, a spread trades as a single underlying and a position in a spread can be established in one transaction. A spread position can be established with respect to location, time, quality and commodity type. In the US, natural gas locational spreads are referred to as basis transactions and trade as a single underlying. The basis is defined as a difference between price at a specific location and the settlement price of the expiring Nymex prompt contract. The price at a specific location is usually one of the monthly indexes published by the industry newsletters. The price of the expiring contract is typically the settlement price on the last day of trading (denoted by LD1 or NX1) or the average of the last three settlement prices (denoted by LD3 or NX3). Initially, in the early 1990s, the industry was using LD3, but the practice at the time of writing favours the LD1 price.

Multiple pricing systems used in the natural gas markets give rise to a variety of swap structures. The most popular structures used in practice are:

❑ fixed-for-floating swaps:

- index swaps with the floating price equal to the monthly index at a specific location;
- gas daily swaps with the floating price equal to gas daily prices at a specific location; and
- basis swaps with the floating price equal to the realised basis – the basis swaps can also be looked at as floating-for-floating swaps (as demonstrated below).

❑ floating-for-floating swaps:

- monthly index price versus gas daily price;<sup>8</sup> and
- basis swaps with Nymex settlement price versus monthly index price plus/minus a differential.

We shall discuss examples of different natural gas swaps and their potential uses in structured transactions.

## Basis swaps

The buyer of the basis swap pays a fixed price (typically at monthly time intervals) and receives the realised monthly basis. Both cash-flows are determined by multiplying the contractually agreed fixed price and the floating price by the underlying notional volume. The example below demonstrates how the fixed price of the contract will be determined in practice. The basic principles are the same as the principles that apply to standard swaps.

Price determination starts with a forward basis curve<sup>9</sup> posted by the basis trader and based on the prices observed in the market. The fixed price of the swap will be determined by equating the values of the floating and fixed legs of the swap. The value of the floating leg is obtained by summing the products of monthly forward basis prices and notional volumes, discounted back to the valuation date. The value of the fixed leg is obtained by summing the products of the fixed price (unknown initially) and monthly volumes, discounted back to the valuation date. The fixed price can be calculated by equating the two legs of the swap.

Floating leg:

$$\sum_{i=1}^n \frac{Volume_{t_i} \times Forward\_Basis\_Price_{t_i}}{(1+r_{t_i})^{t_i}}$$

Fixed leg:

$$\sum_{i=1}^n \frac{Volume_{t_i} \times Fixed\_Price}{(1+r_{t_i})^{t_i}}$$

where  $t_i$  denotes the time of the cash flow ( $i = 1, 2, 3, \dots, n$ ). The interest rate  $r_{t_i}$  denotes a bullet annualised interest rate, used to discount a single cash flow taking place at time  $t_i$ . For example, if a swap has 12 settlement periods,  $n = 12$ , and  $t_6$  denotes the time at which the cash flow corresponding to the sixth settlement will take place. The time is measured, by convention, in years. If  $t_6 = 2.5$ , the sixth cash flow will happen two and a half years from the valuation date, which is assumed to be  $t = 0$ .

This representation of a basis swap can be modified by recognising that the basis is defined as the difference of the index price and the Nymex settlement price of the expiring contract. Under this alternative representation, the counterparty who is long basis receives the

index price and pays Nymex plus a fixed differential (which may be negative). This translates into the following cashflow:

$$\text{Index} - (\text{LD1} + \text{Differential})$$

which corresponds to a basis swap (exchange of two floating cash-flows). Note that the differential may be either positive or negative. This can be rewritten as

$$(\text{Index} - \text{LD1}) - \text{Differential}$$

which is equivalent to receiving basis and paying out a fixed price (fixed differential). The party long benefits from basis increase. Basis is often negative and this is why one should avoid using the term ‘widening basis.’ It’s better to say: ‘basis increased’ or ‘basis decreased.’ If a basis was equal to minus \$1/MMBTU and changed to minus \$0.5/MMBTU, the basis has increased. The basis may increase because the:

- index price increases and the Nymex price drops;
- index price increases and the Nymex price increases less; or
- index price falls and the Nymex price drops even more.

The alternative representation of the basis position is useful for a number of reasons. In many cases, a trader creates a basis position not through a single basis transaction, but by establishing the two legs of the basis separately, often at different times. Even if a basis transaction is established through a single trade, the risk management system may decompose the market exposure of a basis position into two component legs. Finally, a trader may close one leg of the basis position without closing the other leg. For example, a trader who is long basis may decide to close the Nymex leg by buying Nymex contracts. The long basis position is converted to a long index position. Closing the index leg would leave the trader with a short Nymex position. Good risk management software allows for separating how a basis position was established from its risk profile.

### **Swap valuation: basic model**

We will now cover the valuation of a plain vanilla swap. In the example below, we look at a very simple swap transaction with natural gas as the underlying. The tenor of the swap is January 1,

2013 to June 30, 2013. The daily volume is 10,000 MMBtus. We assume for the sake of simplicity the flat term structure of interest rates (identical zero rates for all time periods). The cash-settlement happens on the last day of the calendar month. The forward price curve (in dollars per MMBtu), as of the valuation date, June 30, 2012, is given in Table 11.1.

**Table 11.1** Forward price curve

Jan-13	9.5
Feb-13	8.6
Mar-13	7.1
Apr-13	6.2
May-13	6.1
Jun-13	6.9

The main valuation assumptions are shown in Table 11.2

**Table 11.2** Swap valuation assumptions

Start date	January 1, 2013
End date	June 30, 2013
Valuation date	June 30, 2012
Daily volume (MMBtu)	10,000
Interest rate	5.00%

The day count is assumed to be actual/365 (a year has 365 days and we count actual days). We assume a flat term structure of interest rates ( $r_t = 5\%$ , for all time periods  $t$ ). The calculation is straightforward and uses the formula 4.6 from page 137 (repeated as 11.1).

$$F_f = \sum_{i=1}^n \frac{F_{t_i} \times V_{t_i}}{(1+r_{t_i})^{t_i}} / \sum_{i=1}^n \frac{V_{t_i}}{(1+r_{t_i})^{t_i}} \quad (11.1)$$

It is necessary to multiply the monthly volumes by the corresponding forward prices and discount them back to the valuation date. These discounted products are next summed up. The next step is to calculate the sum of discounted volumes. The ratio of these two numbers gives the fixed price of the swap. Table 11.3 contains the calculations.

In an organisation relying on mark-to-market accounting, the swap will be carried on the books at its fair value. The value of the fixed price leg of the swap is calculated by applying the swap fixed

price to the contractual volumes and discounting the anticipated cashflows back to the valuation date. The value of the floating leg of the swap will be calculated by multiplying the notional volumes by the forward prices and discounting the results back to the valuation date. The discounted products will be summed up for both legs of the swap. Suppose that the swap is valued as of June 30, 2012. Table 11.4 shows the details of these calculations.

A few comments are useful here. First, at the inception, the value of the swap is equal to zero, as both legs have the same value. This means also that buying a swap is equivalent to acquiring a portfolio of forwards, except that one pays a levelised price instead of paying a different price for each component forward position. Of course, a moment after a swap is executed, the market will move and the swap will have either negative or positive value from the point of view of a given counterparty. Over its life, a typical swap will evolve from being an asset one day to being a liability and, depending on the market conditions, its value may fluctuate significantly. This has serious implications for the management of credit risk. Once a swap becomes an asset for a given counterparty, it will be a liability for the other side of the transaction (a swap is a zero-sum game).<sup>10</sup> The other side will have incentives to walk away from the transaction. Our next book will cover the analytical tools required to assess and manage credit risk.

Second, the valuation of the swap above was carried out by ignoring transaction costs. In practice, a trading organisation selling a swap will want to make a profit on the transaction and will usually quote an offer price to a potential buyer. This price will be determined by adding a bid–offer spread to the forward prices used in the valuation, and using the offer side of the forward price curve. When buying a swap, a bid side will be used. The bid–offer spreads are fairly tight and transparent in the most liquid and efficient energy markets. For longer tenors and more opaque markets (less liquid locations or quality grades), the bid–offer spreads are more difficult to gauge, especially for counterparties that are not trading organisations but rather commercial hedgers. The only line of defence for such entities is to use multiple quotes from different dealers. In many cases, a hedger has to decide how many quotes one wants to solicit, especially if the transaction involves a considerable volume and may have a significant market impact. Asking for too

**Table 11.3** Swap pricing

BOM	EOM	No. of days	Volume	Discount factor	Floating leg value	Discounted volume
January 1, 2013	January 31, 2013	31	310,000	0.97167	US\$2,861,567.08	301217.59
February 1, 2013	February 28, 2013	28	280,000	0.96804	US\$2,331,039.51	271051.11
March 1, 2013	March 31, 2013	31	310,000	0.964037	US\$2,121,844.51	298851.34
April 1, 2013	April 30, 2013	30	300,000	0.960178	US\$1,785,931.80	288053.52
May 1, 2013	May 31, 2013	31	310,000	0.956208	US\$1,808,188.97	296424.42
June 1, 2013	June, 30, 2013	30	300,000	0.952381	US\$1,971,428.57	285714.29
Total			1810,000		US\$12,880,000.44	1741312.256
Fixed price					US\$7.3967	

**Table 11.4** Calculating mark-to-market value of a swap

Volume	Fixed price	Discount factor	Forward price	Fixed leg of the swap valuation (1) × (2) × (3)	Floating leg of the swap valuation (1) × (3) × (4)
(1)	(2)	(3)	(4)	(5)	(6)
310,000	US\$7.3967	0.97167	9.5	US\$2,228,022.37	US\$2,861,567.08
280,000	US\$7.3967	0.96804	8.6	US\$2,004,889.33	US\$2,331,039.51
310,000	US\$7.3967	0.964037	7.1	US\$2,210,519.89	US\$2,121,844.51
300,000	US\$7.3967	0.960178	6.2	US\$2,130,651.41	US\$1,785,931.80
310,000	US\$7.3967	0.956208	6.1	US\$2,192,568.66	US\$1,808,188.97
300,000	US\$7.3967	0.952381	6.9	US\$2,113,348.78	US\$1,971,428.57
1,810,000				US\$12,880,000.44	US\$12,880,000.44

Note: The numbers herein are purely hypothetical, to be used as an illustration of the concept.

many quotes will alert the market to the potential transaction and may result in front-running by some dealers. In many trading organisations, the originators dealing directly with the clients are under an obligation to alert the trading desks to potential transactions, especially if they offered an attractive price and still were outbid. Losing a deal typically means that somebody else got it and will come to the market to hedge a the transaction.

Third, the use of interest rates in valuation of swap means that a commodity swap portfolio has sensitivity to interest rates. This may come across as a trivial observation to many market participants, but one should not forget that many managers in merchant energy companies rose through the ranks without ever obtaining even a superficial education in finance. Interest rate sensitivity can be easily assessed using, for example, the finite difference approach.

### **Floating-for-floating swap**

One popular floating-for-floating swap is an exchange of cashflows based on the monthly index versus daily index price at the same or different locations. Final realisations of both prices are unknown at inception of the swap, and forward price curves are used for pricing the swap. An interesting feature of this swap is that, during a given calendar month, the monthly index crystallises, while the gas daily price is still floating and the current portion of the swap is transformed into a fixed-for-floating swap.

### **Examples: Mechanics and uses of natural gas swaps**

The points made above can be illustrated with a few simple examples.

Suppose that two counterparties entered into a swap transaction that can be summarised as follows:

Floating price: IFERC Houston Ship Channel (HSC) Price Index

Fixed price: US\$7/MMBtu

Notional volume: 10,000 MMBtu/day

Term: 12 month

Settlement: Monthly

If the index for a specific month settles at US\$8/MMBtu, the floating payment is calculated as follows, using the notional volume (we assume the month has 30 days):

$$\begin{aligned}\text{Floating payment} &= \text{US\$8/MMBtu} \times 30 \text{ days} \times 10,000 \\ &\quad \text{MMBtu/day} = \text{US\$2,400,000} \\ \text{Fixed payment} &= \text{US\$7/MMBtu} \times 30 \text{ days} \times 10,000 \\ &\quad \text{MMBtu/day} = \text{US\$2,100,000}\end{aligned}$$

In practice, instead of two cashflows going in the opposite directions, the payments are netted and the buyer of the swap receives US\$300,000.<sup>11</sup>

A similar swap transaction can be negotiated in the basis market. An example is shown below, using the same volume assumptions:

Floating price: IFERC Houston Ship Channel Basis  
 Fixed price: US\$0.10/MMBtu<sup>12</sup>  
 Notional volume: 10,000 MMBtu/day  
 Term: 12 month  
 Settlement: Monthly

If the basis for a specific month settles at US\$0.05/MMBtu, the floating payment will be calculated as follows, using the notional volume:<sup>13</sup>

$$\begin{aligned}\text{Floating payment} &= \text{US\$0.05/MMBtu} \times 30 \text{ days} \times 10,000 \\ &\quad \text{MMBtu/day} = \text{US\$15,000} \\ \text{Fixed payment} &= \text{US\$0.10/MMBtu} \times 30 \text{ days} \times 10,000 \\ &\quad \text{MMBtu/day} = \text{US\$30,000}\end{aligned}$$

In practice, instead of two cashflows going in the opposite directions, the payments are netted and the buyer of the swap pays US\$15,000. The basis is calculated by subtracting the Nymex natural gas contract for a specific month from the HSC index, once the index is published.

### An example: Index transactions and swaps

The instruments and transactions described above allow the buyers and sellers of natural gas to customise the ways in which they manage their market exposures. One popular transaction is known as an index deal, and is effectively a floating-price contract. An index deal involves sale/purchase of a given quantity of natural gas every month at the current index price. The current index price is often adjusted by a premium, which is typically quite small (between -2 and +2 cents) and reflects several factors discussed earlier in this chapter.

Index transactions combined with swaps are very popular

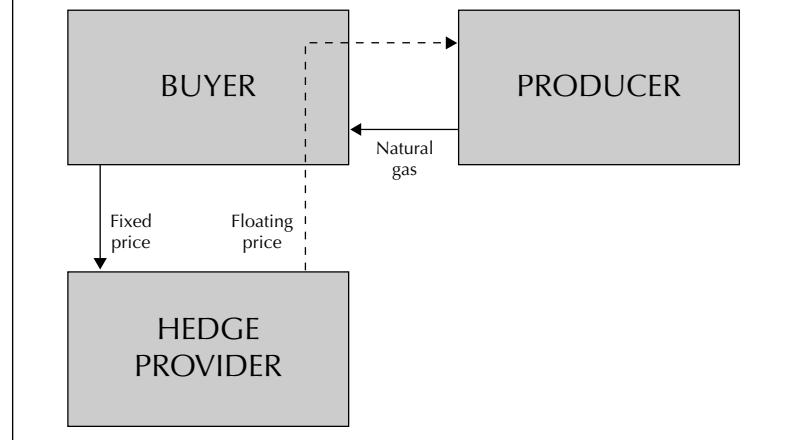
because they give the flexibility to both a buyer and a seller to maintain current relationships with counterparties that prefer to use floating prices, and, at the same time, to hedge index transaction converting floating-price exposure to a fixed-price exposure. For example, a buyer of natural gas can enter into a financial swap transaction, under which they pay the fixed price and receive a floating price (index). This buyer is paying the floating (index price) to the supplier but at the same time is receiving a floating price under the swap contract. These two flows cancel and only the fixed-price exposure is left. This hedge strategy is illustrated in Figure 11.4.

The same result can be accomplished by creating the hedge synthetically, entering into a Nymex transaction and a basis transaction. Under both transactions, the buyer of natural gas pays fixed prices and receives floating prices that add up to a locational index. Again, the buyer of physical gas receives the floating price and pays the floating price and is left with a fixed-price position.

### Futures

The natural gas futures contract launched on Nymex in April 1990 is one of the great success stories of the US energy markets. The contract became a benchmark for natural gas traded across the US and Canada, due to the flexibility of the US pipeline system and the ability to move the molecules from the production area to the market

**Figure 11.4** Converting floating-price exposure into fixed-price exposure



areas most of the time. The second critical ingredient of this success story is the development of the basis market, the market for locational differentials that supplemented the Nymex contract as a hedging vehicle.

The natural futures contract specification is shown in Table 11.5. The contract symbol is NG.

### **Exchange for physicals**

Most entries in Table 11.5 are self-explanatory. However, what does require some explanation are some special provisions related to EFP. EFPs are important as they introduce an element of flexibility into the contract and allow for delivery of natural gas at alternative locations, in volumes roughly equivalent to the futures volume. In principle, an EFP may involve a private contract in a different commodity (for example, natural gas futures may be exchanged for an electricity contract). In practice, EFPs involve the same commodity as the futures contract. An EFP consists of two parallel transactions: a futures transaction and a product transaction. A buyer of the futures is a seller of the product, and vice versa. This means that an EFP can be used to establish a futures position or close an established futures position. An EFP transaction may happen up to two hours after the end of trading in the expiring contract. It is obvious that, during the two hours following contract expiry, no new futures position may be established and one has to rely on a pre-existing futures position to execute an EFP (the same transaction may happen, of course, prior to expiry). The private contract does not have the credit protection that the futures contracts receive. Losing the credit support of the exchange is the cost of flexibility an EFP offers.

By market convention, a buyer of an EFP receives natural gas and pays the invoice price for it.<sup>14</sup> The seller of an EFP delivers natural gas and receives the invoice price. An example of an EFP below is the case of counterparty with an original long futures position exchanging it for physical gas at a specified location. The seller of physical gas receives, from the EFP buyer, the invoice price that is equal to the posted price plus a differential. The differential represents the basis component of the trade and is subject to negotiations between the parties. The buyer completes the transaction by transferring the futures to the EFP seller's account at the posted price. The posted price in principle has no impact on the economic outcome of

**Table 11.5** Nymex natural gas contract specification

<b>Trading unit</b>	10,000 MMBtu.
<b>Price quotation</b>	US dollars and cents per MMBtu.
<b>Trading hours</b>	Open outcry trading is conducted from 9:00 am until 2:30 pm.
	Electronic trading is conducted from 6:00 pm until 5:15 pm via the CME Globex trading platform, Sunday–Friday. There is a 45-minute break each day between 5:15 pm (current trade date) and 6:00 pm (next trade date). Off-exchange transactions can be submitted solely for clearing to the Nymex ClearPort clearing website as an exchange of futures for swaps (EFS) or exchange of futures for physicals (EFP) transaction until 5:15 pm, Monday–Friday, and the day preceding a holiday.
<b>Trading months</b>	The current year plus the next 12 years through December 2020. A new calendar year will be added following the termination of trading in the December contract of the current year.
	On CME Globex: The current year plus the next eight years.
<b>Trading at settlement (TAS)</b>	TAS is available for the front two months except on the last trading day and is subject to existing TAS rules. Trading in all TAS products will cease daily at 2:30 pm Eastern Time. TAS products will trade off of a <i>base price</i> of 100 to create a differential (plus or minus) in points off-settlement in the underlying cleared product on a one-to-one basis. A trade done at the base price of 100 will correspond to a “traditional” TAS trade that will clear exactly at the final settlement price of the day.
<b>Minimum price fluctuation</b>	US\$0.001 (0.1¢) per MMBtu (US\$10.00 per contract).
<b>Maximum daily price fluctuation</b>	US\$3.00 per MMBtu (US\$30,000 per contract) for all months. If any contract is traded, bid or offered at the limit for five minutes, trading is halted for five minutes. When trading resumes, the limit is expanded by US\$3.00 per MMBtu in either direction. If another halt were triggered, the market would continue to be expanded by US\$3.00 per MMBtu in either direction after each

<b>Last trading day</b>	successive five-minute trading halt. There will be no maximum price fluctuation limits during any one trading session.
<b>Settlement type</b>	Trading terminates three business days prior to the first calendar day of the delivery month.
<b>Delivery</b>	Physical. The Sabine Pipe Line Co. Henry Hub in Louisiana. The seller is responsible for the movement of the gas through the Hub; the buyer, from the Hub. The Hub fee will be paid by the seller.
<b>Delivery period</b>	Delivery shall take place no earlier than the first calendar day of the delivery month and shall be completed no later than the last calendar day of the delivery month. All deliveries shall be made at as uniform as possible an hourly and daily rate of flow over the course of the delivery month.
<b>Alternate delivery procedure (ADP)</b>	An ADP is available to buyers and sellers who have been matched by the exchange subsequent to the termination of trading in the spot month contract. If the buyer and seller agree to deliver under terms different from those prescribed in the contract specifications, they may proceed on that basis after submitting a notice of their intention to the exchange
<b>Exchange of futures for physicals (EFP) or swaps (EFS)</b>	The commercial buyer or seller may exchange a futures position for a physical position or a swaps position of equal quantity by submitting a notice to the exchange. EFPs and EFSs may be used to either initiate or liquidate a futures position.
<b>Grade and quality specifications</b>	Pipeline specifications in effect at time of delivery.
<b>Position accountability levels and limits</b>	Any one month/all months: 12,000 net futures, but not to exceed 1,000 in the last three days of trading in the spot month.

---

Source: [www.nymex.com](http://www.nymex.com) (as of July 2, 2008)

the transaction. This can easily be seen by comparing the cashflows related to both legs of the transaction. The buyer gives away futures valued at the current market price and receives the posted price in return. At the same time, they receive natural gas valued at the current location-specific market price and pay the posted price plus a negotiated differential.<sup>15</sup>

Futures leg: Posted price – Futures price

Product leg: Locational price – Posted price + Differential

As one can see, the EFP buyer pays and receives the posted price, a wash. What matters is that the locational price is the futures price adjusted by a differential, following market conventions. In principle, the posted price could be equal to a penny or one billion. There is, however, one good reason to transfer the futures at a posted price as close as possible to the current futures price. The posted price will affect the margin that one counterparty will post and the other will receive. Using the current futures price makes the margin level equitable to both counterparties.

EFPs used to be very popular in the US natural gas markets in the 1990s, and served as a tool of price discovery for physical gas at locations away from Henry Hub. The price discovery process now occurs through direct transactions or physical basis transactions, as explained above. EFPs were dethroned and replaced by Exchange for Swaps (EFS), used by traders who want to adjust their futures position, often without attracting the attention of competitors. The mechanics of the EFS is very similar to the EFPs. EFPs remain important in other markets (especially Brent, as will be explained in Chapter 17).

### ICE swaps

Another development in the natural gas markets in the US has been the growth of Nymex lookalikes. These transactions allow one to acquire economic exposure to the price of the Nymex futures contract, while trading away from Nymex proper, using ICE, the OTC markets or ClearPort. The most popular instrument is the ICE Nymex swap (or Nymex lookalike), based on the specifications shown in Table 11.6. This contract allows market participants to acquire exposure to the final settlement price of the expiring Nymex contract (LD1) through transactions executed on ICE. The details of this contract can be found on the ICE website.<sup>16</sup>

Table 11.6 contains a comparison of the ICE natural gas swap with the Nymex futures contract. At expiration, both the Nymex natural gas contract and the ICE swap have an identical payout profile. Prior to expiry, the swaps trade base on the expectation of the final Nymex settlement price.

Following the 2000 CFMA, ICE was not subject to the rules related

to the position limits (direct position limits and the accountability limits). This changed when the CFTC determined that this contract performs a significant price discovery function (SPDC). As explained in a note from a law firm:

The CFTC's authority to designate SPDCs was introduced in the 2008 Farm Bill, which reauthorized the CFTC and enhanced its authority over Exempt Commercial Markets (ECM), closing the so-called "Enron loophole" that had exempted products traded on ECMs from most substantive CFTC regulation. The law defined SPDCs as contracts that perform a "significant price discovery function" in either the regulated futures markets or the cash markets, based on the extent to which the contracts: (1) are linked to existing exchange-traded contracts; (2) permit arbitrage between ECMs and other markets; (3) are used as a direct price reference for bids, offers or transactions in a commodity; and (4) are traded in a volume that provides material liquidity.<sup>17</sup>

**Table 11.6** Comparison of Nymex and ICE natural gas contract specifications

	Nymex natural gas futures contract	ICE natural gas Henry Hub swap
Trading unit	10,000 MMBtu	2,500 MMBtu
Price unit	US\$/MMBtu	US\$/MMBtu
Last trading date	Three business days prior to the first calendar day of the delivery month.	Three business days prior to the first calendar day of the delivery month.
Settlement type	Physical	Financial
Final settlement price	Volume-weighted average price of trades during the last half-hour of the last trading day (2:00–2:30 pm).	Same as Nymex final settlement price on the last trading day
Delivery location	Henry Hub, LA	N/A
Delivery period	First calendar day of the delivery month through last calendar day of the delivery month	N/A
Trading hours	Open outcry: 10:00 am – 2:30 pm. Electronic trading: 6:00 pm of the prior trading day to 5:15 pm of the trading day.	Electronic trading: 2:30 pm of the prior trading day to 2:30 pm of the trading day.

Source: "Excessive Speculation in the Natural Gas Market, Staff Report with Additional Minority Staff Views," Permanent Subcommittee on Investigations, US Senate, June 25 and July 9, 2007 Hearings

The CFTC finalised its rules regarding SPDC designations in March 2009, and the Henry Hub Financial LD1 Fixed Price contract was designated as an SPDC in June 2009. Other natural gas contracts in this category are: the NWP Rockies Financial Basis Contract, HSC Financial Basis Contract, PG&E Citygate Financial Basis Contract, Waha Financial Basis Contract, Chicago Financial Basis Contract, SoCal Financial Basis Contract, and Alberta Financial Basis Contract.

### **Volumetric production payments<sup>18</sup>**

VPPs are a very popular form of commodity-based financing. Reduced to first principles, these transactions are fairly simple. The buyer of a VPP advances funds to the E&P company (a seller) that uses the proceeds for development of a producing property or general corporate purposes. The buyer is typically a financial institution or merchant energy company with a finance unit,<sup>19</sup> although some VPPs were executed by municipalities and state agencies using proceeds from the issuance of municipal bonds as the source of funds. The seller has the obligation to repay the advanced funds through predetermined volumetric flows of the contracted commodity from producing properties dedicated to the transaction (in some cases, through cashflows generated from the pledged properties). A VPP is often associated with a hedge designed to neutralise price risk by locking-in prices of the commodities underlying a VPP. A hedge may be put on by either a buyer or a seller. A buyer receives predictable cashflows as the commodity volumes are delivered, which helps to reduce the overall risk of the transaction.

Technically, a VPP is a carved-out slice of “working interest”<sup>20</sup> in a number of natural gas or oil properties, transferred to a buyer. A buyer will usually insist on using only proved developed producing (PDP) reserves for the support of a VPP transaction. Flows from such reserves can be predicted with a high degree of precision, and this helps to reduce the volume risk of a VPP. The volumetric flows are typically specified in absolute amounts or as percentages of total flows from the properties allocated to a VPP. A seller will be obligated to allocate a portion of the excess PDP flows (above the VPP volumes) to cover operating costs and taxes. Remaining excess PDP flows, PDP flows after termination of a VPP and the flows from proved undeveloped (PUD) reserves belong to the seller.

VPPs share a number of common features, but in practice may be highly customised. The ability to customise a VPP transaction offers a number of distinct advantages both from the point of view of a seller and a buyer. The benefits to a buyer include:

- ❑ the seller of a VPP continues to operate the properties, covers operating expenses and pays taxes;<sup>21</sup>
- ❑ operational risks are borne by a seller, including potential costs of environmental damage;
- ❑ the buyer acquires control of volumetric flows without having to set up an organisation to manage the properties, an activity that may be outside its core competences;
- ❑ a VPP structure provides solid credit protection as it is considered to be bankruptcy remote; a buyer is protected against default by a seller as the properties supporting a VPP are encumbered (or ring-fenced in a special-purpose entity, SPE);
- ❑ the buyer of a VPP accepts price risk but this risk can be managed through hedges established at the inception of the transaction; the buyer enters into commodity swaps to lock-in the future prices and the interest rate swaps to eliminate the risk of higher interest rates eroding the current value of future nominal cashflows;
- ❑ some residual volumetric risk remains, because the actual delivered volumetric flows can diverge from the committed levels; however, a buyer can negotiate contractual provisions that would allow recouping a production shortfall in a given month by having the volumes made up in subsequent months, and adjusting the volumes for a change in market prices as well as the time value of money; and
- ❑ cashflows from a VPP (or a number of VPPs) can be pooled and securitised by issuing debt instruments sold to individual or institutional investors.

One should keep in mind that a buyer may be exposed to some residual operational and bankruptcy risk. For example, production outages may delay volumetric flows. A financially weak producer may delay field maintenance and this may, in turn, translate into production shortfalls. We shall discuss below how these risks can be further minimised.

---

A VPP is not without risks from the perspective of a buyer. There is reserve risk, the biggest potential risk, which is related to future volumetric flows that may fall short of the levels assumed in a VPP transaction or may materialise with a significant delay. This risk can be very high if the properties underlying a VPP are located in non-conventional resource plays, with limited production history and poorly understood geological properties. A buyer has a number of defences against this risk, including the following.

- ❑ *Overcollateralisation*. A seller may be required to allocate excess reserves to a contract or a contract may contain provisions for drilling additional wells to offset production shortfalls.
- ❑ *Inspection*. The properties allocated to a VPP are assessed by a competent geologist with no conflict of interest (ie, has no preestablished business relationship with a seller).<sup>22</sup>
- ❑ *Diversification*. The properties are diversified, in terms of the number of and locations of wells, fields and basins, and have been in operation for a period of time long enough to establish a production profile.
- ❑ *Contract structure*. The VPP contract contains provisions for make-up volumes in case of field underperformance due to factors such as production outages and other operational problems (such as pipelines' OFOs). Another potential risk is the possibility that a seller may accelerate production flows (when prices are high), impairing field integrity and future production levels. Good lawyers can always come up with contract language to neutralise this risk.

There are many reasons why producers are interested in VPPs.

- ❑ *Cost*: A VPP is effectively a secured loan, and this means that a buyer is willing to accept a lower interest rate.
- ❑ *VPP is non-recourse to the seller* (ie, the lender recovery is limited to the collateral, the properties assigned to a VPP). The royalty interest acquired by the buyer is not treated as part of the seller's estate and is bankruptcy protected. Of course, the implications of this provision are that other creditors are relegated to the ranks of holders of subordinated debts with respect to the buyers of the VPP. Therefore, a VPP transaction increases the

overall risk profile of a given company and reduces its capacity to acquire additional debt. There is no free lunch.

- ❑ Tax treatment: A VPP is treated as a loan for tax purposes.<sup>23</sup> Tax treatment is covered by the Section 636 of the Internal Revenue Code and related Treasury regulations. Specific regulations impose restrictions on VPPs that must be observed to qualify for a loan tax treatment. One important requirement is that the expected life of a VPP should be shorter than the expected economic life of the underlying properties.
- ❑ VPPs are not treated as loans under GAAP accounting: The accounting treatment is used to credit at inception a deferred revenue account and debit cash. As the volumes are delivered, a revenue account is credited and a debit is made to the deferred revenue.
- ❑ Rating agencies treat VPPs as loans:<sup>24</sup> However, they do recognise a number of benefits<sup>25</sup> related to the non-recourse feature and self-liquidating properties.
- ❑ Reserve profile management: Under FAS 69, the reserves and volumetric flows committed to a VPP are excluded from the proved reserves and production numbers.<sup>26</sup> This has an indirect benefit to a producer: the production profile looks flatter or may be increasing, even if the reserves are falling over time. This happens because a portion of early volumetric flows is removed from the reported numbers.

In spite of transferring some significant risks to a buyer, the seller of a VPP continues to be exposed to some risks related to the following.<sup>27</sup>

- ❑ The production cost and taxes. The seller is exposed to the risk of increasing labour and equipment cost over time, and will incur a loss if they cannot control the production cost.
- ❑ In some instances, the seller retains part of the volumetric risk. In case a make-whole provision is embedded in the contract, the seller has to deliver additional volumes if there is a temporary decline in production flows. The additional flows continue until the shortfall is eliminated. This exposes the seller to a cashflow risk and leveraged volume risk. The liability for the volume shortfall may accumulate accordingly to the market prices and if a drop in commodity flows is associated with the period of higher market

prices, significantly higher volumes may be added to the schedule of future commitments. Of course, if the interruption coincides with lower commodity prices, this contractual arrangement will benefit the seller. The need to compensate for prior delivery shortfalls introduces an element of uncertainty, with the potential for the quantity and price risk to compound.

Volumetric Production Payments are a transaction unique to the US. We have seen over time a number of reserve-based financing deals in Canada and Latin America, but they were different in one critical respect from VPPs. The buyers do not acquire “real property interest” in the hydrocarbons outside the US and are effectively a counterparty to a prepaid forward contract. One can, of course, structure a lookalike VPP in Canada, but the legal standing of a buyer would be weaker than in the case of the US-based transaction. This is why Enron used a Cactus vehicle model for the US VPPs and a Caribou vehicle for Canadian deals.<sup>28</sup>

Discussion of VPPs would not be complete without a short review of a version of a pre-paid forward contract used to disguise loans as proceeds from commodity transactions. A prepaid forward accelerates the receipt of cash to the inception of the transaction. In the late 1990s, several companies used such schemes to raise funds and hide debt at the same time. In one case, a transaction attracted the attentions of the criminal justice system and resulted in prison sentences for a number of employees involved in structuring the deal. Most of such transactions shared a number of common features (captured in Figure 11.5).

A financial institution (Bank ABC) sets up a special-purpose entity that it controls but does not technically own.<sup>29</sup> ABC transfers funds to the SPE, which uses the money to prepay XYZ Company for future deliveries of natural gas or oil.

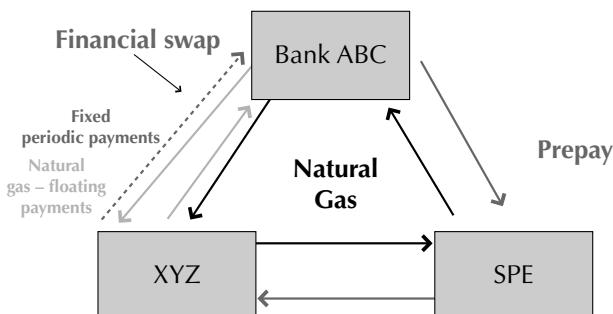
XYZ and SPE enter into a physical commodity transaction under which XYZ delivers over time natural gas to the SPE, according to a predetermined schedule. Simultaneously, the SPE enters into a mirror physical transaction with ABC, under which it “repays” the advanced funds in-kind through volumetric flows of natural gas. ABC enters into a corresponding transaction with XYZ, under which ABC “flows” natural gas to XYZ at current market prices. Under a triangular arrangement natural gas “flows” from XYZ to the SPE,

from the SPE to ABC, and from ABC to XYZ, in exactly the same volume. In practice, this means that no natural gas flows at all: all the flows in this closed system cancel out. Loan repayment happens through a related financial swap between ABC and XYZ. Under a fixed-for-floating swap, with natural gas as the underlying commodity, XYZ is a fixed price payer and receives floating – ie, current market prices. The floating leg offsets what XYZ “pays” to ABC for “deliveries” of natural gas. What remains after all the cancellations is the initial conveyance of funds to XYZ and a series of fixed-price payments under a financial swap going from XYZ to ABC, which are calibrated to cover principal and interest payments. To the outsider, this looks like a legitimate prepaid forward, a perfectly legal transaction. In reality, these related transactions are designed to misrepresent the true leverage of the XYZ Company. Such schemes were totally discredited in the early 2000s following the failure of a number of companies with a proclivity to borrow through such arrangements. But as John Kenneth Galbraith, observed:

Financial operations do not lend themselves to innovation. What is recurrently so described and celebrated is, without exception, a small variation on an established design [...] The world of finance hails the invention of the wheel over and over again, often in a slightly more unstable version.<sup>31</sup>

It is only a matter of time before a similar transaction happens again.

**Figure 11.5** A prepay transaction<sup>30</sup>



Source: US Senate.

### Municipal prepay transactions

Municipal prepay transactions are a special version of a prepayment for future deliveries of natural gas involving the issuance of municipal (tax-exempt) bonds. The legal framework for such transactions was established by the IRS Bulletin 2003-41, issued on October 14, 2003, and subsequently confirmed in the Energy Policy Act of 2005, Section 1327. Under this structured transaction, a municipality or state gas authority<sup>32</sup> issues tax-exempt bonds, with the proceeds being used to prepay the natural gas supplier. The proceeds are placed in a special account and invested at taxable rates exceeding the cost of servicing the debt. The difference is passed to the municipality as a discount to the price of natural gas. The supplier of natural gas (usually a marketer associated with an oil major or a big financial institution) typically receives a profit reflecting multiple services performed for the municipal/state agency:

- structuring fees;
- compensation for managing the issuance of tax-exempt bonds;
- the portion of the interest rate spread (tax-exempt versus taxable) not passed through to the municipality;
- bid–offer spread; and
- other fees that may be built into the structured gas supply contract.

*Leg (1).* The agency representing end users of natural gas issues tax-exempt bonds.

*Leg (2).* The proceeds from the bond issue are passed to the gas supplier. The supplier delivers natural gas to the agency. In this example, we assume gas is delivered at the first-of-the-month index.

*Leg (3)* Natural gas is flown to the end users, who pay index (potentially minus certain discounts representing the benefits of the tax arbitrage to the end users).

*Leg (4).* Municipal/state agency enters into a swap contract with a third party under which it pays floating price (index) and receives fixed. Effectively, the payments received from the end users of natural gas (paying monthly index) are flown to the swap counterparty. The fixed-price cashflows under the swap can be used to service the debt.

*Leg (5).* Gas supplier enters into a mirror swap under which it receives index and pays a fixed price.

*Leg (6).* The gas supplier buys natural gas at the market at the monthly index. Natural gas is flown through the municipal/state agency to the end users.

*Leg (7).* The end users of natural gas or a municipal agency buy protection against non-performance by the gas supplier (and possibly financial swap provider) from an insurance company.

The transaction contains a number of risks to the buyers of natural gas. The credit risk is an obvious one. The natural gas supplier may go bankrupt (a possibility that cannot be dismissed given the history of the financial and energy industries), and so may the insurance company providing protection to the end users. It cannot be excluded that all the financial firms involved in this scheme could suffer a similar fate through exposure to the same markets. For the natural gas supplier, the potential risk is embedded in the way the contracts with the end users are structured. If the contracts are based on take-and-pay provisions, the supplier is vulnerable to volumetric risk should the buyer fail to take all the expected volumes. The volumes of natural gas not taken by the end users under the contract have to be re-marketed, potentially in the volatile daily markets. The fact that the supplier buys natural gas at the floating price (see Leg (6)) offers a degree of protection against this risk.

### **Processing plant contracts**

Natural gas processing plants operate under a number of different contractual arrangements that have important consequences for their profitability under different market conditions, as well as for approaches to hedging.

*Fixed-fee arrangements.* Under fixed-fee arrangements, producers pay a fixed fee to gather and process their natural gas.

*Percentage-of-proceeds (POP) arrangements.* Under POP arrangements, a processing plant receives and processes natural gas on behalf of producers and sells the residue gas and NGL volumes, typically at

index-related prices. The producers receive an agreed upon percentage of the residue gas and NGL proceeds. Under these types of arrangements, the revenues and gross margins of a processing plant increase as natural gas and NGL prices increase, with the opposite happening as natural gas and NGL prices decrease.<sup>33</sup> In other words, a plant is long both natural gas and natural gas liquids.

*Percentage-of-index (POI) arrangements.* POI contracts are divided into two subcategories. Under a POI contract, a processing plant purchases and resells natural gas at index-related prices. The relationship between the index price<sup>34</sup> and the contractual prices (for purchase and resale) is defined in terms of percentages or absolute differentials. The margin is derived from the difference between the prices (percentages or differentials) on both sides. Under a POI with a POP “switch” arrangement, a plant has the option to switch to another, more favourable contractual arrangement if market conditions benefit a plant.

*Keep whole contract.* Under a “keep-whole” contractual arrangement, a processing plant takes title to the NGL volumes separated from the natural gas stream. The producers are compensated for what is called shrink in the industry jargon, the Btu content of the NGLs removed in processing. This means that a processing plant benefits from higher NGL prices and suffers from higher natural gas prices, as it has to purchase dry natural gas needed for replacement of removed liquids (on the Btu basis). In other words, a processing plant has a long position with respect to a basket of natural gas liquids and a short position with respect to natural gas. The spread a processing plant is exposed to under such a contract is known as a frac spread.<sup>35</sup>

The processing plants exposed to the frac spread risk can take a number of defensive measures. Some of these include hedging strategies that utilise different financial instruments, including swaps and options. An option-based strategy consists of purchasing calls on natural gas and/or buying puts on some of the NGLs for which more liquid markets exist. Some energy trading operations offer options on baskets of liquids.

Some processing plants rely on the proxy hedges, which exploit the correlation between prices of different commodities. For

example, ethane prices follow natural gas prices relatively closely, while those of propane and butane are related to the prices of crude oil. These relationships reflect technological interdependencies. As explained in Chapter 8, ethane may be left in the natural gas stream in order to enhance its calorific value. Other NGLs (such as propane or butane) are produced both by refineries and the natural gas industry and may be used as a feedstock in the production or blending of many refined products. Hedging gas liquids with crude oil may, however, be quite risky. For example, suppose that a gas processor is long NGLs and hedges by selling crude oil forward. Oil prices have a propensity to spike without a corresponding spike in the prices of natural gas liquids. Growing output of natural gas from shale formations boosts the production of liquids to levels at which they cannot be absorbed by the chemical industry without lowering prices.<sup>36</sup> The losses on hedge positions (short crude positions) may not be offset by gains (if any) on the hedged positions of natural gas liquids.

Other forms of risk mitigation include the exploitation of optionality embedded in contracts and physical assets. For example, low liquids prices create incentives to leaving some liquids (primarily ethane) in the natural gas stream, increasing the energy content of natural gas exiting the tailgate of a plant. Higher energy content of natural gas delivered from a plant means that a smaller volume of the dry gas has to be acquired to replace the shrink. Some processing plants invested in technology that increases the ability to modify technological process depending on the market conditions. One example of such technological innovation is provided in the filling by Copano Energy LLC:<sup>37</sup>

In 2003, we modified the processing plant to provide natural gas conditioning capability by installing two new 700 horsepower, electric-driven compressors to provide propane refrigeration through the lean oil portion of the plant, which enables us to shut down our steam-driven refrigeration compressor when we are conditioning natural gas. A third 700 horsepower electric-driven compressor has been installed and became operational in January 2007. These modifications provide us with the capability to process gas only to the extent required to meet pipeline hydrocarbon dew point specifications. Our ability to condition gas, rather than to fully process it, provides us with significant benefits during periods when processing is not profitable (when the price of natural gas is high compared to the price of NGLs), including:

- providing us with the ability to minimise the level of NGLs removed from the natural gas stream during periods when natural gas prices are high relative to NGL prices; and
- allowing us to operate our Houston Central Processing Plant more efficiently at a much reduced fuel consumption rate while still meeting downstream pipeline hydrocarbon dew point specifications.

In other words, the investment in the technology allows the plant to switch from the full processing mode to the conditioning mode, on very short notice (about one hour). The conditioning mode allows keeping the heavier hydrocarbons in the natural gas stream to the extent allowed by pipeline tariffs.

The solution used by Copano Energy is an illustration of a real option, a contractual and physical flexibility created through the implementation of forward-looking management and investment strategies. Such strategies form a link between investment decisions and future operational decisions, and represent a creative way of improving energy efficiency. It is important to note that this strategy requires careful coordination of the contractual arrangements and technological solutions. The plant management has to negotiate contracts with customers, allowing the plant to switch to the reconditioning mode at its discretion. A technological solution alone is not sufficient.

The existence of the optionality described above has obvious consequences for the hedging strategy. As one can remember, under keep-whole contracts a plant is short natural gas and can mitigate the risk of higher natural gas prices through the purchase of natural gas swaps and/or calls. When a plant operates in the conditioning mode, the plant requires a different type of protection against high gas prices as more natural gas is delivered at the tailgate. As explained, if the prices of natural gas increase significantly, a plant switches to the reconditioning mode, and more natural gas (on the Btu basis) leaves the plant (some NGLs are left in the gas stream). Risk mitigation can be accomplished by buying a call option on natural gas with a strike price corresponding to the targeted profitability level when the plant still operates in full processing mode. At the natural gas prices corresponding to the cross-over point to the conditioning mode, this protection becomes fully or partially redundant. A risk manager can take advantage of this by selling a call with a higher strike price. Such

a strategy is known as an option spread (not to be confused with a spread option, explained in Chapter 4): buying a call with a low strike and selling an otherwise identical call, but with a higher strike price. This strategy reduces the overall cost of this hedging strategy, as the sale of the higher strike price option allows the cost of acquisition of the low strike option to be partially offset. The calibration of this hedging strategy requires a good understanding of the flexibility of the plant and an understanding of the price ranges at which plant operations will change.

The managers of natural gas processing plants face many challenges in hedging exposures of their operations. One problem is related to the relative lack of depth in the market for natural gas liquids and relatively short tenors of available options and swaps. There are only a few firms with sufficiently big natural gas liquids books or the ability to proxy hedge the baskets of liquids.<sup>38</sup> Proxy hedges require a deep understanding of the market and considerable quantitative skills required for determination of the hedge ratios. Hedging of the exposure, even of a relatively modest gas-processing operation, may have a significant market impact – ie, the market prices may move against the hedger. The hedging firms face a dilemma of dealing with a hedge provider on an exclusivity basis (a provider who may be trusted or not) or asking a larger number of financial firms for competitive bids, exposing the plant and hedge providers to front-running<sup>39</sup> and significant market impact, when the news about an intended hedging transaction spreads across trading floors. The second problem is related to the complexity of the portfolio of contracts of gas-processing plants. A typical operation has multiple contracts with different provisions and exposure to prices at multiple locations. The determination of the exposure profile of a gas-processing firm and the selection of the optimal portfolio of hedges (for example, the level of strike prices for the options) requires considerable quantitative skills.

A typical natural gas-processing company with operations in multiple basins, processing natural gas of varying properties, typically has a portfolio of contracts with different producers. The composition of the portfolio is the result of historical developments, changing the bargaining power of processors and producers, the level of competitive pressures in different regions and the ability of the company to manage price risk. Evolving market conditions

prompt processors to migrate over time to certain contractual structures, which often last longer than the price relationships that were the factors behind them. A good example of the complexity of a portfolio of gas-processing contracts can be found in the financial statements of DCP Midstream, one of the most important companies in this field, and other gas processors.<sup>40</sup> Gross margins on different contracts are shown in Table 11.6.

The differences in gross margins between the different types of contracts and the current price environment – low prices of natural gas, falling prices of ethane, relatively firmer, but still falling, prices of propane and butane – explain the sources of value for DCP (see Figure 11.6) and the disconnect between the composition of a typical NGL barrel and the values that different components represent.

### Volumetric risk management

One of the features of energy markets is the interaction between price risk and volumetric risk. Many contractual arrangements between suppliers and end users contain explicit or implied commitment to respond to short-term fluctuations in demand, often at a flat price to the buyer. Suppliers are, however, exposed to price risk. On a cold winter day, residential users of natural gas respond with increased consumption and acquire additional volumes at a tariff rate. The same is true of electricity demand on a hot summer day in the US: increased use of air conditioning will translate into greater dispatch rates of natural gas-fired power plants. Volumetric risk has two features:

- ❑ it is multiplicative with respect to price risk; additional demand ( $\Delta q$ ) has to be usually satisfied at higher prices ( $\Delta p$ ) than the base load demand; and
- ❑ volumetric risk is often related to weather conditions.

**Table 11.6** Gross margin by contract (DCP Midstream)

Contract type	Unit	Q4 2011	Q3 2011	Q2 2011	Q1 2011
POP	US\$/MMBtu	0.65	0.74	0.76	0.68
Keepwhole	US\$/MMBtu	1.68	1.64	1.65	1.48
Fee: Gas	US\$/MMBtu	0.15	0.14	0.14	0.13
Fee: NGL	US\$/gallon	0.02	0.02	0.02	0.02

Source: <https://www.dcpmidstream.com/pdf/Q4%202011%20MarginbyContract.pdf>

The energy industry has developed a number of solutions to address volumetric risk. For the case of natural gas, we shall discuss several solutions which are used extensively in practice, including:

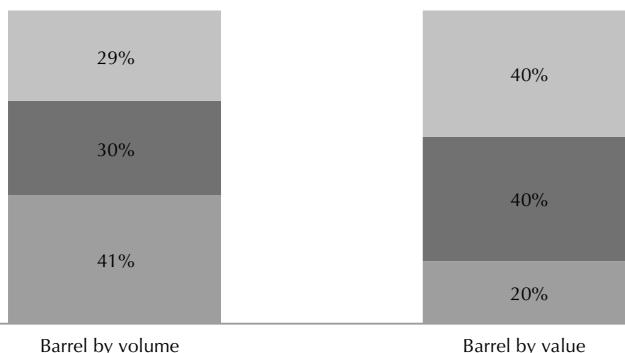
- swing options;
- call options embedded in swaps;
- weather derivatives; and
- tariff provisions.

Tariff provisions (no-notice service or authorised overrun) were covered in Chapter 10, and there is nothing more we can add here.

Swing options are unique in being a topic of more academic papers than the total number of such options ever transacted as standalone contracts. They tend to attract quants, given the many interesting features they share, but they are found in practice only as embedded optionality in physical supply contracts. As such, they are very important and should be valued and managed with care. They are typically hedged by using optionality in physical assets controlled by the supplier, or by buying mirror options from other market participants who own such assets. Valuation of such options is important for two reasons. To the extent they are embedded in structured transactions, a value should be assigned to them in order

**Figure 11.6** Generic NGL barrel profiles

■ Ethane (C2) ■ Propane (C3) ■ Butanes and heavier



Source: [https://www.dcpmidstream.com/pdf/01-13-12\\_2012Plan\\_FINAL%20only.pdf](https://www.dcpmidstream.com/pdf/01-13-12_2012Plan_FINAL%20only.pdf)

to avoid giving them away or overpaying for them. Such options are often hedged by acquiring flexibility in a physical asset through an investment decision, and the valuation of technological flexibility is a necessary step in the process of capital budgeting.

Swing options share features of both American and path-dependent options. Path dependence in this context means that at any point in time the exercise rights depend on decisions related to a given option taken in the past. A typical natural gas swing option gives the holder the right to swing – ie, acquire a certain volume of natural gas  $n$  times during a given month. The volume varies between a minimum amount (which can be equal to zero) and a maximum amount ( $Q_{\min}$  and  $Q_{\max}$ ). The modelling and valuation of such options is often based on the assumption that, if a customer *swings*, ie, exercises the option at all, they will opt for the maximum volume instead of choosing  $Q_{\min}$  or an intermediate level ( $Q_{\min} < Q < Q_{\max}$ ). There may be additional volume restrictions, such as minimum and maximum monthly and/or quarterly takes. Should a buyer exceed the maximum quota or fail to satisfy the minimum take requirements, a penalty may be applied based on current market prices or a negotiated fee. The penalty for a shortfall may be applied to future purchases, after accounting for the time value of money. Figure 11.7 contains a stylised representation of a swing option, with a customer exercising 5 out of 10 exercise rights.

Swaps with embedded options are relatively simple instruments. An embedded option gives a buyer the right to buy additional volume (for example, an additional 50%) at their discretion at the fixed price of a swap. The fixed price of a swap is equal to  $F_p + \Delta F_p$ , where  $F_p$  is the fixed price of the swap without an embedded option and  $\Delta F_p$  is the additional charge for the embedded call option. The additional charge is calculated by “spreading” the option premium over the entire swap volume. Effectively, the premiums for a strip of options are levelised and added to the swap price.

Weather derivatives were discussed in Chapter 5. In this section we shall show a few examples of weather derivative structures designed specifically to address volumetric risk.

A market participant (for example, a utility) who stores natural gas for the winter months faces the risk of exceptionally cold weather in November and December depleting their inventory and forcing them to purchase additional volumes of natural gas.<sup>41</sup> An option with a following payout offers a protection against this risk:

$$\text{Payout} = V \times (\text{HDD}_{\text{Nov-Dec}} - \text{KHDD}_{\text{Nov-Dec}})^+ \times \text{PG}_{\text{Jan}}$$

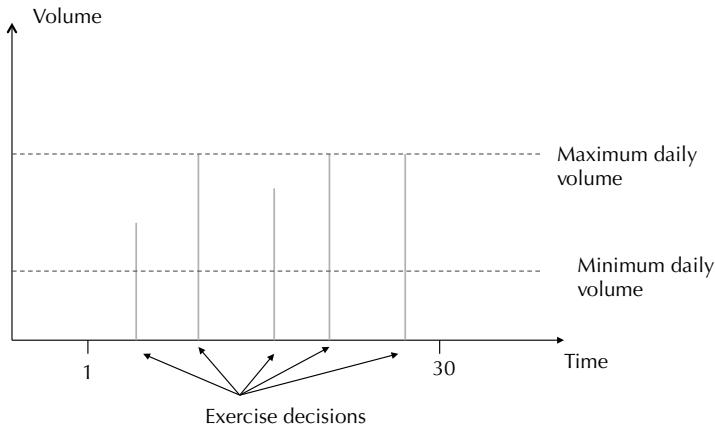
where  $V$  denotes notional volume (for example, 10,000 MMBtus),  $\text{HDD}$  is the cumulative heating degree days number for the November–December time period,  $\text{KHDD}$  is the strike (for example, 950),  $\text{PG}$  is the price of natural gas (first of the month, January index). The superscript “+” (positive part) means that the payment is received by the utility only if the difference  $(\text{HDD} - \text{KHDD})$  is positive. If the difference is negative, the payout is reset to zero. In case the beginning of winter is colder than usual, cumulative heating degree days will likely exceed the strike of 950, and the utility will receive the amount of cash equal to the first of the month price of natural gas multiplied by 10,000 (the notional volume) and by the differential  $\text{HDD}_{\text{Nov-Dec}} - \text{KHDD}_{\text{Nov-Dec}}$ . The payout received by the utility helps either to buy additional gas in the spot markets to support the current load (if the cold weather continues) or to rebuild the inventory of natural gas in storage for the rest of the winter.

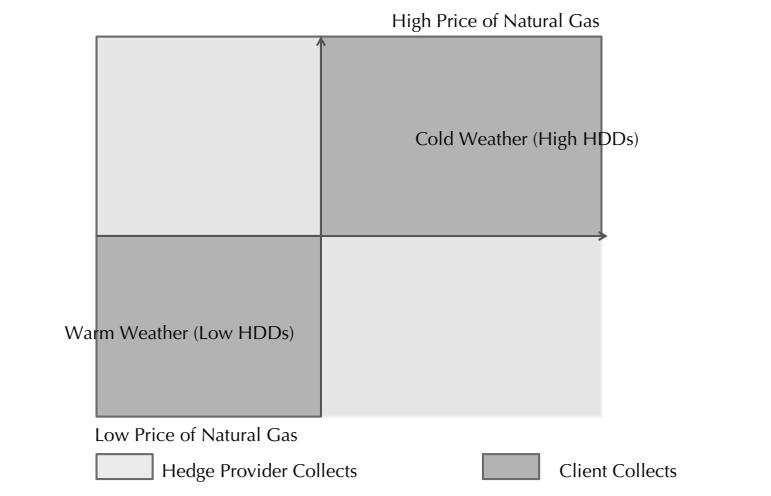
Another example is a swap with two sources of risk, weather (temperature) and the price of natural gas. The payout of the swap is given by:

$$\text{Payout} = V \times (\text{HDD}_{\text{Nov-Dec}} - \text{KHDD}_{\text{Nov-Dec}}) \times (\text{PG}_{\text{Jan}} - \text{KG}_{\text{Jan}})$$

where the symbols are the same as above.  $\text{KG}$  stands for the strike price (natural gas). Figure 11.8 illustrates the payout profile. The

**Figure 11.7** Swing option



**Figure 11.8** Combined weather and natural gas swap

terms “high/low” and “warm/cold” are defined with respect to the strike prices. A utility (in this example a client acquiring weather protection from a hedge provider) receives payment when both terms ( $(HDD_{Nov-Dec} - KHDD_{Nov-Dec})$ ,  $(PG_{Jan} - KG_{Jan})$ ) are either both positive or both negative (this means the product is positive). When the weather is cold (high cumulative HDDs) and prices of natural gas are high, the inventories of natural gas are depleted and the gas purchase contracts may be used to the limit. The utility may be forced to buy additional volumes of natural gas in the spot market. If this coincides with the high price of natural gas, the utility will be negatively affected. A utility also receives a payment when it is warm (low cumulative HDDs) and prices of natural gas are low. The sales to customers are low and revenues are reduced. Low prices of natural gas make it difficult to ask regulators for recovery of the shortfall. In summary, a utility receives payments under the circumstances described by the dark grey quadrants, and makes payments to the hedge provider in grey ( ) quadrants.

## CONCLUSIONS

We have reviewed in this chapter a very complex and mature US natural gas market with its own unique culture and specialised language and conventions. We have covered a number of physical and financial transactions that can be used either for hedging or spec-

ulation. The unique features of this market include the ability to trade locational spreads (which are defined, by convention, with respect to the central price benchmark: futures contract with delivery at Henry Hub, Louisiana) and reliance on price indexes for physical transactions established through price-reporting organisations.

This is a market that has gone out of favour as far as speculators are concerned. The relative oversupply of natural gas, due to expansion of output from shale formations, has suppressed basis differentials and volatility. Our prediction is that it is only a matter of time before price volatility returns to this market – and the speculators will come back as well.

- 1 Short-term price elasticity of demand for natural gas is of the order of magnitude of 0.1, and is about 0.5 in the long run. Price elasticity of demand is the percentage change in demand divided by the percentage change in price. See, Patrick L. Anderson, Richard D. McLellan, Joseph P. Overton and Dr. Gary L. Wolfram, 1997, "Price elasticity of demand," Mackinac Center for Public Policy, November 13. An in-depth study of the price elasticity of demand, including a survey of the literature, can be found in M. A Bernstein and J. Griffin, 2006, "Regional differences in the price-elasticity of demand for energy," RAND Corporation Santa Monica, California, Battelle Contract No. DE-AC36-99-G010337.
- 2 James T. Jensen, 2004, "The development of a global LNG market. Is it likely? If so When?" Oxford Institute for Energy Studies. See also excellent presentations by Jensen Associates (<http://www.jai-energy.com/index.php?page=pubs>).
- 3 <http://www.investopedia.com>.
- 4 Elyse Tanouye, 1992, "'Gunning' plays can claim victims in the futures pit," *Wall Street Journal*, June 17.
- 5 The name survives though most market activity is now often squeezed into a period of a few hours in the Nymex contract expiration day.
- 6 "Supplemental Comments of Platts, Transparency Provisions of the Energy Policy Act of 2005" FERC, Docket No. AD06-11-000, February 23, 2007.
- 7 Transactions executed on Fridays are for Saturday, Sunday, Monday delivery.
- 8 This swap mutates periodically into fixed-for-floating swap. Once the monthly index settles, it becomes effectively a fixed price for the current month, with the *Gas Daily* price floating from day to day.
- 9 The forward basis curve reflects market expectations of future basis differentials (ie, the difference between future monthly prices at a specific location and the final settlement price of the expiring Nymex contract).
- 10 This assumes that both sides use the same forward price curves and interest rate assumptions. In less-transparent markets, the price curves used by both counterparties may vary. The lack of market transparency is often abused and may be a source of non-existent profits if the price curves are manipulated.
- 11 Under the conventions used in the financial markets, the payer of the fixed price is called a buyer of the swap.
- 12 We are using an example of positive basis, for simplicity. The basis for the HSC is typically negative.
- 13 We will assume the month has 30 days.
- 14 The discussion below follows Fletcher J. Strum, 1997, *Trading Natural Gas: Cash, Futures, Options and Swaps* (Tulsa, OK: PennWell Books), pp 90–101.
- 15 The differential may be either positive or negative.

- 16 <https://www.theice.com/productguide/ProductDetails.shtml?specId=693>.
- 17 [http://www.bracewellgiuliani.com/index.cfm/fa/news.advisory/item/274abfe0-bde1-43f7-9326-79628c57e959/CFTC\\_Designates\\_Seven\\_ICE\\_Natural\\_Gas\\_Contracts\\_as\\_Significant\\_Price\\_Discovery\\_Contracts\\_Subjecting\\_Them\\_to\\_Regulatory\\_Oversight.cfm](http://www.bracewellgiuliani.com/index.cfm/fa/news.advisory/item/274abfe0-bde1-43f7-9326-79628c57e959/CFTC_Designates_Seven_ICE_Natural_Gas_Contracts_as_Significant_Price_Discovery_Contracts_Subjecting_Them_to_Regulatory_Oversight.cfm).
- 18 Jesus Melendrez from W&T Offshore offered many comments on this section.
- 19 Enron was one of the pioneers of VPP transactions.
- 20 "A percentage of ownership in an oil and gas lease granting its owner the right to explore drill and produce oil and gas from a tract of property. Working interest owners are obligated to pay a corresponding percentage of the cost of leasing, drilling, producing and operating a well or unit. After royalties are paid, the working interest also entitles its owner to share in production revenues with other working interest owners, based on the percentage of working interest owned." (<http://www.glossary.oilfield.slb.com/Display.cfm?Term=working%20interest>).
- 21 A typical VPP contract may contain the following language: "The Overriding Royalty conveyed hereby is a non-operating, non-expense-bearing limited overriding royalty interest free of all cost, risk, and expense of production, operations, and delivery to the Delivery Point. In no event shall Grantee ever be liable or responsible in any way for payment of any costs, expenses, or liabilities attributable to the Subject Interests (or any part thereof) or in connection with the production, saving or delivery of Overriding Royalty Hydrocarbons to the Delivery Point."
- 22 Enron relied on a team of in-house geologists to evaluate properties considered for VPPs.
- 23 Lance W. Behnke and Elizabeth L. McGinley, 2011, "Understanding critical tax aspects of volumetric production payments," *Oil & Gas Financial Journal*, April 1 (<http://www.ogfj.com/articles/print/volume-8/issue-4/features/understanding-critical-tax-aspects-of-volumetric.html>).
- 24 See Moody's Investors Services, "Volumetric production payments – analytical implications and adjustments for E&P companies." Moody's treats VPP's as collateralised borrowings for analytical purposes.
- 25 Paul Riddle, 2006, "Volumetric production payments. An effective monetization strategy," IPAA, Private Capital Conference, January 19.
- 26 This means that VPPs distort some statistics related to company performance. The income statement includes the related cost and cashflows, and this may convey a distorted picture of the exploration and production effort.
- 27 A perspective on a 2012 Chesapeake Energy VPP transaction is provided in a *Forbes* article: Christopher Helman, 2012, "Chesapeake Energy's new plan: Desperate measures for desperate times," *Forbes*, February 13. The article is an indication that in a VPP transaction division of risks between a seller and a buyer may vary depending on negotiated contracts.
- 28 See Michael E. Humphries, 1995, "The competitive environment for oil and gas financing," *Energy Policy*, (23)11, November, pp 991–1,000.
- 29 In practice, there could be more degrees of separation between ABC and an SPE.
- 30 See "Governance, accounting, and auditing reform, Post-Enron," in Leonard J. Brooks, 2007, *Business & Professional Ethics for Directors, Executives & Accountants* (4e) (Mason, OH: Thomson/South-Western); Robert Roach, counsel and chief investigator, 2002, "The accounting treatment of prepay," US Senate's Permanent Subcommittee on Investigations, July 27.
- 31 John Kenneth Galbraith, 1994, *A Short History of Financial Euphoria* (Knoxville, TN: Whittle Direct Books).
- 32 In practice, a special non-profit agency will usually be established as the natural gas buyer. The agency will act as a go-between for the natural gas supplier and the final buyers of natural gas.
- 33 The description of the contracts is based on the information included in the 10-K of Copano Energy, March 2007, p 52 (available at <http://yahoo.brand.edgar-online.com/fetchFilingFrameset.aspx?dcn=0000950129-07-001429&Type=HTML>).

- 34 Index prices were discussed earlier in this chapter.
- 35 See 10-K Form of Markwest Hydrocarbon (available at <http://yahoo.brand.edgar-online.com/fetchFilingFrameset.aspx?dcn=0001104659-07-017189&Type=HTML>).
- 36 See the Credit Agricole Securities USA newsletter of 8 July 2010. “The explosion of drilling activity in NGL-rich plays such as the Granite Wash and Eagle Ford Shale has enabled producers to generate high rates of return even as gas prices hover at sub-economic levels. However, it appears that this success may be catching up on them. The price of ethane, which typically accounts for about half of the NGL stream, now trades at 25% of the price of WTI, matching its lowest level of the last ten years. Furthermore, ethane’s price premium to natural gas on a Btu basis has descended at US\$1.50/MMBtu[...]. This is approaching its lows over the last five years. It is important to distinguish between pure oil producers and liquids producers as NGL prices decline. The pricing of our “blend” of NGLs has fallen from 60% of Nymex–WTI in 1Q10 to just over 50% in 2Q10 and currently sits at 48%. This movement in relative price could be detrimental for any producer who has attempted to hedge their NGL production with crude oil hedges.” See also <http://www.bnenergy.com/Not-Gonna-Lie-NGL-Production-Up> for more recent price developments.
- 37 Form 10-K of Copano Energy, March 2007.
- 38 Proxy hedges, as explained above, are indirect hedges of gas liquids with a basket of crude, natural gas and possibly heating oil, gasoline and residual fuel oil.
- 39 Once the information about imminent execution of a hedging programme in an illiquid market gets out, many traders will take a position in anticipation of a hedging-induced market move.
- 40 “DCP Midstream LLC, with headquarters in Denver, Colo., leads the midstream segment as the second-largest natural gas gatherer and processor, the largest natural gas liquids producer, and one of the largest marketers of natural gas and natural gas by-products in the United States. DCP Midstream is a 50:50 joint venture between Spectra Energy and ConocoPhillips.” (<https://www.dcpmidstream.com/ABOUTUS/Pages/AboutUs.aspx>).
- 41 Based on a RenRe presentation, with the transaction structure modified by the author (changed to an option from a swap).



## *International Natural Gas Markets*

As mentioned in Chapter 8, one can identify several natural gas pricing regimes worldwide, with huge disparities in terms of price levels and price variability. This is likely to change over the next decade, creating many trading opportunities as well as potential risks for those who find themselves on the wrong side of history. The future of this market will be defined by changes in the technology of natural gas production and distribution, the inevitable march towards market-based solutions across the globe and growing concerns about the environmental consequences of climate change. Natural gas markets will become more integrated, with links created through LNG and, possibly, carbon trade. As always, it is easier to see the final outcome of many interacting forces. It is the timing and path of transition that remain unknown. Any trader or analyst who tries to anticipate future moves in such a complex game of multidimensional chess, with many participating players, should think through the potential consequences of the following developments:

- ❑ the impact of the shale revolution on the global natural gas market;
- ❑ the drive towards liberalisation of energy markets and introduction of competition and market-based solutions;
- ❑ the future of carbon trade and the controversy surrounding climate change;
- ❑ economic growth of India and China and other emerging countries;
- ❑ growing concerns regarding energy security and diversification of sources of energy; and
- ❑ progress in the development of renewable energy sources.

This list is more than sufficient to keep one awake at night, and it is only likely to grow with the passage of time.

This chapter will focus on two rapidly evolving markets: the European natural gas market and the world LNG market. Both markets provide a good illustration how geopolitical and economic forces interact with technological progress to change commodity flows, price levels and transaction structures. Technological revolution related to shale gas production has to be analysed in the context of growing concern about the security of energy supplies. Potential flows of natural gas from shale formations are likely to have huge geopolitical impact, reducing the dependence of many countries on imports from a small group of producers. The shale gas revolution is also likely to contribute to the evolution of contractual arrangements for natural gas flowing through pipelines into Europe or transported as LNG. We are likely to witness a shift from long-term contracts using pricing formulas based on indexation to oil, towards spot trading and the possible emergence of an integrated North Atlantic gas market.

### **EUROPEAN NATURAL GAS MARKETS<sup>1</sup>**

The European natural gas industry is a tale of two markets. The UK market is vibrant, open, supported by highly sophisticated market participants, and generally seen as mature. The continental European Union (EU) market, on the other hand, is a system in transition, undergoing transformation under the impact of political and market forces. The salient features of the old continental system include:

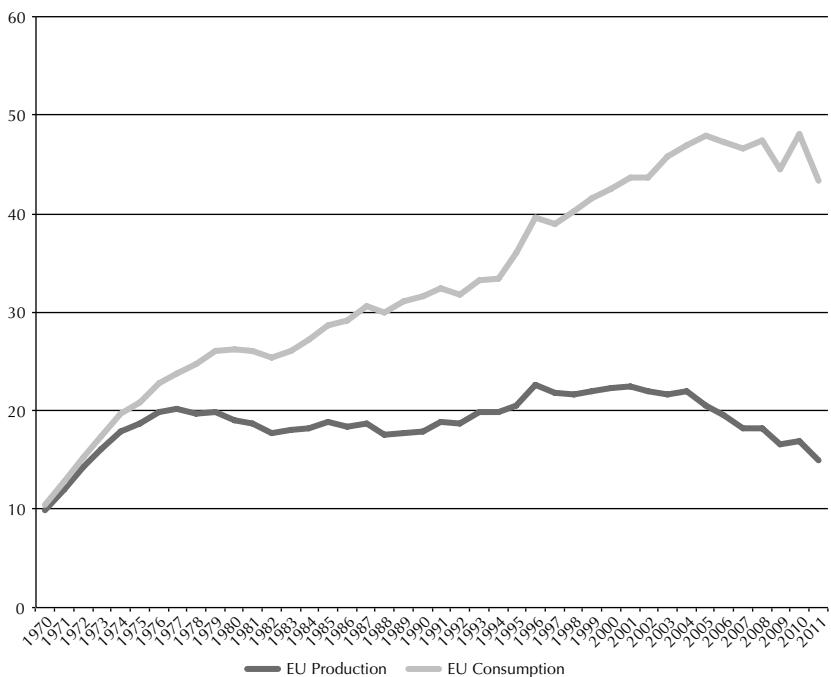
- ❑ reliance on long-term supply contracts with gas prices indexed to oil and refined products prices; such contracts are opaque and discourage efficient use of natural gas – we shall discuss the specific historical circumstances that led to the choice of this market design; and
- ❑ domination of the market by a number of incumbents (large utility companies controlling transportation and distribution infrastructure) and a few foreign suppliers.

The existing market structure is rapidly evolving and we may see a reversal of fortunes in the future. The 2008 recession reduced

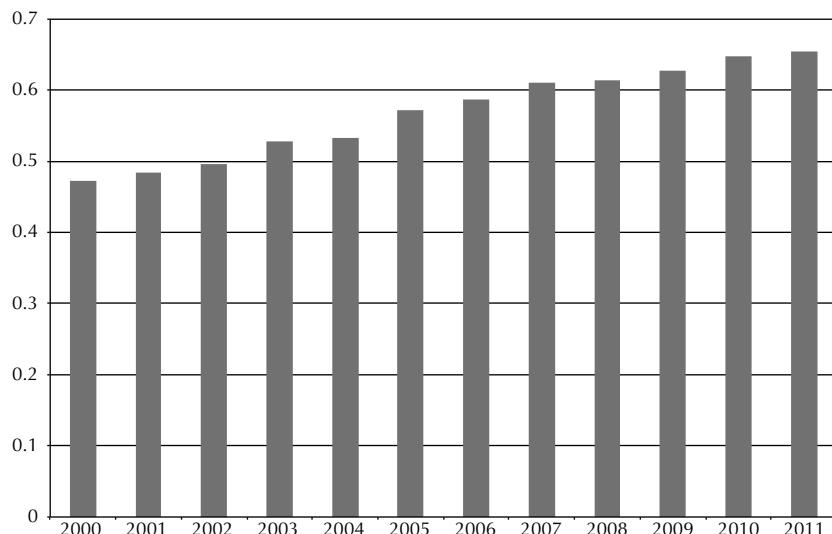
demand for natural gas and foreign suppliers came under pressure to offer discounts to contract prices, which were increasing following the lead from crude prices. The EU is committed to market liberalisation and the process of regulatory change is accelerating. What will eventually emerge is a system sharing many features of the US and UK markets.

All the EU countries (with the exception of the Netherlands) share dependence of foreign supplies of natural gas, imported over pipelines and as LNG. The supplies of natural gas from the North Sea and on-shore fields are dwindling, and this limits the growth of the spot and forward markets supported by domestic production as the volume of potential sales is falling. Figure 12.1 shows the opening wedge between the EU production and growing consumption, while Figure 12.2 illustrates the increasing dependence of the EU on imports of natural gas (100% = 1 in this graph), which is due to

**Figure 12.1** European Union, production and consumption of natural gas (Bcf/day)



Source: BP, "Statistical Review of World Energy" (June 2012)

**Figure 12.2** Imports dependency rate, EU-27, natural gas (2000–11)

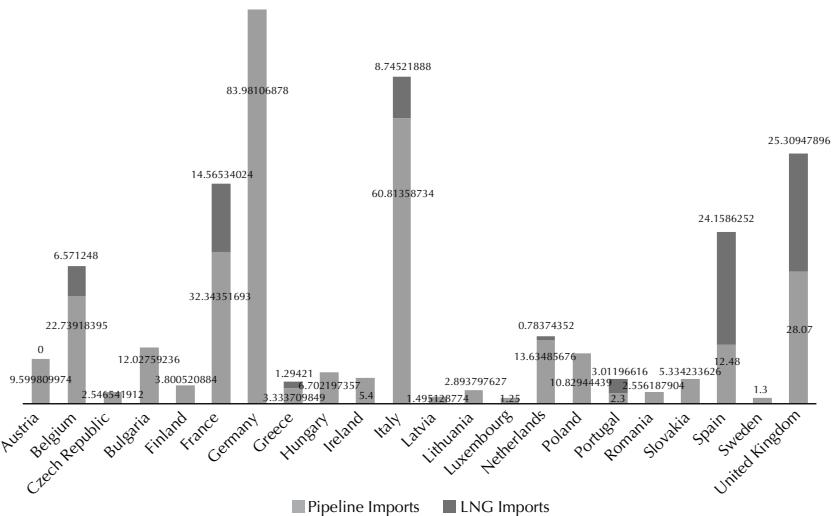
Source: BP, "Statistical Review of World Energy" (June 2012)

increase to over 70% by 2020.<sup>2</sup> Figure 12.3 provides information about imports of natural gas for a number of the biggest EU consuming countries.

In 2009, import dependency was 64.2% for the 27 member countries, varying from a near 100% dependency on imports for certain countries to being a net exporter (for example, the Netherlands). The imports of natural gas are highly concentrated in terms of the countries of origin, creating a serious energy security challenge from supply disruptions due to economic or geopolitical factors. One can easily envisage a scenario of serious economic downturn caused by the curtailment of natural gas supplies to Europe, with some countries being much more vulnerable than others.

### The UK market

The UK market was dominated through the 1990s by British Gas (BG), a state-owned, vertically integrated monopoly which owned some gas-producing fields in the North Sea (many fields were owned by other companies), as well as the UK gas transmission network. Following the Gas Act of 1986, the company was priva-

**Figure 12.3** EU Natural Gas Imports (billion cubic feet), 2011

Source: BP, "Statistical Review of World Energy" (June 2012)

tised, with shares floated on the stock market on December 8, 1986. In 1997, BG was broken up into two companies. Centrica Plc, which inherited the business of gas supply<sup>3</sup> and gas production in the North and South Morecambe field, and BG Plc, which inherited the natural gas transportation business and international exploration and production (E&P) and trading business. In 1999, BG Plc was further restructured by its demerging into two companies: BG Group Plc and Lattice Group Plc. The Lattice Group, which contained Transco, merged later with National Grid Company to form National Grid Transco Plc. Following a few more reorganisations, the latter company became National Grid Gas Plc. To make things more complicated, Centrica has the rights to the British Gas name inside the UK, and BG Group outside the UK.<sup>4</sup>

The privatisation of British Gas was followed by a reorganisation of the market into three segments. The wholesale market is based on a competitive model, with a number of participants including producers (British Gas among them), traders and suppliers. The second segment included big consumers (with consumption exceeding 2,500 therms per year) who negotiated directly with the providers (including British Gas). The third segment included small

customers with annual volumes falling below this threshold, with British Gas still enjoying a monopolistic position in this area. By spring 2012, every end user in the UK had freedom of choice over their natural gas provider.

The grid company operates as a common carrier (public gas transporter in the UK), offering services to the shippers operating under licences from Office of Gas and Electricity Markets (OFGEM). The relationship between the grid company and the shippers is regulated by the Uniform Network Code (UNC; prior to 2005, it was just Network Code). UNC is a set of rules regulating third-party access to the National Transmission System (NTS). The points of entry into the grid are called aggregated supply points, and most of them correspond to beachhead terminals, which connect the UK to natural gas fields in the North Sea. National Grid high-pressure pipeline tariffs are based on a three-component system, which includes:

- entry charges;
- exit charges; and
- commodity charges.

The tariff design varies from the solutions adopted in the US and many other countries, which rely on rates that are distance-specific. UNC uses the concept of entry and exit capacity. The exit charges were based until 2009 on the long-run marginal cost,<sup>5</sup> and targeted a 50/50 split of the costs between entry and exit points. In 2009, the system was modified with two types of exit capacity being defined: flat capacity (for a daily off-take<sup>6</sup> quantity) and flexibility capacity (allowing intraday variations in off-take), allocated through auctions. Entry capacity into the system is allocated through a number of auctions covering quarter, month and day.<sup>7</sup> The shippers who did not succeed at buying capacity at an auction, still have a fallback to acquire it on the secondary market or through participation in the short-term next day auctions for residual capacity (firm or interruptible). Other options the shippers have is to suspend the flow of natural gas from the producing fields or sell the volumes at the beachhead terminals to the shippers who had secured transportation capacity. Injecting gas into the grid without having sufficient transportation capacity results in heavy penalties.

Commodity charges cover the system operator (SO) costs and are distance-insensitive.

Gas injected into the system becomes perfectly fungible and can be traded in the paper market corresponding to the National Balancing Point (NBP), which plays the role of Henry Hub in the UK market. One difference is that the NBP does not correspond to a physical location, but is rather a virtual point. This design is made viable by making the exit charges independent from the specific point at which the molecules are injected into the system. Natural gas becomes perfectly homogeneous once it enters into the grid and a daily system balance can be established. In other words, for every user, the daily volumes injected into the system and taken out can be determined and potential imbalances can be calculated. The same calculation is carried out at the system level.

In addition to moving natural gas between entry and exit points, the users of the system can trade gas. This is done through acquiring and disposing trade nominations submitted to the grid operator. A seller submits a disposing nomination and a buyer an acquiring nomination. The quantity in the acquiring trade nomination is credited to the user, and the quantity in a disposing trade nomination represents a debit to the user.

The way to think about this mechanism is to use the analogy of a bank branch system with an electronic network. Once money is injected into the system at one branch it can be extracted at a different branch, and all currency units are treated as homogeneous. A acquiring nomination is an equivalent of making a deposit, and an disposing nomination is the equivalent of withdrawing funds. A deposit and withdrawal can be made at different branches or at the same branch: the transaction costs will be the same.

Gas day in the UK starts at 6:00 am. The shippers have the obligation to communicate the volumes they wish to withdraw from the National Grid transportation system after 13:00 pm on the previous day, and the volumes they plan to inject after 16:00 pm, including the entry points. The shippers in this market do not have to be volumetrically balanced, but the rules are designed to create economic incentives to submit balanced schedules. The incentives are created for the system of prices at which the imbalances are settled (see below), and which reward those shippers who are long and punish the shorts. Daily imbalance for a given shipper can be calculated as:

### Deliveries into the system – Sales + Purchases – Off-take

This formula calculates the aggregate amount of gas (in kWh) a shipper puts into the pipeline system (deliveries) less the amount taken out (off-take). Sales are equivalent to taking gas out, purchases equivalent to injecting gas. In the case of an imbalance, the total is different from zero, and the shipper has to settle by effectively buying natural gas from, or selling it to, the system operator. Again, a banking analogy may be useful. A bank customer starting with a zero balance, and making deposits and withdrawals during the day will end up either with a zero or positive balance, or an overdraft. A positive balance will earn interest; an overdraft will require the imposition of charges a bank will collect from the customer.

The imbalances are managed through the automatic cash-outs taking place at the system buy or sell price for the day. Such prices are referred to as on-the-day commodity market (OCM) prices. Trading at NBP is managed by APX-ENDEX.<sup>8</sup> The minimum transaction volume in the OCM market is 4,000 therms. APX supports screen-based trading (both for day-ahead and intraday transactions), which is anonymous and fully cleared. National Grid may participate in this market buying natural gas from, and selling to, the system users. The UNC defines three prices (as explained on the Transco website):

- SMP is the price in pence per kWh calculated as the sum of all Market Transaction charges divided by the sum of the Trade Nomination Quantities for all transactions effected in respect of that day, subsequently adjusted to account of any bids which are to be excluded in association with resolving constraints.
- SMP buy: the higher of the SAP (adjusted to exclude bids associated with a constraint) plus 0.0287 or the highest price (excluding bids associated with constraints) of any trade to which Transco as system balancer is party, irrespective of whether the trade was for a buy or a sell.
- SMP sell: the lower of the SAP (adjusted to exclude bids associated with a constraint) minus 0.0324 or the lowest price (excluding bids associated with constraints) of any trade to which Transco as system balancer is party, irrespective of whether the trade was for a buy or a sell.

**Table 12.1** ICE Futures UK natural gas futures contract

DESCRIPTION	
Contracts are for physical delivery through the transfer of rights in respect of natural gas at the national balancing point. Delivery is made equally each day throughout the delivery period.	
CONTRACT SPECIFICATIONS	
<b>Trading period/strip</b>	78–83 consecutive month contracts 11–12 consecutive quarters.  Quarters are strips of three individual and consecutive contract months. Quarters always comprise a strip of Jan–Mar, Apr–Jun, Jul–Sep or Oct–Dec.  13 consecutive seasons.  Seasons are strips of six individual and consecutive contract months. Seasons always comprise a strip of Apr–Sep or Oct–Mar. Months, quarters and seasons are listed in parallel.
<b>Expiration date</b>	Trading shall cease at the close of business two business days prior to the first calendar day of the delivery month, quarter or season.
<b>Contract security</b>	ICEU acts as central counterparty to all trades thereby guaranteeing the financial performance of ICE Futures Europe contracts registered in the name of its Members up to and including delivery, exercise and/or settlement.
<b>Trading hours</b>	07:00–17:00, local London time (LLT), Monday–Friday.
<b>Contract size</b>	Multiples of five lots of 1,000 therms per lot of natural gas per day.
<b>Units of trading</b>	One lot equals 1,000 therms of natural gas per day (1 therm = 29.3071 kilowatt hours)
<b>Quotation</b>	The contract price is in sterling and in pence per therm
<b>Minimum price flux</b>	Futures/blocks – 0.01 pence/therm EFPs/EFSs – 0.005 pence/therm
<b>Maximum price flux</b>	There are no limits
<b>Settlement price</b>	The weighted average price of trades during a 15-minute settlement period from 16:00:00, London time.

<b>Daily margin</b>	All open contracts are marked-to-market daily.
<b>position limits</b>	There are no position limits.
<b>Trading methods</b>	Electronic futures, exchange for physical (EFP), exchange for swap (EFS) and block trades are available for this contract.
<b>Delivery/settlement basis</b>	Matching acquiring and disposing trade nominations (buyer from ICEU, seller to ICEU) are input by buyer and seller to National Grid Transco via Gemini before 18:30 on the business day prior to delivery. Delivery takes place in kilowatt-hours (29.3071 kilowatt hours per therm).

---

Positive daily imbalances are settled at the SMBP, negative imbalances at SMSP. To be precise, the shippers who are short will pay SMBP, the shippers who are long will receive SMSP. Trading at NBP is carried out under a standardised contract known as the “Short Term Flat NBP Trading Terms & Conditions” (NBP97). The contract is based on a number of principles, including:

- equality of traded volumes and delivered volumes, meaning that volumes delivered are guaranteed to equal volumes traded;
- constant delivery rate through the gas day; and
- limited force majeure (only the events outside the control of the affected party).

The ICE Futures UK natural gas futures contract is defined with respect to NBP (see Table 12.1). The contract specification is available from the ICE website.<sup>10</sup>

ICE also publishes the Month Ahead Index, which is used as a benchmark for OTC swaps and options.

The NBP price is important not only from the point of view of the spot market and exchange-based instruments used for hedging. Most long-term contracts, which still represent about 50% of natural gas UK supplies, reference the NBP price. This is the main difference with respect to the continent, where long-term contracts are not only longer (20–30 years, compared to 8–12 years in the UK), but are also indexed to oil.

Understanding the UK natural gas market is important, in spite of

its relatively small size, as it provides a template that may be replicated by other countries. For example, the concept of virtual balancing points has already been copied in Italy, Spain and France.

### **Continental Europe**

The discussion of the continental gas markets is equivalent, for all practical purposes, to a discussion of the EU. The EU relies on natural gas for about 25% of its primary energy consumption (24.47% in 2009),<sup>11</sup> and over 60% of the total supply is imported,<sup>12</sup> primarily from Russia, Norway and Algeria. The dependence on foreign supplies of natural gas (given the importance of natural gas to winter heating and power generation in some countries)<sup>13</sup> is a critical issue for EU energy security, although the expansion of shale gas production in Europe and growing imports of LNG may help to alleviate this problem.

The continental markets are characterised by a high concentration of supply and distribution, and low levels of transparency. In spite of repeated initiatives undertaken by the European Commission, recurring assessments produce roughly the same disappointing description of the level of market development:<sup>14</sup>

- ❑ “Unbundling and the dismantlement of vertically integrated large incumbents has not fundamentally occurred, there is still a high level of market concentration;
- ❑ illiquid markets and a lack of infrastructure limit the access of new entrants to the market for gas in the EU;
- ❑ cross-border sales do not generate significant competitive pressure;
- ❑ there is no truly reliable information on gas markets in the EU; and
- ❑ higher transparency is needed.”

The markets develop at the national frontiers and the points of interface between the systems of the biggest distributors, and it is not an accident that the biggest physical market hubs evolved at the locations where different pipelines meet at the borders.<sup>15</sup>

The continental market for natural gas is also characterised by an oligopolistic structure with a limited number of suppliers and a few large scale distributors. At the beginning of the 21st century, each national market (with the exception of Germany) was dominated by

a single company. A liberalisation programme, launched by the EU through a number of directives,<sup>16</sup> led to the paradoxical outcome of increasing market concentration – as the European utilities concluded that increased size (in terms of assets and overall customer base) is the best strategy to meet the challenges of a liberalised gas market, and went through a number of mergers and acquisitions, evolving into multinational companies, active in many different energy commodity markets.

One of the most important features of the natural gas market in continental Europe is the reliance on long-term supply contracts, with maturities extending up to 20 and more years, and the use of pricing formula linking natural gas to prices of oil, lagged by a few months.<sup>17</sup> These contracts typically have a “take-or-pay” component, which means that there are certain minimum volume requirements and some volumetric flexibility, with the option of rolling some volumes into the future (although the payment has to happen in a current year). The use of this approach is a legacy of the solutions implemented in the early days of the development of natural gas industry in Europe, which can be explained by some unique historical circumstances. In the US and Canada, the industry grew gradually through the discovery of many natural gas fields dispersed over many different regions. The development of the long-haul pipeline system, which began during World War II, was motivated by national security objectives. Once the backbone of the pipeline system was created, commercial interests could take over. In Europe, however, growth occurred in discrete steps, with milestones marked by the discovery or connection of major fields.<sup>18</sup> The producers required long-term contracts to justify development costs and the construction of infrastructure required to take natural gas to market. The sheer scale of the enterprise required major changes in energy consumption patterns and the displacement of other fuels. One cannot ignore also the rent-seeking behaviour of the producers: newly discovered fields were characterised by very low production costs and the “cost plus” pricing formula was not very attractive.<sup>19</sup> The utilities, which took the other side of the long-term supply contracts, had the market power to force the price terms on end users, as they were monopolistic suppliers. There were many other reasons to relate the prices of natural gas to the prices of competing fuels. It seemed reasonable to link these two markets through a

common pricing regime. The oil market was more mature at the time, and the use of oil as a benchmark was an easy solution to what would be a rather insurmountable problem in contract negotiations.

However, the rationale for this pricing policy has now disappeared. The period of high oil prices following the first two oil shocks in the 1970s resulted in the decommissioning of power plants relying on oil – and the rationale for using oil as an anchor for pricing natural gas, a competing fuel, was lost, at least from the point of view of the end users of electricity. The producers see it, understandably, in a different light. High prices of oil translate into high prices of natural gas, resulting in significant economic rents. On the other hand, there are growing pressures from end users to de-link these markets. Financial crisis and ensuing recession reduced demand for natural gas and resulted in an oversupplied market. The contrast between depressed spot prices and high long-term contract prices amplified pressure to revise the system for pricing natural gas. The producers agreed to renegotiate some supply contracts, and to the introduction of an element of spot pricing. In 2010, Gazprom agreed to link up to 15% of gas volumes to spot prices, and push into the future some contracted volumes, while insisting that the overall volumetric commitment by the buyers be maintained. Statoil has been reported to have made similar price concessions.

### **Long-term supply contracts**

The first big gas field to be discovered was Groningen in the Netherlands (1959), with exclusive production rights extended to Nederlandse Aardolie Maatschappij (NAM, a 50/50 joint venture between Shell and Exxon). The principles of natural gas policy were outlined in a famous memo from the Dutch Minister of Economic Affairs, Jan Willem de Pous (*Nota de Pous*). Under the principles promulgated in the memo, natural gas would be priced in terms of the alternative fuels that would be displaced, such as heating oil or residual fuel oil.<sup>20</sup>

Subsequent negotiations with the European buyers of Groningen gas led to the development of a pricing system based on netback from the domestic prices in different countries back to the Dutch borders. Given that different countries had different levels of internal natural gas prices, and that transportation costs differed, this solution resulted in multiple prices of molecules leaving the Dutch

territory. To avoid the potential for re-sale of gas between low and high price countries, destination clauses were inserted into the contracts to prevent re-marketing of natural gas. A unique feature of the early Dutch contracts was the capacity charge, a component of the sale price, added to compensate the Dutch sellers for the management of volumetric risk – ie, providing a flexible response to changes in the levels of natural gas required by the buyers. Volumes would fluctuate due to changes in weather conditions and the levels of economic activity. The Dutch sellers could provide this service due to proximity of producing fields to their customers, and were willing to accept volume flexibility in their contracts. Another principle of the Note was the “harmonisation” of natural gas production and sales to end users (ie, avoiding excess supply levels that would undermine the adopted link to oil prices).

Subsequent contracts signed by European buyers with other producers (Norway, Algeria, the Soviet Union) followed a similar pattern, but with some differences. Longer transportation distances made it difficult for the sellers to offer the same volumetric flexibility as the Dutch producers. The contracts were structured as take-or-pay arrangements, with volumetric risks shifted to the buyers, who committed to paying for a certain volume, irrespective of circumstances. The minimum take was defined as a percentage of the maximum annual quantity (take-or-pay contracts are a very important transaction structure used in many markets, and will be covered in more detail later). Most long-term contracts were indexed to oil or some refined products (such as gasoil), with some other variables, ranging from prices of other energies to inflation indexes.

The information about the level of natural gas prices under long-term contracts with oil indexation is not easily available. The best indicator in public domain is the German Border Price, published every month by Bundesamt für Wirtschaft und Ausfuhrkontrolle (BAFA). This is an implicit price that can be calculated by dividing the value of natural gas imports at the German border by the corresponding imports volume. This price is still dominated by long-term contracts, with the order of magnitude of the spot transactions being about 10%.

The long-term pricing formulas reflect the objectives of both sides, and are a result of long negotiations designed to reconcile the conflicting interests of the counterparties.<sup>21</sup> A buyer of natural gas

wants to be assured that the natural gas purchased under the contract remains competitive with respect to other fuels (if used in power generation or for space heating), and can compete against natural gas from other sources. The sellers want to be assured of a level of prices sufficient to cover the cost of developing new fields, transportation, domestic taxes, etc. The price formulas used on the continent for long-term contracts are based on an additive design, such as, for example.<sup>22</sup>

$$P = P_0 + 0.4 \times F_1 \times (GO - GO_0) + 0.6 \times F_2 \times (LSFO - LSFO_0)$$

where  $GO$  ( $GO_0$ ) stands for the price of gasoil (the base price of gasoil),  $LSFO$  ( $LSFO_0$ ) for low sulphur fuel oil (the base price of low sulphur fuel oil),  $F_1$  ( $F_2$ ) are pass-through factors. These factors determine the extent to which increases (decreases) in absolute prices included in the formula are translated into the current prices of natural gas. The UK price formulas for long-term contracts are based mostly on multiplicative design.

$$P = P_0 \times \left( 0.2 \times \frac{GO}{GO_0} + 0.3 \times \frac{FO}{FO_0} + 0.5 \times \frac{PPI}{PPI_0} \right)$$

In this example, the contract price  $P$  is escalated based on the prices of gasoil, fuel oil and the producer price index. The rates of change for these variables are weighted by the coefficients negotiated by the buyer and seller and written into the contract. The contracts determine the choice of the base periods for the variables used in the formula and the frequency of re-pricing. Given that the updated levels of the prices used in the formula should be available when re-pricing occurs (on a monthly, quarterly or annual basis), the formula-based prices will adjust to the new market conditions with a lag. The variables included in the price formulas include potential prices of gasoil and fuel oil at ARA (Antwerp–Rotterdam–Amsterdam), prices of electricity, coal, spot prices of natural gas, consumer or producer price index. The design of the pricing formula, including the choice of variables and of the base period, is very important and requires a great amount of experience, insight and analytical work.

We have seen in our career a number of long-term contracts with a catastrophically poor design of pricing formulas, which resulted in serious losses to one side and great riches to the other side. The

biggest risk is that the dynamics of market prices of a given commodity (natural gas in this case) decouples from the dynamics of the variables used in the pricing formula. The famous Enron J-block contract is the best example. Enron acquired a significant volume of natural gas from a field in the North Sea (called J-block), priced under a long-term contract using a complex formula like the one above. As soon as the ink on the contract was dry, the prices of natural gas in the UK dropped, following deregulation and a significant increase in production. Unfortunately for Enron, the purchase formula-based price continued to climb. A protracted dispute followed, ending in a significant loss to Enron.<sup>23</sup> Another long-term supply contract Enron negotiated relied on a formula producing reasonable prices only for a specific narrow range of prices included in the pricing formula. Relatively minor departures from the price band for the formula inputs resulted in a significant drop in the prices received by Enron. A simple three-dimensional graph demonstrated this to anybody who was not blind, but the contract had been already signed. Another risk in long-term contracts is that competitors will negotiate at a later date contracts giving them a considerable price advantage, should the market prices change to a significant degree.

The risks of long-term contracts are managed through carefully crafted contracts that contain provisions for contract renegotiation or for contract termination. Renegotiations are usually envisaged at fixed time intervals (every three to five years) or in the case of a significant change in market conditions. Some contracts have the so-called “joker” provision, to allow for an extraordinary renegotiation of terms at the request of one counterparty (typically, once during the term of the contract). The structure of contracts for natural gas in the period preceding the 2007–08 financial crisis is best summarised as follows:<sup>24</sup>

“The final report of the 2007 sector inquiry by the European Commission’s DG Competition shows a very similar pattern of average indexation for exports from the Netherlands, Norway and Russia to EU 25 countries with indexation to gas-oil between 52 and 55 percent and indexation to heavy fuel oil between 35 and 39 percent, the total pegging to fuel oil products being between 87 and 92 percent, with the rest more individually linked to inflation, coal, crude oil or fixed.”

According to the same source, Algerian gas was pegged more heavily to crude (70%), followed by gasoil (9%) and heavy fuel oil (6%).

As signalled earlier, there are many reasons to believe that this pricing mechanism will soon come to an end. Given the risks of long-term contracts with formula-based prices, their popularity hinged upon the ability to design tariffs for end users with the terms representing a mirror image of the structure of long-term contracts. The right to choose suppliers promoted by the EU undermines this approach to management of the risk of long-term contracts. At this point (the spring of 2012), it seems likely that the factors mentioned above will force a shift to a different pricing regime over time. One likely development is a growing reliance on the spot prices at the major European trading hubs.

### **Regulatory developments**

EU countries recognise that the continental market for natural gas faces a number of challenges, similar to many other maturing energy markets. Some of the critical problems include:

- ❑ limited access by new market participants to the physical infrastructure (transportation, storage) combined with privileged access granted to market incumbents;
- ❑ different operational procedures for pipeline nominations, balancing, border connection practices;
- ❑ lack of transparency and of uniform price formation practices at emerging market hubs;
- ❑ limitations in the interconnection infrastructure, hampering the unrestricted flow of natural gas across the continent; and
- ❑ a high degree of concentration both in production/imports of natural gas and retail distribution of natural gas (Table 12.2 contains information about the market shares of the three largest companies in production/imports for EU countries, and if the data is available, the corresponding HHI indexes; it also contains data about the market shares of the three largest companies in retail gas markets; it is worth remembering that the FERC in the US uses the 1,800 level of HHI index as the cut-off level for indication of high market concentration).

**Table 12.2** Share of three largest companies (production and imports) and retail, the HHI for production and imports concentration (2009)

Country	Production/imports (%)	HHI	Retail (%)
Austria	88	3926	80
Belgium	89.6	5115	89.6
Bulgaria	89.5	0	50.42
Estonia	99	0	99
Finland	100	0	NA
France	88	4399	97
Germany	68.11	1886	30.1
Great Britain	36	688	NA
Greece	100	10000	100
Hungary	82.3	2875	68.2
Ireland	77.3	2861	96
Italy	80.14	3313	57.5
Latvia	100	10000	100
Lithuania	100	5800	100
Luxembourg	100	9189	74.7
Poland	100	9608	100
Romania	89.1	0	54.78
Slovak Republic	99.6	9010	99.6
Slovenia	100	6791	82
Spain	69	2396	62
The Netherlands	91	5814	79.4

Source: "2009–2010 report on progress in creating the internal gas and electricity market. Technical annex," European Commission, commission staff working paper, Brussels, June 9, 2011

The European Commission is firmly committed to a vision of integrated, liberalised natural gas markets in Europe, with their reform programme being embodied in increasingly aggressive energy directives. The EU has issued a number of directives aimed at liberalisation and opening up the continental natural gas market. The first directive (98/30/EC), dated June 22, 1998, promulgated rules for transportation, distribution supply and storage of natural gas. The guiding principles of the directive were non-discrimination, establishing a competitive market, interconnection and interoperability (common standards and rules for access to the system). One of the main issues addressed was access to transmission, storage and LNG infrastructure. The member countries were given the option of basing third-party access (TPA) on regulatory approach (through relevant tariffs) or under a negotiated regime. The first directive was followed by two more that sought to promote similar principles,

with limited success so far. The Third Energy Package<sup>25</sup> adopted in 2009 included a number of far-reaching provisions, including:<sup>26</sup>

- ❑ regulations defining conditions for access to natural gas transmission grids (gas regulation);<sup>27</sup>
- ❑ the establishment of the Agency for the Cooperation of Energy Regulators (ACER);
- ❑ directives concerning common rules for the internal market in natural gas;
- ❑ separation between transmission/transportation activities and production/supply activities (unbundling), following one of three models:
  - ownership unbundling;
  - Independent System Operator (ISO); or
  - Independent Transmission System Operator (ITSO).
- ❑ rules to operate markets based on common procedures and principles; and
- ❑ the choice of suppliers for retail and commercial customers.

Unbundling is the key to the further liberalisation of the markets. There can be no competition among the suppliers without non-discriminatory access to a transportation network. A critical provision of the Gas Regulation can be found in Section 19:

“To enhance competition through liquid wholesale markets for gas, it is vital that gas can be traded independently of its location in the system. The only way to do this is to give network users the freedom to book entry and exit capacity independently, thereby creating gas transport through zones instead of along contractual paths. [...] Tariffs should not be dependent on the transport route. The tariff set for one or more entry points should therefore not be related to the tariff set for one or more exit points, and vice versa.”

The Third Energy Package envisaged adoption of European Network Codes, which would define the detailed rules for natural gas trading and transportation arrangements, including network tariff design, arrangements for allocating network capacity, congestion management procedures, balancing and settlements. The Council of European Energy Regulators (CEER) played a significant role in developing consensus for the final design of the European natural gas market, with a working name of the Target Gas Model. The model proposed by the CEER's was endorsed on March 23, 2012,

---

by the Madrid Forum,<sup>28</sup> a semi-official body representing stakeholders in the emerging integrated European gas market, including regulators, TSOs, suppliers, consumers, traders, trading exchanges, member state governments and the Commission itself. The path taken with respect to market design is a good illustration of the realities of the EU and the need for very long and laborious consultations between the interested parties to build consensus. The EU operates under the principle of unanimous agreement of all concerned parties.

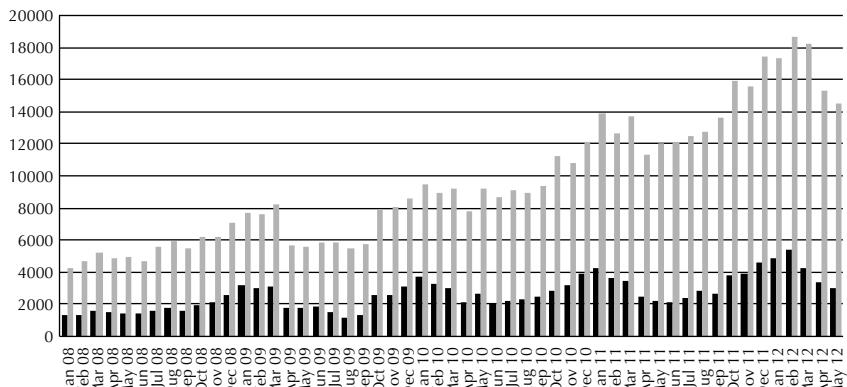
The critical component of the proposed market structure is the concept of an entry–exit zone, with tariffs applied only at the entry and exit points, and independent of the distance travelled, an arrangement in the spirit of the current UK model. Trading inside the zone would happen at a virtual hub. Some of such hubs are already in existence on the continent (see the next section). It remains to be seen if the concept of a virtual hub will triumph over the US model of a real hub, a point through which significant volumes of natural gas flow and are transferred between pipelines and storage facilities. Most likely, some physical hubs will continue to exist, especially at the perimeter of the EU. More information on the European natural gas market design and the entry / exit zones can be found in the paper by Miguel Vazquez, Michelle Hallack, Jean-Michel Glachant, "Designing the European Gas Market: More Liquid & Less Natural," *Economics of Energy & Environmental Policy*, 2012, Volume 1, Issue 3, pp. 35–38.

### **European natural gas hubs**

At this point, there are several natural gas trading hubs in Europe and every energy trader should follow closely the developments related to this rapidly evolving business. At this point, there is no indication that one of these hubs will evolve into a European Henry Hub, but it does create the potential for a more complex market with many lucrative arbitrage opportunities. Statistical evidence is available that the prices of natural gas (at least in northwest Europe) are converging;<sup>29</sup> a clear indication that a budding integrated European gas market is becoming a reality.

The most important natural gas hubs in Europe include the following.

**Figure 12.4** Monthly volumes TTF (January 2008–February 2012), million cubic metre (traded volumes and net volumes (in black))



Source: <http://www.gastransportservices.nl/en/transportinformation/ttf-volume-development>

The Title Transfer Facility (TTF) is a virtual trading hub established in 2003 by Gasunie. The TTF is operated by an independent subsidiary of Gasunie, Gas Transport Services (GTS). The hub supports transactions for natural gas already present in the Gasunie system. Transfer of title to natural gas happens through “nominations,” electronic messages containing the key details of a transaction. Physical short-term gas and gas futures contracts are traded and handled by APX-ENDEX.

Figure 12.4 illustrates the growth of monthly volumes on TTF. As explained on the TTF website, “for a specific period, the traded volume is the sum of the nominated volumes on TTF made by shippers and confirmed by GTS. The net volume for the specified period is the sum of the net hourly positions per shipper. It exclusively concerns the values observed by GTS, based on nominations.”

Zeebrugge. The terminal at Zeebrugge has developed around a number of critical infrastructure assets, with overall throughput capacity of 48m<sup>3</sup> billion per year (about 10% of the EU border capacity). Assets include:

- ❑ Interconnector Zeebrugge Terminal, an interface point between a bi-directional pipeline connecting the Bacton Terminal in the

UK and the Fluxys transmission network; the pipeline is 235 kilometres (146 miles) long and can transport 25.5m<sup>3</sup> billion of natural gas per year in the direction of the UK, and 20m<sup>3</sup> billion from the UK to Europe;

- Zeepipe Terminal, connecting to an underwater pipeline from the Norwegian offshore Troll and Sleipner fields; and
- Zeebrugge LNG Terminal, a facility both for unloading and loading LNG tankers.

The Zeebrugge hub is operated by a company called Huberator, created in 1998. Huberator currently has 78 members, including both physical and financial players. The company offers a number of services, including:

- title tracking;
- back-up services to ensure the firmness of bilateral contracts; and
- supporting spot and financial trading of natural gas; the transactions can be executed directly between members, through brokers or on the AXP Gas Zee platform, which offers two contracts: within-day and day-ahead.

The day-ahead market offers contracts for individual days, weekend strips, balance-of-the-week, working days for the next week. The within-day market offers trading in 24 continuous hourly blocks on a 24/7 basis.

The Zeebrugge hub has grown very quickly since the start of the 2000s, both in terms of physical throughput and financial trading. The main challenge for the future growth of the hub is related to different specifications for UK and European gas. Fluxys has previously addressed this problem through gas swaps. The growth in LNG imports and imports from Russia (due to the construction of the Nordstream pipeline under the Baltic Sea) will reduce this flexibility.

*Central European Gas Hub (CEGH)* at Baumgarten, Austria. CEGH is owned by OMV Gas International and is managed in cooperation with Gazprom. It was established in 2002 as a physical hub for trading, but its pipeline connections (Slovakia, Hungary, Italy, Slovenia) make it suitable for other functions, such as balancing. One

limitation of CEGH is its reliance primarily on one supplier (Gazprom). CEGH offers the following services:

- title transfer service;
- wheeling service;
- no-notice storage nomination service;
- gas auctions; and
- nomination service.

*Punto di Scambio Virtuale (PSV).* The PSV is a virtual hub, operated by Snam Rete Gas, the Italian TSO. The primary purpose of the short-term bilateral physical transactions, supported by a bulletin board, is to facilitate volume balancing on a daily basis. Since the inception of the pool, longer-term physical transactions have also been executed.

*Centro de Gravedad (CDG).* The salient features of the Spanish natural gas market include its physical isolation from the rest of Europe and a heavy dependence on imports of LNG. Given the dependence of the electricity sector on hydropower, an occasional drought increases demand for LNG, underlining the importance of this energy source to the country's economy. The wholesale market for natural gas is organised around a single virtual balancing point (CDG). Trading is carried out on a platform designed by ENAGAS<sup>30</sup> called MS-ATR. Transactions on this platform include intraday, day-ahead and month-ahead contracts.

*NetConnect Germany (NCG)/GASPOOL.* GASPOOL is a result of merger of five major German pipeline operators who merged their market areas and set up a joint company: Gasunie Deutschland, ONTRAS–VNG Gastransport, Wingas Transport, StatoilHydro Deutschland and DONG Energy Pipelines. The company is based in Berlin and its market area covers most of north Germany. GASPOOL is sub-divided into two market areas for transporting and trading H-Gas (high calorific) and L-Gas (low calorific).<sup>31</sup> EEX supports GASPOOL trading with a number of spot and futures contracts. Spot contracts include intraday, day and weekend contracts. Futures are available for balance-of-month (BOM), month, quarter and year maturities. NetConnect Germany (NCG) in its current form was

---

formed on October 1, 2011, and coordinates natural market operations of the following companies: Bayernets GmbH, Fluxys TENP GmbH, GRTgaz Deutschland GmbH, Terranets bw GmbH, Open Grid Europe GmbH and Thyssengas GmbH. Geographically it translates into the south and west of the country and its control area connects to the Netherlands, Belgium, France, the Czech Republic, Austria and Switzerland. The company operations include balancing services, operation of a virtual trading point, and the provision of energy and billing data. Most transactions take place in the spot and OTC markets.

NCG and GASPOOL-related trades executed on the EEX platform are used to calculate the price index called the European Gas Index (EGIX), available since January 17, 2011 from the EEX website. The index is calculated separately for NCG and GASPOOL and aggregated into the value for Germany.<sup>32</sup> The EGIX index is calculated as the arithmetic average of the daily values (calculated separately for each hub and for Germany):

$$\text{Daily Value} = \frac{\sum_{i=1}^N (P_i V_i)}{\sum_{i=1}^N V_i}$$
$$\text{Daily value (Germany)} = \frac{(\text{Daily value}_{\text{NCG}} \times V_{\text{NCG}} + \text{Daily Value}_{\text{GASPOOL}} \times V_{\text{GASPOOL}})}{V_{\text{NCG}} + V_{\text{GASPOOL}}}$$

where  $N$  stands for the number of transactions in the front months relevant for the index on the trading day (the value of  $N$  will be usually different for the two hubs)<sup>33</sup>

$P$  is the price of a transaction in front months relevant for the index on the trading day;

$V$  is the volume of a transaction in front months relevant for the index on the trading day.

EGIX is calculated as (only the formula for the case of the German index is shown):

$$\text{EGIX(Germany)} = \frac{\sum_{j=1}^n \text{Daily Value}_j^{\text{Germany}}}{n}$$

where  $n$  stands for number of trading days which have already been closed with an identical front month relevant for the index.

The EGIX index is an important step in development of the

European gas market based on gas-on-gas competition, with prices reflecting the realities of supply and demand.

*Points d'Echange Gaz (PEG).* The French territory is divided into three balancing (entry-exit) zones, with two notional exchange points (PEG) in the principal Northern and Southern Zones.<sup>34</sup> In the Northern Zone, exchange is limited to the same type of natural gas (H for H or B for B).<sup>35</sup> The suppliers rely on long-term contracts executed in the OTC markets (*il marché de gré à gré*), with the usual indexation to oil. In addition to long-term contracts, Powernext, an exchange based in Paris, offers spot and futures contracts.

The maturity and efficiency of a trading hub can be measured in a number of ways. A frequently used index is known as a churn ratio. The churn ratio is defined as the overall volume of trades divided by physical gas deliveries at the hub, measured over a period of time. In other words, it is a measure of how many times a molecule of gas changes hands. On the financial exchanges, where only a small percentage of the contracts traded result in physical delivery (most positions are closed before expiration by entering into offsetting transactions), the churn ratio may be very high.<sup>36</sup> However, this should be distinguished from the churn rate, “the ratio between the total volume of trades and the physical volume of gas consumed in the area.”<sup>37</sup> ICIS developed a “tradability index” based on the width of the bid-offer spread observed in a given market during a quarter.<sup>38</sup> The following contracts are included in the calculations of the index: within-day, day-ahead, balance-of-month, month-ahead, next quarter, next season, two-seasons ahead, one-year ahead, two-years ahead, three-years ahead. Given that there are 10 contracts included in the index, and a maximum score per contract is 2, the maximum composite index for any hub is 20. The features of the merging European natural gas hubs are summarised in Table 12.3.

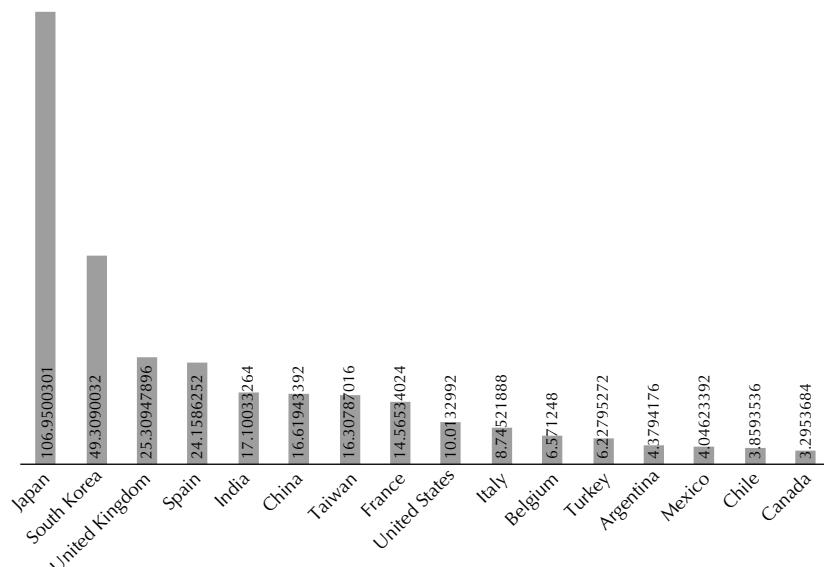
## LNG MARKETS AND TRANSACTIONS

The overall size of the LNG market is estimated to have reached the level of 224 million tonnes per annum (MMtpa) in 2010 (an increase of 41 MMtpa, or 22%, over 2009).<sup>39</sup> The details of the LNG markets are illustrated in Figures 12.5 and 12.6, which show the biggest exporters and importers of LNG in 2009. As one can see, the market is dominated by a relatively small number of big buyers and sellers.

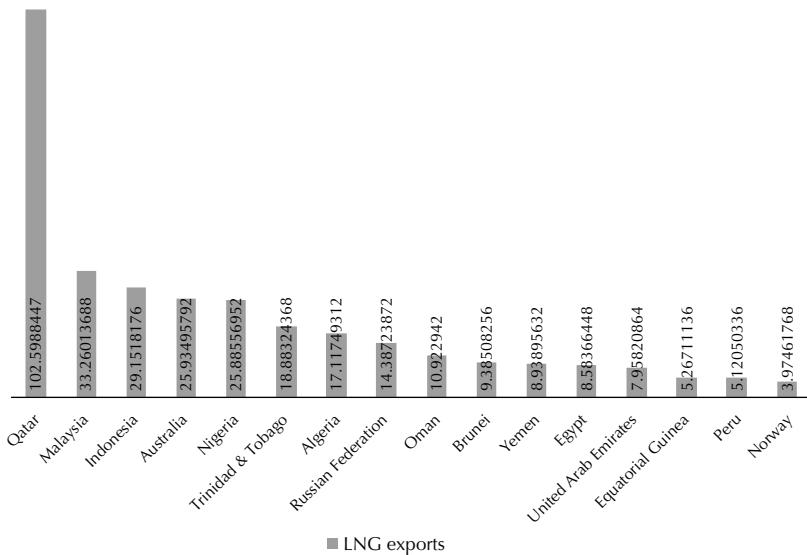
**Table 12.3** European natural gas hubs

	NBB	TTF	NCG	ZEE	GPL (BEB prior to October 2009)
Market Location and Date					
Started Trading	Britain 1996	Holland 2003	Germany 2009	Belgium 2000	Germany 2009
Market Type	Traded	Traded	Transition to Traded	Physical Transit	Physical Transit
Volume Increase S09--> 01	+20%	+100%	+100%	Flat (+/-10%)	S09-W09: +300% W09-W10: +35%
Average Daily Volume: Q4-10 (GWh)	41,500	3,550	2,550	2,250	2,150
Measure of Liquidity and Activity in the:	Spot Prompt Curve	Very Good Very Good Very Good	Very Good Very Good Good	Very Good Very Good Poor	Good Good Poor
Heren Tradeability Index (x/20)		19	17	12	9
Price Correlation to NBP (%)		100	93	89	94
Estimated Churn Ratio	19	4.5	2	4	<1

Source: Patrick Heather, 2011, "The recent evolution of the European gas market. Towards oil-gas decoupling," Florence School of Regulation, Florence, March 21

**Figure 12.5** LNG imports by country (2011, billion m<sup>3</sup>)

Source: BP, "Statistical Review of World Energy June 2012"

**Figure 12.6** LNG exports by country (2011, billion m<sup>3</sup>)

Source: BP, "Statistical Review of World Energy June 2012"

Historically, the solution to complicated logistics of LNG production and transportation, high capital costs, long investment gestation periods and long time periods required for the amortisation of the capital costs of facilities was solved through business arrangements based on the control of the entire value chain under a set of related contracts. Production and liquefaction facilities were financed by a syndicate of major companies (including national oil and gas companies, and major integrated oil and gas companies). Syndication was seen as a solution to high capital costs and as a risk-sharing arrangement. Buyers were usually the large regulated utilities, which could transfer the cost of purchasing LNG to their rate payers. The tankers were either owned by the supply syndicate or chartered. In both cases, the tankers would be dedicated to a given contract, shuttling between the liquefaction and regasification facilities. This arrangement was supported by a contract known as a sale and purchase agreement (SPA). Such contracts had a number of distinct features.<sup>40</sup>

- ❑ Long duration: the contracts extended typically for 20 or more years, a time period necessary to amortise the capital cost of the

entire LNG value chain. The contracts were usually over-collateralised with reserves extending within a comfortable safety margin beyond the horizon of the contract.

- ❑ The prices of LNG were indexed to oil. This practice was an outcome of the early years of the LNG industry, when natural gas competed against oil as the fuel for electricity generation. The economic rationale for this practice survived as the importance of oil for power generation diminished. LNG prices in the Far East are typically based on the JCC (known as a Japanese Crude Cocktail in the trading community).<sup>41</sup> Prices for Japanese LNG imports are based on a general formula:

$$\text{LNG} = a + b \times \text{JCC}$$

with the coefficients  $a$  and  $b$  negotiated contract by contract.<sup>42</sup> The pricing formula may contain terms which result in a lagged and smoothed reaction of LNG prices to the prices in the crude market. Prices negotiated for the European contracts are typically based on Brent crude (see the chapters on the oil markets).

- ❑ The contracts are usually negotiated using CIF (cost, insurance, freight) prices, with FOB (free on board) contracts growing in popularity in recent years. A FOB contract means that the buyer takes possession of LNG at the loading terminal.
- ❑ The contracts are typically structured as take-or-pay. This very important contract structure, used in many different commodity markets, effectively shifts the volumetric risk to the buyer (with other aspects of the contract shifting the price risk to the seller). A buyer is obligated to take a significant percentage of the negotiated annual volume (often as much as 85–90%), irrespective of the actual needs. This means that the buyer has to absorb the fluctuations of demand due to weather or economic conditions. Should the buyer be unable to take all the volumes, they would have to pay for the volumes they agreed to buy. Under some contracts, the buyer could take the delivery at a later date, after making a payment according to schedule.
- ❑ Most contracts had destination restriction provisions which limited the ability of a buyer to divert LNG to other locations in case of a temporary drop in their needs.

Long-term pricing contracts, indexed to oil prices, often include a modifier designed to protect the buyer when prices of crude increase and reach historically high levels, and to protect the seller when prices drop. Such modifiers are often referred to as an S-curve. The S-curve modifier is an example of a traditional approach to risk management, with safeguards designed for the protection of both sides of a transaction.

High capital costs and complicated logistics explain why most projects are vertically integrated, with facilities constructed and tankers built under long-term contracts, with tenors extending to 20 or more years. An alternative is to achieve vertical integration through contractual arrangements. The owners of liquefaction plants procure regasification capacity and arrange shipping under long-term leases.

The spot market for LNG cargos, estimated at 20% (47 Mmtpa) of the total world output,<sup>43</sup> exists on the margins of this system and relies on supplies from two sources:

- production from the liquefaction plants in excess of long-term contractual obligations; and
- diversion options embedded in the supply contracts which allow redirecting cargos to destinations with higher prices.

According to BENTEK Energy:<sup>44</sup>

"In 2005, as many as 20 ships had flexible contract characteristics, such as delivery options. Today, [September 2010] as many as 117 vessels have some form of flexibility – nearly one-third of the world fleet." This flexibility comes at a price and will require in turn more proactive risk management by the market participants."

The market will likely evolve from long-term contracts with pricing linked to oil prices (and possibly other variables and escalation factors) to an increased reliance on spot transactions, responsive to the current prices of natural gas. This trend will be supported not so much by the availability of excess, uncommitted liquefaction capacity, but primarily by new or renegotiated contracts that will give the parties (mostly sellers) the option to redirect cargoes from the initial destinations to more attractive locations.<sup>45</sup>

Some regasification terminals have been constructed or are under construction in the expectation of the growth of the spot LNG market

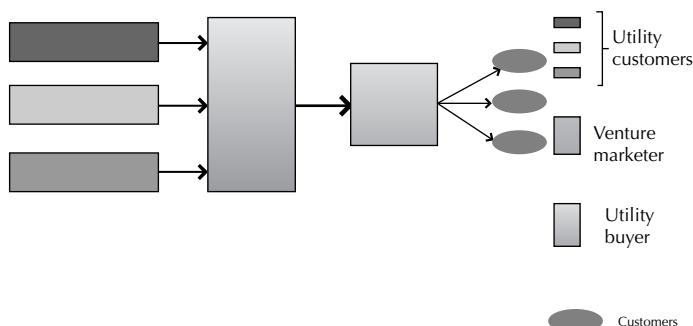
both in absolute and relative terms, without securing future LNG supply to guarantee high utilisation margins. This may be a risky strategy, as there is no evidence of a parallel strategy of developing merchant, ie, market-oriented and otherwise uncommitted, LNG facilities.

The development of a vibrant spot market for LNG is one possible direction in which the industry can evolve. Since the early 2000s, two alternative forms of market organisation have emerged<sup>46</sup> that represent hybrid solutions: a progress toward more flexible arrangements while preserving many elements of the traditional vertically integrated supply chain. Figure 12.8(a) illustrates a traditional LNG contract, while Figure 12.8(b) depicts a self-contracting arrangement. One of the LNG venture partners is exclusively responsible for the marketing of natural gas, entering into contracts with multiple end users and with embedded options giving them the ability to divert tankers to the most profitable destinations, as described below:<sup>47</sup>

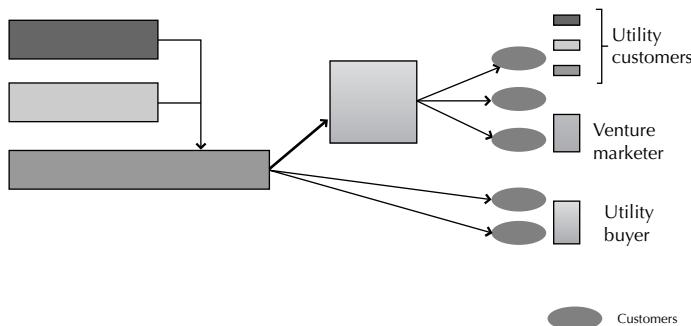
“Self-contracting gives suppliers destination flexibility that was not available under the traditional contracting system. The ultimate market destinations are defined, not by the terms of the contract, but by the best netbacks available to the supplier, given his portfolio of liquefaction and re-gasification assets.”

LNG aggregators take this model one step further by creating flexibility along the entire supply chain. They create portfolios of supply, transportation and regasification contracts, which allow them to respond quickly to evolving market conditions:<sup>48</sup>

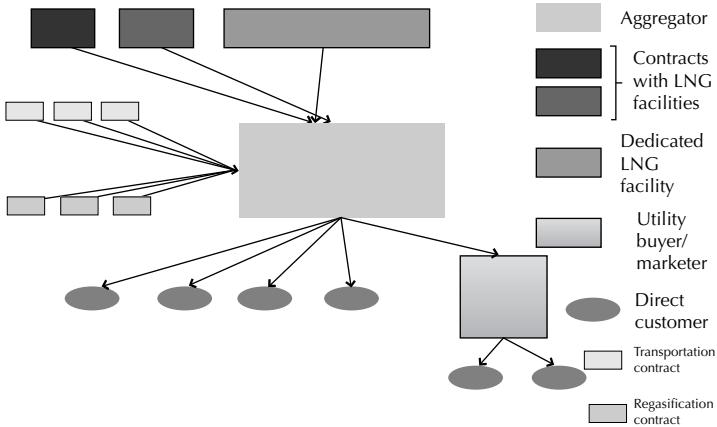
**Figure 12.8a** Traditional LNG contract



Source: “Putting a Price on Energy. International Pricing Mechanisms for Oil and Gas,” Energy Charter Secretariat, 2007.

**Figure 12.8b** Self-contracting arrangement

Source: "Putting a Price on Energy. International Pricing Mechanisms for Oil and Gas," Energy Charter Secretariat, 2007.

**Figure 12.8c** LNG aggregators

"Aggregators acquire a portfolio of LNG sources (such as short-term and long-term LNG purchases, or equity on LNG production projects), a portfolio of LNG sales (short-term and long-term contracts and capacity in import terminals within liquid markets) and a shipping fleet, and aim to deploy these to access the most lucrative gas market opportunities from time-to-time."

The aggregators may contract both with affiliated and unaffiliated companies. For example, an oil major may enter into a long-term contract with a liquefaction capacity venture in which it has an interest or with an independent producer. Such purchases are made

on an FOB basis, and then LNG may be sold either to unaffiliated utilities or gas marketers, including marketers affiliated with the oil major.

The emergence of the aggregators changes the traditional allocation of risks in the LNG contracts, and modifies the historical mutual dependence between buyers and sellers who used to be connected at the hip. The traditional view was that the sellers were shielded from volumetric risk in return for accepting the price risk through indexation and escalation clauses. In the new LNG universe, the buyers have to worry about additional issues (as explained in this *Energy Risk* chapter).

- ❑ Exercise of diversion options by the seller will result in less regular deliveries and cargos of varying sizes and quality of gas.
- ❑ The sellers are likely to show a preference for the use of a liquidated damages<sup>49</sup> approach in contract design, instead of the traditional take-or-pay provisions. The end users who value reliability and continuity of supplies are likely to object. Aggregators may engage in behaviour known as commercial default – ie, the failure to perform by an otherwise creditworthy entity if profits of the alternative course of action exceed the amount of liquidated damages.
- ❑ The probability of volumetric shortfalls increases as the aggregators are likely to run a tight ship with limited spare capacity (ie, they are unlikely to over-collateralise their contracts with reserves and excess liquefaction and regasification potential).

It is expected the LNG market will be shaped by a number of technological and commercial trends, such as the continued shale natural gas revolution in the US. The surge in natural gas production in the US has again proved how dangerous excessive reliance on the extrapolation of currently observed trends is. The US demand for LNG was greatly overestimated as far back as 2006. The US is now exporting natural gas indirectly, through redirection of the supplies intended just a few years ago for the US markets. The US-based regasification facilities are underutilised, with the owners scrambling to find a solution to their mal-investments. The owners of some facilities explored the potential for the re-export of LNG imported under long-term contracts and stored on-site at the terminal.<sup>50</sup> Some

operators of US-based LNG regasification facilities are exploring the opportunities for the development of brownfield liquefaction facilities.<sup>51</sup> Table 12.4 lists the US-based liquefaction projects under consideration.

Market conditions at the time of writing are characterised by a surplus of LNG capacity. According to Cheniere, global LNG capacity in 2011 was equal to about 37.5 Bcf/day, with regional LNG demand adding up to 30 Bcf/day.<sup>52</sup> This surplus can be explained by two factors. First, the expectations of growing US demand for LNG did not materialise (as mentioned above), and the producers had to find alternative markets. Second, the recession in western Europe has reduced demand for LNG, primarily due to the decrease in demand for natural gas by the industrial sector. These market conditions are unlikely to persist for an extended period of time. This conjecture is supported by a number of factors, including:

- ❑ western European countries seeking to diversify the sources of supply of natural gas to reduce their dependence on pipeline supplies;
- ❑ economic growth is likely to continue unabated in Asia after the current slowdown; and
- ❑ new markets for LNG are growing in Latin America; according to a study by BENTEK Energy, the balance between demand and supply of LNG is likely to swing to a net short position (demand exceeding available capacity) by 2015.<sup>53</sup>

The statement above may need to be revised, given a number of developments that took place in 2011 and 2012. The earthquake in Japan in March 2011, and the subsequent shutdown of nuclear power plants, complicated the picture. Japanese imports of LNG increased between March 2011 and January 2012 by 3.1 Bcf/day.<sup>54</sup> Continued uncertainty with respect to the strength of economic recovery in the US and western Europe further muddies the picture.

## CONCLUSIONS

The markets we covered in this chapter – the European natural gas markets and the world LNG markets – are in transition, creating a unique set of opportunities and challenges for market participants. The push towards more flexible, market-based arrangements, and

**Table 12.4** Proposed US liquefaction plants<sup>55</sup>

<b>Location</b>	<b>Export capacity (Bcf/day)</b>	<b>Export capacity (MM tons/year)</b>	<b>Target date (as of 12/6/11)</b>	<b>Export licence</b>	<b>Developer</b>
Sabine Pass, LA	2.2	16	2015	Unrestricted	Cheniere Energy
Lake Charles, LA	2	15	2015	FTA <sup>56</sup> licence, unrestricted licence under review	Southern Union and BG
Cameron, LA	1.7	12	2017	FTA licence under review	Sempra Energy
Freeport, TX	1.4	9	2015	FTA licence, unrestricted licence under review	Macquarie Group, ConocoPhilips
Kitimat, BC	1.3	10	2015	NEB export licence	Apache, Encana, EOG Resources
Jordan Cove, OR	1.2	9	2017	FTA licence under review	Fort Chicago Energy Partners and Energy Projects Development
Cove Point, MD	1	8	2016	FTA licence, unrestricted licence under review	Dominion
Hackberry, LA	1.7	12	2016		Sempra
Lake Charles, LA	2	15	2018		Energy Transfer Equity

Source: Argus, press reports

market liberalisation and integration, are, in our view, unstoppable. This offers an opportunity for agile and innovative trading organisations, and forces the incumbents to reconsider their business models. This inevitable market evolution is happening against a background of weak economic growth in Europe and the US, and slowing growth in Asia – adding additional uncertainty to an already complicated picture.

- 1 This section is based on a number of sources, including “Lo sviluppo del trading del gas naturale nel mercato Italiano. Proposte per una possibile evoluzione normativa e strutturale del PSV,” Master Medea 2006/2007, Project Work in Collaborazione con ENI Divisione Gas & Power; “From regional markets to a single European market”, report by Eneris and Mercados Energy Markets International, 2010 ([http://ec.europa.eu/energy/gas\\_electricity/studies/doc/2010\\_gas\\_electricity\\_markets.pdf](http://ec.europa.eu/energy/gas_electricity/studies/doc/2010_gas_electricity_markets.pdf)).
- 2 [http://www.energy-regulators.eu/portal/page/portal/EER\\_HOME/EER\\_PUBLICATIONS/CEER\\_PAPERS/Gas/Tab/C11-GWG-82-03\\_GTM%20vision\\_Final.pdf](http://www.energy-regulators.eu/portal/page/portal/EER_HOME/EER_PUBLICATIONS/CEER_PAPERS/Gas/Tab/C11-GWG-82-03_GTM%20vision_Final.pdf).
- 3 Gas supply business in the UK (and in Europe for many countries) is equivalent to the local distribution business in the US.
- 4 <http://www.nationalgrid.com/uk/About/history/>.
- 5 “Charges reflect the estimated long run marginal cost (LRMC) of reinforcing the system to transport additional gas between entry and exit points” (see [http://www.gasgovernance.co.uk/sites/default/files/TPD%20Section%20Y%20-%20Charging%20Methodologies\\_5.pdf](http://www.gasgovernance.co.uk/sites/default/files/TPD%20Section%20Y%20-%20Charging%20Methodologies_5.pdf) for more details).
- 6 Gas consumed by a site or customer.
- 7 The following auctions are anticipated for year 2011/2012: Quarterly System Entry Capacity (QSEC), the Annual Monthly System Entry Capacity (AMSEC), Rolling Monthly Trades & Transfer System Entry Capacity (RMTNTSEC) and Day Ahead Daily System Entry Capacity (DADSEC). See National Grid, “Explanation of the NTS SO and TO Commodity Charges for the formula year 2011/12.”
- 8 The information about the company is available from its website. <http://www.apxindex.com/index.php?id=150>.
- 9 BBL – a gas pipeline from Balgzand (the Netherlands) to Bacton (UK). “Interconnector (UK) Limited operates a sub sea gas pipeline and terminal facilities to provide a strategic bi-directional link between the UK and continental European energy markets. The Interconnector system comprises compression terminals at IBT in UK and IZT in Belgium connected by a 40 inch 235 kilometre pipeline.” (see, <http://www.interconnector.com/>).
- 10 <https://www.theice.com/productguide/ProductSpec.shtml?specId=910>.
- 11 <http://epp.eurostat.ec.europa.eu/tgm/refreshTableAction.do?tab=table&plugin=1&pcode=tsdcc320&language=en>.
- 12 To be precise, 62.38% in 2010. A link to data can be found at <http://epp.eurostat.ec.europa.eu/tgm/refreshTableAction.do?tab=table&plugin=1&pcode=tgigs360&language=en>.
- 13 The aggregate numbers do not convey the seriousness of European dependence on imported natural gas. Some countries depend 100% on imports, making them very vulnerable, especially during winter. Commercial disputes between Gazprom and its clients in Belarus and Ukraine have in the past resulted in drastic curtailments of gas flows into the EU.
- 14 Koen Rademaekers, Allister Slingenbergh and Salim Morsy, 2008, “Review and analysis of EU wholesale energy markets: Historical and current data analysis of EU wholesale electricity, gas and CO<sub>2</sub> markets,” Final report, ECORYS Nederland BV, Rotterdam, December 9.

- 15 For example, the Central European Gas Hub AG (CEGH, formerly known as Gas Hub Baumgarten) at the Austrian–Slovak border.
- 16 The following European Commission gas directives were issued: 98/30/EC in June 1998, 2003/55/EC, 2009/73/EC. The 1991 Gas Transit Directive mandated gas transportation companies to allow third-party access to their pipeline networks.
- 17 Some pricing formulas may be very complex. We have seen contracts linking North Sea natural gas to prices of electricity, gasoil and inflation in the UK.
- 18 Energy Charter Secretariat, 2007, “Putting a price on energy: International pricing mechanisms for oil and gas.”
- 19 For example, the cost of production from the Groningen field in the Netherlands was a few cents at the time the field was discovered.
- 20 The task of commercialisation of the Dutch natural gas was entrusted to Gasunie, a joint venture of Dutch State Mines (DSM) (40%), the Dutch government (10%) and Exxon and Shell (each 25%).
- 21 The discussion of long-term pricing formulas relies on Michael Polkinghorne, 2008, “Predicting the unpredictable: Gas price reopeners,” White and Case LLP, Paris Energy Series No. 2. The first of the formulas quoted below has been corrected for a typo.
- 22 Ibid.
- 23 For further details, see <http://www.bg-group.com/MediaCentre/PressArchive/1997/Pages/pr-010.aspx>.
- 24 Energy Charter Secretariat, 2007, “Putting a price on energy: International pricing mechanisms for oil and gas.”
- 25 The Third Directive (2009/73/EC) replaced the second directive (Directive 2003/55/EC); see [http://ec.europa.eu/energy/gas\\_electricity/legislation/legislation\\_en.htm](http://ec.europa.eu/energy/gas_electricity/legislation/legislation_en.htm) for more information.
- 26 “The entry into force of the EU third energy package” ([http://ec.europa.eu/energy/gas\\_electricity/legislation/doc/20110302\\_entry\\_into\\_force\\_third\\_package.pdf](http://ec.europa.eu/energy/gas_electricity/legislation/doc/20110302_entry_into_force_third_package.pdf)).
- 27 <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:211:0036:0054:EN:PDF>.
- 28 The official name is the Gas Regulatory Forum (see [http://ec.europa.eu/energy/gas\\_electricity/forum\\_gas\\_madrid\\_en.htm](http://ec.europa.eu/energy/gas_electricity/forum_gas_madrid_en.htm) for more information).
- 29 See a study by Rudolph Harmsen and Catrinus Jepma of six gas hubs: the NBP (United Kingdom), TTF (Netherlands), Zeebrugge (Belgium), NCG (Germany), Gaspool (Germany) and PEG (France), over a three-year period, from April 23, 2007, to May 7, 2010 (Rudolph Harmsen and Catrinus Jepma, 2011, “North West European gas market: Integrated already,” *European Energy Review*, January 27). The authors concluded that “The results of [...] analysis show that in the long run the relative LOP [law of one price] holds for all 15 pairs of hubs. Thus, it appears that the gas hubs in North-West Europe form one integrated market for natural gas. As a result of arbitrage, hub prices never drift too far apart.”
- 30 Enagás is the technical manager of the Gas System and common carrier for the high-pressure gas network in Spain.
- 31 The natural gas market imported by European countries varies in terms of physical and chemical qualities. Gas from Norway has the highest calorific value (43.96 MJ/m<sup>3</sup>), while gas from the Netherlands has the lowest (35.80 MJ/m<sup>3</sup>), with Russian gas between these levels (see <http://www.ommi.co.uk/PDF/Articles/75.pdf>). In Germany, two types of natural gas are marketed: low calorific and high calorific. In the Netherlands and northwest Germany, low-calorific gas is produced generally on-shore and high-calorific gas is produced offshore.
- 32 [http://cdn.eex.com/document/86102/EGIX-Info-Sheet\\_EN.pdf](http://cdn.eex.com/document/86102/EGIX-Info-Sheet_EN.pdf).
- 33 According to the index documentation, “if less than three transactions in front months contracts relevant for the index are concluded on a given trading day in the NCG or GASPOOL market area, the daily values for NCG and GASPOOL correspond to the settlement prices of the front months established. If less than three transactions each in the front month contracts relevant for the index are concluded in the NCG market area and in the

- GASPOOL market area on a given trading day, the daily value for Germany corresponds to the arithmetic mean of the daily values for NCG and GASPOOL.”
- 34 West/North/East zones merged on January 1, 2009, into Northern Zone.
- 35 In France, gas B is low-calorific gas from the Netherlands. H stands for high-calorific gas.
- 36 According to Andrey A. Konoplyanik, the churn ratio of the Nymex WTI contract was between 1,680 and 2,240, and for the ICE Futures Brent contract was equal to 2,014 (February 2010). The average for the Nymex natural gas contract in 2009 was 377. See Andrey A. Konoplyanik, 2011, “How market hubs and traded gas in European gas market dynamics will influence European gas prices and pricing,” presentation at the European Gas Markets Summit, London, UK, February 15–16.
- 37 Rudolph Harmsen and Catrinus Jepma, 2011, “North West European gas market: integrated already,” *European Energy Review*, January 27. The terms “churn rate” and “churn ratio” should be used with caution, because the assumptions underlying their calculations may vary from source to source.
- 38 According to ICIS, “if a bid–offer spread of less than €0.30/MWh is assessed to have been available every working day during the quarter for a particular contract, it will score two in the index. If the contract is assessed to have been available at a spread of less than €0.50/MWh every day, but not less than €0.30/MWh every day, it will score one in the index. If the contract is assessed to have not been available at a spread of less than €0.50/MWh every day, it will score zero in the index. The maximum score for any hub is 20 and the minimum is zero.” (see <http://www.icis.com/energy/gas/europe/hub-report-methodology/>).
- 39 International Gas Union, <http://www.igu.org/igu-publications-2010/IGU%20World%20LNG%20Report%202010.pdf>.
- 40 James T. Jensen, 2004, “See the development of a global LNG market. Is it likely? If so, when?”, Oxford Institute for Energy Studies.
- 41 JCC stands for “Japanese Customs-Cleared Crude,” an index of the Japanese import prices of crude. See, <http://www.paj.gr.jp/english/> for more information.
- 42 The coefficients for the Japanese LNG negotiated for contracts prior to 2002 were estimated to average 1.2367, 0.1226, for a and b, respectively (See Gary Eng, “A Formula for LNG Pricing,” June 16, 2006. Subsequent update available at <http://www.lngpedia.com>). The contracts details are not in the public domain.
- 43 International Gas Union, <http://www.igu.org/igu-publications-2010/IGU%20World%20LNG%20Report%202010.pdf>.
- 44 BENTEK Energy, op.cit.
- 45 One example with more specific information is available from a recent industry report: “To the extent LNG cargos have flexibility in delivery locations, supplies are instead heading to Europe and Asia, where LNG prices remain higher than those that have prevailed in US markets.” See “Natural Gas Outlook: Spot Prices Increase on Electric Power Demand,” September 16, 2010, <http://www.cattlenetwork.com/Natural-Gas-Outlook—Spot-Prices-Increase-On-Electric-Power-Demand/2010-09-16/Article.aspx?oid=1238635&fid=1>.
- 46 See Energy Charter Secretariat, 2007, “Putting a price on energy: International pricing mechanisms for oil and gas,” for a discussion of self-contracting arrangements. LNG aggregation arrangements are described by Susan Farmer and Ben Smith, 2009, “LNG buyers, beware,” *Energy Risk*, February, pp 60–2.
- 47 *ibid*, p 183.
- 48 Susan Farmer and Ben Smith, op.cit.
- 49 “Monetary compensation for a loss, detriment, or injury to a person or a person’s rights or property, awarded by a court judgment or by a contract stipulation regarding breach of contract. Generally, contracts that involve the exchange of money or the promise of performance have a liquidated damages stipulation. The purpose of this stipulation is to establish a predetermined sum that must be paid if a party fails to perform as promised.” (see <http://law.jrank.org/pages/8310/Liquidated-Damages.html>).

- 50 "Sabine Pass and Freeport – currently the only US terminals that can re-export LNG – have together exported 9.7 Bcf of gas since they received the approval last year," according to Waterborne Energy analysts in Houston. Sempra Energy's Cameron LNG terminal in Louisiana has also applied for a re-export licence (see, <http://af.reuters.com/article/energyOilNews/idAFN2221699920100922?pageNumber=2&virtualBrandChannel=0>).
- 51 The FERC has approved (April 2012) Texas-based Cheniere Energy's plan to build a liquefaction and export terminal at Sabine Pass (see also, [http://www.freeportlng.com/The\\_Project.asp](http://www.freeportlng.com/The_Project.asp)).
- 52 Cheniere Energy, Citi Global Energy Conference, June 5, 2012 (<http://phx.corporateir.net/phoenix.zhtml?c=101667&p=irol-presentations>).
- 53 BENTEK Energy, 2010, "The LNG market: New world order," September 15.
- 54 Javier Diaz and Kelly Bennett, 2012, "The global LNG market post Fukushima: Why the world needs North America's future LNG exports," BENTEK Energy Benposium.
- 55 <http://www.argusmedia.com/pages/NewsBody.aspx?id=777612&menu=yes>,  
<http://af.reuters.com/article/commoditiesNews/idAFL3E8FH58V20120417>,  
[http://markets.on.nytimes.com/research/stocks/news/press\\_release.asp?docTag=201204020830BIZWIRE\\_USPRX\\_\\_\\_\\_BW5438&feedID=600&press\\_symbol=102500](http://markets.on.nytimes.com/research/stocks/news/press_release.asp?docTag=201204020830BIZWIRE_USPRX____BW5438&feedID=600&press_symbol=102500).
- 56 Free Trade Agreement.

Section 4

## **Oil Markets**



## *Oil Markets: Properties, Production and Reserves*

This chapter will explore some general topics related to the properties of crude oil. One of the biggest differences between natural gas (covered in the previous section) and oil is that the latter is a commodity with multiple grades, with different physical and chemical properties, which translate into differences in market values. A trader always buys or sells a specific grade at a specific location. Even in the case of financial futures with standardised properties of the underlying crude, multiple grades can be usually delivered into a contract. The same is true of refined products for which acceptable specifications vary from location to location and from season to season. The physical components of the industry infrastructure (refineries, pipelines, tankers) are optimised for processing and transporting certain types of crude and products. A typical refinery, for example, is configured for processing certain types of crudes and, if a basket of inputs with different qualities has to be used, profits will be reduced. A trader can create a lot of value if they can match the technological parameters of a refinery with a portfolio of crudes they can acquire in the market.

On many occasions we have seen traders buying presumably attractively priced crude, only to discover it is landlocked, with limited transportation options, with only one local refinery equipped to process it. In such situations, a trader becomes a captive supplier of a refinery or a captive customer of a company controlling transportation. This chapter will hopefully provide useful advice to traders and analysts embarking on a study of this market. As in the case for all commodities, understanding the physical layer of the industry is a precondition to being a successful trader and marketer.

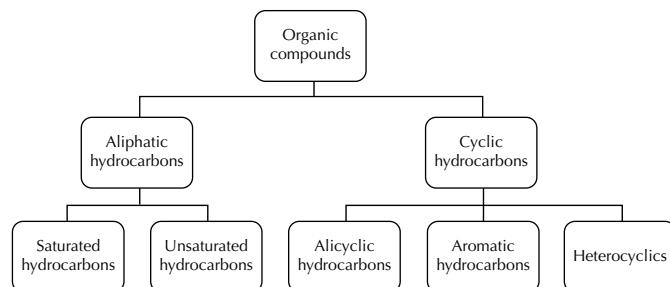
The second topic covered in this chapter will be a definition of reserves. This is a critical topic for anybody involved in the acquisition of oil-producing properties, or stock analysts covering companies that produce oil and natural gas. Reserves are an important metric in the valuation of companies, and it helps to realise that it can sometimes be a very fuzzy and imprecise concept.

### CRUDE OIL AND ITS PROPERTIES

Unlike natural gas, coal of a given variety and electricity, oil is not a homogenous substance – but rather a mixture of different hydrocarbons, with physical and chemical differences related to the number of atoms of carbon and hydrogen in each molecule, and the structure of the molecules (the way the atoms of carbon and hydrogen are bonded together). Crudes produced in different locations have qualities dependent on the proportions of different hydrocarbons and the presence of other substances, which are usually seen as contaminants that often require complex and expensive processes of removal and disposal.<sup>1</sup>

Given that hydrocarbons are made of only two different types of atoms, the number of different compounds and the variety of their properties are astounding. Anybody active in the oil business has to revisit at some point their school textbooks on organic chemistry. The organic compounds can be classified as aliphatic and cyclic hydrocarbons.<sup>2</sup> Aliphatic<sup>3</sup> hydrocarbons contain only hydrogen and carbon.

**Figure 13.1** Classification of hydrocarbons



Source: Donald L. Burdick, William L. Leffler, "Petrochemicals in Non-Technical Language," PennWell, Tulsa 2001

When carbon atoms are linked by a single bond, they are called saturates. In the case of multiple bonds, the term non-saturates is used.

The hydrocarbons found in oil are classified based on their molecular structures, and include the following.

- ❑ Paraffins are compounds which have a molecular structure given by  $C_nH_{2n+2}$ , and can be gases or liquids at room temperature. For example:<sup>4</sup>
  - ❑  $CH_4$ ,  $C_2H_6$ ,  $C_3H_8$ , and  $C_4H_{10}$ , are gases that liquefy at or below 32°F;
  - ❑  $C_6H_{14}$ , through  $C_9H_{20}$ , are liquids boiling at 150–300°F; and
  - ❑  $C_{30}H_{62}$  and higher are solids melting above 150–300°F and boiling above 500°F.

Examples of paraffins include methane ( $CH_4$ ), ethane ( $C_2H_6$ ), propane ( $C_3H_8$ ), butane ( $C_4H_{10}$ ), iso-butane (also known as methyl propane), which is isometric<sup>5</sup> to butane, pentane ( $C_5H_{12}$ ), hexane, undecane ( $C_{11}H_{24}$ ) and cetyl ( $C_{30}H_{62}$ ).

- ❑ Aromatics<sup>6</sup> are compounds that have a molecular structure given by  $C_6H_{5-Y}$ , where Y is a longer molecule that attaches to the benzene ring. The examples include benzene ( $C_6H_6$ ), xylene ( $C_8H_{10}$ ), toluol ( $C_7H_8$ ), cymol ( $C_{10}H_{14}$ ), naphtalene ( $C_{10}H_8$ ).
- ❑ Naphthenes (or cycloalkanes), which have a molecular structure given by  $C_nH_{2n}$ , are typically liquid at room temperature. The molecules have a ringed structure with one or more rings (the rings have only single carbon bonds); examples include cyclobutane ( $C_4H_8$ ), cyclopentane ( $C_5H_{10}$ ) and cyclohexane ( $C_6H_{12}$ ).
- ❑ Olefins (also known as alkenes) contain one or more pairs of carbon atoms that are linked by a double bond. The olefins are usually classified as cyclic (a double carbon bond is a part of a ring structure) or acyclic (with an open carbon chain containing double bond). An alternative classification is based on the number of double carbon bonds per molecule (monoolefins, diolefins and triolefins). Acyclic monoolefins have a molecular structure given by  $C_nH_{2n}$  (what makes them different from naphthenes with the same general molecular formula is the presence of a double carbon bond), and acyclic diolefins (also known as acyclic dialkenes), or acyclic dienes, have the general formula  $C_nH_{2n-2}$ . Examples of lower monoolefins include ethylene ( $C_2H_4$ ),

propylene ( $C_3H_6$ ), butylene ( $C_4H_8$ ). Examples of acyclic diolefins include butadiene ( $C_4H_6$ ) and isoprene ( $C_5H_8$ ).

### Physical properties and classification of crudes

The most important physical and chemical characteristics of oil can be summarised through:

- gravity (density);
- sulphur content;
- total acid number (TAN);
- viscosity; and
- the percentage of vacuum residue (VR).<sup>7</sup>

All these jointly determine how effectively oil can be processed into refined products. Many of the properties discussed briefly below may not seem like essential information, but be mindful that many traders have been fired because they ignored some specifications of crudes they transacted.

Density is defined as the weight of a substance per unit of volume. Specific density is defined as the ratio of densities of two substances. If the reference density is that of water, the term gravity is used. Specific gravity is defined as the ratio of the weight of a compound to the weight of water. In the oil industry, the API gravity (following the standards established by the American Petroleum Institute) is used, and its definition<sup>8</sup> is given by:

$$\frac{141.5}{\text{Specific gravity}} - 131.5 \quad (13.1)$$

where specific gravity is measured at 60°F. If API gravity is greater than 10, the liquid floats on water – this can be easily seen from equation 13.1. If the specific gravity is one (ie, the liquid has the same density as water), the API gravity will be equal to 10.<sup>9</sup> Heavy oils have low API gravity, and the higher the number, the lighter is the specific type of oil or refined product. Asphalt, for example, has a gravity of around 11, gasoline of about 60. Heavy oil is associated with gravity of less than 22.3°, light oil with gravity of exceeding 31.1°, typically in the range of 32–36.<sup>10</sup> Hydrocarbons with an API gravity less than 10 will be covered in the section on non-conventional oils.

Crude oils contain varying quantities of sulphur. Oil with a low sulphur content is referred to as sweet crude,<sup>11</sup> while oil with a high sulphur content is sour crude. This is an important distinction, because sulphur creates complications in the refining process (corrosion of equipment, poisoning of catalysts). Many countries, including the US, have strict regulations regarding the sulphur content in refined products. This explains why sulphur has to be removed at a significant cost. Sweet crude contains less than 0.5% sulphur, sour crude more than 2.5%.

A combination of these two characteristics leads to the 3 x 3 classification of different oil produced worldwide shown in Table 13.1. One troubling trend has been the drop in relative and absolute levels of light and sweet crudes production and the growing relative output of heavier and sourer crudes. This trend has very serious consequences for the relative prices of different refined products. Light crudes yield a basket of refined products with a higher content of highly valued outputs, and can be processed in relatively simple refineries. Heavy crudes run through simple refineries yield baskets of products that are much less attractive and have to be processed further, using complex, expensive and energy-intensive technologies to create a more attractive combination of outputs (this topic will be discussed in more detail in the section on refinery configuration).

The TAN measures the acidity of crude due to the presence of naphthenic acids. These acids can be neutralised by adding potassium hydroxide (KOH). Crudes requiring more than 1.0 mg KOH/g are considered to be high TAN grades. The TAN, in conjunction with the sulphur level, defines the corrosivity of oil, the property that makes transportation and processing of crudes difficult.<sup>12</sup> Typically, high TAN crudes tend to be on the heavy side (under 29 API) and typically low in sulphur content.

**Table 13.1** Classification of different crudes

Crude	Light	Intermediate	Heavy
Sweet	S<0.5%, API>31.1°	S<0.5%, 22.3°<API<31.1°	S<0.5%, API<22.3°
Intermediate	0.5%<S<2.5%, API>31.1°	0.5%<S<2.5%, 22.3°<API<31.1°	0.5%<S<2.5%, API<22.3°
Sour	S>2.5%, API>31.1°	S>2.5%, 22.3°<API<31.1°	S>2.5%, API<22.3°

S stands for the percentage of sulphur, API for the API gravity

Other important properties of crude oils and refined products include:

- water content;
- flash point;
- fire point;
- evaporation point;
- Reid vapor pressure (RVP);
- pour point; and
- volatile organic compounds (VOCs).

There are many other chemical and physical properties of crude that we cannot cover due to space limitations. A discussion of some properties requires a background in science, and this is not what we claim. It is important to remember, however, that some properties may become very important under some circumstances. The evaporation rate may, for example, be of critical importance during an oil leak or spill, as many volatile fractions of oil may evaporate and dissipate within days after an accident.

Flash point is the lowest temperature at which fuel produces a vapour/air mixture that ignites when exposed to flame.<sup>13</sup> Fire point is the lowest temperature at which the vapour of a liquid ignites and burns for at least five seconds. Ignition temperature, known also as the auto ignition point, is the lowest temperature at which a liquid will ignite without the presence of flame. Pour point<sup>14</sup> is the minimum temperature at which a liquid will flow under well-defined test conditions. The problem with measuring the pour point is that it may be path-dependent – ie, dependent on how long the liquid was heated or cooled before the test was performed. Viscosity and certain other physical and chemical characteristics of oil and refined products are covered in the section on fuel oils and gasoline.

One can find multiple classifications of crudes in the literature, which vary depending on the type of analysis (for example, engineering or trading). Most classifications are based on gravity and sulphur content, or these two variables and, additionally, TAN and VR. Chemical engineers focus on the properties of crudes and distinguish between:

- paraffinic crudes;
- naphthenic crudes; and
- asphaltic crudes.

The classification is based on the ratio of paraffinic (naphthenic, polynuclear aromatics) relative to other hydrocarbons.

Another classification of crudes used by the industry is based on its origin. Crude oil from a specific field tends to be homogeneous with respect to its physical and chemical properties. Examples of the most important crudes (which will be discussed later in Chapter 17) include Brent and WTI.

The chemical and physical characteristics of crude oil certified by a testing laboratory are summarised in a document known as an assay.<sup>15</sup> One can find examples of assays by visiting oil company websites. The assay of the Ekofisk crude (an important crude discussed in detail in the section on oil pricing) can be found on the Statoil website.<sup>16</sup> Many terms used in the assay will become clear after reading the section on refining.

One can ask if a trader or a fundamental analyst working for an energy trading operation should worry about the physical properties of oil and about refinery operations (covered in Chapter 15). The answer is that they should know enough to ask the right questions, seek advice from the experts and connect the dots.

The Trans-Alaska Pipeline (TAP) provides a good example of complications arising from certain physical properties of crude, in combination with the constraints imposed by the physical infrastructure. The 800-mile pipeline extends from the producing region of Prudhoe Bay to the port of Valdez in Alaska. When the pipeline was first put into operation, it transported two million barrels a day,<sup>17</sup> and took about three days to move the oil across. In the middle of 2011, volumes dropped to one third of initial levels, and the time required to move the oil across increased five times.<sup>18</sup> Naturally warm oil stays above ground for a much longer time, meaning that it cools down before reaching its destination. Ice crystals and waxes may form inside the pipeline, potentially affecting the ability to operate the system.<sup>19</sup> This problem can be compared to blocked arteries due to build-up of plaque along their walls. There are several measures the pipeline can take to address this problem. One potential technological solution is to heat oil transported in the pipeline.

Increasing the volume is more difficult, as there is a strong opposition to granting permits to oil companies wishing to expand drilling in Alaska. An alternative is to decommission the pipeline, a step that would be very expensive to the owners. Under law, the pipeline has to be dismantled.<sup>20</sup>

An analyst, who can leverage their understanding of the technological issues related to the production and transportation of oil, and anticipate, for example, the potential impact of the challenges related to TAP operations on market prices and the involved energy companies, can make a significant contribution to their company. This is another illustration of the central message of this book: to function in the commodity markets, one has to understand the details of the physical layer of the industry. One cannot trade with any degree of success just by watching the prices flashing on the screen, and relying on technical models or statistical algorithms that are blind to the underlying reality.

### **UNITS OF MEASUREMENT**

A barrel of oil is probably one of the most widely accepted Anglo-Saxon units of measurement across the world. The world oil output is typically reported in millions of barrels per day (bbl/day). The abbreviation bbl stands for blue barrel. According to energy industry folklore, in the early days of the industry oil was transported by truckers along the dirt roads in Pennsylvania in old wine barrels with a capacity of 50 gallons. The barrels were not covered and a lot of oil was spilled on occasions (this was in the days long before the Environmental Protection Agency). John Rockefeller was a man disinclined to overpay for his oil, so he decided to standardise the barrel as equal to 42 gallons in order to account for the spills. Although this is a good story, it is likely to be an exaggeration. "Blue" comes from the colour used in the logo of his company (Standard Oil of Ohio), a colour now associated with the corporate logo of Chevron. This story is probably an exaggeration as well. According to Paul H. Giddens, an oil business historian, in the early days of the industry kerosene was shipped in blue barrels, gasoline in red barrels. Kerosene was initially more valuable (there were no cars on the roads), so the blue became the colour of choice for defining the measurement unit. Another reason to doubt the Standard Oil roots of the bbl definition is that this abbreviation was used for a long time

before John Rockefeller descended on this earth – in order to avoid confusion with bale, an English unit of weight, for which the abbreviation bl had been used.

Barrels per day can be abbreviated as BPD, BOPD, bbl/d, bbl/day, bbls/d, bbls/day, bpd, bd or b/d (in order to make everybody's life easier). One thousand barrels is typically denoted by Mbbl, one million by MMbbl (see the chapter on natural gas for an explanation). 1 barrel equals 158.984 litres (according to BP, "Statistical Review 2006").<sup>21</sup>

Countries relying on the metric system use tons or m<sup>3</sup> to measure the volumes of oil and refined products. The conversion of barrels into metric tons is somewhat tricky, given that oil is not a homogeneous substance, with the same physical and chemical properties independent of the source. According to the EIA, the conversion factor for the US in 2006 was 7.333 barrels to a metric ton. For example, for Saudi Arabia the corresponding conversion factor was 7.403.<sup>22</sup>

The conversion factors for products are shown in Table 13.2:<sup>23</sup>

**Table 13.2**

Products	To convert			
	Barrels to tonnes	Tonnes to barrels	Kilolitres to tonnes	Tonnes to kilolitres
	Multiply by			
LPG	0.086	11.6	0.542	1.844
Gasoline	0.118	8.5	0.740	1.351
Kerosene	0.128	7.8	0.806	1.24
Gas oil/ diesel	0.133	7.5	0.839	1.192
Residual fuel oil	0.149	6.7	0.939	1.065

Source: BP

The conversion factors of barrels of crude into energy units depend on the quality of the specific type of oil. For the US, the conventional conversion factor is 5.8 MMBtus per barrel. For Saudi Arabia, the conversion factor is 5.91 MMBtu per barrel.

As with natural gas, a standard recommendation applies: the units and conversion factors should be specified in the contract (unless one is engaging in a highly standardised transaction). The conversion factors should be listed with reference to the specific conditions under which they apply (temperature and pressure).

## OIL OUTPUT AND DEMAND STATISTICS

Any fundamental analyst and trader engaged in the oil markets faces the unenviable task of working with oil statistics that are plagued with a number of problems. Gathering information about oil output and consumption from over 100 countries is a formidable task, given the differences in definitions, inevitable errors, omissions and double counting. The problem is amplified by the shift of demand and production to the countries with less than perfect reporting systems, and is further complicated by differences in the quality and energy content of different crudes. Any aggregation will create unavoidable distortions, which have to be recognised in market data analysis. There are several private and public data sources about the output of, and demand for, crude and refined products that will be briefly reviewed in this section.

A good starting point is the crude oil output information from the EIA, which publishes one of the most important statistical revues of the oil industry, *International Petroleum Monthly*.<sup>24</sup> The US petroleum products output data contain the wealth of information, including:<sup>25</sup>

- Crude Oil Production, 48 States (including lease condensates);
- Crude Oil Production, Alaska;
- Natural Gas Plant Liquids Production;
- Renewable Fuels and Oxygenate;
- Petroleum Processing Gain;
- Petroleum Imports;
- Petroleum Exports;
- Petroleum Net Imports;
- Petroleum Stock Change; and
- Petroleum Adjustments.

Crude oil reported along with other liquids is often referred to as other conventional crude. Conventional crude is typically defined as crude with an API gravity exceeding 10°. Lease condensate, with API gravity of 50–85°, liquid under normal pressure and temperature conditions, is produced along with natural gas. In the case of associated gas, lease condensate is separated at the wellhead and usually comingled with the oil flow. Lease condensate from non-associated gas is included in the natural gas liquids. Refinery processing gains correspond to the volume increase resulting from breakdown of

bigger hydrocarbon molecules into smaller ones, which are not as densely packed and therefore have larger volume. Ethanol blended with gasoline (as opposed to ethanol consumed in more conventional ways) and MTBE are also included in the petroleum statistics.

This template is followed by other reporting entities with some variations, such as:

- other biofuels, in addition to ethanol may be included (such as, for example, biodiesel);
- liquids produced through the conversion of other hydrocarbons than crude oil:
  - natural gas-to-liquids (GTL) technology); and
  - coal-to-liquids (CTL) technology).
- non-conventional crudes, which may include ultra-heavy crudes and syncrude.<sup>26</sup>

Table 13.3 contains a summary of the reporting conventions followed by the most widely used sources. A quick look at the data demonstrates the difficulty of following statistics related to oil output, given overlapping and/or incomplete coverage.

One of the data sources included in Table 13.3, and used widely by the industry is the BP “Statistical Review,”<sup>27</sup> available free of charge for download on an annual basis. It is highly regarded, and a frequently quoted source of data.

The conventional way of reporting the production of oil is in (millions) of barrels per day. A barrel is equal to 42 gallons or roughly 159 litres. Many countries measure oil output in tons. The conversion of barrel into tons varies depending on the origin of crude and its physical and chemical properties. As mentioned, the conversion factor for the US was equal to 7.333 barrels per metric ton,<sup>28</sup> while the conversion factor for Venezuela (for example) was equal to 6.685 (oil from Venezuela is much heavier).

The poor quality of production and demand data is an obvious impediment to the development of more transparent and efficient energy markets, and a major cause of significant price fluctuations and market instability. One joint initiative undertaken to address this problem is the Joint Oil Data Initiative (JODI), a project carried out under the auspices of seven international organisations involved in oil statistics:

**Table 13.3** Comparison of global data sources on oil production and reserves

Source	Reserves Data	Grouping of liquids in production of data	Liquids excluded
IEA Oil Market Report	No	<ul style="list-style-type: none"> <li>• Crude oil, condensate, NGLs<sup>1</sup>, non-conventional<sup>2</sup></li> <li>• Refinery gains</li> <li>• Other Biofuels<sup>3</sup></li> </ul>	None
EIA International Petroleum Monthly	No	<ul style="list-style-type: none"> <li>• Crude oil, condensate, NGLS, other liquids<sup>4</sup>, refinery gains</li> <li>• NGLs reported separately</li> </ul>	None
BP Statistical Review	Proved	<ul style="list-style-type: none"> <li>• Crude oil, oil shale, oil sands, condensate, NGLs (aggregated)</li> </ul>	CTLs, GTLs, biofuels
Oil and Gas Journal	Proved	<ul style="list-style-type: none"> <li>• Crude oil, condensate, syncrude (aggregated)</li> </ul>	NGLs, CTLs, GTLs, biofuels
World Oil Magazine	Proved	<ul style="list-style-type: none"> <li>• Crude oil, condensate, syncrude (aggregated)</li> </ul>	NGLs, CTLs, GTLs, biofuels
OPEC Annual Statistical Bulletin and Oil Market Report	Proved	<ul style="list-style-type: none"> <li>• Crude oil, condensate, NGLs<sup>5</sup> (Aggregated)</li> <li>• Refinery gains</li> </ul>	CTLs, GTLs, biofuels
IHS Energy PEPS database	Proved and probable	<ul style="list-style-type: none"> <li>• Liquids (crude oil, condensate, NGLs, LPG, heavy oil, syncrude)</li> </ul>	CTLs, GTLs, biofuels

Source: U.S. Energy Information Administration (2009b)

Notes: Precise definition and coverage of liquids is not always made clear.

1. NGLs reported separately for OPEC.

2. Including biofuels, oil sands, oil shales, CTLs, GTLs and blending components such as MTBE.

3. Biofuels from sources outside Brazil and US.

4. Biofuels, CTLs, non-oil inputs to MTBE, orimulsion and other hydrocarbons (EIA, 2008b).

5. NGLS reported separately for OPEC.

Source: "Global Oil Depletion. An assessment of the evidence for a near-term peak in global oil production," UK Energy Research Centre, August 2009, p 29 (<http://www.ukerc.ac.uk/support/Global%20oil%20Depletion>)

- the Asia Pacific Energy Research Centre (APERC);
- the Statistical Office of the European Commission (Eurostat);
- the International Energy Agency (IEA–OECD);
- the International Energy Forum (IEF);
- the Latin American Energy Organization (OLADE);
- the Organization of the Petroleum Exporting Countries (OPEC); and
- the Energy and Industry Statistics Section of the United Nations Statistics Division (UNSD).

These organisations held a meeting in Paris in November 2000, which led to the launching of Joint Oil Data Exercise<sup>29</sup> in April 2001. The most important result of JODI is the creation of a publically available database, which currently contains the following information:<sup>30</sup>

*Products:*

- crude oil (including lease condensate, excluding NGL);
- LPG (propane and butane);
- gasoline (motor and aviation gasoline);
- kerosene (jet and other kerosene);
- gas/diesel oil;
- heavy fuel oil (resid and boiler oil, including bunker oil); and
- other oil (categories of products listed above and other products such as refinery gas, ethane, naphtha, petroleum coke, white spirit and SBP, paraffin, waxes, bitumen, lubricants, etc).

*Flows:*

- production (marketed production, after removal of impurities but including quantities consumed by the producer in the production process);
- imports/exports (goods having physically crossed international boundaries, excluding transit trade, international marine and aviation bunkers);
- closing stock (the primary stocks level at the end of the month within national territories; includes stocks held by importers, refiners, stock-holding organisations and governments);
- stock changes;
- refinery intake (observed refinery throughputs);
- refinery output (gross output (including refinery fuel)); and
- demand (deliveries or sales to the inland market (domestic consumption) plus refinery fuel plus international marine and aviation bunkers).

The information is submitted by over 90 countries, with timeliness, coverage and reliability of the data at reasonable levels for the top 30 countries. One can expect that JODI will at some point evolve into a critical data source for the energy industry.<sup>31</sup>

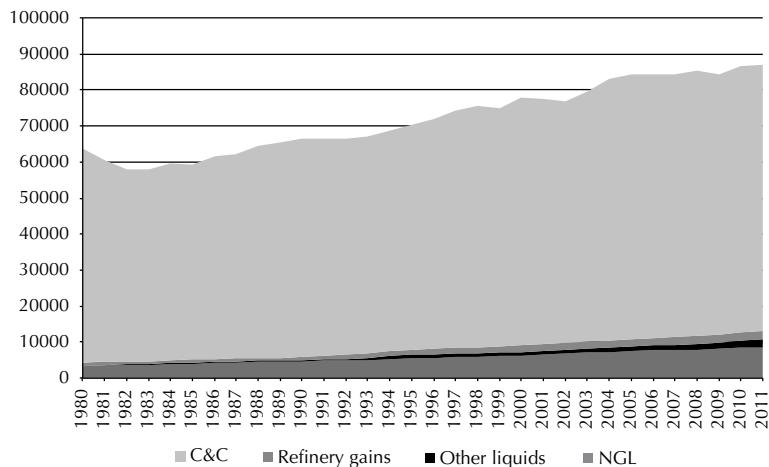
### Production trends

Figure 13.2 illustrates the worldwide production trends of crude oil and the growing dependence on natural gas liquids and unconventional oil.<sup>32</sup> Conventional oil production<sup>33</sup> worldwide (Figure 13.3) reached a plateau sometime in 2005/2006, with small production gains over 2005 levels in 2010 and 2011, so the world has to rely increasingly on other liquids (such as NGLs) – which are, at best, imperfect substitutes for conventional oil, both from the point of view of their physical and chemical qualities and the cost of extraction.

Total world oil supply, as reported by the EIA, reached the level of 87,040 thousand barrels per day in 2011, with its components: crude and condensates (C&C), NGLs, other liquids and refinery gains representing respectively 74,050.7, 8,577.3, 2,119.5 and 2,292.7 thousand barrels per day. NGLs and other liquids accounted for 77% of overall annual production gains between 2005 and 2011. This is not a surprising conclusion, given that production levels of conventional oil are stagnating (as shown in Figure 13.3).

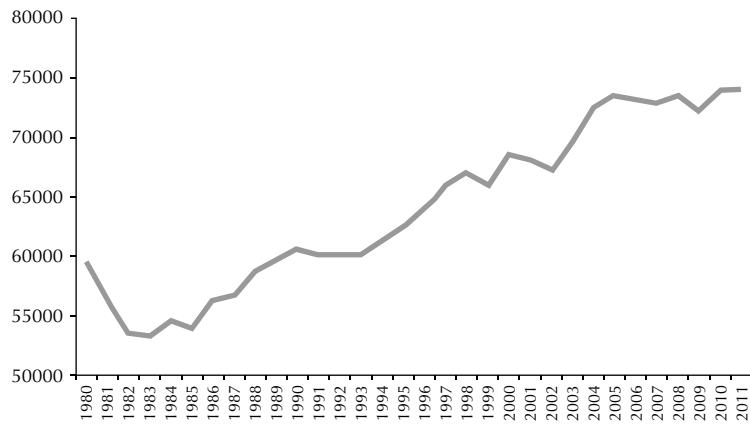
Production trends in the US are illustrated in Figure 13.4. A salient feature is a significant increase in the output of crude and condensates, starting in 2009. This is related to the growth in exploration activities in several unconventional oil regions in the US (which will be covered in the next chapter). Over the longer time period, NGLs and other liquids still represented 66.5% of total production gains between 2005 and 2011. The practice of comingling the production of conventional oil (by both the IEA and the US EIA) with NGLs and other liquids has many critics. The most frequent objection raised by critics who subscribe to the peak oil theory (covered in the next chapter) is that reporting the growing output of NGLs and biofuels along with conventional crude oil masks the fact that the peak of narrowly defined oil production has been reached, or is nigh. Some industry observers criticise reporting practices, using sometimes very strong language.<sup>34</sup> Accusations of conspiracy<sup>35</sup> to hide the truth from the public seem to be pushing things too far. A simpler explanation is more likely to be the answer: it is fairly difficult to do a good job reporting oil output statistics, across many countries, various grades and inconsistent national classifications of different types of crude. The Occam's razor principle (mentioned in Chapter 4) should suffice: why look for a plot if the limitations of our capabilities provide the answer.

**Figure 13.2** World total oil supply (EIA) (1980–2011, thousand barrels per day, annual data)



Source: U.S. Energy Information Administration

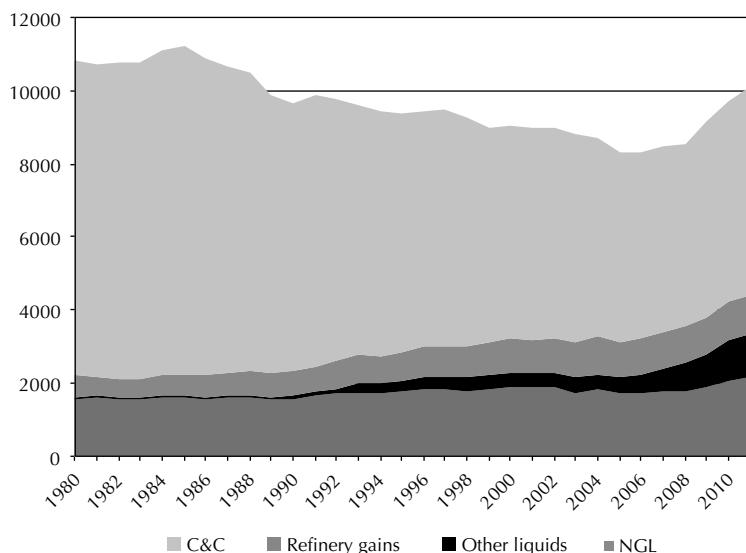
**Figure 13.3** World supply of conventional oil supply, 1980–2011, thousand barrels per day, annual data



Source: U.S. Energy Information Administration

Another objection is related to the lack of homogeneity of the reported aggregates. Natural gas liquids are not an effective substitute for conventional oil. Including biofuels such as ethanol may amount to double counting some volumetric flows. Farmers burn

**Figure 13.4** US total oil supply (1980–2011, thousand barrels per day, annual data)



Source: U.S. Energy Information Administration

diesel produced in refineries in order to produce corn, which is transformed into ethanol, which counts as part of the production level of liquid fuels. It is likely that upgrading refineries to produce more gasoline and less diesel out of a given basket of crudes would be more socially beneficial. The production of petroleum as reported by the EIA and IEA would, however, fall.

Some critics point out the problems related to transparency of the reported numbers and the difficulty of navigating the statistical tables given less than perfect documentation. For example, the “Monthly Oil Market Report” by the IEA reported in June 2010 “other biofuels” at around 460 kb/d, with the following explanation: “Other biofuels are from sources outside Brazil and US. US and Brazil oil supply includes ethanol.”<sup>36</sup> A month later the corresponding number increased to 1,800 Kb/d and the footnote mutated to “As of July 2010 OMR, Global Biofuels comprise all world biofuel production including fuel ethanol from the US and Brazil.” Tracking changes in the underlying definitions requires sometimes the equivalent of detective work. Another problem is somewhat liberal use (guilty as charged) of the terms such

as: petroleum, liquids, crude, oil. Be mindful to always check what underlying definition is being used.

In our view, one should use the term "liquids" when referring to a basket of conventional crude, NGLs, biofuels, etc.

## **RESERVES**

This section offers a very general review of the different definitions of reserves used in practice, and provides a warning against excessive reliance on any published oil and natural gas reserve statistics. Hopefully, the references provided here will help the reader to explore this topic in depth.

The concept of reserves is very important for a number of reasons. At a company level, reported reserves and the ability to replace reserves is a very important valuation metric behind the stock price, and a critical criterion used to evaluate the management.<sup>37</sup> Many mergers and acquisitions in the energy industry are driven by a desire to obtain control of the reserves at attractive prices compared to the exploration effort.<sup>38</sup> The wisdom of such corporate combinations hinges on the ability to assess the information provided by the acquisition target and third-party experts. At the national and international levels, the reserves are critical to the ability to predict production trends and price dynamics. In the case of OPEC countries, the reserves are not only the source of national pride, but also are used as the benchmark for allocating production quotas.<sup>39</sup>

There is no concept more important and more vague when it comes to the exploration and production of hydrocarbons than reserves.<sup>40</sup> The confusion often arises from different systems used to define and classify reserves, and overlap between terms used by reservoir engineers, economists and accountants. Economists and market analysts can make up their own definitions; accountants have to follow generally accepted accounting principles. Reserves should not be conflated with the total volume of hydrocarbons existing at a given location: they represent the volumes that can be extracted economically, given the current state of the technology. The reported reserves will usually evolve over time as the technology changes, and additional information is obtained in the course of exploration and production. Another source of confusion arises from inconsistent definitions of crude used by different authors.

The quality of reported reserve numbers varies from company to

company and from country to country. Figure 13.5 illustrates a well-known fact: that the oil reserves of some OPEC countries underwent a one-time change in the past, and have not been revised since. Given ongoing production and exploration activities, constancy of reserves is simply impossible and is a source of frequent complaints about the poor quality of data available to oil traders. Understanding the reported numbers and their limitations is a critical skill for any participant in the energy markets.

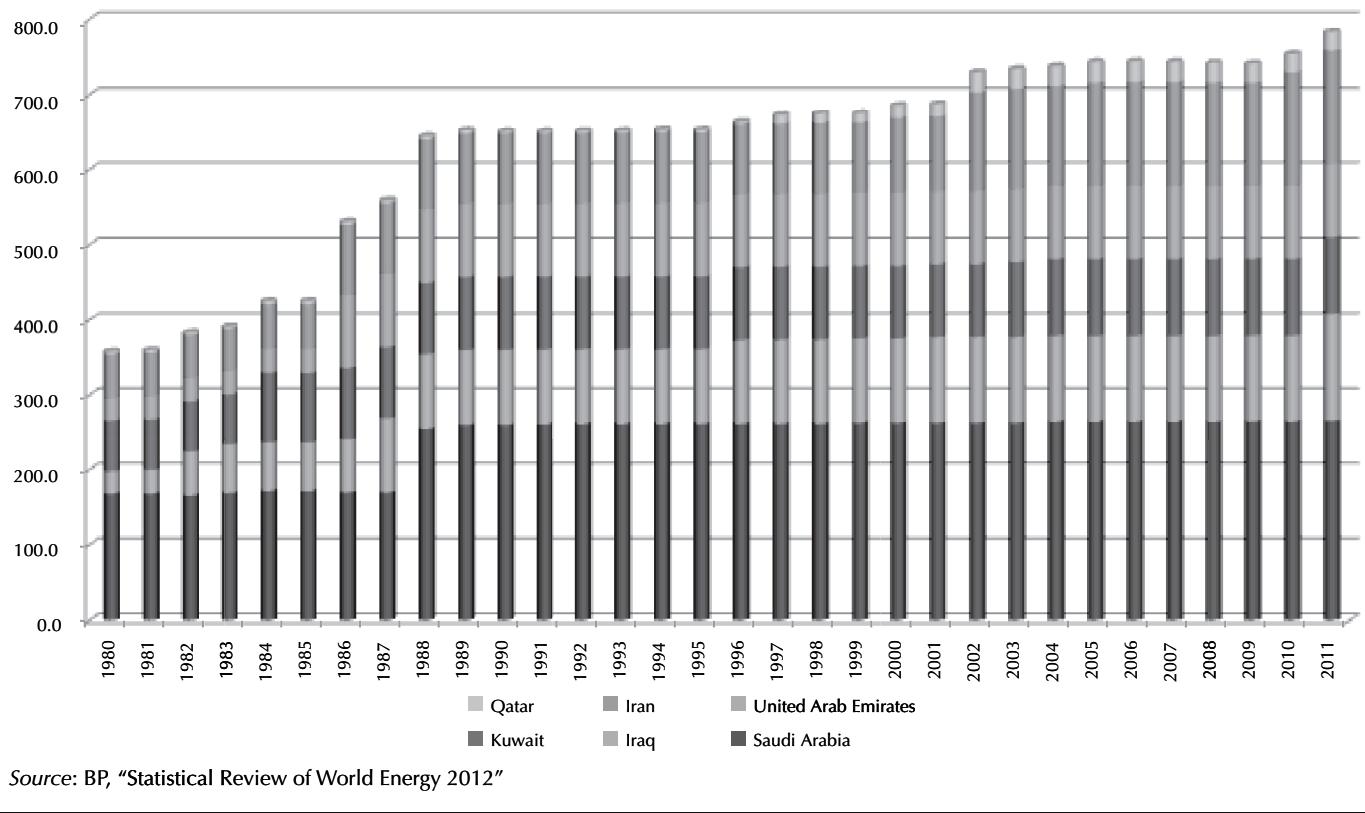
The industry in the US relies on two sets of guidelines for reserve recognition and reporting. Given the importance of the US economy, US energy companies and the fact that many international oil companies are listed on US stock exchanges, these rules have universal significance. One set of guidelines is the result of the efforts of Oil and Gas Reserves Committee of the Society of Petroleum Engineers and the World Petroleum Congress, which published independently in 1987 a set of definitions related to oil reserves. In the following decade, both organisations decided to join forces to develop a resource classification system. In March of 1997, a task-force produced "SPE/WPC Petroleum Reserves Definitions." The work on this project was resumed in 1999, with the additional participation of the American Association of Petroleum Geologists (AAPG). The updated reserve definitions were approved by these three organisations in February 2000.<sup>41</sup> The definitions were further revised in 2007, and constitute what is referred to as the Petroleum Resources Management System (PRMS).<sup>42</sup>

Figure 13.6 shows the breakdown of reserves between recoverable reserves, contingent resources and prospective resources, with a further breakdown based on the probabilistic estimates of the actual volumes in place, which underlies the PRMS.<sup>43</sup>

The PRMS distinguishes between reserves and contingent resources.<sup>44</sup>

- ❑ Reserves are those quantities of petroleum anticipated to be commercially recoverable by application of development projects to known accumulations from a given date forward under defined conditions. Reserves must further satisfy four criteria: they must be discovered, recoverable, commercial, and remaining (as of the evaluation date) based on the development project(s) applied. Reserves are further categorised in accordance with the level of certainty associated with the estimates and may be sub-classified based on project maturity and/or characterised by development and production status.

**Figure 13.5 Middle East OPEC oil reserves**



Source: BP, "Statistical Review of World Energy 2012"

**Figure 13.6** PRMS reserve definitions

		Production				
		Reserves				
Total petroleum initially-in-place (PIIP)	Discovered PIIP	Commercial	1P	2P	3P	
		Sub-commercial	Probable	Possible		
			Contingent resources			
			1C	2C	3C	
		Unrecoverable				
Undiscovered PIIP		Prospective resources				
		Low estimate	Best estimate	High estimate		
		Unrecoverable				

↔ Range of uncertainty ↔

↑ Increasing chance of commerciality

Source: <http://www.spe.org/spe-app/spe/industry/reserves/prms.htm>

- Contingent resources are those quantities of petroleum estimated, as of a given date, to be potentially recoverable from known accumulations, but the applied project(s) are not yet considered mature enough for commercial development due to one or more contingencies. Contingent resources may include, for example, projects for which there are currently no viable markets, or where commercial recovery is dependent on technology under development, or where evaluation of the accumulation is insufficient to clearly assess commerciality. Contingent resources are further categorised in accordance with the level of certainty associated with the estimates and may be sub classified based on project maturity and/or characterised by their economic status.

The reserves are classified as following.<sup>45</sup>

- Developed reserves are expected quantities to be recovered from existing wells and facilities.
- Developed producing reserves are expected to be recovered from completion intervals that are open and producing at the time of the estimate.

- Developed non-producing reserves include shut-in and behind-pipe reserves.
- Undeveloped reserves are quantities expected to be recovered through future investments.

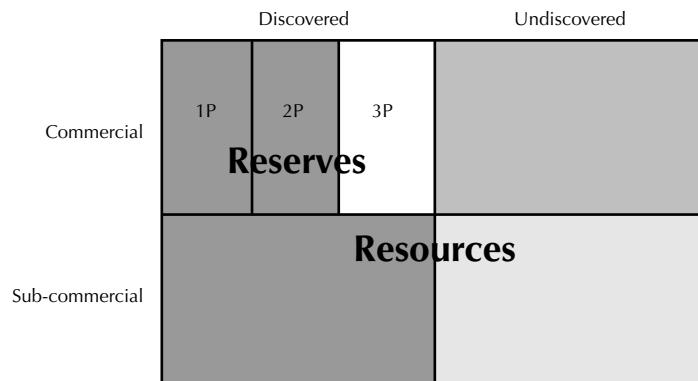
The UN proposed a system based on a different set of definitions, which distinguishes between economic viability (E), field project status (F) and geological assessment (G), with further sub-classes for each.<sup>46</sup>

Most reserve classification systems allow the use of principles, including probabilistic approaches, developed originally by Vincent Ellis McKelvey (illustrated in Figure 13.7).<sup>47</sup> This diagram helps to avoid a frequent semantic confusion between reserves and resources. The popular press often fails to distinguish between resources and reserves, and this leads to excessively optimistic projections and misleading headlines.

In this diagram, reserves represent a fraction (upper left quadrant) of the total existing resources,<sup>48</sup> which may be discovered or undiscovered at a time. The reserves are classified as

- proved (1P);
- proved and probable (2P); and
- proved, probable and possible (3P).

**Figure 13.7** McKelvey Box



Source: UKERC Review of Evidence on Global Oil Depletion, July 2009

These terms are used extensively across the industry, but the interpretations, as noted above, differ from source to source and country to country. One possible approach to make the interpretations of these definitions less arbitrary is to use a probabilistic approach that frequently associates the 1P reserves with a 90% likelihood of being exceeded (P90), the 2P reserves with 50% likelihood (the acronym P50 is used) and the 3P reserves with 10% likelihood (P10). This corresponds, respectively, to the 10th, 50th and 90th percentile of the probability distribution.<sup>49</sup> Some authors use the acronyms P1, P2 and P3, and we shall use these intermittently.

One frequently overlooked issue related to the probabilistic definitions of reserves is aggregation of the reserves. Most companies and analysts add 1P or 2P reserve estimates for different producing properties. This approach is not justified, as one can see from the following simplified example. Suppose that two properties have reserves that can be described by the same normal probability distributions, with the mean of 1,000,000 barrels and the standard deviation (SD) of 600,000 barrels. The P1 reserve level is given by  $1,000,000 - 1.282 \times 300,000 = 230,800$ .<sup>50</sup> The addition of P1 reserves for both properties produces a P1 reserve of 461,600 barrels. Treating both properties as a portfolio requires estimating the parameters of a probability distribution applicable to this case. The mean of a new (portfolio) distribution (which is still normal) is equal to the sum of the means of the individual probability distributions – ie, 2,000,000 barrels. The new SD is calculated by adding the *variances* and taking the square root of the sum.<sup>51</sup> The SD is ~848,528 barrels, and the new P1 reserve is 912,187 barrels, a much higher number than 461,600.<sup>52</sup> The addition of P90 reserves is a very conservative procedure, and every analyst should be aware of this.

The difficulty of comparing estimates from different sources is compounded by other factors, in addition to the varying and inconsistent definitions of reserves. Most publically available sources use competing definitions for oil. Christophe McGlade compiled a table containing a comparison of the most widely used groupings (see Table 13.4).

The source article for Table 13.4 is recommended reading for everybody. It contains a number of additional clarifications and comments that we have to omit in the interest of space. Some of the definitions used above will become clear after reading the section on non-conventional oil.

**Table 13.4** Crude oil definitions: comparisons

Reporting source	Source groupings in reserve data
OGJ	Conventional crude oil, extra heavy oil, condensate, shale oil and oil sands
World Oil	Conventional crude oil, extra heavy oil, condensate, shale oil and oil sands, oil sands from Canada
BP	Conventional crude oil, extra heavy oil, condensate, shale oil, oil sands and NGL, oil sands from Canada
OPEC	Conventional crude oil, extra heavy oil, condensate, shale oil and oil sands, extra-heavy oil from Venezuela
BGR	Conventional crude oil, condensate and NGL, extra-heavy oil, shale oil and oil sands
WEC	Conventional crude oil, condensate and NGL (where it cannot be separated), NGL, extra-heavy oil, oil sands, shale oil
IEA	Conventional crude oil, condensate and NGL, extra heavy oil, oil sands
EIA	Different sources are used
Energy Watch Group	Conventional crude oil, condensate and NGL
Campbell	Crude oil > 17.5° API excluding shale oil, CTL, deep water oil, polar oil and NGL from gas plants, crude oil < 15° API, shale oil, CTL, deep water oil, polar oil, and NGL from gas plants, USGS
USGS	Conventional crude oil > 17.5° API, NGL
IHS	Conventional crude oil > 17.5° API, NGL
	Conventional crude oil, extra heavy oil, NGL, condensate, and oil sands

Source: Christophe McGlade, "Uncertainties in estimating remaining recoverable resources of conventional oil" ([http://www.iaee.org/en/students/best\\_papers/mcglaide.pdf](http://www.iaee.org/en/students/best_papers/mcglaide.pdf))

Another important issue that are discussed at length in the cited sources is reserve growth. Reserve growth is defined as the changes in reserves unrelated to new discoveries. This may happen for a number of reasons, including better seismic data that become available over time, changes in market prices (which may both increase or decrease reserves) and technological progress (enhanced oil recovery, secondary oil recovery) that increases the recovery factor.<sup>53</sup> Reserve growth is reported sometimes in the year when the adjustment is made. Sometimes, the adjustment is pushed back to the discovery year. The underlying logic here is that the hydrocarbons were in the ground on day one and were just not recognised at the time.

### SEC rules

In the US, the classification of reserves is based on SEC regulations and applies to publicly owned companies. Given that US-domiciled companies and companies with securities traded in the US financial markets have to comply with SEC rules (as well as some non-US companies that follow these rules on a voluntary basis for reserve reporting but without audit), the SEC-supported classification system has a significant practical importance, covering about 25% of the world crude oil output.

The SEC revised its reserves related rules, with the changes becoming effective on or after January 1, 2010.<sup>54</sup> Many changes were designed to align SEC definitions with the standards endorsed in the PRMS. The most important changes to reporting requirements were:

- ❑ the use of 12-month average prices (unless prices are based on the contractual provisions) to determine if the resources can be considered as commercial; this approach replaces the use of a single year-end spot price;
- ❑ definitions of oil and gas producing activities were extended to include new sources of hydrocarbons, such as oil sands, oil and gas produced from coal and shales; the critical criterion is the final product and not, as previously, modes of extraction;<sup>55</sup>
- ❑ “reasonable certainty” used in the SEC rules (with respect to the volume of resources) is defined as a “high degree of confidence,”<sup>56</sup> following the PRMS definitions (both probabilistic and deterministic methods can be used);
- ❑ reporting entities have the option of reporting “probable” and “possible” reserves;
- ❑ the definition of “proved reserves” was broadened to allow the use of new advanced technologies in oil and gas exploration;<sup>57</sup>
- ❑ the qualifications of the person responsible for the reserve estimation process must be disclosed; and
- ❑ a tabular format is required for reserve estimation disclosures (see Table 13.5).

### Reserves data

Information about oil reserves is available from the BP “Statistical Review of World Energy.” The most recent issue (June 2012) reports total world reserves of oil as of the end of 2011 at 1652.6 thousand

million barrels. The data is compiled from a variety of sources, including primary official sources, third-party data from the OPEC Secretariat, *World Oil, Oil & Gas Journal* and independent estimates for Russia and China. Reserves include crude oil, gas condensate and NGLs, and are defined as “those quantities that geological and engineering information indicates with reasonable certainty can be recovered in the future from known reservoirs under existing economic and operating conditions.” Table 13.6 contains estimates of reserves for a group of countries accounting for about 95% of the world reserves. A few comments are required. Venezuela has increased its reported reserves from 99.4 to 296.5 thousand million barrels through inclusion of the Orinoco Belt resources (20 thousand million barrels in 2007, 220 in 2011). Canadian oil sands reserves are

**Table 13.5** Summary of oil and gas reserves as of fiscal-year end based on average fiscal-year prices

	Reserves				
	Oil (mbbls)	Natural Gas (mmcf)	Synthetic Oil (mmbls)	Synthetic Gas (mmcf)	Product A (measure)
Reserved category					
<b>PROVED</b>					
<b>Developed</b>					
Continent A					
Continent B					
Country A					
Country B					
Other Countries in Continent B					
<b>Undeveloped</b>					
Continent A					
Continent B					
Country A					
Country B					
Other Countries in Continent B					
<b>TOTAL PROVED</b>					
<b>PROBABLE</b>					
Developed					
Undeveloped					
<b>POSSIBLE</b>					
Developed					
Undeveloped					

Source: SEC

**Table 13.6** Oil: proved reserves (thousand million barrels)

<b>Country</b>	<b>2011</b>
Venezuela	296.5
Saudi Arabia	265.4
Canada	175.2
Iran	151.2
Iraq	143.1
Kuwait	101.5
United Arab Emirates	97.8
Russian Federation	88.2
Libya	47.1
Nigeria	37.2
US	30.9
Kazakhstan	30.0
Qatar	24.7
Brazil	15.1
China	14.7
Angola	13.5
Algeria	12.2
Mexico	11.4
Azerbaijan	7.0
Norway	6.9
Sudan	6.7
<b>Total</b>	<b>1,576.2</b>

Source: BP, "Statistical Review of World Energy 2012"

estimated at 169.2 million thousand barrels, with 25.9 under active development, according to official estimates.

## CONCLUSIONS

Crude oil, unlike natural gas,<sup>58</sup> is a highly heterogeneous commodity, a mixture of many different hydrocarbons and impurities. Its physical and technical properties vary depending on origin, and this in turn impacts the way it is transported, stored and processed. Understanding the differences between different grades of crude is of critical importance, not only in trading crude, but also in hedging and optimising refinery operations. Widely followed statistics of oil output at global and national levels often obscure the fact that the reported totals increasingly refer to flows of varying composition and characteristics. In the case of conventional oil, the proportion of heavy and sour crude is increasing, and this requires (as we shall discuss in the chapter on refining) an overhaul of the

existing – and construction of new – refineries designed for processing crude of lower quality. In the case of liquids, we can see a growing share of NGLs, which are not a perfect substitute for conventional oil.

- 1 "Out of the ground, crude oil can be as thin and light-colored as apple cider or as thick and black as melted tar" (see [http://www.chevron.com/products/prodserv/fuels/documents/Diesel\\_Fuel\\_Tech\\_Review.pdf](http://www.chevron.com/products/prodserv/fuels/documents/Diesel_Fuel_Tech_Review.pdf)).
- 2 Donald L. Burdick and William L. Leffler, 2001, *Petrochemicals in Non-Technical Language* (Tulsa, OK: PennWell Books). Other useful books about energy chemical properties of oil and refining include Xavier Normand, 1978, *L'industrie du Raffinage du Petrol* (Paris, France: Editions Technip), and Surinder Parkash, 2003, *Refining Processes Handbook* (Amsterdam, the Netherlands: Elsevier).
- 3 Coming from fat (from the Greek *aleiphar*, fat or oil).
- 4 See footnote 2.
- 5 Isomers are compounds with the same molecular formula, the same numbers of different atoms, but a different molecular structure – ie, the way in which different atoms are linked together – and, by extension, different properties.
- 6 Many compounds in this group have a distinct odour (hence the name).
- 7 The terms *vacuum residue* and *viscosity* will be explained in the chapter on the processing of crude oil.
- 8 The formula uses the number of 141.5 because, in the beginning of the 20th century, US instruments used to measure specific density (hydrometres) were incorrectly calibrated to the modulus of 141.5, instead of 140. When this was discovered, the US Bureau of Standards had no choice but to bless the existing practice.
- 9 To be precise, the density of water is 0.9990 grams per cubic centimetre ( $\text{g}/\text{cm}^3$ ) at  $60^\circ \text{F}$  ( $15.6^\circ \text{C}$ ).
- 10 William L. Leffler, 1985, *Petroleum Refining* (Tulsa, OK: PennWell Books). Also consult the Canadian website devoted to energy information: <http://www.centreenergy.com>.
- 11 In the early days of the oil industry, crude, as an input to the production of kerosene, used in lamps, was competing against whale oil. The high content of sulphur was resulting in unpleasant odour and smoke emitted by the lamps. The oilmen used to put a drop of oil on their tongues to determine the quality of crude (high sulphur level was associated with a sour taste).
- 12 This is a rule of thumb. There is no good, generally accepted method of measuring the corrosivity of crude oil.
- 13 The flash point is tested using an open cup or closed cup method, such as Cleveland Open Cup, and Pensky-Martens or Setaflash.
- 14 The pour point is very important for the oils used in engines or crude oil transported over pipelines under conditions of low ambient temperatures.
- 15 One can learn more about assays from "Strategic petroleum reserve crude oil assay manual" (<http://www.spr.doe.gov/reports/docs/CrudeOilAssayManual.pdf>).
- 16 <http://www.statoil.com/en/OurOperations/TradingProducts/CrudeOil/Crudeoilassays/Pages/EkofiskBlend.aspx>.
- 17 2,033,082 barrels per day in 1988, 619,655 in 2010 (<http://www.alyeska-pipe.com/Pipelinefacts/Throughput.html>).
- 18 Russell Gold, 2011, "Shrinking oil supplies put Alaskan pipeline at risk," *Wall Street Journal*, May 11 (<http://online.wsj.com/article/SB10001424052748704570704576274682119735102.html#articleTabs%3Darticle>).
- 19 As reported by the *Wall Street Journal*, "The problems facing the pipeline were made very clear in January [2011], when a leak on the North Slope forced two back-to-back winter shut-

downs for a total of 148 hours. Temperatures inside the pipeline dropped by almost two degrees a day. Much longer, says E. G. "Betsy" Haines, Alyeska's oil movement director, and wax in the crude would have begun congealing, potentially turning TAP into the world's largest tube of Chap Stick."

- 20 The TAP is owned by BP Pipelines (Alaska) (46.93%), ConocoPhillips Transportation Alaska (28.29%), ExxonMobil Pipeline Company (20.34%), Unocal Pipeline Company (1.36%) and Koch Alaska Pipeline Company (3.08%).
- 21 The most recent BP statistical review uses the conversion factor of 159 litres.
- 22 <http://www.eia.gov/iea/convheat.html>.
- 23 BP, "Statistical Review, 2010."
- 24 A note from EIA: "The Energy Information Administration is discontinuing the publication of the *International Petroleum Monthly* in its current format. The December 2010 issue will be the last one to be available as a separate publication. Beginning with the January 2011 issue, you may access the data in this report, as well as energy data for all fuels for over 200 countries, on our International Energy Statistics website."
- 25 EIA, *Monthly Energy Review* statistical tables.
- 26 Syncrude is explained in the chapter on non-conventional oil (Canadian oil sands section).
- 27 Available from the BP website at <http://www.bp.com/sectionbodycopy.do?categoryId=7500&contentId=7068481>.
- 28 <http://www.eia.doe.gov/ipm/appe.html>.
- 29 Renamed JODI at the 8th International Energy Forum in Osaka in 2002.
- 30 <http://www.jodidata.org/Qnr.shtm>.
- 31 The current state of the database has been assessed succinctly by Eithne Treanor in the article "Return of the Jodi?" (*Energy Risk*, July 2008): "Jodi supporters claim it has already done much to reduce the opacity of worldwide oil data and has prepared the ground for better data collection and analysis. If some key concerns can be addressed, Jodi's organisers say it could yet fulfill its lofty ambition to be the central hub for worldwide oil data. [...] However, there is a long way to go. The consensus across Jodi's user base is that the quality of data needs to be vastly improved and that the project needs greater visibility."
- 32 IEA identifies the following categories of unconventional oil: Canadian oil sands, Venezuelan extra-heavy, oil shales, coal-to-liquids, gas-to-liquids, other (for example, blending additives such as butyl ether (MTBE), ethyl tertiary butyl ether (ETBE) or tertiary amyl methyl ether).
- 33 Conventional oil production numbers are based on the EIA data (crude + condensates).
- 34 See, for example, <http://www.theoildrum.com/node/7349#more>.
- 35 See, for example, Terry Macalister, 2009, "Key oil figures were distorted by US pressure, says whistleblower," *Guardian*, Monday, November 9. "The world is much closer to running out of oil than official estimates admit, according to a whistleblower at the International Energy Agency who claims it has been deliberately underplaying a looming shortage for fear of triggering panic buying. The senior official claims the US has played an influential role in encouraging the watchdog to underplay the rate of decline from existing oil fields while overplaying the chances of finding new reserves."
- 36 See <http://www.crudeoilpeak.com/?p=1876>.
- 37 One of the growing concerns with respect to the big integrated international oil companies is their inability to replace reserves, given that most oil resources are controlled by national oil companies restricting access to foreign-based entities. The push to develop deep ocean resources is dictated by necessity, not by the attractiveness of these opportunities.
- 38 An old adage in the energy industry is exploring for Btus or barrels of reserves on Wall Street.
- 39 The link between production quotas and stated reserves (as well as the element of national pride) lead some analysts to coin the term "political reserves," the hydrocarbons existing in the minds of political leaders. "Fossil fuels public data are unreliable, in particular on

- reserves, confused with resources. Even production data are flawed, because publishing data is a political act and depends upon the motive of the author." (see Jean Laherrere, 2007, "Oil and gas: what future?" <http://www.oilcrisis.com/laherrere/groningen.pdf>).
- 40 The main source for this section is the report by the technology and policy assessment function of the UK Energy Research Centre, 2009, "Global oil depletion: An assessment of the evidence for a near-term peak in global oil production," August. A very good summary of the issues related to reserve estimation is available in the paper by Christophe McGlade, 2010, "Uncertainties in estimating remaining recoverable resources of conventional oil" ([http://www.iaee.org/en/students/best\\_papers/mcGlade.pdf](http://www.iaee.org/en/students/best_papers/mcGlade.pdf)).
- 41 See Society of Petroleum Engineers, 2001, "Guidelines for the evaluation of petroleum reserves and resources: A supplement to the SPE/WPC petroleum reserves definitions and the SPE/WPC/AAPG petroleum resources definitions."
- 42 "Petroleum Resources Management System," Prepared by the Oil and Gas Reserves Committee of the Society of Petroleum Engineers (SPE); reviewed and jointly sponsored by the World Petroleum Council (WPC), the American Association of Petroleum Geologists (AAPG) and the Society of Petroleum Evaluation Engineers (SPEE). 2007.
- 43 See <http://www.spe.org/spe-app/spe/industry/reserves/prms.htm>. According to the information available on the website of the SPE "A new Petroleum Resources Management System was approved by the Society of Petroleum Engineers (SPE) Board of Directors in March 2007, culminating two years of intense collaboration by SPE, the World Petroleum Council (WPC), the American Association of Petroleum Geologists (AAPG), and the Society of Petroleum Evaluation Engineers (SPEE). The system also was approved by the boards of the other societies following a significant industry review and comment period."
- 44 Ibid.
- 45 Ibid.
- 46 PowerPoint presentation, "Activities of the ad hoc group of experts and the UNFC. United Nations framework classification system for fossil energy and mineral reserves/resources" (see [http://www.spee.org/images/PDFs/2009Convention/4\\_Smith\\_PPT2009AnnualMeetingPresentation.ppt](http://www.spee.org/images/PDFs/2009Convention/4_Smith_PPT2009AnnualMeetingPresentation.ppt), accessed on November 27, 2009).
- 47 V. E. McKelvey, 1972, "Mineral resource estimates and public policy," *American Scientist*, pp 32–40.
- 48 "The sum of cumulative production and reserves is commonly referred to as *cumulative discoveries*." "*Ultimately recoverable resources* (URR) represent the amount of oil that is estimated to be economically extractable from a field or region over all time – from when production begins to when it finally ends. [...] For individual fields, the URR represents the sum of cumulative discoveries and estimates of future reserve growth." Steve Sorrell, Jamie Speirs, Roger Bentley, Adam Brandt, Richard Miller, "Global Oil Depletion: An assessment of the evidence for a near-term peak in global oil production," A report produced by the Technology and Policy Assessment function of the UK Energy Research Centre, August 2009, <http://www.decc.gov.uk/assets/decc/what%20we%20do/global%20climate%20change%20and%20energy/international%20energy/energy%20security/1791-uk-erc-report-global-oil-depletion.pdf>.
- 49 This mapping is by no means a generally followed convention. There are indications that the US and Canadian E&P companies use a more cavalier approach to reserve classification, and 1P reserves are closer to P60–P65. See H. Jung, 1997, "Reserves definitions – clarifying the uncertainties," *Journal of Canadian Petroleum Technology*, (36)4, pp 26–30; Jean Laherrère, 1999, "Reserve growth: Technological progress, or bad reporting and bad arithmetic?" *Geopolitics of Energy*, pp 7–16.
- 50 The number used in the calculations, -1.282, corresponds to the 10% probability level in the standardised normal distribution.
- 51 This is an illustration of a well-known fact that the variances are additive but standard deviations are not. We assume also that the two distributions are independent (in the probabilistic sense) and, therefore, the correlation coefficient applicable in this case is zero.

- 52 See Richard Pike, 2006, "Have we underestimated the environmental challenge," *Petroleum Review*, June. According to the author, "Merely adding the individual P90 figures arithmetically underestimates progressively the true P90 of the combined distributions. What is actually being derived is the P99.99... at the global scale."
- 53 The estimates of the global recovery factors cluster around 30%, although they may be much higher in the case of individual oil fields.
- 54 The rules are available at <http://www.sec.gov/rules/final/2008/33-8995.pdf>.
- 55 Oil sands may be exploited through a mining process, which does not resemble conventional ways of producing oil.
- 56 This is also explained as "more likely to be achieved than not."
- 57 Previously fluid contact information obtained in a well penetration was used.
- 58 Even in the case of natural gas, one has to deal in some cases with quality differences (for example, H and L gas in continental Europe), but these complications pale in comparison with crude.

## *Non-Conventional Oil*

The concept of non-conventional oil is not something that is cast in stone and immutable: it is likely to evolve over time, as is the concept of non-conventional natural gas.<sup>1</sup> What was previously considered cutting-edge technology, becomes the standard technology of today. This term is used often without offering a precise definition and, like beauty, is in the eye of the beholder. For the purpose of this chapter, we shall adopt the definition of non-conventional oil corresponding to the classification used by the IEA in “World Energy Outlook” (2010).<sup>2</sup> The following resources are included:

- ❑ Canadian oil sands;
- ❑ Venezuelan extra heavy oil (Orinoco Belt deposits);
- ❑ oil shales;
- ❑ coal-to-liquids;
- ❑ gas-to-liquids; and
- ❑ refinery additives and gasoline blending additives (MTBE).

We discuss also oil from shale rock formations (which is different from oil shales).

Additionally, we include biofuels, following the conventions applied by the EIA in its “Oil Market Report”. We shall cover in detail: oil sands, heavy oil and oil from the shale formations, as well as biofuels. Coal-to-liquids and gas-to-liquids technologies, and conversion of coal or natural gas to liquids, are unlikely to affect energy trading in the near future, and will not be discussed here in detail.

This topic is very important to energy traders and analysts for a number of reasons. Supplies of liquid fuels will be increasingly important for satisfying global energy needs, primarily as a result of economic growth in the emerging economies and hundreds of

millions of people aspiring to reach the lifestyle typical of middle classes in developed countries. Most forecasts point to high decline rates from conventional energy sources, and to a growing dependence on developing existing fields and new discoveries. A significant portion of future supplies will come from the resources classified today as non-conventional. Even more importantly, the prices of crude, like the prices of other commodities, are set at the margin. Economic rationality dictates a time sequence of exploitation of natural resources, starting with the deposits that are most accessible and cheapest to extract. Of course, geography, history and chance may intervene, but the general tendency is difficult to deny. At any point in time, the most expensive source of supply required to satisfy demand determines the floor price to which the market prices will gravitate, subject to constant distortions caused by speculation, perceptions of geopolitical risks and market imperfections.

Fluctuations in demand and technological breakthroughs will constantly shift the marginal price, but a trader should have a good understanding of the region of the supply curve that sets the price. The section of the supply curve setting the market price will be likely associated with some of the technologies discussed in this chapter. Finally, the US is currently in the early stages of the same technological revolution that redefined the natural gas industry. Application of horizontal drilling and hydraulic fracturing technology is behind a significant increase of crude oil and condensate production in some locations in the US, with more potential for additional output gains.

The chapter will end with a review of the very controversial peak oil theory. Non-conventional oil is very important in this context: the opponents of the peak oil theory see oil from shale formations as the ultimate repudiation of alarmists predicting the end of our civilisation when the oil runs out.

### CANADIAN TAR SANDS<sup>3</sup>

The term “Canadian oil sands” corresponds to heavy oil and bitumen (also called tar<sup>4</sup> or asphalt) saturating sand or carbonate rock formations. Bitumen is defined as:<sup>5</sup>

“[C]rude oil with a dynamic viscosity at reservoir conditions of more than 10,000 centipoise.<sup>6</sup> Canadian “bitumen” supply is more loosely accepted as production from the Athabasca, Wabasca, Peace River and Cold Lake oil-sands deposits. The majority of the oil produced

from these deposits has an API gravity of between 8° and 12° and a reservoir viscosity of over 10,000 centipoise although small volumes have higher API gravities and lower viscosities.”

The Province of Alberta Energy Resource Conservation Board estimates that the crude oil reserves in the tar sands are equal to 27m<sup>3</sup> billion (170 billion barrels). These estimates are based on the remaining volumes of 270m<sup>3</sup> billion and a very conservative recovery rate of 10%. If technological progress is able to increase the recovery rate to 25% (a fairly reasonable number), the crude oil potential of Alberta will exceed that of Saudi Arabia (42m<sup>3</sup> billion).

Chemically bitumen is composed primarily of highly condensed polycyclic aromatic hydrocarbons.<sup>7</sup> In terms of its history, bitumen is oil that has been degraded by bacteria as it migrated to the surface. Naturally occurring bitumen has to be distinguished from what is sometimes called refining bitumen, which is the lowest fraction obtained from distillation of crude – and which represents what the industry calls the bottom of the barrel. Two principal techniques for industrial exploitation of tar sands are open surface mining or in situ production. Strip (surface) mining applies to deposits which are no deeper than 75 metres and is carried out through the removal of oil sands which are then treated with hot water to separate the bitumen. The scale and the invasive nature of this process is difficult to imagine unless one has visited the production sites or has seen the pictures of open pits invoking the descriptions of the Saruman’s mines in the Tolkien’s books. As described by Robert Kunzig:

“The dump trucks that rumble around the mine, hauling 400-ton loads from the shovels to a rock crusher, burn 50 gallons of diesel fuel an hour; [...] And every day in the Athabasca Valley, more than a million tons of sand emerges from such crushers and is mixed with more than 200,000 tons of water that must be heated, typically to 175°F, to wash out the gluey bitumen.”<sup>8</sup>

In situ production is based on steam injections into the deposit in order to reduce the viscosity of oil and make it flow to the surface. Two principal techniques involve cyclic steam simulation (CSS) and steam-assisted gravity drainage (SAGD).<sup>9</sup> In the case of CSS, the same well is used for steam injections and oil removal, cycling between these two functions. In the case of SAGD, two horizontal wells are drilled in the same vertical plane. Steam<sup>10</sup> is injected from the well at the lower depth and oil released through heating seeps

into the deeper well from which it is removed to the surface. Oil from the tar sands requires further processing before it can be used as a refinery feedstock. The dilution process consists in mixing oil with lighter hydrocarbons (for example, natural gas liquids) that results in a mixture called dilbit (diluted bitumen). Diluted bitumen can be transported by pipeline to a refinery, which has to be configured to accept this specific type of crude. Upgrading is a process which can be described as highly specialised refining, producing syncrude (synthetic crude), which is chemically close to conventional heavy crudes. Upgrading happens by removing carbon (coking), or adding hydrogen (hydrocracking).<sup>11</sup> Syncrude can be used as diluent, in which case the product is called syndilbit.

The production of oil from the tar sands offers many technological challenges and environmental risks related to:<sup>12</sup>

- energy input;
- CO<sub>2</sub> emissions and pollution concerns;
- water usage; and
- land usage.

Production of oil from the tar sands is energy-intensive. The efficiency of this process is measured using the concept of energy return on investment (EROI), the ratio of energy obtained from a given process to energy used in all the stages of production:

$$\begin{aligned} EROI &= \frac{\text{Energy}_{\text{output}}}{\text{Energy}_{\text{input}}} \\ &= \frac{\text{Energy}_{\text{output}}}{\text{Energy}_{\text{direct}} + \text{Energy}_{\text{indirect}} + \text{Energy}_{\text{labor}} + \text{Energy}_{\text{environment}}} \end{aligned}$$

In the formula above, indirect energy input is related, for example, to the production of capital equipment required in producing a given form of energy. Environmental input is related to the legal requirements to mitigate the environmental impact of the production processes.

In the case of surface mining, energy input is related, for example, to removal of the top layer of soil, excavating source rock, transporting it to the separation and upgrading facilities and then to a refinery, processing in a refinery and making delivery to the final user. In the case of the in situ process, the cost will include fuel

required to produce steam. The environmental costs include land reclamation and the externalities related to emissions of greenhouse gases. The estimates of EROI vary from source to source, and tend to be very sensitive to the assumptions and definition of the boundary of the system. The summary results can be found in an article by Charles Hall of the SUNY College of Environmental Science and Forestry<sup>13</sup>

The EROI of tar sands oil is lower than that of conventional oil. One has to recognise that the EROI of conventional oil is not constant and changes over time due to technological progress (EROI increases) and because the industry migrates to more challenging and less-accessible fields (EROI decreases). For example, in the 1930, the EROI of oil produced in the US was over 100; in 2005, the EROI of oil imported to the US was about 18.<sup>14</sup>

CO<sub>2</sub> emissions are related to large volumes of fuels (primarily natural gas) required to produce steam, and emissions during the technological processes related to upgrading, refining and transportsations of syncrude to and from processing plants. The emissions range from 9.2–26.5 grams of CO<sub>2</sub> per mega joule of syncrude for the surface mining process, and 16.2–28.7 for the in situ process. This compares with 4.5–9.6 grams per mega joule for conventional oil production. The “well to wheels” emissions (or the emissions generated through the entire supply chain) are 260–320, 320–350, 270–240 for surface mining, in situ with and without upgrading, respectively, against 250–280 for conventional oil production (the units are, as above, grams of CO<sub>2</sub> per mega joule of energy obtained from oil). Dilbit and syncrude contain relatively high levels of sulphur and heavy metals compared with conventional crude, which creates safety concerns among environmental groups.<sup>15</sup>

Heavy water usage is related to the processes of separation of bitumen from the rock extracted in the surface mining operation. According to the environmental impact study cited earlier, it takes 11 tons of oil sands to produce one cubic metre (6.3 barrels) of syncrude. The water usage (net of recycling) is estimated to range between two and three barrels of water per barrel of oil. In situ production technology requires water for production of steam resulting in net usage (after recycling) of one barrel of water per barrel of oil. Without recycling, the water consumption may reach eight barrels per barrel of oil. Additional water is consumed also in the upgrading process (two

to three barrels of water per barrel of oil). Water is coming primarily from the local rivers, an activity stressing a very fragile ecological system. These concerns can be addressed in the future through more effective recycling of water and extracting water from saline aquifers.

The volume of water used in the production process represents only one potential environmental threat. The process of separation of bitumen from the solids leaves about  $3.3\text{m}^3$  of fluid/solid waste (called raw tailings) per one cubic metre of oil. Tailings are stored in the retention ponds, where fine solids form over time what is called mature fine tailings (MFT), which represent a major problem as they settle very slowly. To illustrate the dimensions of the problem, in a typical plant producing  $47,700\text{m}^3$  per day ( $\sim 300,000$  barrels/day) of oil,  $250,000\text{m}^3$  per day (540,000 tons per day) of sands are excavated, producing  $1,000,000\text{m}^3$  of tailings. As of 2009, the tailing ponds (oil sands process wastewater, OSPW, and MFT) contained  $720,000,000\text{m}^3$  and cover an area of 130 square kilometres. The bursting of a retention wall would have a devastating impact on the environment of Alberta, especially on her rivers.

### **ORINOCO BELT HEAVY OIL**

Orinoco Belt Heavy Oils are deposits of oil with an API gravity under  $10^\circ$ . The deposits are located between 500 and 1,000 metres<sup>16</sup> and, as the reservoir temperatures are higher than in the case of Canadian oil sands, viscosity is lower and the extraction of oil through conventional techniques is somewhat easier. The primary recovery rates for vertical and horizontal wells are, respectively, five and 10–15%. Higher recovery rates can be achieved through production techniques such as CSS or SAGD. According to a USGS assessment of technically recoverable heavy oil and associated gas resources:<sup>17</sup>

“The mean of the distribution of heavy oil resources is about 513 BBO [billion barrels of oil], with a range from 380 to about 652 BBO. The mean estimate of associated dissolved-gas resource is 135 trillion cubic feet of gas (TCFG), with a range from 53 to 262 TCFG.”

No assessment of economically recoverable resources was made, but most experts agree that a significant fraction of the resources can be profitably extracted at current prices. Expansion of output from the Orinoco fields will encounter similar technological and environmental difficulties as for Canadian oil sands. In addition, the

somewhat unstable economic and energy policies of the current Venezuelan government make it difficult to attract the foreign capital and know-how required for addressing challenges faced by the national oil company in addressing the formidable technological problems related to exploitation of the Orinoco belt oil.<sup>18</sup>

### OIL SHALES AND SHALE OIL

Oil shales have to be distinguished from oil extracted from shale rock formations. Oil from shale formations (which we shall call shale oil) is produced through application of the same techniques that brought the shale natural gas revolution to the US: horizontal drilling and hydraulic fracturing. It is recommended to use consistently the term “oil produced from shale formations” to avoid confusion with oil shales. It is a bit confusing, but this is the jargon used in the industry. We are in the very early stages of a technological revolution that could expand the output of oil in the US by 3–4 million barrels a day by 2020. Seven shale formations are being actively explored for oil in the US: Avalon, Bakken, Barnett, Eagle Ford, Monterey, Mowry and Niobrara. The information about technically recoverable US shale oil resources is shown in Table 14.1.

Oil shales are much less promising at this point. It is possible to characterise oil shales as oil that Mother Nature did not have sufficient time to cook. For humans to finish this project requires a very significant expenditure of energy.

The term oil shales applies to sedimentary rocks (clays, marls or carbonates) with a high content of kerogen. The critical difference

**Table 14.1** US shale oil resources

Play	Technically Recoverable Resource	Area (sq. miles)	Average EUR
	Oil (BBO)	Leased	Oil (MBO/well)
Eagle Ford	3.35	3,323	300
Avalon & Bone Springs	1.58	1,313	300
Bakken	3.59	6,522	550
Monterey/Santos	15.42	1,752	550

Source: U.S. Energy Information Administration

between tar sands and oil shale is that, in the latter case, kerogen bonds more closely to the surrounding substrate.<sup>19</sup>

The technologies for extracting oil from shale are similar to methods used in the exploitation of tar sands. The in situ process relies on injecting heat source into the rock formation, releasing oil, which is pumped to the surface. A mining approach (surface or underground) requires crushing extracted ore and effectively cooking oil out of it in a retort. Shale oil produced through both processes is then upgraded. Shale oil can still not be economically produced, but any energy professional should occasionally monitor developments in this area: one never knows when a sudden technological breakthrough will happen.

### **Shale oils**

As mentioned above, shale oil (ie, oil from the shale rock formations) is produced using the same technologies that are behind expansion of production of shale gas: horizontal drilling and hydraulic fracturing. The epicentre of US developments related to shale oil is the Williston Basin, covering the North and South Dakotas and Montana, which include such formations as Bakken Shale, Three Forks, Tyler and Spearfish. A very rapid development of Bakken explains oil production growth in North Dakota: a jump from 110,000 barrels a day in 2006 to about 530,000 barrels of oil equivalent at the end of 2011. Other locations important for the production of shale include oil include Eagle Ford (Texas), Niobrara/Codell (Colorado, Kansas, Nebraska, Wyoming) and the Permian Basin.

Bakken was assessed in the late 1990s by Leigh Price and the US Geological Survey (USGS), who estimated that OIIP amounted to 413 billion barrels (an average). Another assessment (by Continental Resources) put the combined Bakken/Three Forks resources at 900 billion barrels OIIP. The estimate of recoverable oil by the same company was 20 billion barrels for Bakken, with an equivalent target for Three Forks. In January 2011, the North Dakota Industrial Commission estimated recoverable reserves in Bakken/Three Forks at 11 billion barrels. Leonardo Maugeri in his widely read paper (from which the information about Bakken above is derived) projects total Bakken/Three Forks production at over three million barrels a day by 2020.<sup>20</sup> His forecast is based on a number of assumptions, including a price of oil (WTI) equal to or greater than US\$70 per

barrel through 2020, 200 drilling rigs in operations per week, an ultimate recovery rate of 10% per well and, most importantly, “a combined average depletion rate for each producing well of 15% over the first five years, followed by a 7% depletion rate.” Another optimistic and widely reported forecast came from an analyst from Raymond James:<sup>21</sup>

“We’re now forecasting that US oil production (excluding NGLs) will grow from 5.6 MMBpd in 2010 to a whopping 9.1 MMBpd in 2015. Including natural gas liquids, total U.S. petroleum liquid production grows 60% from 7.7 MMBpd in 2010 to 12.2 MMBpd in 2015.”

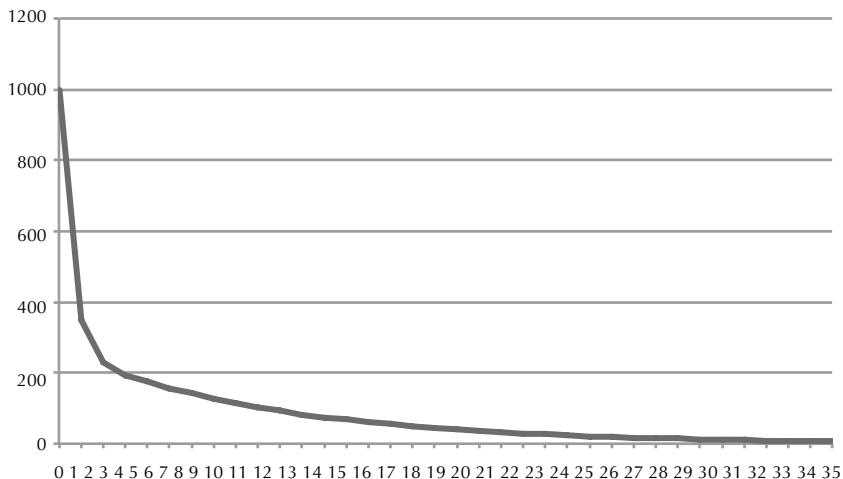
This is forecast for the US, not just for Bakken, but the shale developments are the driving force behind this optimistic prediction. Art Berman, a Houston-based geologist, is far more sceptical. According to him:<sup>22</sup>

“The total contribution of shale oil to US supply is presently 900,000 bo/d and will probably not increase to more than 2 mmbopd (14% of consumption) or less because of high decline rates.”

Oil production in Bakken (and in other locations) faces a number of challenges. One is related to severe price pressures, arising from transportation constraints. The development of rail and pipeline transportation has not kept up with a rapid growth of crude production. There are two primary pipeline exit corridors from Bakken. Crude oil from Bakken flows either east through Clearbrook, Minnesota, to the Midwest refinery centres – or south, through Guernsey, Wyoming, to the refineries in the Midwest and the Rockies. The main oil pipelines in North Dakota include Enbridge, Tesoro, Bridger, Plains, Butte and Belle Fourche. The capacity out of Bakken is limited and oil that leaves Bakken proper competes for constrained pipeline capacity with production from the Rockies. When price spreads justify such a solution, oil may be even trucked to the Gulf Coast. An alternative solution is rail transport.<sup>23</sup>

The second challenge facing the producers in the Bakken and other plays are high decline rates. Figure 14.1 shows a typical decline curve for a Bakken well. It is obvious that decline rates are quite high and higher than assumed in many forecasts of Bakken (and other future plays) oil production. The key to an ultimate resolution of the debate between shale oil sceptics and enthusiasts is analysis of the decline curves once more data has become available.

The developments related to non-conventional oil have changed

**Figure 14.1** Typical Bakken well production

Source: North Dakota Department of Mineral Resources, Activity and Projections Williston (August 3, 2010)

Note: First-year decline rate: 65%; second-year: 35%, third-year: 15%, 10% the following years.

the tone of discussion about future oil supplies and injected fresh optimism into the market. The study by Leonardo Maugeri mentioned above contains a forecast of liquids production capacity for 2020. Potential gains come from different sources, including new projects and growth of production from existing fields, but non-conventional oil plays a significant part.<sup>24</sup> The main conclusion is:

“In any case, the single most important issue that emerges from my analysis is that, from a purely physical and technical point of view, oil supply and capacity are not in any danger. On the contrary, they could significantly exceed world consumption needs and even lead to a phase of oil overproduction if oil demand does not exceed a compounded rate of growth of 1.6 percent each year to 2020.”

The forecast in Table 14.2 is derived through a bottom-up, field-by-field analysis, which results in additional capacity of 49 million barrels per day of oil and NGLs in 2020 (more than half current capacity). Adjusting this forecast for risk factors renders a revised forecast of 29 million barrels per day. Further adjustment for depletion and reserve growth<sup>25</sup> results in a net capacity gain of 17.6 million barrels per day (as shown in Table 14.2).<sup>26</sup>

**Table 14.2** World's oil production capacity to 2020 (mbd)

Country	Production capacity 2011-end	Additional unrestricted production	Additional adjusted production	Net production additions or losses*	Production capacity 2020
<b>Saudi Arabia</b>	<b>12.3</b>	0.9	0.9	0.9	<b>13.2</b>
<b>United States</b>	<b>8.1</b>	7.6	4.7	3.5	<b>11.6</b>
<b>Russia</b>	<b>10.2</b>	1.2	0.8	0.4	<b>10.6</b>
<b>Iraq</b>	<b>2.5</b>	10.4	5.1	5.1	<b>7.6</b>
<b>Canada</b>	<b>3.3</b>	6.8	3.4	2.2	<b>5.5</b>
<b>Brazil</b>	<b>2</b>	6	3.3	2.5	<b>4.5</b>
<b>China</b>	<b>4.1</b>	0.7	0.5	0.4	<b>4.5</b>
<b>Iran</b>	<b>3.8</b>	0.5	0.2	-0.4	<b>3.4</b>
<b>Kuwait</b>	<b>3</b>	1	0.4	0.4	<b>3.4</b>
<b>UAE</b>	<b>2.7</b>	0.86	0.8	0.7	<b>3.4</b>
<b>Venezuela</b>	<b>2.7</b>	2.3	1.2	0.5	<b>3.2</b>
<b>Nigeria</b>	<b>2.4</b>	1.7	0.8	0.4	<b>2.8</b>
<b>Angola</b>	<b>1.9</b>	1.38	1	0.7	<b>2.6</b>
<b>Kazakhstan</b>	<b>1.6</b>	1.6	0.9	0.9	<b>2.5</b>
<b>Qatar</b>	<b>2.1</b>	0.7	0.5	0.3	<b>2.4</b>
<b>Mexico</b>	<b>3</b>	0	0	-0.7	<b>2.3</b>
<b>Algeria</b>	<b>2.1</b>	0.7	0.5	0.2	<b>2.3</b>
<b>Libya**</b>	<b>1</b>	1.2	1.2	1.2	<b>2.2</b>
<b>Norway</b>	<b>2.3</b>	0.4	0.2	-0.4	<b>1.9</b>
<b>Azerbaijan</b>	<b>1.1</b>	0.4	0.3	0.1	<b>1.2</b>
<b>India</b>	<b>0.9</b>	0.6	0.3	0.2	<b>1.1</b>
<b>Indonesia</b>	<b>1</b>	0.4	0.3	0	<b>1</b>
<b>UK</b>	<b>1.2</b>	0.2	0.1	-0.5	<b>0.7</b>
<b>Sub-total</b>	<b>75.3</b>	47.54	27.4	-18.6	<b>93.9</b>
<b>Others</b>	<b>17.7</b>	2	1.2	-1	<b>16.7</b>
<b>World total</b>	<b>93</b>	49.54	28.6	17.6	<b>110.6</b>
<i>Of which:</i>					
Crude oil		78		86	
NGLs		15		24.6	

\*Including depletion and reserve growth

\*\*Libya's 2011 production capacity was curtailed by the Civil War. Before the Civil War, its output capacity stood at 1.9 mbd.

Source: Leonardo Maugeri, 2012, "Oil: The next revolution. The unprecedented upsurge of oil production capacity and what it means for the world," Belfer Center for Science and International Affairs, John F. Kennedy School of Government, June

## BIOFUELS

Biofuels is a term used to describe fuels produced from biomass of different kinds. In most cases, this term applies to liquids, although solid (charcoal, dried manure, sawdust, etc) and gaseous (biogas from landfills) fuels are also used. Most classifications of biofuels encountered in practice evolve around the term "generation," and

unfortunately remind one of press releases and marketing brochures. They are, however, quite popular and we shall use them as well. One classification distinguishes between two generations of biofuels:

Generation 1:

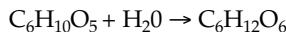
- bio alcohols;
- biodiesel;
- green diesel;
- vegetable oil; and
- bio ethers.

Generation 2: Advanced biofuels

Bio alcohols include ethanol ( $C_2H_5OH$ ) and, less commonly, propanol (propyl alcohol  $CH_3CH_2CH_2OH$ ) and butanol (butyl alcohol  $C_4H_9OH$ ).

Table 14.3 shows data about production of biofuels in 2011.

Ethanol is currently produced from corn (primarily in the US) and sugar cane (primarily in Brazil), with these two countries accounting for most of the world production (as shown in Table 14.3). Production of ethanol from corn or sugar cane involves similar technological processes, which start with the separation of starch from other components of the corn.<sup>27</sup> The next stage is hydrolysis (addition of a water molecule to starch ( $C_6H_{10}O_5$ )<sub>n</sub>) which produces glucose:



Ethanol is produced by fermentation from glucose, as summarised by the following reaction:



Production of ethanol from corn is a very controversial issue, with several complex moral, technical and economic policy problems that would require a long treaty to address at length. The moral issue is related to the indisputable fact that ethanol (used as a liquid fuel) competes against food when it comes to the allocation of arable land and the efforts of most productive and experienced farmers in the world, operating in the US farm belt and Western Europe.<sup>28</sup> A very complex technical problem is related to the estimates of the net energy value (NEV), the energy surplus arising from production of

**Table 14.3** Biofuels production (2011, thousand barrels daily)

Country	Biofuels production
US	<b>567</b>
Brazil	<b>265</b>
Germany	<b>57</b>
Argentina	<b>45</b>
France	<b>35</b>
China	<b>23</b>
Canada	<b>19</b>
Thailand	<b>18</b>
Spain	<b>16</b>
Belgium	<b>10</b>
Netherlands	<b>9</b>
Italy	<b>9</b>
Austria	<b>9</b>
Colombia	<b>8</b>
Poland	<b>8</b>

Source: BP "Statistical Review of World Energy June 2012"

ethanol over the energy expensed in the production of ethanol. A related question is the impact of ethanol plants on CO<sub>2</sub> emissions. Finally, the question one has to address is whether ethanol production should be supported by policy measures such as subsidies and import duties and quotas.

Net energy obtained from ethanol is difficult to estimate, as it requires addressing two difficult questions:

- ❑ definition of the boundary of the system; and
- ❑ the energy value of ethanol co-products.

Running an ethanol plant requires energy. One should recognise that additional amounts of energy are required to grow biomass and distribute produced ethanol to end users. One has to recognise that "[f]ossil energy is essential to industrial agriculture. The following are the major energy inputs to industrial corn farming:

- ❑ nitrogen fertilisers (all fossil energy);
- ❑ phosphate, potash and lime (mostly fossil energy);
- ❑ herbicides and insecticides (all fossil energy);
- ❑ fossil fuels: diesel, gasoline, liquefied petroleum gas (LPG) and natural gas (NG);

- electricity (almost all fossil energy);
- transportation (all fossil energy);
- corn seeds and irrigation (mostly fossil energy);
- infrastructure (mostly fossil energy);
- labour (mostly fossil energy).<sup>29</sup>

We can also argue that energy used to manufacture farm machinery, build access roads, etc, should be incorporated. Depending on where we stop, the net energy gain from ethanol can vary a lot from one estimate to another.<sup>30</sup> One of the biggest challenges is valuation of co-products of ethanol production (see below and footnote 32).

A useful summary can be found in the book by Michael McElroy, who examines the two most exhaustive studies of energy efficiency of corn ethanol.<sup>31</sup>

Pimentel and Patzek argue that the production of corn ethanol uses more energy than that obtained from the end-product. According to Shapouri and McAloon, it is marginally energy efficient. Of course, the inclusion of ethanol co-products<sup>32</sup> in the analysis changes the conclusions. According to McElroy, given all the "the uncertainties inherent in all of these analyses, we conclude that the energy balance for corn produced ethanol is marginally positive: the

**Table 14.4** Energy required to produce a gallon of ethanol

<b>Process</b>	<b>Shapouri and McAloon (2001)</b>		<b>Pimental and Patzek (2005)</b>	
	<b>Energy<sup>a</sup> (BTU/gallon)</b>	<b>Energy<sup>b</sup> (BTU/gallon)</b>	<b>Energy<sup>c</sup> (BTU/gallon)</b>	<b>Energy<sup>c,d</sup> (BTU/gallon)</b>
Corn production	18,713	20,725	37,870	2,751
Corn transport	2,120	2,348	4,830	4,830
Ethanol production	49,733	55,079	56,415	56,415
Ethanol distribution	1,487	1,487	1,487	1,487
Total	72,052	79,639	100,602	90,242
Net energy value	+3,948	-3,639	-24,602	-14,242
Percentage gain or loss	+5.2%	-5.0%	-32%	-19%

<sup>a</sup> Corresponding to a yield of 2.66 gallons per bushel.

<sup>b</sup> Assuming a yield of 2.4 gallons per bushel rather than the 2.66 gallons per bushel assumed by Shapouri and McAloon (2004).

<sup>c</sup> Estimate of energy consumed in ethanol distribution taken from Shapouri and McAloon (2004)

<sup>d</sup> Revisions to the Pimentel and Patzek (2005) budget as discussed in the text.

Source: Michael McElroy

energy captured in the ethanol is greater than the fossil energy employed in its production by about 20 to 30%.”<sup>33</sup>

The issue of energy efficiency of corn ethanol is further complicated by the logistical complications related to ethanol distribution. Blending of ethanol can take place at local terminals, located close to the final users. This is required because ethanol binds with water and cannot be transported across long distances in the product pipelines.

Questionable energy efficiency, the impact on food prices, complications related to logistics and, above all, the new era of fiscal discipline seem to have lowered the initial enthusiasm for ethanol. In June 2011, the US Senate voted 73–27 to repeal the 45 cent per gallon tax credit for blending corn ethanol with conventional fuel and the 54 cent per gallon tariff on imports.<sup>34</sup> At the time of writing, the subsidies ended effectively in 2011. The ethanol industry still benefits from the mandate to use a certain amount of ethanol as fuel (13.2 billion gallons to be blended into gasoline in 2012 and 36 billion gallons by 2022).

The change in the US policy with respect to ethanol subsidies is likely to benefit the sugar cane-based ethanol industry in Brazil. The process of harvesting sugar cane starts with setting the field on fire. This serves a number of objectives, including removal of weeds and purging the area from snakes and rodents, which run for their lives when the blaze starts. The pollen of sugar cane flowers is very hard and sharp, and blindness is a typical affliction of older workers in the developing countries who harvest cane without protective gear. Fire softens the stalks of sugar cane and helps the machete-wielding labourers who cut the cane. After the cane stalks are crushed to extract the juices, the remaining fibrous matter (called bagasse) can be used as a source of energy for producing heat (used to boil off water contained in sugar cane juice) or electricity.

The unquestionable success of the ethanol industry in Brazil can be explained by two factors. The most important is the cost, as exemplified in Table 14.5. Most cars in Brazil (over 80%) can run on any mixture of gasoline and ethanol, and this percentage will only increase as the older vintages of cars are retired.<sup>35</sup>

A discussion of ethanol would not be complete without mention of cellulosic ethanol. Cellulose is the substance (chemically it is a polysaccharide, similar to starch, but with a more complex molecular structure,  $(C_6H_{10}O_5)_n$ ) which provides material for the primary walls

**Table 14.5** Production costs for ethanol in main producing countries

	Worldwatch Institute, 2006 (euro/l Gasoline Eq.)	OECD Directorate for Trade and Agriculture (\$/l Gasoline eq.)	OECD-FAO, 2008 (US\$/IGE)
Ethanol from sugar cane (Brazil)	0.21–0.3 0.263–0.375 (US\$/IGE)	0.331	0.29
Ethanol from corn (US)	0.33–0.52 0.413–0.65 (US\$/IGE)	0.437	0.75
Ethanol from grain (EU)	0.41–0.66 0.513–0.825 (US\$/IGE)	0.869	1.25
Ethanol from sugar beet (EU)		0.848	0.5
Ethanol from cellulose	0.66–0.99 0.825–1.24 (US\$/IGE)		

Source: Worldwatch Institute 2006 (data extracted from graph), OECD 2008 (data extracted from graph). For comparability, production costs are converted to gasoline/diesel equivalent by dividing production costs per litre of fuel by the energy content relative to gasoline and diesel (0.66 for ethanol, 0.89 for diesel).

Exchange rate Euro/\$ = 1.25 in 2005–2006.

of the cells of plants. It is very abundant in nature and its use as a raw material in the production of ethanol would not compete directly with production of food. The problem is that commercially viable technology for production of cellulosic ethanol does not yet exist.<sup>36</sup> Even if a commercially viable technology is eventually developed, cellulosic ethanol is unlikely to address the challenge of the US liquid fuel supply. According to Tad Patzek:<sup>37</sup>

“Suppose that one would like to replace 18% of the current 20 EJ  $y^{-1}$  the US uses as automotive gasoline. [...] If the switchgrass ethanol cycle described here [in the quoted paper] were used to achieve this goal with close to 50% probability, one would need [...] 140 million hectares of switchgrass, or the entire area of active U.S. cropland, using the mean of the energy efficiency distribution [...]. The U.S. agricultural area is from [www.ers.usda.gov/AmberWaves/July06/SpecialIssue/pdf/BehindDataJuly06.pdf](http://www.ers.usda.gov/AmberWaves/July06/SpecialIssue/pdf/BehindDataJuly06.pdf). According to USDA, the total US cropland area (harvested, summer fallow, and failed) was 140 million hectares in 2006. Another 40 million hectares were devoted to pastures and idle cropland.”

The cited paper, like many other studies, points out that ethanol, irrespective of the source of feedstock biomass, is unlikely to make a dent in the demand of our energy-thirsty society for liquid fuels.

Biodiesel is a term used with respect to different to several different types of fuels, produced with different technologies. They are similar in one sense: they are seen as substitutes for diesel oil.

First-generation biodiesel is produced through a chemical reaction in which triglycerides<sup>38</sup> are treated with alcohol (ethanol or methanol are usually used) in the presence of a catalyst. The product is not an equivalent of petroleum diesel, but is the alkyl ester or methyl ester, depending on what type of alcohol was used. Biodiesel is typically used by blending with diesel (up to 20%). Biodiesel is non-toxic and biodegradable. Its disadvantage, compared to petroleum diesel, is that it has much higher cloud and pour points and loses viscosity and gels in low temperatures.

Second-generation biodiesel (so-called green diesel) is chemically equivalent to petroleum diesel, and can be blended with it in any proportion. It is produced by reacting the feedstock (biomass) with hydrogen (hydro-processing) or by combusting biomass to produce syngas (a mixture of carbon monoxide and hydrogen), and then using the Fischer–Tropsch synthesis (FTS) reaction to produce hydrocarbons.

FTS technology should be on the radar screen of every energy professional, as it is behind a number of technologies which are likely to reshape our energy future: coal-to-liquids (CTL), gas-to-liquids (GTL) and biomass-to-liquids (BTL). A few words on the history of this process will help to explain its current state and future promise.<sup>39</sup> The precursor technology to FTS was developed in 1902 by P. Sabatier and J. D. Senderens, who invented a process for producing methane ( $\text{CH}_4$ ) from carbon monoxide (CO) and hydrogen. In 1923, three chemists (Hans Tropsch, Franz Fischer and Helmut Pichler) came up with a technology for producing middle distillates (diesel, light and heavy gasoil) from synthesis gas (a mixture of carbon monoxide and hydrogen: syngas) in the presence of iron–cobalt or copper–iron catalysts.<sup>40</sup> The patents had been acquired by the German concern I. G. Farben, which built the first pilot plant in 1926 in Ludwigshafen, Germany. Frank Howard, head of the Standard Oil New Jersey Research and Development unit, who visited the plant in the same year, grasped immediately the commercial implications of this development. In a letter addressed to the president of SO of NJ, Walter Teagle, he wrote:

“Based upon observations and discussions today, I think that this matter is the most important which has ever faced the company since the dissolution [of the John Rockefeller’s company]. This means absolutely the independence of Europe in the matter of gasoline supply”.

In 1929, Standard Oil decided to enter into an agreement with I. G. Farben to acquire the right to use its all hydrogenation processes outside Germany in return for 2% of SO of NJ stock, worth US\$35 million at the time. Both firms effectively agreed to divide the turf, with the US firm staying out of the chemical business and the German firm out of the oil business. During the Second World War, the FTS process became key to the German military operations, with the production of synthetic fuel reaching 124,000 barrels per day, and providing 57% of total fuel supply and 92% of aviation fuel supply. Allied air raids reduced the output to less than 3,000 barrels per day by 1944, effectively crippling the German air force. After the war, a commercial scale plant utilising FTS technology was built in Brownsville Texas, with smaller plants supported by the US government in Louisiana and Texas. The era of low oil prices, following discovery of very productive oil fields across the globe, reduced the interest in this technology practically to zero in most developed countries. The centre of gravity shifted to South Africa, primarily for geopolitical reasons (growing isolation and international boycotts) and lack of domestic oil reserves. South African Coal, Oil and Gas Corporation (renamed later as Sasol) started construction of a liquids plant relying on FTS technology, with production starting in 1955 in Sasolburg. The second Sasol plant in Secunda started operations in 1975.

The oil price shock in the 1970s revived interest in FTS technology outside South Africa. By 2004, about 50 projects were announced with total projected capacity of 900,000 barrels of liquids per day. The most important projects included:

- ❑ Chevron/Sasol: 100,000 BPD GTL Qatar;
- ❑ Shell: 140,000 BPD GTL Qatar;<sup>41</sup>
- ❑ Conoco/Phillips: 130,000 BPD GTL Qatar;
- ❑ Exxon-Mobil : 150,000 BPD GTL Qatar;<sup>42</sup>
- ❑ Chevron/Sasol: 33,000 BPD GTL Nigeria;
- ❑ Shell: Six projects worldwide @ 72,000 BPD each;
- ❑ Sasol 2: 80,000 BPD CTL China; and
- ❑ Shell: 70,000 BPD CTL China.

The FTS process is divided into three steps. The first step involves generation of syngas (a mixture of CO and H<sub>2</sub>) from biomass, coal or

natural gas. In the second step, long paraffin chains ( $C_1$  to  $C_{100+}$ ) are formed from syngas in the presence of different catalysts. In the third step, the long paraffin chains are hydrocracked to produce kerosene ( $C_{14}-C_{20}$ ), diesel ( $C_{10}-C_{13}$ ), gasoline ( $C_5-C_{10}$ ), jet fuel ( $C_{10}-C_{13}$ ) and naphtha ( $C_4-C_{10}$ ).

## PEAK OIL THEORY

The peak oil strategy is associated with the name of Marion King Hubbert (1903–1989), who worked as a geologist for Shell in Houston. In 1956, Hubbert formulated a very simple model for the cumulative production of an exhaustible commodity, using the logistic growth curve:

$$Q(t) = \frac{Q_{\max}}{1 + ae^{bt}} \quad (14.1)$$

where  $Q_{\max}$  stands for the ultimate recovery of the resource,  $t$  stands for time,  $Q(t)$  denotes the cumulative production as of time  $t$ , and  $a$  and  $b$  are parameters of the function. Figure 14.2 shows a plot of a logistic function.

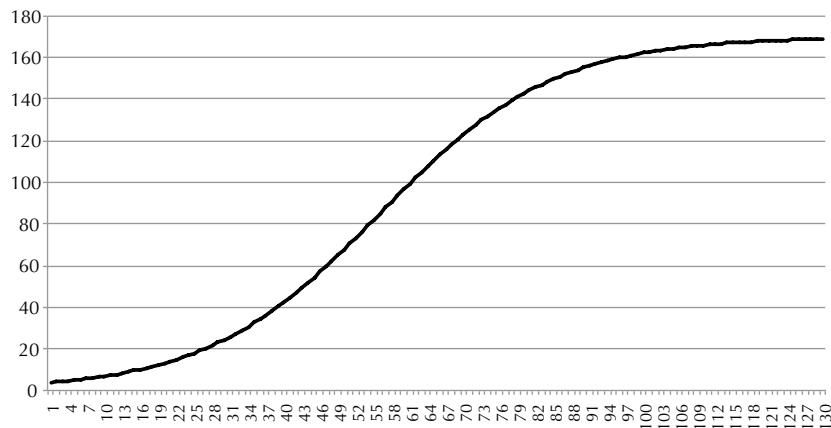
Maximum annual production is given by:

$$P_{\max} = \frac{Q_{\max}|b|}{4} \quad (14.2)$$

and the year of the maximum production is:

$$t_{\max} = \frac{1}{b} \ln\left(\frac{1}{a}\right) \quad (14.3)$$

Hubbert's estimate of the peak of the US oil production, based on the following assumptions:  $Q_{\max} = 170$  billion barrels,  $a = 46.8$ ,  $b = -0.0687$ , and the base year of 1910.5, was placed in year 1967.<sup>43</sup> Alternative estimates based on the improved ultimate reserve assumptions yielded estimates of peak between 1967 and 1973. At the time of the prediction, the estimated oil output was 7.1 million barrels, with a shut-in capacity of 3.5 million barrels a day. Given the US capacity for oil production, it was a very bold assertion at the time, but it proved quite prescient (as illustrated by Figure 14.3). An alternative representation of the Hubbert's theory is the bell-shaped curve showing production rate over time, starting and ending at zero, and peaking at some point in time.

**Figure 14.2** Cumulative production, Hubbert's Model

Note: Original Hubbert's estimates used for the parameters ( $a = 46.8$ ,  $b = -0.0687$ ,  $Q_{\max} = 170$  billion barrels)

### Extensions of Hubert's Model

The formula (14.1) is used in many empirical research papers to plot the ratio of current production to cumulative production against the ratio representing the remaining resource (it takes a few simple mathematical transformations). More information can be found at: <http://www.theoildrum.com/story/2006/8/16/102942/337>.

It is important to recognise that, over time, alternative models were developed which do not imply symmetry of production curve and rely on more complex mathematical modelling. One example is a paper by Ugo Bardi and A. Lavacchi.<sup>44</sup> Relying on the Lotka–Volterta model,<sup>45</sup> they fitted a curve to the historical data of oil production in the lower 48 United States, achieving a remarkable fit. Al Fattah and Startzman came up with a multi-cyclic Hubbert equation given by:<sup>46</sup>

$$q(t) = \sum_{i=1}^k 4(q_{\max})_i \left\{ \frac{e^{-a(t-t_{\max})}}{\left[1+e^{-a(t-t_{\max})}\right]^2} \right\}_i \quad (14.4)$$

where  $k$  is the total number of production cycles,  $q_{\max}$  and  $t_{\max}$  are the peak oil production rate of each cycle and the corresponding time of occurrence, respectively, and  $a$  is a constant. Ibrahim Sami Nashawi,

**Figure 14.3** Hubbert's estimate and actual US oil production 1911–2011

Adel Malallah and Mohammed Al-Bisharah applied this model to a number of countries.<sup>47</sup>

Another formulation of the peak oil theory can be found in a 2011 article by Chris Skrebowski, in which a distinction between geological and economic peak oil was made:<sup>48</sup>

“Thus the geologists are right that the depletion of low-cost oil will produce Peak Oil but it will not be caused by a shortage of oil resources. The economists are right that there is no shortage of oil resources or oil substitutes but have so far failed to recognise that there is an oil price which cannot be afforded and this constraint will create and define an economic Peak Oil to be differentiated from a geological Peak Oil.”

One of the mutations of peak oil theory is the *export land model* proposed by Jeffrey Brown.<sup>49</sup> The idea behind it is so simple and obvious that the term “model” seems an exaggeration: if an oil-producing country is an oil consumer and the rate of growth of consumption exceeds the rate of growth of output, the country will eventually run out of oil available for exports. “Peak export,” as obvious as it is, should not be ignored. Many oil-producing countries face a toxic mix of high oil production decline rates (for existing

fields), policies that hamper additions to reserves, high population growth and price subsidies fostering wasteful consumption. Some former net oil exporters have become net importers, and more cases may be on the way.

A recent extension of the export land model is the distinction between global net exports (GNE) and available net exports (ANE), global exports adjusted for imports of Chindia (China + India, a term coined by an Indian politician). ANE is oil supply left to satisfy the needs of importers after the thirst of Chindia is quenched. ANE have been falling consistently over the last few years, pointing to higher prices in the future. The critics of this theory point out to rapid expansion of production of non-conventional oil, going as far as predicting self-sufficiency for the US or North America.

The success of the Hubbert's forecast lead to the development of a cottage industry of experts coming up with alternative predictions of peak of oil production for individual countries and the rest of the world. The supporters of this view organised themselves into a number of different organisations, including:

- ❑ the Association for Study of Peak Oil & Gas (ASPO);<sup>50</sup>
- ❑ the Oil Depletion Analysis Centre (ODAC); and
- ❑ the Post Carbon Institute.

The debate about peak oil is followed by a number of industry participants blogging on TheOilDrum.com, where the reader can obtain updates on the most recent publications and controversies.

The best and most general definition of peak oil theory comes from Colin Campbell, ASPO founder and honorary chairman: "the term peak oil refers to the maximum rate of the production of oil in any area under consideration, recognising that it is a finite natural resource, subject to depletion."<sup>51</sup> This is a very general definition, which does not imply any specific mathematical model or any prediction when peak production will be reached. The followers of peak oil theories represent a very broad spectrum of opinions with respect to a number of critical questions:

- ❑ the timing of the peak;
- ❑ the widely defined social and economic consequences; and
- ❑ policy recommendations.

A summary of the projections of the timing of the peak (please, note that this is a moving target and the estimates may change) can be found in the well-known Hirsch (although somewhat dated) report.<sup>52</sup>

<input type="checkbox"/> 2006–07:	Bakhitari
<input type="checkbox"/> 2007–09:	Simmons
<input type="checkbox"/> after 2007:	Skrebowski
<input type="checkbox"/> before 2009:	Deffeyes
<input type="checkbox"/> before 2010:	Goodstein
<input type="checkbox"/> around 2010:	Campbell
<input type="checkbox"/> after 2010:	World Energy Council
<input type="checkbox"/> 2010–20:	Laherrere
<input type="checkbox"/> 2016 EIA:	(Nominal)
<input type="checkbox"/> After 2020:	CERA
<input type="checkbox"/> 2025 or later:	Shell
<input type="checkbox"/> no visible peak:	Lynch

The Hirsch report remains a starting point for the study of peak oil.<sup>53</sup> It makes a number of important observations.

- Peak oil is not an energy crisis: it is a crisis of liquid fuels. The crisis is exacerbated by a number of factors, including:
  - Dependence of highly developed economies on a transportation system designed around truck transportation and a highly energy-intensive consumption model based on moving daily a large number of people from the suburbs to place of employment and moving goods to the suburbs.
  - Inertia of the existing vehicle fleet. In 2003, the US fleet of automobiles and light trucks (vans, pick-ups and SUVs) was about 210 million. Replacement of only half the automobile fleet would take ~10–15 years. The average age of light trucks was seven years at the time. Under normal conditions, replacement of one-half of the stock of light trucks would require 9–14 years.
- There are several mitigation approaches including:
  - improved oil recovery (IOR);
  - heavy oil/oil sands;
  - coal liquefaction;
  - clean substitute fuels (GTL technology); and

- exploration for new oil fields.

### The future of oil: The controversy

The effectiveness of different mitigation technologies depends on when the programme to address peak oil begins with respect to its onset. According to the Hirsch report, delaying mitigation programmes until the onset of the oil peak exposes oil-consuming countries to major macroeconomic and social shocks. The implementation of mitigation programmes 10–20 years earlier leaves enough time to make a transition to an oil-constrained world. Given the endemic myopia of corporate boards and the political establishment, it is not difficult to predict the likely course of events.

The positions with respect to the future social and economic consequences of peak oil vary from dire predictions of the breakdown of civilisation and economic collapse<sup>54</sup> to more balanced views that the transition to a world freed from the reliance on oil will be manageable, although challenging.<sup>55</sup> Such views, disseminated by a group consisting mostly of retired geologists and freelance journalists, were met often with ridicule and dismissed as senile rants of octogenarian Cassandras. This tide of dismissive comments has been receding in the last few years, with several respected institutions publishing in-depth studies that warn about the social and geopolitical consequences of peaking oil production. In addition to the excellent and balanced Hirsch report, some other contributions will be discussed below.

A thinktank known as the Future Analysis department of the Bundeswehr Transformation Center, in a study led by Thomas Will, issued a number of warnings related to:<sup>56</sup>

- ❑ the growing importance of oil as a determinant of political (and possibly military) power, the potential for new resurgent regional or even global powers;
- ❑ increasing economic leverage and influence of oil exporters, the potential for promotion of different ideologies through control of oil supplies;
- ❑ the evolution of the oil industry from arrangements based on free markets to government-negotiated supply contracts, leading to Balkanisation and the fragmentation of supply chains;<sup>57</sup>
- ❑ the potential for the breakdown of many trading arrangements

based on low transportation costs that favoured in the past countries with low labour costs;

- ❑ potential relapse into planned economy, with governments relying on rationing and direct allocation schemes, and other coercive measures;
- ❑ the potential for a growing number of failed states that cannot cope with increasing oil prices, with the unavoidable breakdown of social and political structures; and
- ❑ a global chain reaction leading to the breakdown of the current international order, with a growing threat of extremist movements.

It was reported<sup>58</sup> that the UK government had created a working group including the representatives of the Department of Energy and Climate Change, the Bank of England and the Ministry of Defence. The deliberations of the group are confidential.<sup>59</sup>

At the other end of the spectrum, there are some who believe that the market mechanism will trigger additional supplies of oil, which effectively will defer the peak oil for a very long time. The debates between the two camps are often acrimonious, reflect differences of political and social philosophies, and sometimes represent an academic version of a bar brawl.

The finite nature of oil supplies is recognised even by the most respected sceptics of peak oil, although it is admitted that oil may cease to be used an energy source long before resources run out. This latter position was expressed most succinctly by the former Saudi oil minister, Ahmed Zaki Yamani, who quipped that the Stone Age did not end because mankind ran out of stones. According to Dan Yergin, there is an alternative to the peak oil theory:<sup>61</sup>

“[T]here is another way to visualize the future availability of oil: as a ‘plateau.’ In this view, the world has decades of further growth in production before flattening out into a plateau – perhaps sometime around midcentury – at which time a more gradual decline will begin. And that decline may well come not from a scarcity of resources but from greater efficiency, which will slacken global demand.”

We can understand scepticism with respect to peak oil. The end of oil has been foretold many times in history,<sup>62</sup> but we can point out that many failed predictions of our death do not prove that we are immortal. Having said this, the assumption of an undulating plateau

seems as arbitrary as the assumption of a sharp, one-time peak. Why should we assume that the finality of resources will be exactly offset through human ingenuity and the deployment of resources? Fluctuations fed by business cycles and waves of new technologies are more likely.

The position the author takes can be summarised as follows:

Peak oil is a tautology and it is difficult to understand the passion of opponents of this theory, which often border on the negation of simple arithmetic. If the supply of a given resource is finite (or, in general, the extraction rate exceeds the rate at which the resource can be replenished), the flow rate has to reach a maximum at some point and eventually fall to zero. This is a simple manifestation of the old rule that if something cannot go on forever, it eventually has to stop.<sup>63</sup> The population of whales, which came close to extinction in the mid-19th century (and was probably saved by the transition to oil), and the trees of the Easter Islands demonstrate this point.

One exception would have to be granted if it can be proved that oil is a renewable resource (see the section below on the abiotic theory of the origins of oil). The urge to deny the obvious is deeply engrained in our culture and psyche: we tend to be an optimistic species (if we were not, it would be difficult to explain how we muster the strength to get up in the morning) and do not like to see things fall apart. Ugo Bardi referred to this trait of our collective personality as the Seneca Effect. The Roman philosopher wrote to his friend:<sup>64</sup>

“It would be some consolation for the feebleness of ourselves and our works if all things should perish as slowly as they come into being; but as it is, increases are of sluggish growth, but the way to ruin is rapid.”

Hubbert’s peak oil model is a stylised formulation of the obvious truth of finality of oil resources, expressed in the language of mathematics. It cannot be treated as a forecast (in spite of an initial successful test of its predictive power): most models based on the techniques borrowed from physical science fail when applied to complex socioeconomic systems. Such models fail to recognise and account for the response of the system to adverse developments. Adjustments depend on unpredictable configurations of the political and social forces and often on the personalities of our leaders. The conclusions about a unique peak of a symmetric curve or other well-defined and pleasing patterns seem to be a bit simplistic. Our position may be explained using the analogy of human life. We

know with absolute certainty our final destination, but we do not know how and when we are going to reach it. We may be cut down in our prime, when we are still strong and resourceful, or we can last until few sparks of life are left in us. Similarly, we may abandon oil when better, cleaner and cheaper sources of energy are developed. It is also likely we will be squeezing the last drops of oil from rocks located in some wretched and inhospitable places at a very high cost. What is certain, given the history of oil so far, is that the journey will not happen over a mild plateau. We can expect spasms, crises, conflicts and – this should be music to traders' ears – high price volatility.

The arguments about the ability of free markets to postpone the peak indefinitely are absurd and pointless. Even the most efficient market mechanism cannot produce something that does not exist or an infinite flow from a finite resource. If this was not the case, we would have ample supplies of the elixir of youth and pixie dust. This does not mean, however, that we do not have to rely on the market to manage the transition to a world without oil.

### **Abiotic oil theory**

One set of arguments denying the possibility of peak oil based on the abiotic (as opposed to biogenic) theory of oil formation deserves additional attention. Whether one agrees with this position or not, it is based on solid science – although, as with any scientific hypothesis, it may be found wrong. The abiotic theory is sometimes called the Russian–Ukrainian theory, as its origins can be traced to the famous Russian chemist Dmitri Mendeleev and the work of several Soviet geologists, including the most prominent: Nikolai Alexandrovich Kudryavtsev (1883–1971).<sup>65</sup> The original theory can be traced back to the writings of Aristotle and Agricola, and can was succinctly summarised by Mikael Höök and others:<sup>66</sup>

“The authors favouring the abiotic theory claim that the hydrogen–carbon system generates hydrocarbons under pressures found in the mantle of the Earth and at temperatures consistent with that environment[...]. There exists experimental proof that, in some conditions of high pressure and temperature (eg, under a diamond anvil) it is possible to combine carbon and hydrogen to generate hydrocarbons[...]. Also the Fischer–Tropsch (1930) process – developed in the 1920s – is proof that it is possible to create long-chain, petroleum like, hydrocarbons starting from inorganic reactants.”

The arguments used by the proponents of the theory include:

- ❑ perceived insufficiency of source rock to generate super-giant deposits;
- ❑ occurrence of petroleum at great depth; and
- ❑ columns of flames have been seen during the eruptions of some volcanoes.

The field explorations in places such as the Siljan Ring, offshore Vietnam, Eugene Island in the Gulf of Mexico and the Dnieper-Donets basin did not produce support for the abiotic theory. On the other hand, exploration techniques based on the biogenic theory consistently lead to discovery of oil in commercial quantities, confirming the old industry adage about the “truth of the borehole.”

## CONCLUSIONS

Forecasting a specific number is an exercise fraught with danger, as any quantitative analyst knows, and this is especially the case for complex systems characterised by non-linearities and feedback loops. Small differences in the initial assumptions compound quickly into significant effects. The only saving grace is a short collective and institutional memory. The history of the energy industry can be written as a history of failed forecasts of doom or immediate age of abundance. This should teach humility to every person attempting to make bold forecasts and bet billions based on this. The ability to come up with “the dog-ate-my-homework” excuses also helps. It is easy to blame sinister dictators and impoverished people punching holes in the pipelines under the scorching sun for the forecast repeatedly falling off the mark. Having said this, we can still risk one prediction. What seems to be emerging in the case of natural gas and oil is a highly volatile pricing regime that only a trader could dream of. High production decline rates from shale wells place the industry on a proverbial treadmill, and make it highly dependent on capital inflows. A minor disturbance will translate into big price swings. What will become critically important to trading is fundamental analysis combined with the ability to preserve cool heads in an environment characterised by alternating waves of excessive pessimism and optimism.

- 1 In the section on natural gas, we used the analogy of an unconventional lifestyle, for which there are as many definitions as the number of people in the room.
- 2 This section relies heavily on the IEA WEO 2010.
- 3 This section is based on number of sources. The reserve and environmental data are quoted after the updated and exhaustive report, Pierre Gosselin, Steve E. Hrudey, Anne Naeth, André Plourde, René Therrien, Glen Van Der Krak and Zheghe Xu, 2010, "Environmental and health impacts of Canada's oil sands industry," The Royal Society of Canada Expert Panel, December.
- 4 Tar is produced through the distillation of coal and is chemically different than bitumen. The term "tar" used with respect to rocks containing bitumen is technically incorrect, but so widely used that is pointless to fight this specific usage.
- 5 Ian C. Walker, 1998, "Marketing challenges for Canadian bitumen," Imperial Oil, Calgary, Alberta, Canada.
- 6 10,000 cP means that oil is has 10,000 viscosity of water (water at 20° C has the viscosity of 1 centipoise).
- 7 Historically, bitumen has been used for thousands of years and its applications range from waterproofing (for example, in boats) to mummification in ancient Egypt. The chances are that Moses floated to his rescue on a boat made of reeds covered with bitumen. The first (but not last) battle over control of hydrocarbons was fought in 312 BC between the Seleucid Syrians and the Nabatean Arabs living around the Dead Sea (see Arie Nissenbaum, 1978, "Dead Sea asphalts –historical aspects," *AAPG Bulletin*, pp 837–44).
- 8 Robert Kunzig, 2009, "Scraping Bottom," *National Geographic Magazine*, March.
- 9 Other techniques include cold heavy oil production on site (CHOPS), used for very heavy oils, and toe-to-heel air injection (THAI). The latter technology uses two wells: a vertical one for air injection and a horizontal production well. A portion of the bitumen is burned underground.
- 10 Steam is typically produced using natural gas, although other solutions are being explored, given the falling production of natural gas in Canada. Alternatives include the construction of a nuclear power plant near the tar sands location, underground combustion of the deposits.
- 11 See the section on refining for details.
- 12 The discussion of these issues can be found in the report by Pierre Gosselin, *et al*, op.cit.
- 13 <http://www.theoldrum.com/node/3839>. The data in this post (Table 14.1) are based in the research by M. C. Herweyer and A. Gupta.
- 14 <http://www.theoldrum.com/node/3810>.
- 15 TransCanada's "proposed Keystone Gulf Coast Expansion Project is an approximate 2,673-kilometre (1,661-mile), 36-inch crude oil pipeline that would begin at Hardisty, Alberta and extend southeast through Saskatchewan, Montana, South Dakota and Nebraska. It would incorporate a portion of the Keystone Pipeline (Phase II) through Nebraska and Kansas to serve markets at Cushing, Oklahoma before continuing through Oklahoma to a delivery point near existing terminals in Nederland, Texas to serve the Port Arthur, Texas marketplace." The project is actively contested (summer of 2012) for a number of reasons, including pollution risk and threat to the aquifers along the route of the pipe. According to the critics, a rupture of the pipeline, given the quality of transported crude, would have a serious environmental impact.
- 16 WEO 2010, IEA.
- 17 World Petroleum Resources Project, 2009, "An estimate of recoverable heavy oil resources of the Orinoco Oil Belt, Venezuela," USGS (<http://pubs.usgs.gov/fs/2009/3028/pdf/FS09-3028.pdf>).
- 18 "FDI in the Bolivarian Republic of Venezuela went from US\$349 million in 2008 to a negative inflow of US\$3.105 billion in 2009, mainly as a result of nationalizations during the year." "Foreign Investment in Latin America and the Caribbean," 2009 ([http://www.eclac.org/publicaciones/xml/2/39422/2010-414-LIEI-Book\\_WEB.pdf](http://www.eclac.org/publicaciones/xml/2/39422/2010-414-LIEI-Book_WEB.pdf)).

- 19 A. K. Gupta, M. C. Herweyer and C. A. S. Hall, 2008, "Appendix E: Oil shale: Potential, EROI and social and environmental impacts," *The Oil Drum*, April 15.
- 20 Leonardo Maugeri, 2012, "Oil: The next revolution. The unprecedented upsurge of oil production capacity and what it means for the world," Geopolitics of Energy Project, Belfer Center for Science and International Affairs, John F. Kennedy School of Government, June. This is an unrestricted forecast – ie, the forecast is not adjusted for different potential risks.
- 21 Quoted after Trisha Curtis, 2012, "Keeping our options open: Markets for Canadian crude and the pipeline dilemma," Energy Policy Research Foundation (EPRINC), CERI Oil Conference April 23. This presentation contains alternative forecasts for shale oil production a reader could find useful.
- 22 Arthur E. Berman, 2012, "Oil-prone shale plays: Expectation & experience," ASPO USA webinar, May 17.
- 23 There are several sources a reader can use to get updates on the Williston Basin oil transportation situation. North Dakota's Pipeline Authority has regular webinar presentation about current constraints and projects under way (<https://www.dmr.nd.gov/pipeline/>). See, for example, Justin J. Kringstad's presentation (July 29, 2011). Rusty Braziel's blog is another valuable source of information (<http://www.rbnenergy.com>).
- 24 According to Leonardo Maugeri, "the shale/tight oil boom in the United States is not a temporary bubble, but the most important revolution in the oil sector in decades."
- 25 Reserve growth refers to upward creep of initial reserve estimates. Initial estimates are often conservative (given accounting rules which require caution) and are increased as more information becomes available and as technology improves.
- 26 See above for a discussion of this paper.
- 27 There are two basic technological processes used in production of ethanol: dry- and wet-milling processes. About 80% of ethanol produced in the US is from dry-grind technology. One has to observe that both processes use water. Wet-milling process requires soaking corn in water for about 48 hours.
- 28 The factors behind price spikes and food price volatility are discussed in "A report by the high level panel of experts on food security and nutrition of the Committee on World Food Security," Rome 2011 ([http://www.fao.org/fileadmin/user\\_upload/hlpe/hlpe\\_documents/HLPE-price-volatility-and-food-security-report-July-2011.pdf](http://www.fao.org/fileadmin/user_upload/hlpe/hlpe_documents/HLPE-price-volatility-and-food-security-report-July-2011.pdf)). The authors see biofuel as a main factor behind increasing food prices. See also "Famine : la FAO pointe les biocarburants et la surexploitation des sols," *Le Monde*, August 3, 2011.
- 29 Tad W. Patzek, 2004, "Thermodynamics of the corn-ethanol biofuel cycle," *Critical Reviews in Plant Sciences*, pp 519–67.
- 30 David Pimentel and Tad W. Patzek, 2005, "Ethanol production using corn, switchgrass, and wood; biodiesel production using soybean and sunflower," *Natural Resources Research*, (14)1, March; Hosein Shapouri and Andrew McAlloon, 2004, "The 2001 net energy balance of corn ethanol," US Department of Agriculture.
- 31 Michael B. McElroy, "Energy. Perspectives, Problems, & Prospects," Oxford University Press, 2010.
- 32 The information on co-products is available from the National Corn Growers Association. "A bushel of corn used in the dry grind ethanol process yields 2.8 gallons of ethanol, 17 pounds of carbon dioxide, and 16 pounds of distillers grains. These residual grains are used as a quality source of energy and/or protein in beef, dairy, swine, and poultry rations. If fed in local livestock markets, the product is normally sold as wet distillers grains (WDGs). If transported to distant markets or exported, the product is dried at the plant and becomes dried distillers grains (DDGs). In many cases, the solubles stream, or syrup, from the fermentation process is mixed in with the distillers grains, resulting in WDGs or DDGs." (See <http://ncka.com/coproducts>).
- 33 Michael B. McElroy, "Energy. Perspectives, Problems, & Prospects," Oxford University Press, 2010.

- 34 Ed Crooks and Joe Leahy, 2011, "Senate vote hits US ethanol industry," *Financial Times*, June 17.
- 35 In the US, the typical mix is 10% of ethanol (E10). On July 7, 2011, the House Energy & Environment Subcommittee of the Committee on Science, Space, and Technology held the hearing to examine the scientific and technical issues related to EPA's recent waiver permitting mid-level ethanol blends of up to 15% ethanol. Many experts are concerned that the waiver was granted without sufficient tests and studies.
- 36 As Robert Bryce observed in his 2010 book, *Power Hungry: The Myths of "Green" Energy and the Real Fuels of the Future* (New York, NY: Public Affairs), cellulosic ethanol has been promoted as a solution to the US energy problems at least since 1921, when the cost was about US\$24 per gallon (at today's prices).
- 37 Tadeusz W. Patzek, 2010, "A probabilistic analysis of the switchgrass ethanol cycle," *Sustainability*, pp 2734–69.
- 38 Triglycerides are main components of animal fat and vegetables oils.
- 39 Based on "The GTL/CTL/BTL process. Fischer-Tropsch (F-T) fuels. Historical review" (<http://www.angtl.com/pdfs/Historyoff-T.pdf>). A good review of the science behind FTS can be found in M. Ojeda and S. Rojas, 2010, *Biofuels from Fischer-Tropsch Synthesis* (New York, NY: Nova Science Publishers).
- 40 The process for hydrogenation of lignite (brown coal) was invented in 1913 by Friedrich Bergius, who did not get full recognition for his contribution, as a countless number of inventors before and after him.
- 41 The plant became operational in mid-2011, at the overall cost of US\$18–19 billion. It produces GTL Naptha, GTL Kerosene, GTL normal paraffins, GTL gasoil or diesel and GTL base oils. See Sylvia Pfeifer, 2011, "Pearl crowns Shell's Qatari gasoil hopes," *Financial Times*, June 5.
- 42 This project was cancelled in 2007. According to the *Financial Times*, "Exxon said the decision to abandon it was consistent with its investment approach, "which focuses on maximising the value of the resources for both the host government and our shareholders". The Qatar project was Exxon's only active involvement in a GTL plant. However, it said it would continue to look for alternative opportunities for the technology in which it had invested US\$600m over two decades and taken out 3,500 patents." See, Ed Crooks, 2007, "Exxon cancels gas-to-liquids project in Qatar," *Financial Times*, February 20.
- 43 Alfred J. Cavallo, 2004, "Hubbert's petroleum production model: An evaluation and implications for world oil production forecasts," *Natural Resources Research*, December.
- 44 Ugo Bardi and Alessandro Lavacchi, 2009, "A simple interpretation of Hubbert's model of resource exploitation," *Energies*, (23), pp 646–61.
- 45 The famous Lotka-Volterra model is a system of non-linear differential equations describing the dynamics of the bio system in which two types of species interact as a prey and a predator. The modified version of the model used by Bardi and Lavacchi assumes the prey reproduces very slowly.
- 46 S.M. Al-Fattah and R.A. Startzman, 2000, "Forecasting world natural gas supply," *Journal of Petroleum Technology*, 52, May, pp 62–72.
- 47 Ibrahim Sami Nashawi, Adel Malallah and Mohammed Al-Bisharah, 2010, "Forecasting world crude oil production using multicyclic Hubbert Model", *Energy and Fuels*, March, pp 1788–1800 (<http://pubs.acs.org/stoken/presspac/presspac/full/10.1021/ef901240p>).
- 48 <http://www.theoildrum.com/node/8410#more>.
- 49 Jeffrey J. Brown, Samuel Foucher and Jorge Silveus, 2010, "Peak oil versus peak net exports – which should we be more concerned about?" [http://www.aspousa.org/2010presentationfiles/10-7-2010\\_aspousa\\_TrackBNetExports\\_Brown\\_J.pdf](http://www.aspousa.org/2010presentationfiles/10-7-2010_aspousa_TrackBNetExports_Brown_J.pdf).
- 50 ASPO International is an umbrella for a large number of international organisations.
- 51 See the ASPO website.
- 52 Robert L. Hirsch, Roger Bezdek and Robert Wendling, 2005, "Peaking of world oil produc-

- tion: Impacts, mitigation, & risk management," Science Applications International Corporation/US Department of Energy, National Energy Technology Laboratory.
- 53 Other recent reports on the subject include "The oil crunch. A wake-up call for the UK economy," Second report of the UK Industry Taskforce on Peak Oil & Energy Security (ITPOES), February, 2010; Jaromir Benes, Marcelle Chauvet, Ondra Kamenik, Michael Kumhof, Douglas Laxton, Susanna Mursula and Jack Selody, 2012, "The future of oil: geology versus technology," International Monetary Fund, May; "Transport energy futures: long-term oil supply trends and projections," Report 117, Department of Infrastructure, Transport, Regional Development and Local Government, Canberra, Australia, Commonwealth of Australia, 2009. Additional reports are mentioned later in this chapter.
- 54 "War, starvation, economic recession, possibly even the extinction of homo sapiens," from Daniel Yergin, 2011, "There will be oil," *Wall Street Journal*, September 18.
- 56 Based on Stefan Schultz, 2010, "Peak oil and the German government. Military study warns of a potentially drastic oil crisis," *Spiegel Online*, September 1.
- 57 "Bilateral, conditioned supply agreements and privileged partnerships, such as those seen prior to the oil crises of the 1970s, will once again come to the fore." See the footnote above.
- 58 Terry Macalister and Lionel Badal, 2010, "Peak oil alarm revealed by secret official talks. Behind government dismissals of 'alarmist' fears there is growing concern over critical future energy supplies," *Observer*, Sunday August 22.
- 59 In response to the FoI Act inquiry by the *Observer*, the official government letter stated: "We recognise the public interest arguments in favour of disclosing this information. In particular we recognise that greater transparency makes government more open and accountable and could help provide an insight into peak oil. However any public interest in the disclosure of such information must be balanced with the need to ensure that ministers and advisers can discuss policy in a manner which allows for frank exchanges of views and opinions about important and sensitive issues." See the footnote above.
- 60 "Joint Operating Environment 2010," United States Joint Forces Command.
- 61 Daniel Yergin, 2011, "There will be oil," *Wall Street Journal*, September 18.
- 62 Speaking about peak oil predictions Daniel Yergin observed: "This is actually the fifth time in modern history that we've seen widespread fear the world was running out of oil. The first was in the 1880s, when production was concentrated in Pennsylvania and it was said that no oil would be found west of the Mississippi. Then oil was found in Texas and Oklahoma. Similar fears emerged after the two world wars. And in the 1970s, it was said that the world was going to fall off the "oil mountain." But since 1978, world oil output has increased by 30%." See the footnote above.
- 63 This is known as Herbert Stein's (1916–99) law, a highly regarded economist chairman of the Council of Economic Advisers under President Nixon and President Ford. Herbert Stein was known for his sense of humour.
- 64 Lucius Anneaus Seneca, c-65AD, *Seneca's Letters to Lucilius*, no. 91.
- 65 The best-known paper by Kudryavtsev published in Russian was "Against the organic hypothesis of the origin of petroleum," *Petroleum Economy*, 1951, 9, pp 17–29.
- 66 Mikeal Höök, op.cit.

## *Oil Processing*

Understanding the processing of oil is very important for any participant in the energy markets. The demand for oil is derived from demand for refined products. At any given point in time, a refinery can be best described as using what is known in economics as a fixed coefficient model (or a fixed coefficient production function). The chemical characteristics of oil available to a refinery are given, and the way a refinery is configured determines the basket of outputs. Over a longer time period, measured in months, a refinery can acquire additional flexibility by rearranging its operations (US refineries go through seasonal reconfiguration twice a year), and / or modifying the mix of crudes it buys. The ability to modify a production profile is often limited and varies from refinery to refinery. The technological flexibility of a refinery can be enhanced in the long-run through investments in additional specialised units, which will be described below. It is important to recognise that, while any refinery has limited ability to modify its production profile, the industry collectively has much more ability to adjust to fluctuations in demand through changes in the utilisation levels of refineries, trade flows between different regions and reliance on inventories. An analyst supporting a trading desk should understand not just the operations of a typical refinery, but also changes in refining capacity worldwide, global trends in demand for refined products and changes in oil output, both in terms of level and quality.<sup>1</sup> For example, the addition of refinery units that can enhance the industry's capacity to process sour and heavy crudes increases the demand for lower quality crudes, and this in turn reduces price differentials between different grades of oil. Shocks to prices of one grade of crude typically propagate over time to other crudes,<sup>2</sup> creating opportunities for forward-looking traders to design creative trading strategies.

We will start the chapter with a walk through a modern refinery and a review of the different processes that convert crude into a basket of commercially valuable products. (It is worth noting that this section, and the rest of the chapter, has failed to satisfy any of our friends who kindly helped me with their comments. We received comments that the chapter dangerously oversimplifies many of the technological issues related to oil processing, or that it is too sophisticated, technical and heavy. The solution has been to leave the chapter as is.)<sup>3</sup> We follow this section with a discussion of the tools used to characterise and describe refinery operations, such as the Nelson's Complexity Index, refining margins and tools used for managing refinery operations (different optimisation packages). The second part of the chapter will cover properties and technological processes related to certain refined products, such as fuel oil and gasoline. Understanding the issues related to gasoline blending and the use of oxygenates is critical to understanding seasonality and the dynamics of gasoline prices – and, by extension, crude prices and prices of gasoline additives. Most car drivers in the US do not even realise that they use two different blends of gasoline: summer and winter gasoline, and this is the best testimony to the efficiency of the distribution system of this important commodity.

## REFINING

As explained above, crude oil is a mixture of different hydrocarbons and other chemical substances – including water, sediments, heavy metals and sulphur compounds. Oil extracted from underground rock formations has practically no commercial value (it has marginal uses as a fuel<sup>4</sup> or primitive lubricant) and has to be transformed through a number of processes, jointly known as refining, into commercially useful products.<sup>5</sup> For history buffs, the first refinery was opened in Poland in 1854 (or 1856 according to some sources), followed by the first commercial large-scale refinery in Romania in 1858.

Crude oil is the most important (although not unique) input used by refineries. Other inputs include natural gas liquids and other hydrocarbons (such as synthetic crude, which was described in the section on non-conventional oil), additives and blending components (such as oxygenates) and the so-called refinery backflows from petrochemical plants. Petrochemical plants are important buyers of

refined products that are used as inputs in the production of thousands of different products critical to a modern society, including plastics, fertilisers, paints and lubricants. A portion of the feedstocks received from refineries is returned back to refineries after processing and extracting the inputs required for production of different chemicals. Refineries are also important users of electricity and natural gas, which is used as fuel for inside-the-fence power plants or to produce the heat necessary to support refining processes. The list of refined products is a long one, with the most important categories being:<sup>6</sup>

- refinery gas;
- ethane;
- liquefied petroleum gases;
- naphtha;
- aviation gasoline;
- gasoline type jet fuel;
- unleaded gasoline;
- leaded gasoline;
- kerosene type jet fuel;
- other kerosene;
- transport diesel;
- heating and other gasoil;
- resid fuel: low-sulphur content;
- resid fuel: high-sulphur content;
- white spirit<sup>7</sup> + specific boiling point (SBP) spirits;
- lubricants;
- bitumen;
- paraffin waxes;
- petroleum coke; and
- other products.

The objective of refining is to convert a given basket of crudes that are available to a refinery into products with the highest commercial value. The composition of a basket of refined products is a function of the crudes used as feedstocks and the design of a refinery, combined with the skills of the chemical engineers who seek to maximise the output of highly priced products. The main challenge faced by the petroleum industry is the falling output of the most

---

valued crudes (light and sweet varieties) that can be processed into baskets of highly valued components without the need to use more expensive technologies. Growing a relative supply of heavy and sour crudes requires the construction of more complex and costly refineries, with equipment supporting processes that go beyond simple distillation.

Refining is a sequence of very complex and interacting processes that will be briefly described below. This section should be viewed as merely an introduction to a very complex topic, and the author does not claim to be a subject expert. What is presented has to be treated as an absolute minimum of what an energy trading professional has to master. Any trader or quantitative analyst responsible for hedging a specific refinery should educate themselves about its configuration and operations, and seek help from chemical engineers and plant managers. One has to understand how easily a refinery can be reconfigured and how much time it takes, what is the range of crudes it can use, what are the products it can deliver and what constraints are created by existing supply and sales contracts, as well as its storage and transportation infrastructure. Many refinery managers are ideologically opposed to hedging because they worry about the potential risk of outages. These may be very costly if a refinery is hedged and then it goes down. What follows is a double blow. Not only the hedges are not being offset by operations-related flows, but prices may change to the disadvantage of a refiner. For example, an outage may cause the spike in the local prices of refined products. A refinery selling forward products could have to cover at a loss.

A full schema, even of a relatively simple refinery, looks like a chart of the motherboard of a personal computer. The main processes taking place in a refinery include:

- distillation;
- sulphur removal;
- conversion;
- upgrading; and
- product blending.

Other sources include additionally preprocessing and certain support processes (such as, for example, generation of electricity and heat, production of steam and hydrogen, waste water treatment).

## Preprocessing

Crudes often arrive at a refinery from many destinations and vary with respect to sulphur content and the degree of contamination with water and other substances, such as H<sub>2</sub>S or mercaptans (organic compounds containing sulphur). Sweet and sour crudes are stored and processed separately to avoid contamination. The separation process involves removal of salt, water and other impurities at its initial stage. The separation processes also involve the removal or neutralisation of such harmful substances as sulphur and metals, especially as metals like copper, vanadium, nickel and gases like nitrogen can damage the catalysts used in many processes after distillation takes place. This is also true of sulphur that has to be removed for additional reasons (restrictions on the content of sulphur in many refined products, such as diesel). Sulphur removal happens also after crude is distilled into different fractions. Two techniques frequently used are Merox treatment (neutralisation of sulphur by converting it into sulphides) and hydro treatment, this involves mixing refined products with hydrogen (hence the name) and passing it over a catalyst (cobalt or molybdenum) under high pressure and temperatures.

## Distillation

Distillation is a process of separating different components of a liquid by taking advantage of differences in their boiling points. This process has been known since time immemorial and is not technologically very complicated, as any moonshiner might admit. The challenge consists in carrying out this process in the most cost-efficient way and avoiding mixing different compounds with close physical characteristics. This is usually accomplished by increasing the height of the distillation columns.

In a refinery, distillation involves the separation of different components of oil (called product fractions or cuts). The cuts are not homogeneous and represent mixes of different hydrocarbons with close physical and chemical qualities that will be further transformed using processes best adapted to their specific characteristics.

The process of evaporation is followed by condensation and removal of different fractions from a distillation column. The process of distillation may be carried continuously or in batches. In the case of the continuous process, oil is heated in a furnace and passed to a

distillation column, which is the tallest structure in a refinery. The fuel used to heat oil may be either natural gas (this is why many refineries consume a lot of natural gas) or certain products of the refining process which have relatively low market value compared to other outputs, or which may be difficult to store and transport (ethane, propane, butane, resid). The most volatile fractions rise to the top of a tower, with other fractions being captured at the lower levels. The heaviest fractions collect at the bottom and are known as a straight-run residue. As the vapours rise through the tower, they are captured in horizontal plates (trays) and removed from the tower for further processing. The trays, with their mushroom-shaped caps, are designed to catch the liquids but allow the gases to pass (bubble up). They are used to control the heat-exchange process and make it efficient enough to reduce the height of distillation towers to manageable dimensions.

Distillation is effectively a process of heat and mass transfer, with lighter fractions rising to the top, collecting lighter molecules, and with descending liquids collecting heavier molecules as they drop to lower trays. This process may be repeated for higher efficiency with some of separated products (called reflux) returned back to the tower to cool down the vapours and facilitate condensation. Given the differences in the boiling points, different fractions may be removed from the tower at different points, from the bottom to the top.<sup>8</sup> The process of continuous distillation happens typically at atmospheric pressure. An alternative is to run distillation at reduced pressure (vacuum distillation). This technology, used primarily to process further resid – ie, the heaviest fractions obtained from atmospheric distillation tower (long resid) – has many advantages as it may require less energy. Boiling happens when the pressure of the liquid exceeds ambient pressure. Lowering the ambient pressure allows the reduction of the temperature to which oil has to be heated and the number of stages in the distillation process. Lower temperatures not only save energy and reduce the cost of the process, but also prevent chemical changes in some components of the feedstock: the break-up (cracking) of certain molecules. When cracking happens, it is more difficult to control the distillation process. Another reason for vacuum distillation is the ability to reduce the height of the distillation tower. Under atmospheric pressure, resid would have to be heated up to 1,000°F (~540°C), and this would require a very tall

distillation tower. Vacuum distillation requires temperatures of about 400°F.<sup>9</sup>

Flashing is a specialised distillation process applied to straight-run residue under low pressure (4.5–5.5 psi) and temperatures in the range of 1,000–1,100°F (~540–590°C). The outputs from this process are called flasher tops (light flashed distillate and heavy flashed distillate, used as lubricating oil and feedstocks for other processes). Flasher bottoms are used in production of asphalts and/or blending components for residual fuel oil. Stripping is the process of further separation of fractions removed at the different levels of the distillation column. Strippers are columns next to the main distillation tower where lower fractions are removed using steam injected against the flow of the hydrocarbons.

In the jargon of the industry, the products of the distillation process are called straight-run oils and gases (only physical separation processes are involved).<sup>10</sup> What follows is a brief review of the main fractions, which will explain why refining seldom stops at simple distillation. The boiling ranges of different products vary and also overlap:<sup>11</sup>

- LPG: <99°F
- gasoline: 90–400°F (32–204°C)
- naphtha: 150–300°F (66–149°C)
- kerosene: 300–480°F (149–249°C)
- No. 2 fuel oil: 325–750°F (163–399°C)
- No. 5 and 6 fuel oil: 600–1,000°F (316–538°C)

The industry also uses the generic term distillates. The light distillates include, for example, LPG, naphtha and gasoline. Middle distillates include kerosene and diesel.

LPG fractions include hydrocarbons with one to five carbon atoms, ranging from methane, through ethane, propane, butane to pentanes. These hydrocarbons are sold to the chemical industry or are used inside the refinery as fuel or inputs to other processes. Fuel is required as a source of heat for electricity production and for a number of other processes (preheating oil, steam generation, etc). Butanes and iso-butane is sent to the alkylation unit. The separation of LPGs takes place in the distillation columns operating on the same principle as the main crude distillation unit. The gas plant also

receives inputs from the catalytic cracker and alkylation unit (see below for an explanation of this).

Gasoline produced through the distillation process (also called straight-run gasoline, SRG) represents only a fraction of that produced in a modern refinery. It is typically low octane (see the next section for a discussion of the octane number) and requires upgrading. Increasing the output of gasoline beyond simple distillation requires additional processing and investment in fairly capital- and energy-intensive units.

Naphtha is used as a blending component of gasoline but its quality is generally lower than SRG. An alternative is to use it as a feed into catalytic reformers. The output, called reformate, can have an octane number exceeding 90, compared to 35–40 in the case of naphtha. Kerosene is used as jet fuel or is blended into diesel. Potential sulphur contamination becomes more serious as we reach this or higher fractions. Sulphur tends to attach to larger molecules, so it is necessary to run naphtha through a desulphurisation unit. Fuel oils and gasoline will be discussed in more details later, given their commercial importance. What is left at the bottom of the distillation column is called resid, which can be used as asphalt or roofing tar. It can be further processed through the unit called a coker, which converts resid into higher-value products, such as gasoline or middle distillates.

The distillation process separates feedstock crudes into different components but does not change their chemical properties. This means that the chemical characteristics of the crude delivered to a refinery determine entirely what is produced. Product yield depends on the crudes used as inputs to the process, with light sweet crudes producing the most valuable baskets of hydrocarbons.<sup>12</sup> As a general rule, higher gravity is associated (lighter crude) with a lower percentage of heavy fractions. Unfortunately, given upstream production trends, the structure of demand increasingly diverges from the proportions of different hydrocarbons available from produced crudes through a simple distillation process (not enough light fractions, too many heavy fractions).

### **Desulphurisation**

Hydrodesulfurization (HDS) units remove sulphur from the products of the distillation process. Hydro treatment<sup>13</sup> is a process used

for cuts containing naphtha, kerosene, gas oils and atmospheric resids and its objective is removal of impurities, primarily sulphur, but also nitrogen, oxygen, olefins, halogens and heavy metals (lead, copper, nickel), which can damage catalysts used in the refining processes (described below). The process uses hydrogen and takes place in the presence of a catalyst, under conditions of high pressure and temperature (625–698°F (~330–370°C)). The upgraded products are raw distillates that are marketed as kerosene, jet fuel, low sulphur and ultra-low sulphur diesels, or can be used for further processing. The by-products are hydrogen sulphide (from hydrogen reacting with sulphur), ammonia (from hydrogen reacting with nitrogen) and free metals. Hydrocracking is an advanced version of hydro treatment that involves some conversion (explained below). The feedstocks to this process have to be hydro treated first. The catalysts used include palladium, tungsten and platinum placed on the bed made of aluminium silica.

Hydro treatment and hydrocracking are refining processes of increasing importance. This can be explained by the growing relative supply of sour crudes, the push towards low sulphur products (ultra-low sulphur diesel, low sulphur resid) and the increasing importance of conversion processes that require expensive catalysts that can be easily poisoned by sulphur or nitrogen. The weak link in this process is hydrogen supply. Although hydrogen is produced by a number of refining processes, most refineries run a deficit of this feedstock.

Ultra-low sulphur diesel (ULSD) has been mandated by a number of countries, although definitions vary (typically, ULSD is defined as diesel with a sulphur content between 5–50 ppm). In the US, as of December 1, 2010, all highway diesel which is allowed is low sulphur (15 ppm), with the same standard being extended to railroad locomotives and marine engines by 2012. Some exemptions that have been granted will expire by December 1, 2014.<sup>14</sup> The rules promulgated in the US, the EU and many other countries also apply to the level of aromatics. Aromatics are associated with low quality of fuel (poor ignition, low cetane number, increased smoke point of jet fuel). The aromatics contribute to emission of particulates that are carcinogenic.

### Conversion and upgrading

Conversion and upgrading include processes that change the chemical structure of certain hydrocarbons. This may include breaking up big molecules into smaller ones (cracking), combining smaller molecules into bigger ones (for example, alkylation) and isomerisation, which leaves the number of the atoms of carbon and hydrogen in a molecule the same, but changes the way they combine.

More complex refineries go beyond simple distillation and use cracking, by breaking carbon–carbon bonds under conditions of high pressure and temperature, sometimes in the presence of a catalyst (such as zeolite-based catalysts or alumina). There are several types of cracking technologies: thermal cracking, hydro cracking, steam cracking and catalytic cracking. Thermal cracking is a residue-conversion process carried out at temperatures exceeding 750°F (400°C), leading to the break-up of some longer molecules into smaller ones. An additional objective of this process is reducing the viscosity of the residue (this is why these cracking units are sometimes referred to as visbreakers).

Catalytic (or cat) cracking is the process of cracking taking place in the presence of catalysts such as zeolite and alumina. The catalyst is introduced to the reactor vessel, where it is kept suspended by pressurised air pumped into the unit creating what is called a fluidised bed (this is why this unit is called a fluid catalytic cracker, FCC). Otherwise, the catalyst would settle at the bottom and its exposure to the feedstocks would be limited. This process produces heat that can be used elsewhere in the refinery. FCC chemical reactions occur at temperatures of 950–1,020°F (510–550°C) and pressures of 10–30 psi. The feeds to the process include atmospheric and vacuum gasoils, such as heavy distillates, flasher tops and cycle oil.<sup>15</sup> The outputs are cat-cracked gasoline, cat-cracked fuel oils, C4 and lighter fractions, as well as coke.<sup>16</sup> Hydro cracking is thermal or catalytic cracking with hydrogen added. This process is used to maximise the yield of jet fuel and diesel. One of the advantages of this process is lower sulphur level, due to the presence of hydrogen, which binds with sulphur.

The outputs from the FCC unit are used as inputs in the isomerisation, alkylation and hydro-cracking units. The distillates and gasoline have to be typically passed through sulphur removal units. Given the range of high-value products that can be produced

through the addition of the FCC, it is one of the critical units of a modern refinery. Another aspect of the cracking process is volume gain (also called refinery gain). As big molecules are broken into smaller ones, they are packed less effectively and the produced volumes expand. The mass remains unchanged, as the law of preservation of matter still holds. Refineries sell products by volume – therefore, the volume gain works to their advantage.

Coking is a process of severe thermal cracking, under high temperatures and high-pressure conditions. The objective is to increase the output of lighter oils and petroleum coke. The latter product is a solid substance consisting of relatively pure carbon that can be used as a fuel or as an input to a number of manufacturing processes (dry cells, electrodes). Petroleum coke can be used in coal-fired boilers installed in power plants or industrial plants. Unfortunately, it typically has a high sulphur content that poses environmental problems.

Alkylation is the opposite of cracking and consists in compressing smaller molecules into bigger ones. The reason for using this process is to convert products that are too light to be used in the gasoline pool. Its primary objective is production of alkylate, a valuable blending component (with an octane number of about 95) for higher-grade gasoline. The process combines propylene, butylenes and sometimes pentylenes with iso-butane in the presence of an acid catalyst (either hydrofluoric or sulphuric acid), forming branched, saturated molecules, primarily isopentane and iso-octane. The use of two different catalysts explains why the alkylation units may be referred to as a sulphuric acid alkylation unit (SAAU) or a hydrofluoric alkylation unit (HFAU). The output from the alkylation unit has to be distilled to separate alkylate from other hydrocarbons, such as propane and butane.

Reforming units are ubiquitous in modern refineries, and it is difficult to find one that does not support this process. Reforming is a process of converting naphthenes to aromatics, increasing the aromatic content of naphtha and improving the octane number of the output. There are many different reforming technologies: plating, hydroforming, power forming, ultra forming and catalytic reforming. Isomerisation changes paraffinic molecules to isomeric forms for better gasoline properties (primarily a better octane number). Isomerisation results in products that have a high Reid

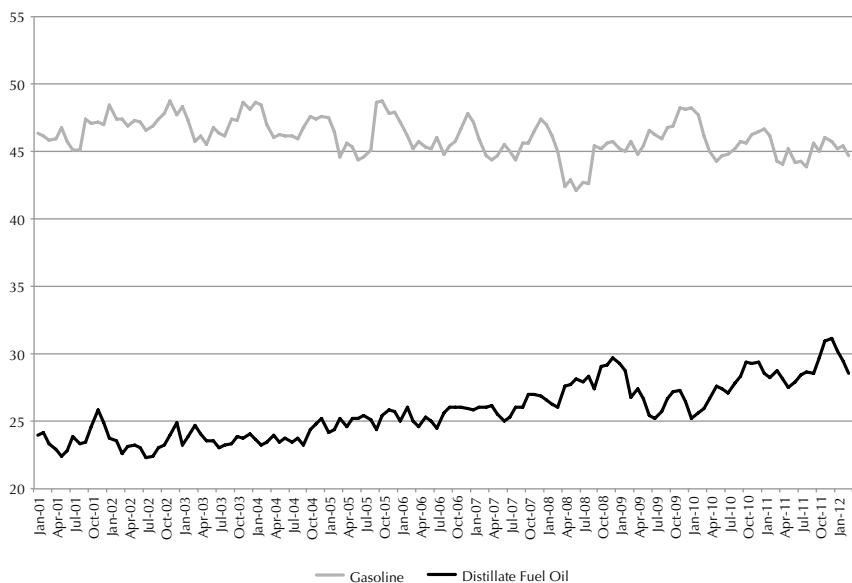
Vapor Pressure (RVP) level (see the section on gasoline blending later in this chapter for an explanation), and limits their potential use in gasoline blending. Both processes described above result in volume reduction.

The yields of US refineries are reported by EIA. Figure 15.1 shows the evolution of percentage yields over the period 2001–12 (March) for gasoline and distillate fuel oil. The numbers demonstrate the results of using the more complex processes described above, not just simple distillation.

### Refinery complexity index

As one can see, refineries can vary from relatively simple installations relying only on distillation to very complex operations comprising multiple technological processes. They can be ranked using a complexity index designed in 1960 by Wilbur L. Nelson. This index is calculated annually by the *Oil & Gas Journal*. Wilbur Nelson assigned a complexity factor of one to the basic atmospheric distilla-

**Figure 15.1** US refinery yields (2001–12, gasoline and distillate fuel oil, percentages)



Source: U.S. Energy Information Administration

tion unit, with additional complexity factor assigned to each major refining unit, based on its complexity and cost relative to plain vanilla distillation. This number is designed to capture both the cost and value addition. The higher the unit complexity number, the more technologically advanced is the unit, and the more valuable its products. (It is, however, important to avoid simplifications and identify more complex refineries with more profitable ones. Many other factors, including the cost of crude oil and access to attractive markets, also matter.) The complexity of each refining unit is estimated by multiplication of its complexity number by its throughput ratio (expressed as percentage of basic distillation capacity). Table 15.1, adopted from a sample produced by the *Oil & Gas Journal*, shows the calculation of the index for a specific refinery.

Column (2) in the table contains complexity factors assigned to each unit. The calculation of the Nelson Complexity Index is carried out in last column on the right. Complexity factors are multiplied by unit capacity, the products are summed up and the sum divided by

**Table 15.1** Nelson Complexity Index calculation

Unit (1)	Complexity factor (2)	Capacity (3)	= (1) X (2)
Distillation	1	<b>403</b>	
Vacuum distillation	2	200	400
Viscraking	2.75	0	0
Delayed and fluid coking	6	27	162
Catalytic cracking	6	90	540
Catalytic reforming	5	60	300
Catalytic hydro cracking	6	20	120
Catalytic hydro refining	2.5	250	625
Catalytic hydro treating	2.5	0	0
Alkylation/polymerisation	10	0	0
Isomerisation	15	21	315
Lubes	60	0	0
Oxygenates	10.0	0	0
H2	1.0	0	0
Coke	6.0	1.6	9.6
Sulphur	240.0	0.46	110.4
Asphalt	1.5	0	0
Total			2582
Nelson Complexity Index			7.406948

Source: 2010 OGJ Worldwide Refining Survey

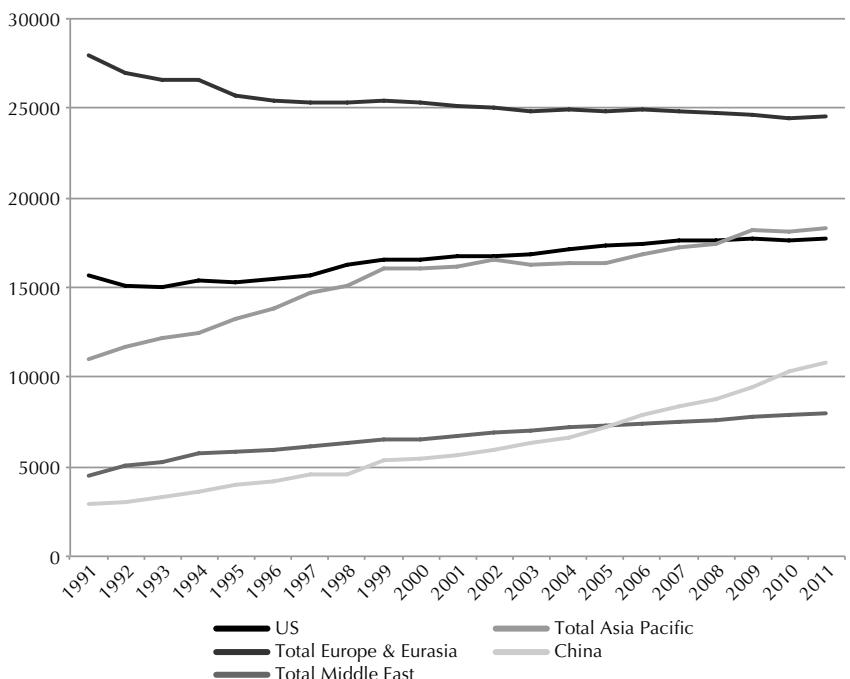
distillation capacity (403), yielding 6.4. By convention, one is added to this result, hence the result of 7.4 in Table 15.1.

### World refining capacity

This section contains some basic facts about existing refining capacity and prevailing trends in terms of trends in overall levels and geographical distribution. Figure 15.2 shows the evolution of refining capacity in the US, China, Middle East, Europe and Eurasia, Asia and Pacific (excluding China) from 1991 to 2011. The shortcoming of this information is that it is based on atmospheric distillation capacity only.

The reader may find more information on this topic in a study by Purvin & Gertz about world refining capacity available as of end 2007.<sup>17</sup> The report contains data on the refining capacity breakdown

**Figure 15.2** Refining capacity (1991–2011, thousand barrels daily)



Source: BP "Statistical Review of World Energy 2012"

by region (Europe, Russia/CIS,<sup>18</sup> Middle East, Africa, North America and Latin America). There are several trends here that require a brief comment. The number of refineries decreased considerably between 1995 and 2007, primarily in Europe and North America, with a small increase in the Middle East and India. At the same time, all regions except for Russia witnessed an increase in average refinery capacity (as demonstrated in Table 15.2). This trend reflected the efforts to exploit economies of scale by concentrating production in the most efficient refineries and closing old, inefficient and environmentally obsolete operations. A corresponding trend was the so-called capacity creep observed in the US and Europe: the ability to increase production through expansion and upgrades of existing refineries, without a greenfield expansion. In Russia, some old refineries were closed as the region went into an economic freefall following the collapse of communism. In some refineries, excess distillation capacity was reduced without decommissioning a refinery. This resulted in the reduction of the average refinery size, with a corresponding increase in the complexity of an average plant.

Another important difference between the regions is the varying level of refinery capacity structure. In North America and Europe, coking and cracking refineries represent 99% and 83% of the overall capacity, respectively. In Russia/CIS, this percentage is only 62%.

Refinery economics studies rely on the examination of the margins. The generally accepted approach to the calculation of refinery margins is based on a method developed by the IEA.<sup>19</sup> The calculations start with the gross product worth (GPW), obtained by multiplying the spot price of each of the refined products (less an allowance for refinery fuel and loss) by the yield obtained from the total barrel of crude. This number is adjusted by the following cost components:

- cost of crude;
- transport costs:
  - marginal crude freight, insurance and ocean loss (in case of an FOB crude),<sup>20</sup> and
  - applicable fees and duties.<sup>21</sup>
- credit allowance (the cost of financing crude in transit, between the purchase date and the receipt date);
- variable refinery operating costs:

- electric power;
  - water;
  - chemicals, additives, catalysts; and
  - refinery fuels (beyond own production).
- fixed refinery costs:
- labour;
  - maintenance;
  - taxes; and
  - overhead costs.

The calculations are based on many assumptions and calculated refinery margins may diverge from actual results of a specific operation (especially in the case certain exposures are hedged). Nevertheless, the study of refining margins is a useful exercise for any trader, as it helps to identify the stress points in the system. Figure 15.3 shows refining margins for the US Gulf Coast refineries for different crudes. It is difficult to find a better illustration of a very well-known fact: refining is a feast-and-famine business.

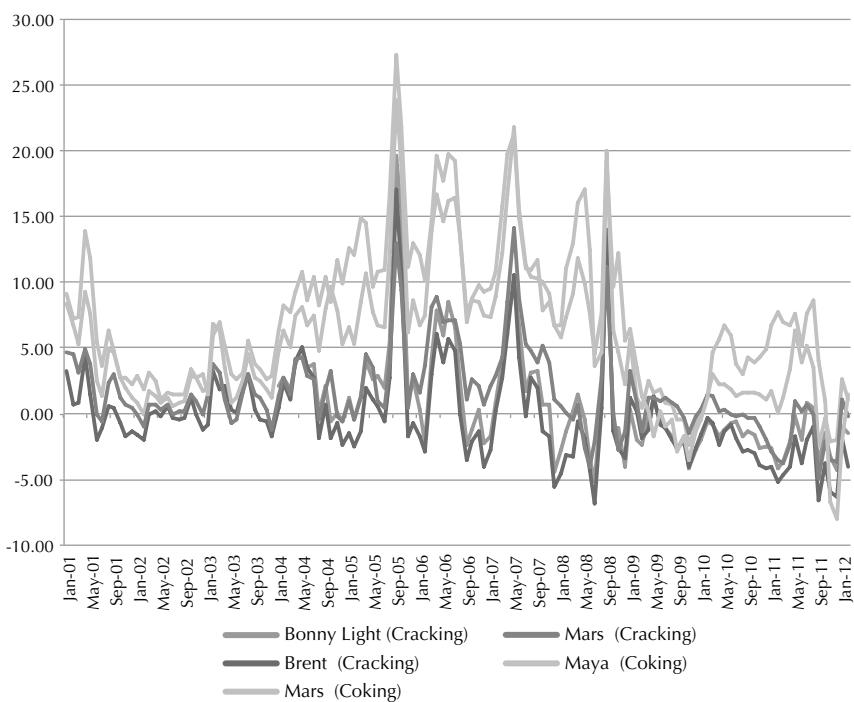
### Refinery optimisation

Refining has a bad reputation as a business. The normal refinery cycle can be described as many lean years followed by a few years of high refining margins. Refiners are trapped in a vice of more stringent regulations related to product quality, environmental impact of their operations and changing crude supply structure, with a shift towards the less attractive (heavier and more sour crudes). Adjusting to revised product specifications and a less-attractive crude slate requires significant investment outlays, lowering profit margins. The challenging environment in which refineries now operate enforces a

**Table 15.2** Average refinery capacity (thousands of tons/year)

Region	Europe	Russia/ CIS	Middle East	Africa	North America	Latin America	India	China	Other Asia
1995	5,220	8,722	6,266	2,976	4,410	4,510	4,226	3,394	6,155
2007	6,063	6,834	7,376	3,392	5,909	5,589	8,561	6,651	7,897
	843	(1,888)	1,110	416	1,499	1,079	4,335	3,257	1,742

Source: Study On Oil Refining and Oil Markets, Purvin&Gertz, January 2008. Study prepared for European Commission. [http://ec.europa.eu/energy/oil/studies/oil\\_en.htm](http://ec.europa.eu/energy/oil/studies/oil_en.htm)

**Figure 15.5** Refinery margins, US Gulf Coast (USGC)

Source: oilmarketreport.org (IAE)

need for disciplined, effective management, exploiting every opportunity to increase profitability. Several operational strategies have been identified in the literature.<sup>22</sup>

- ❑ Exploiting the economies of scale (typically, achieving size in excess of 300,000 barrels per day).
- ❑ Increasing complexity factor (as explained in the previous section, more complex refineries have the enhanced ability to convert less-valuable hydrocarbons into higher-priced products, such as gasoline, jet fuel and petrochemical feedstocks).
- ❑ Developing a link to a petrochemical plant in the proximity of the refinery. Such links can be established in huge industrial complexes – for example, the Houston Ship Channel.
- ❑ Acquiring the ability to process multiple crudes and produce a diversified basket of products. This allows a refinery to take

advantage of changing price relationships to optimise the mix of inputs and outputs, and to avoid outages related to the disruption of the supply of crudes from certain locations. A refinery configured to process very narrowly defined crude (for example, the light sweet crude of very high quality) would have to shut down if supplies were curtailed or would suffer losses if the prices of the crude spiked.

- ❑ The installation of a sophisticated controls system for the management of physical flows between different units and/or storage tanks, optimisation of the physical processes and monitoring in real time of the pressure and temperatures inside the refinery units. Such systems are known in the industry as:
  - distributed control system (DCS);
  - internal model control (IMC);
  - model variable control (MVC); and
  - model predictive control (MPC).
- ❑ Use of optimisation techniques to select the best mix of inputs and outputs, subject to the operational constraints, determined by the design of the refinery, safety considerations and environmental laws. The remaining part of the section will review the use of optimisation techniques in refinery operations.

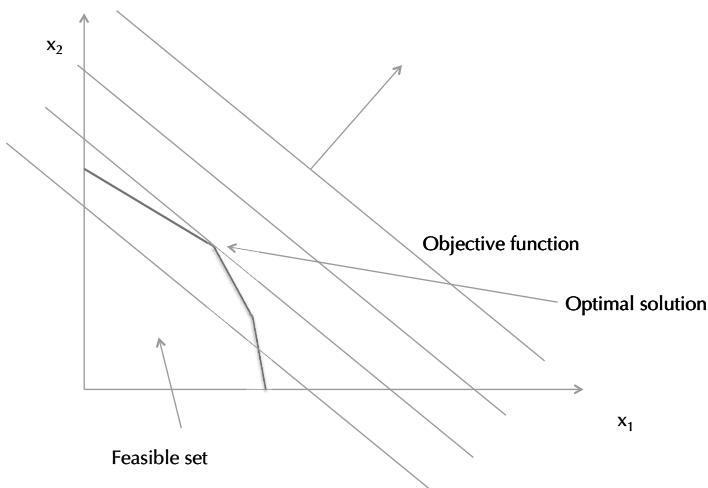
The first optimisation technique applied in practice to the solution of large-scale problems was linear programming (LP). The boost to the development of LP was the need to optimise the allied convoys during the Second World War, details of which remained classified until 1947.<sup>23</sup> An LP problem can be summarised as follows:

$$\max c = c_1x_1 + c_2x_2 + \dots + c_nx_n \quad (15.1)$$

subject to the following  $m$  constraints:

$$\begin{array}{l} \frac{a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \leq b_1}{a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \leq b_m} \end{array} \quad (15.2)$$

In addition, a solution to the LP programme obeys non-negativity constraints (ie, all the  $x_i$  ( $i = 1, \dots, n$ ) are positive). Equation 15.1 is the objective function which is maximised,<sup>24</sup> with the constants  $c_i$  called the weights, and  $x_i$ 's being the decision variables. In practical applications, the weights can be, for example, prices or unit costs, and the objective function can represent the profit or the total cost. The set of  $m$  inequalities captures the operational constraints specific to a given

**Figure 15.4** An LP problem in two dimensions

problem.<sup>25</sup> A standard LP problem can be illustrated graphically in a two-dimensional case (two decision variables). Figure 15.4 shows the feasible set (ie, the combinations of decision variables that are viable given the constraints of the specific problem). A set of parallel lines corresponds to different levels of the objective function, with the arrow pointing in the direction of its increase. The optimal solution happens typically at one of the vertices of the feasible set (a set of optimal solutions may happen along one of the edges of the feasible set). The most widely used algorithm for finding a solution to an LP problem, the simplex method, is a procedure for moving from one vertex of the feasible set to another, always in the direction of increasing value of the objective function.

For those who are familiar with this technique, this is trivial. For those introduced to this method for the first time, it may seem very abstract and detached from practical problems. It is, however, impossible to run a modern refinery without extensive use of this tool. A typical LP programme for a refinery may have as many as 3,500 decision variables and 1,500 constraints. The field is dominated in the US by a number of firms that developed sophisticated optimisation programmes customised for refinery use. The most important contributors include:

- ❑ RPMS<sup>26</sup> (Honeywell's Refinery and Petrochemical Modeling System);
- ❑ PIMS<sup>27</sup> (Aspen Technology); and
- ❑ GRTMPS<sup>28</sup> (Haverly Systems).

In the past, refineries would use a number of complimentary models to optimise different segments of their operations: crude acquisition and transportation, refining, product blending and product distribution. Several trends in this area merit closer attention.

- ❑ The trend towards more complex models covering jointly different components of the refining value chain. Implementation of the comprehensive models is made possible by increasing computing power and progress in information technology: the ability to collect in real time information related to the levels of inventories, the status of different refinery units, location of the tankers and other relevant information. The model formulations extend over multiple time periods, in contrast to the early models that were static – ie, covered a single period, without dynamic links from one time interval to another.<sup>29</sup>
- ❑ Application of more sophisticated solution algorithms, including integer programming and non-linear formulations. Integer programming restricts certain decision variables to assume integer values only. For example, a given delivery truck may be assigned to just one route between a refinery and a distribution terminal. The corresponding decision variables are restricted to a set of integers. Non-linear optimisation recognises that many chemical and physical processes taking place in a refinery are non-linear.<sup>30</sup> More specialised algorithms are used to accelerate computations, in view of certain structural properties of the refinery optimisation models.<sup>31</sup>

A good example of an advanced refinery model can be found in a paper by Adebayo Alabi and Jordi Castro (see footnote 29). The model is somewhat abstract, but it illustrates well the trend towards integrated refinery modelling signalled above. A person interested in optimisation limited to refining processes should consult the book by J. Aronofsky *et al* (cited also in footnote 29), which remains the best

source on the topic in spite of having been published more than 30 years ago.

## REFINED PRODUCTS

This section will provide a review of certain refined products, important from the point of view of a trading operation. This is by no means an exhaustive analysis of highly complex and specialised markets for this class of energy commodities, but hopefully will prove to be a good starting point for an interested analyst or a trader.

### Fuel oils<sup>32</sup>

Fuel oils are, next to gasoline, the most important of refined products. The details of the refining process were covered earlier. Fuel oils have been classified by the American Society for Testing and Materials (ASTM) as fuel oil No. 1, 2, 3, 4, 5 and 6. The physical characteristics used in classification of different fuel oil include:

- flash point;
- water and sediment content;
- physical distillation or simulated distillation;
- kinematic viscosity;
- Ramsbottom carbon residue;
- ash;
- sulphur;
- copper strip corrosion;
- density; and
- pour point.<sup>33</sup>

No. 1 fuel oil is used for domestic heating and as a light diesel fuel oil. It is a distillation fraction produced after gasoline fraction is obtained. No. 1 and kerosene are marketed as one product and are practically undistinguishable.<sup>34</sup> No. 2 fuel is called heating oil in the US and gasoil in Europe. The term gasoil is a historical artifact going back to the times when this was an input to production of gas used for home heating. No. 2 can be produced through distillation (in this case it is called straight-run distillate) or through the catalytic-cracking process (catalytically cracked distillate). No. 2 is used as a fuel for domestic and industrial boilers or as diesel (hence the term road diesel). The industry distinguishes between domestic and industrial No. 2 fuel

oils. Domestic fuel oil is typically obtained from the straight-run stream, is lighter, and is used for home heating or as diesel oil. Industrial distillate is typically the product of catalytic cracking.<sup>35</sup>

No. 3 is seldom used. No. 4 is used primarily for small boilers serving schools, hospitals and apartments, and in the industrial boilers that cannot use high-viscosity oils such as No. 5 and 6. No. 4 is produced from straight-runs, through cracking or by blending residual (No. 5 and 6) with distillate (No. 1 and 2).

No. 5 and 6 fuel oils are called residual fuel oils (shortened as resid), or heavy fuel oil. No. 5 is typically obtained through the blending of No. 6 with distillate oils. Light No. 5 has low viscosity and can be used without preheating (unlike heavy No. 5). Given that much more No. 6 oil is produced than No. 5, the term resid is often used exclusively in reference to the first type of fuel oil. No. 6 is also known as Bunker C fuel.<sup>36</sup>

### **Physical and chemical properties of fuel oils<sup>37</sup>**

API gravity is reported by convention at the standardised temperature of 60°F. If the actual temperature of fuel oil is different, one can translate the measured API gravity into a standardised value. One way to accomplish this is to use a translation table available for API which is included in the *Manual of Petroleum Measurement Standards*.

The API gravities (defined above) of fuel oil decrease as their numbers go up, with No. 2 ranging from 26–39 API degrees, No. 4 from 24–32, No. 5 from 16–22 and No. 6 from 10–15. The API gravity is one of the most important characteristics of fuel oils, and can be used to infer many other properties. For example, the hydrogen content of fuel oil can be calculated using the formula:<sup>38</sup>

$$\text{Hydrogen content (\%)} = a - 15 \times d \quad (15.3)$$

where  $a$  is a number ranging from 24.50 to 25.45.<sup>39</sup> The  $d$  number in the formula is specific gravity which can be calculated from the API gravity using the formula:

$$\text{Specific gravity} = \frac{141.5}{\text{API gravity at } 60^{\circ}\text{F} + 131.5} \quad (15.4)$$

For example, if the API gravity of fuel oil is 15, the specific gravity is equal to 0.9659 and the hydrogen content (in percentages) is 10.51%. The remainder is mostly carbon, as the percentages of sulphur,

oxygen and nitrogen, ash and non-combustibles are quite low (and can be obtained from the product specification).

The percentages of carbon and hydrogen are important for the amount of heat released (or “liberated” in the industry jargon) during the combustion process. One pound of hydrogen produces about 62,000 Btus when burned, one pound of carbon: 14,600 Btus. Fuel oil containing a higher percentage of hydrogen will produce more heat from the same volume.

It is important to recognise that higher API gravity results in less heat value per unit of volume and higher heat value per unit of weight. Higher API gravity corresponds to lower specific gravity and higher hydrogen content (as one can easily see from formulas 15.3 and 15.4). This results in higher heat value per pound. However, higher API gravity lowers specific gravity, the weight per unit of volume decreases faster than the weight increases.

One can calculate an approximate heat value for a given type of fuel oil using the Dulong formula:<sup>40</sup>

$$V = 14600C + 62000\left(H - \frac{O}{8}\right) + 4000S \quad (15.5)$$

where  $V$  is the heating value per pound,  $C$  is the percentage of carbon,  $H$  is the percentage of hydrogen,  $O$  is the percentage of oxygen and  $S$  is the percentage of sulphur. In this formula, the implicit assumption is that oxygen contained in the fuel is combined with hydrogen having mass equal to  $\frac{1}{8}$ th of that of oxygen. For example, if the carbon content is 88%, hydrogen content is 10%, and oxygen and sulphur percentages are equal, the formula gives a higher approximate heat value than is the case, but is useful in practice for back-of-the-envelope calculations.

Viscosity is defined as the resistance of fluid to flow. This physical characteristic of fuel oils is important in practice, especially when it comes to transportation<sup>41</sup> and pumping. Viscosity is also important because industrial and residential equipment is often designed to operate with fuel of a certain viscosity. The use of fuel of an improper level of viscosity in a burner may result in efficiency losses. For example, a fuel oil with low viscosity is easier to pump but may flow through a burner at a high rate – and this can result in inefficient combustion, soot accumulation and smoke. Viscosity of fuel oil varies with temperature: at low temperatures fuel oil become

semisolid or even solid. Car tanks may be difficult to load and unload in winter as cool oil becomes more viscous.

### Other refined products

Kerosene (or kerosine) is a fraction heavier than naphtha, composed of paraffins, cycloparaffins and aromatics, produced through distillation and also in the cracking units. In the UK, as with many countries in the former English sphere of influence, it is called paraffin. Kerosene is used in many developing countries for cooking in portable stoves and as a source of light in portable lanterns. It is used as fuel to power small motors and heavy equipment such as tractors, and it is also used for space heating. The main use of kerosene is the production of jet fuel, after processing it to adjust the freezing point and burn point (jets fly at the altitudes where ambient temperature drops to minus 40°F).

Gasoil is a distillation fraction heavier than kerosene but lighter than lubricating oils, and in addition to atmospheric (atmospheric gas oil, or AGO) or vacuum distillation (VGO), it can be obtained from cracking or hydro-cracking units. In the US, gasoil is frequently referred to as fuel oils – and their properties have been described in more details above. European gas oil is a European and Asian designation for #2 heating oil and #2 diesel fuel. The futures contracts for gasoil include the ICE Futures Europe contract listed in London and the Nymex heating oil contract listed in New York.

Gasoline is a refined product critical to the modern road transportation system and, by extension, to our civilisation and standard of living. In the US, the transportation system depends practically entirely on liquid fuels, dominated by gasoline, diesel, jet fuel and biofuels such as ethanol.

Gasoline (called petrol in the UK and some other countries)<sup>42</sup> is a mixture of many hydrocarbons and additives that will be discussed below, and comes from many different stages of the refining process. Depending on the source, the terms in Table 15.3 are used in the industry.

The production and delivery of gasoline to the final user is a process with a very complex logistics that the public at large is almost completely unaware of. An average driver takes for granted that gasoline will be always available and will have the required quality. Cheap gasoline is also seen as a birthright. Periods of rapidly

increasing prices trigger a very adverse reaction, affect the results of elections and sometimes produce a competition in assigning blame to a wide range of the usual suspects. Technical specifications of gasoline are determined through the interaction of several constantly evolving factors, including:

- chemical and physical properties of hydrocarbons which are blended into gasoline;
- design of car engines;
- environmental regulations;
- regulations regarding car engine mileage efficiency; and
- local environmental conditions – such as, primarily, temperature and also pressure.

These different factors interact and sometimes improvement with respect to one criterion translates into lower scores with respect to other required qualities. A person without a background in science or engineering who condemns themselves to reading many technical documents on this subject develops a deep respect for energy industry professionals. We have to wonder why this story remains largely untold, and why the industry fails to educate its customers about the very complicated machinery that runs so smoothly.

At this point, we should just mention a number of environmental

**Table 15.3** Industry terms

Gasoline feedstock	Properties
Straight-run naphtha	Low octane, no olefins and low level aromatics
Reformate	Output from catalytic reformer. High octane number, low olefin level, high aromatic level
Cat cracked gasoline	Output from catalytic cracker. Intermediate octane number, high in olefins, medium level of aromatics
Hydrocracked gasoline	Output from hydrocracker, medium and below octane number, average level of aromatics
Alkylate gasoline	Output from the alkylation unit
Isomerate gasoline	Product of isomerisation of pentane and hexane

Source: <http://en.wikipedia.org/wiki/Gasoline>

regulations affecting gasoline blending and distribution in the US. Practically every developed country has similar rules on their books. Some of the terms mentioned here, such as different types of emissions, are discussed in more detail later in this section (RVP, MTBE) or in the chapter on environmental products.

- ❑ The Tier 2 Gasoline Sulphur<sup>43</sup> requires refiners to cap the sulphur content of gasoline at 30 parts per million (ppm) on average, with no individual batch exceeding 80 ppm. This programme was phased in from 2004–07, with increasingly stricter limits.
- ❑ The Mobile Source Air Toxics (MSAT)<sup>44</sup> rules apply to air toxics emitted by cars and trucks. The substances covered include benzene and other hydrocarbons, such as 1,3-butadiene, formaldehyde, acetaldehyde, acrolein and naphthalene.
- ❑ Reformulated gasoline (RFG)<sup>45</sup> is blended to reduce emissions of air toxics and smog forming pollutants. RFG has been mandated for a number of metropolitan areas – beginning with a first stage in 1995, and a second stage in 2000. California has introduced its own separate RFG programme (CaRFG). Phase 1, implemented in 1991, eliminated lead from gasoline and established standards for deposit control additives and RVP. Phase 2 set limits for sulphur, aromatics, oxygen, benzene, T50, T90,<sup>46</sup> olefins and RVP. Phase 3 eliminated methyl-tertiary-butyl-ether from California gasoline.
- ❑ EPA regulates the volatility/ RVP<sup>47</sup> of conventional retail gasoline during the summer (June 1–September 15) to reduce emissions which contribute to ozone formation. This and other reasons for seasonal adjustments to RVP are discussed below.
- ❑ Winter oxygenated fuel<sup>48</sup> programmes are designed to increase fuel oxygen content to enhance fuel combustion. This programme is mandated in areas that exceed carbon monoxide levels that are deemed acceptable. The oxygenated fuel season varies depending on location, but is generally October through February or March.<sup>49</sup>
- ❑ E15<sup>50</sup> is a mixture of gasoline (85%) and ethanol (15%) by volume ethanol and 85 volume. The EPA has granted a partial waiver<sup>51</sup> to allow the use of E15 in light-duty motor vehicles, model year 2001 and newer, subject to several conditions.

One of the most important facts of the US gasoline markets is the seasonal price spike during the summer driving season. In the public mind, this is quite correctly associated with higher demand. There is, however, an additional important factor related to relative curtailment of supply, resulting from the switch from winter gasoline to summer gasoline. This switch is dictated by the need to meet certain performance specifications of gasoline that have a seasonal component. These characteristics are volatility and the octane number.

The volatility of gasoline (its tendency to vapourise) is measured by RVP, the vapour pressure of the specific gasoline blend, measured at 100°F. Atmospheric pressure varies with elevation, but it can be assumed to be about 14.7 psi for most of the US. If the vapour pressure of gasoline exceeds atmospheric pressure, it will vapourise and mix with air. The trade-off is that low RVP gasoline will not escape easily from the tanks and contribute to pollution but may not produce enough vapour for a car engine to start. Vapourisation of gasoline is a bigger potential problem in the summer, as in many places in the US the temperatures can exceed 100 degrees on a hot day in a place exposed to the sun, such as a hot highway or an open car park. This is especially important in places where a combination of high temperatures and high elevations can cause vapour lock: where a fuel pump starts pumping a mixture of liquid and vapour, which may cause the engine to choke.<sup>52</sup> Another potential problem is that gasoline in a tank or open container may reach a boiling point and contribute to pollution. This is why summer RVP has to be lower relative to the winter levels, and for most US locations at or below 8.5 psi.<sup>53</sup> Winter RVP is typically around 13 psi. On a cold start, about 10% of gasoline has to vapourise to start an engine. RVP that is too low would make it impossible to start a car in the morning.

Seasonal adjustments to the RVP level are dictated by two factors described above: engine design and environmental concerns. For the summer, gasoline is blended to vapourise less easily. The RVP of a specific gasoline blend is dependent on the percentages of different hydrocarbons included in the mix. A typical summer gasoline can include:<sup>54</sup>

- 40% fluid catalytic cracker gasoline;
  - 25% straight run gasoline;
  - 15% alkylate gasoline;
-

- 18% reformate gasoline; and
- 2% butane.

Butane is relatively the cheapest component but, unfortunately, has the highest RVP, of about 52.<sup>55</sup> This is why its percentage has to be reduced during the summer but may be expanded during the winter months. This, in turn, allows the blenders to expand the supply and lower the cost. Lower seasonal demand combined with increased supply and lower cost, produces lower prices. Table 15.4 shows an example of a simple calculation (that may be carried in a spreadsheet) illustrating how much butane may be used for a specific RVP target. In order to reach the target RVP level of 10, the initial basket of different gasoline components (19,000 barrels) has to be augmented by an additional 3,081 barrels of butane. The calculation assumes that the RVP of blended gasoline is equal to the volume-weighted average of RVPs of different components. In practice, one can expect that the relationship will be characterised by non-linearities that have to be incorporated in the calculations of optimal blending proportions.

Octane number is a measure of propensity of gasoline to cause engine malfunction known as knocking. This phenomenon, noticed by engineers from the very beginning of spark-ignition engines, is caused by the self-ignition of gasoline and is related to the propagation of flame inside the cylinder of an engine. The ideal process would consist of the smooth propagation of flame across the gasoline/air mix starting from the spark plug towards the other end of the cylinder. Such a process would optimise engine performance, gradually increasing the pressure on the piston. What happens in practice is that partial combustion of fuel mix increases the pressure and temperature inside the cylinder, leading to explosive instantaneous ignition, causing an effect known as the knock (due to the audible effects). Knocking reduces engine efficiency, increases wear and tear and, in extreme situations, may cause significant damage.

Car manufacturers and the oil industry sought to address the problem by setting a joint committee to research the issue. A Cooperative Fuel Research Committee was set up in 1927, with representatives from the American Petroleum Institute, the American Manufacturers Association, the National Bureau of Standards and the Society of Automotive Engineers. John Campbell

of General Motors built a test engine which was run with different fuel mixes to determine the conditions under which knocking takes place. This test engine became a standard tool for rating different blends of gasoline. The numbers that are posted at gas stations are calculated as:

$$\frac{\text{RON} + \text{MON}}{2}$$

where RON stands for Research Octane Number and MON stands for Motor Octane Number.<sup>56</sup> This average called an anti-knock index (AKI).<sup>57</sup> Both techniques use the mixture of iso-octane (trimethyl pentane, C<sub>8</sub>H<sub>18</sub>) and normal heptane (C<sub>7</sub>H<sub>16</sub>) for calibration of a specific gasoline blend. If a mixture of 80% iso-octane and 20% normal heptane matches the test performance of a specific gasoline, the gasoline is 80 octane. Pure iso-octane was defined as 100 octane. The motor method is used for testing at high engine speeds and loads, and the research method is used at low speeds. The use of both methods is necessary, as a single test will not capture engine performance under different operational conditions.

The use of oxygenates (see Table 15.5) as gasoline additives is a somewhat complex topic, as these additives have a double function of improving the octane rating and controlling emissions of different types. Historically, the industry used lead compounds as octane enhancers. Tetra ethyl lead (TEL) was developed by a General Motors (GM) research unit in December 1921, and introduced as a commercial product in Ohio. The additive was marketed by a joint

**Table 15.4** Gasoline blending example

Component	Barrels (1)	RVP (2)	(1) × (2)
Straight-run naphtha	4,000	1.0	4,000
Reformate	6,000	2.8	16,800
Alkylate	1,000	4.6	4,600
Cat-cracked gasoline	8,000	4.4	35,200
	19,000		60,600
Normal butane	x	52	52x
10 psi RVP			
X = 3,081			

Source: Donald L. Burdick and William L. Leffler, 2001, *Petrochemicals in Non-technical Language* (Tulsa, OK: PennWell Books)

venture of GM and Standard Oil of New Jersey (a predecessor of Exxon). As GM was 38% owned by E. I. Du Pont de Nemours, there were actually three partners in this operation. The use of lead as a gasoline additive was very controversial from the start<sup>58</sup> and, eventually, the incontrovertible scientific evidence of exceptionally harmful health effects of lead resulted in its phase-out in the US in the late 1970s<sup>59</sup> and later in other countries.<sup>60</sup> The lead additives were replaced with oxygenates for the purpose of improving octane ratings. In the late 1980s<sup>61</sup> and after the promulgation of Clean Air Act, oxygenates became important for control of carbon monoxide (CO) pollution in cold weather in certain parts of the US failing to attain the National Ambient Air Quality Standard (NAAQS). The programme administered by the EPA required a minimum of 2.7% oxygen in gasoline (measured by weight). In 1991, the EPA designated 412 non-attainment areas, with a population of about 80 million. Over time, the advancements in engine design technology (fuel injection systems, the use of catalysts) reduced the number of non-attainment areas to eight in 2012 (in six states). The CO emissions result from incomplete combustion of gasoline. The winter gasoline programme should not be confused with the reformulated gasoline programme designed to reduce emissions contributing to ozone formation in the summer months and other toxic substances during the entire year. This requirement goes back to 1995, when nine ozone non-attainment areas were identified, and requires 2.1% oxygen by weight. Table 15.7 shows the comparative statistics of different liquid fuels, including Phase One reformulated gasoline.

The oxygenates included the following substances:

- methyl tertiary butyl ether (MTBE);
- ethyl tertiary butyl ether (ETBE);
- tertiary amyl methyl ether (TAME);
- diisopropyl ether (DIPE);
- tertiary-butyl alcohol (TBA);
- methanol; and
- ethanol.

The target of 2.7% of oxygen could be reached with 15% MTBE by volume and 7.4% ethanol (11.7% MTBE and 5.8% ethanol for reformulated gasoline). MTBE was very controversial from the very

**Table 15.5** Typical properties of oxygenates

	<b>Ethanol</b>	<b>MTBE</b>	<b>ETBE</b>	<b>TAME</b>
Chemical formula	CH <sub>3</sub> CH <sub>2</sub> OH	CH <sub>3</sub> OC(CH <sub>3</sub> ) <sub>3</sub>	CH <sub>3</sub> CH <sub>2</sub> OC(CH <sub>3</sub> ) <sub>3</sub>	(CH <sub>3</sub> ) <sub>3</sub> CCH <sub>2</sub> OCH <sub>3</sub>
Oxygen content, percent by weight	34.73	18.15	15.66	15.66
Octane, (R+M)/2	115	110	111	105
Blending vapour pressure, RVP	18	8	4	1.5

Source: U.S. Energy Information Administration (March 2000)

beginning, with serious concerns related to its health effects. The concerns were related to the impact of MTBE (and some other oxygenates) fumes inhaled by drivers, causing nausea, headaches and coughing. The more serious concerns regarding MTBE were related to the safety of drinking water. Water contamination happens primarily through gasoline leaking from underground and surface tanks. Another point of MTBE entry into the water systems is related to exhaust emissions that persist in the atmosphere and eventually enter the hydrological cycle by being dissolved in rainwater.

MTBE was effectively discontinued as an oxygenate as a result of issues related to the potential contamination of ground water. California has very strict limits on MTBE levels in water (5 mg/L for taste and a 13 mg/L health-based threshold). In 1996, MTBE reached the level of 610 mg/L in the water system of Santa Monica. In 2003, the Santa Monica City Council finalised a settlement with three oil companies,<sup>62</sup> which put them under an obligation to clean the wells and make an additional payment to the municipality.

As a result of growing concerns, California and New York states banned MTBE as of January 1, 2004, and many other states have followed suit. The final nail in the coffin of MTBE came with the Energy Policy Act of 2005, which did not contain a provision sought by the industry, shielding manufacturers and distributors of MTBE from suits related to ground water contamination. This provision was being actively promoted by a number of the members of the House, but did not survive the conference. Litigation related to the water-contamination claims continues.<sup>63</sup>

Reformulated gasoline blendstock for oxygen blending (RBOB) is non-oxygenated gasoline traded in the New York Harbor barge market. It is used for blending with 10% denatured fuel ethanol (92% purity). Nymex has switched its unleaded gasoline futures contract to this product in 2006. CARBOB (AZ RBOB) is the unfinished motor gasoline that meets the requirements promulgated by the California Air Resources Board (Arizona Cleaner Burning Gasoline regulations).

**Table 15.6** US gasoline specifications

	Fuel parameter values (national basis)				
	Conventional gasoline		Gasohol	Oxyfuel (2.7 wt% oxygen)	Phase I RFG
	Avg <sup>(i)</sup>	Range <sup>(ii)</sup>	Avg	Avg	Avg
RVP <sup>(iii)</sup> (psi)	8.7-S	6.9-15.1 11.5-W	9.7-S 11.5-W	8.7-S 11.5-W	7.2/8.1-S 11.5-W
T50 (°F)	207	141-251	202	205	202
T90 (°F)	332	286-369	316	318	316
Aromatics (vol%)	28.6	6.1-52.2	23.9	25.8	23.4
Olefins (vol%)	10.8	0.4-29.9	8.7	8.5	8.2
Benzene (vol%)	1.60	0.1-5.18	1.60	1.60	1.0 (1.3 max)
Sulphur (ppm)	338	10-1170	305	313	302 (500 max)
MTBE <sup>(iv)</sup> (vol%)	–	0.1-13.8	–	15	11 (7.8-15)
EtOH <sup>(iv)</sup> (vol%)	–	0.1-10.4	10	7.7	5.7 (4.3-10)

(i) As defined in the Clean Air Act

(ii) 1990 MVMA survey

(iii) Winter (W) higher than summer (S) to maintain vehicle performance

(iv) Oxygenate concentrations shown are for separate batches of fuel; combinations of both MTBE and ethanol in the same blend can never be above 15 volume percent total.

Source: <http://www.epa.gov/otaq/rfgnew.htm>

LPG are a mixture of propane and butanes, and were covered in the chapter on natural gas liquids. It is important to recognise that production of natural gas and refineries are two competing sources of these hydrocarbons. It explains why the prices of propane and butanes correlate well with the prices of crude oil and why WTI futures are used often in hedging exposure to a subset of the basket of natural gas liquids. Over the last few years, the growing production of natural gas has increased the supply of NGLs and suppressed their price relative to crude.

Straight-run naphtha (SRN) contains primarily straight and branched paraffins, naphthenes and aromatics (molecules containing between five and 12 carbon atoms), and is also characterised by the absence of olefins. Light SRN has a boiling range of about 95–194°F (35–90°C), heavy SRN of about 176–392 °F (80–200°C). It is important to recognise that naphtha may be produced through other refinery processes, described above, such as catalytic cracking, hydrocracking and coking. Naphtha obtained from these processes varies from SRN as it contains olefins, higher ratios of aromatics and branched paraffins. Naphtha is used primarily for producing gasoline. Light naphtha is typically blended with reformate gasoline (gasoline produced from catalytic reforming units). Heavy naphtha is used as a feedstock to catalytic reforming units, as it has a rather low octane number and is not a good blending component. Another important use of naphtha is a feedstock to steam cracking units for production of olefins. Industrial uses of naphtha include feedstock in production of industrial solvents and shoe polish, fuel used in turbines, lighters, portable stoves and lanterns (as any camping enthusiast knows well).<sup>64</sup>

## CONCLUSIONS

This chapter has covered refinery operations and properties of certain refined products, which are critical to the understanding of the price dynamics of crude oil, gasoline and middle distillates. This is, by necessity, a very general introduction to an extremely specialised field, critical to energy trading. Every energy market participant should make an investment in mastering the details of this industry. Making market bets based on technical price charts, without a deep understanding of the technology underlying the

market, is popular among many traders, but in the long run is a road to perdition. There is a lot of evidence, derived both from experience acquired through participation in the oil and products markets and academic studies that the supply of refinery capacity and rates of refinery utilisation affect oil prices. In the next chapter, we will cover the infrastructure that connects oil producers, processors and end-users of refined products. This is the link in the supply chain through which shocks propagate and equilibrium is restored.

- 1 See, for example, Stéphane Dées, Audrey Gasteuil, Robert K. Kaufmann and Michael Mann, 2008, "Assessing the factors behind oil price changes," European Central Bank, working paper series, No 855, January, for empirical evidence of refinery capacity and utilisation rates on the oil markets.
- 2 See, for example, <http://www.mit.edu/~jparsons/Presentations/Heavy%20Light%20Spreads%20-%20CEEPR%20Dec%2007%20-%20Slides.pdf>.
- 3 As Bill Cosby said: "I don't know the key to success, but the key to failure is to try to please everyone."
- 4 Heavy oil from Venezuela mixed with water and surfactants (emulsifiers) is used as a fuel call orimulsion.
- 5 The section of refining relies to a large extent on the book by William L. Leffler, "Petroleum Refining," Pennwell Books, 1985.
- 6 OECD/IEA, 2005, *Energy Statistics Manual*.
- 7 White spirit is used as a solvent for paints and varnishes.
- 8 The distillation tower is effectively a heat exchanger, cooling vapours of different hydrocarbons in a controlled way.
- 9 Other separation processes include solvent extraction and molecular sieves methods.
- 10 Other terms used are *topping* or *skimming*.
- 11 Paul F. Schmidt, 1985, *Fuel Oil Manual* (New York, NY: Industrial Press), p12.
- 12 An example of the yields from different types of crudes under simple distillation can be found at <http://www.bp.com/extendedsectiongenericarticle.do?categoryId=9011824&contentId=7022702>
- 13 Surinder Parkash, 2003, *Refining Processes Handbook* (The Netherlands: Gulf Professional Publishing), Chapter 2.
- 14 European Union emission standards can be found at [http://europa.eu/legislation\\_summaries/environment/air\\_pollution/128186\\_en.htm](http://europa.eu/legislation_summaries/environment/air_pollution/128186_en.htm). In Europe and Japan ULSD is defined 10 ppm sulphur maximum. The EU implemented this standard in 2009.
- 15 Cycle oil is the residue unbroken through catalytic cracking that may be run through the reactor again.
- 16 Coke amounts produced by this process are small but important as a source of heat for the cracking processes. Another product from the FCC unit is called *slurry*, a low-value heavy residue that may be either blended into bunker fuel or hydro-cracked.
- 17 Purvin & Gertz provides a database of existing refineries, reported in a study commissioned by European Commission (January 2008): [http://ec.europa.eu/energy/oil/studies/oil\\_en.htm](http://ec.europa.eu/energy/oil/studies/oil_en.htm).
- 18 CIS stands for the Commonwealth of Independent States, the territory of the former Soviet Union.
- 19 See [http://omrpublic.iea.org/glossary\\_pop.asp?GLOSSARY\\_ID=219](http://omrpublic.iea.org/glossary_pop.asp?GLOSSARY_ID=219).
- 20 The losses are related to loading, transit and ballasting. Some crude is lost due to emissions. See <http://www.epa.gov/ttnchie1/ap42/ch05/final/c05s02.pdf>.

- 21 Assumes a single voyage for an appropriately sized tanker chartered on the spot market. The FOB term is explained in the section on transportation.
- 22 See Ahmed Sirag Khogeer, 2005, "Multi objective multi refinery optimization with catastrophic failure effects objectives," Colorado State University, PhD dissertation.
- 23 There were many independent contributions to this field of applied mathematics, dating back to the 19th century. War and the availability of computing power accelerated progress in this area. The history of LP can be found in George B. Dantzig, "Linear programming" ([http://www2.informs.org/History/dantzig/LinearProgramming\\_article.pdf](http://www2.informs.org/History/dantzig/LinearProgramming_article.pdf)).
- 24 The decision whether an LP problem is formulated as a maximisation or minimisation task is arbitrary. Maximisation of an  $f$  function is equivalent to minimisation of function  $-f$ .
- 25 Some constraints can be specified as equalities.
- 26 <http://hpsweb.honeywell.com/Cultures/en-US/Products/OperationsApplications/PlanningScheduling/RPMS/default.htm>.
- 27 <http://www.aspentechnology.com/core/aspen-pims.aspx>.
- 28 <http://www.haverly.com/grtmps.htm>.
- 29 Adebayo Alabi and Jordi Castro, 2007, "Integrating crude supply, blending and product distribution models in refinery planning using a mathematical programming language," Universitat Politècnica de Catalunya, working paper, May 21. The authors rely in their model on the contributions made by a number of researchers in this area over the last few years. Some of the basic contributions in this area include Julius S. Aronofsky, John M. Dutton and Michael T. Tayyabkhan, 1978, *Managerial Planning with Linear Programming* (Hoboken, NJ: Wiley); Jehoon Song, Hyungjin Park, Dong-Yup Lee and Sunwon Park, 2002, "Scheduling of actual size refinery processes considering environmental impacts with multi-objective optimization," *Industrial and Engineering Chemistry Research*, pp 41, 4794–806.
- 30 One example is gasoline blending. "It is well known that the octane blending is not linear. In other words, the octane number of the gasoline is not equal to the volumetric average of the octane numbers of the blending stocks. One way to account for the nonlinearity is using interaction parameters." Capturing interactions is done through multiplicative terms (two decision variables multiplied, resulting in a term  $a_{ij}x_i x_j$ ) and this is a non-linear model, which requires more complex algorithms to solve. See Xuan Li, 2000, "Refinery-wide optimization. A dissertation in chemical engineering," Texas Tech University, May.
- 31 The matrices corresponding to a given LP implementation have often a special structure, arising from the properties of the underlying problem. Mathematicians can exploit the properties of the matrices to accelerate calculations.
- 32 This section is based primarily on the book by Paul F. Schmidt, op.cit.
- 33 <http://www.astm.org/Standards/D396.htm>.
- 34 Paul F. Schmidt, op.cit., p 138. Kerosene has a higher gravity and a lower end point than No.1.
- 35 Ibid, p 19.
- 36 Technically, a bunker fuel is any fuel that can be used to run ships. Bunker A fuel is equivalent to No.2, bunker B fuel is either No.4 or No.5.
- 37 Based Paul F. Schmidt, op.cit..
- 38 The formula developed through empirical tests by the National Bureau of Standards.
- 39 For API gravity between 0 and 9 degrees a number is 24.50, increasing to 25 for the range of 10 to 20 degrees, 25.20 for the range 21 to 30 degrees, and 25.45 for the API gravity starting at 31.
- 40 After French physicist Pierre Louis Dulong (1785–1838).
- 41 For example, pipelines should be designed taking into account viscosity of transported products. Pipeline pressure depends on its diameter and viscosity of transported fuel.
- 42 Physicists distinguish between dynamic and kinematic viscosity. This discussion of the differences between these two concepts is beyond the scope of this book.
- 43 <http://www.epa.gov/tier2/finalrule.htm>.
- 44 <http://www.epa.gov/otaq/fuels/gasolinefuels/MSAT/index.htm>.

- 45 <http://www.epa.gov/otaq/fuels/gasolinefuels/rfg/index.htm>.
- 46 T50 and T90 (in degrees Celsius or Fahrenheit) are the distillation temperatures where 50% and 90% of the gasoline is evaporated.
- 47 <http://www.epa.gov/otaq/fuels/gasolinefuels/volatility/index.htm>.
- 48 <http://www.epa.gov/otaq/fuels/gasolinefuels/winterprograms/index.htm>.
- 49 See <http://epa.gov/otaq/regs/fuels/420b08006.pdf> for more information.
- 50 <http://www.epa.gov/otaq/regs/fuels/additive/e15/index.htm>.
- 51 The waiver was requested by Growth Energy (<http://www.growthenergy.org>) and 54 ethanol manufacturers under the Clean Air Act.
- 52 Modern cars with fuel-injection technology do not have this problem.
- 53 This is location specific. EPA requires that summer gasoline blends should not exceed 7.0 psi in some locations, and 9.0 psi in others (see <http://www.epa.gov/oms/regs/fuels/420b10018.pdf>). As explained by the EPA, “[T]his guide is intended for quick reference purposes only. Federal volatility regulations (40 CFR 80.27) apply to ‘designated volatility nonattainment areas’ and to ‘designated volatility attainment areas,’ as defined in 40 CFR 80.2(cc) and 80.2(dd), respectively.”
- 54 See <http://www.theoildrum.com/node/5858>.
- 55 Butane has also lower volumetric energy density than other gasoline components; winter gasoline may cost less, but may offer lower mileage per gallon.
- 56 The motor method can be found at ASTM d 357, and the research method at ASTM d 908.
- 57 “RON correlates best low-speed, mild-knocking conditions; MON correlates best high speed and high- temperature knocking conditions and with part-throttle operation. For a given gasoline, RON is always greater than MON.” The same Anti Knock Index (AKI) number may be associated with different combinations of RON and MON. See “Motor gasolines technical review” ([http://www.chevronwithtechron.com/products/documents/69083\\_MotorGas\\_Tech\\_Review.pdf](http://www.chevronwithtechron.com/products/documents/69083_MotorGas_Tech_Review.pdf)).
- 58 See William Kovarik, 2003, “The 1920s environmental conflict over leaded gasoline and alternative fuels,” paper to the American Society for Environmental History Annual Conference, March 26–30.
- 59 EPA started the programme designed to reduce the use of lead in gasoline in 1973. At the time, the level of lead in gasoline was about 2–3 grams per gallon (200,000 tons per year). See <http://www.epa.gov/history/topics/lead/02.htm>.
- 60 Many historians believe that the use of plates made of lead and the fondness of Romans for sweetened grapes, seasoned in lead vats, and grape syrup, resulted in widespread poisoning, contributing to the decline and fall of the Roman Empire. Leaded gasoline is still marketed in some developing countries.
- 61 The first winter oxygenated gasoline programme was implemented in Denver, Colorado, in 1988.
- 62 ExxonMobil, Chevron and Shell (see “MTBE at center stage,” <http://www.eprinc.org/download/mtbeatcenterstage.pdf>).
- 63 <http://www.mtbelitigationinfo.com/go/site/942/>.
- 64 Many historians believe that naphtha was a main component of Greek fire, an incendiary weapon using a napalm-like substance, burning on water, invented by the Greeks under the Byzantine empire.

## *Oil Transportation and Storage*

Oil and refined products are transported over long distances and in large volumes. This is due to a number of factors. Oil fields are mostly located in places removed from the consuming markets where the end users are concentrated. The refineries, for economic and historical reasons, are clustered closer to the markets, mostly around seaports and along major transportation routes, such as rivers and canals, or major rail lines. The availability of crudes and refinery configuration often result in a composition of regional supply that does not correspond to the structure of final demand. This requires exports of final surplus products and imports of deficit products. This is true, for example, of Western Europe, where the automobile transportation fleet is heavily dependent on diesel. Europe often has a surplus of gasoline, which is shipped to the US and other markets, sometimes with significant consequences for refinery margins in the importing countries.

Unlike for natural gas, the shippers of oil and refined products may have to rely on many different transportation modes. Oil may be shipped by the Baku–Tbilisi–Ceyhan pipeline to a port in the Mediterranean, loaded on a tanker, and then unloaded and shipped by railway to a refinery. Managing the shipments is a complex process that requires the cooperation of professionals with multiple skills and experience. In our career, we have often been asked by traders to use our language skills to call different ports and inquire about draught limits and bunkering requirements. We have woken up port managers and conferred with them in half forgotten languages, asking about technical details we hardly understood. That our employers did not experience an Exxon Valdez problem only demonstrates the importance of luck in trading.

As in the case of other commodities, understanding logistics and

storage is key to many successful trading strategies, which evolve around arbitraging locational price differences (for example, North Atlantic arbitrage between the US and West European refined products markets). A necessary precondition of such transactions is the ability to arrange at short notice access to transportation and storage services at both the source and destination. Data related to inventories of crude and refined products play a role similar to storage reports in the US natural gas markets, and are closely followed by the trading community.

This chapter will detail the logistics of transporting crude oil and other refined products (tankers, pipelines and other modes of transportation). The review of different types of charters is an important introduction to the section on freight derivatives in the next chapter. We will end with a discussion of the importance of oil inventory data.

## TANKERS

The business of shipping crude and refined products in tankers and on barges is a critical component of the global economy. Table 16.1 illustrates the growth of shipment of crude oil since 1970.

The history of the oil tanker business is as fascinating as the history of the oil business itself. The first transatlantic shipment of oil we are aware of took place in 1861, when 1,329 barrels of oil were shipped from Philadelphia to London on the *Elizabeth Watts* of 350

**Table 16.1** World seaborne trade (*billions of ton-miles*)

Year	Crude	Products	Total
1970	5597	890	6487
1980	8385	1020	9405
1990	6261	1029	7290
2000	8180	1319	9499
2001	8074	1345	9419
2002	7848	1394	9242
2003	8390	1460	9850
2004	8795	1545	10340
2005	8875	1652	10527
2006	8983	1758	10741
2007	9214	1870	11084
2008	9300	1992	11292

Source: Source: United Nations Conference On Trade And Development, Geneva, Review of Maritime Transport 2009 "Report by the UNCTAD Secretariat," United Nations New York and Geneva, 2009, Chapter 1

dwt (deadweight tonnage).<sup>1</sup> The development of specialised tankers was closely related to the rivalry in the late 19th century between two emerging centres of oil production: North America and Baku in Azerbaijan, which was controlled at that time by Russia.

Tankers are usually classified by size and cargo type. Dirty tankers transport crude and resid, clean tankers carry products such as gasoline, jet fuel, kerosene, naphtha and gas oil. There is no generally accepted classification of clean tankers. One standard, proposed by Drewry Shipping Consultants,<sup>2</sup> breaks down the tanker fleet into four major classes:<sup>3</sup>

- ❑ “LR2 (long range 2 tankers), with a product cargo carrying capacity in excess of 80,000 dwt. LR2 tankers typically operate on long-haul voyages, although port constraints limit their trading routes. LR2s generally trade on long-haul routes from the Middle East to Asia, Europe and the Gulf of Mexico or the Caribbean.
- ❑ LR1 (long range 1 tankers), with an oil cargo carrying capacity of approximately 50,000 to 79,999 dwt. LR1 tankers are engaged in a range of product trades, generally from Europe to the United States, the Gulf of Mexico, or back. They also trade within the Mediterranean, or within Asia as well as between the Middle East and Asia.
- ❑ MR2 (medium range 2 tankers), with an oil cargo carrying capacity of approximately 30,000 to 49,999 dwt. MR2 tankers are employed in shorter regional trades, mainly in North West Europe, the Caribbean, the Mediterranean and Asia. A typical cargo size would be between 45–50,000 tons.
- ❑ Handysize/MR1 (medium range 1 tankers), with an oil-carrying capacity of 10,000 to 29,999 dwt. MR1 tankers trade on a variety of regional trade routes carrying refined petroleum products on trade routes not suitable for larger vessels.”

Dirty tankers are classified usually into several classes, known in the industry jargon as:

- ❑ ultra-large crude oil carriers (ULCC);
- ❑ very-large crude oil carriers (VLCC);
- ❑ Suezmax;
- ❑ Aframax;
- ❑ Panamax; and
- ❑ Handymax.

Definitions for these vary slightly from source to source. ULCCs were designed in the 1970s to carry between 2.5 and 4 million barrels,

but the evolution of the industry reduced the demand for their services. They were either decommissioned or used as floating storage facilities, when market conditions (strong contango) favoured storage-oriented strategies.

The SEC filing cited above by the New Lead Holdings contains a very useful description of different types of vessels:

- “VLCCs, with an oil cargo carrying capacity in excess of 200,000 dwt. VLCCs carry the largest percentage of crude oil, typically on long-haul voyages, although port constraints limit their trading routes. For example, only a few US ports, such as the Louisiana Offshore Oil Port, are capable of handling a fully laden VLCC. VLCCs generally trade on long-haul routes from the Middle East to Asia, Europe and the US Gulf or the Caribbean [...].”
- Suezmax tankers, with an oil cargo carrying capacity of approximately 120,000 to 200,000 dwt. Suezmax tankers are engaged in a range of crude oil trades, most usually from West Africa to the United States, the Gulf of Mexico and to the Caribbean; from the Middle East to Europe, within the North Sea, the Mediterranean and within Asia.
- Aframax tankers, with an oil cargo carrying capacity of approximately 80,000 to 120,000 dwt. Aframax tankers are employed in shorter regional trades, mainly in North West Europe, the Caribbean, the Mediterranean and Asia.
- Panamax tankers, with an oil carrying capacity of 50,000 to 80,000 dwt. Panamax tankers represent a more specialised trading sphere by generally taking advantage of port restrictions on larger vessels in North and South America and, therefore, generally trade in these markets.
- Handy tankers, comprising both Handysize tankers and Handymax tankers, with an oil cargo carrying capacity of less than 50,000 dwt but more than 10,000 dwt. Handy tankers trade on a variety of regional trade routes carrying refined petroleum products and crude oil on trade routes not suitable for larger vessels.”

**Table 16.2** Average tanker dimensions

	VLCC	Suezmax	Aframax/LR2	Panamax/LR1
Length (metres)	331.4	272.8	243.2	229.2
Breadth (metres)	58.2	46.3	41.6	33.3
Draft (metres)	21.2	16.5	14.1	13.1
Barrel intake (bbls)	2,044,000	1,008,000	690,000	473,000
Speed (knots)	15.2	14.9	14.7	14.9

Source: <http://www.skibskredit.dk/Shipping-Research/Tanker-Ships/Crude-Tankers/Segments.aspx>

Figure 16.1 provides information about the size of the tanker fleet worldwide in 2009.

### Charter contracts

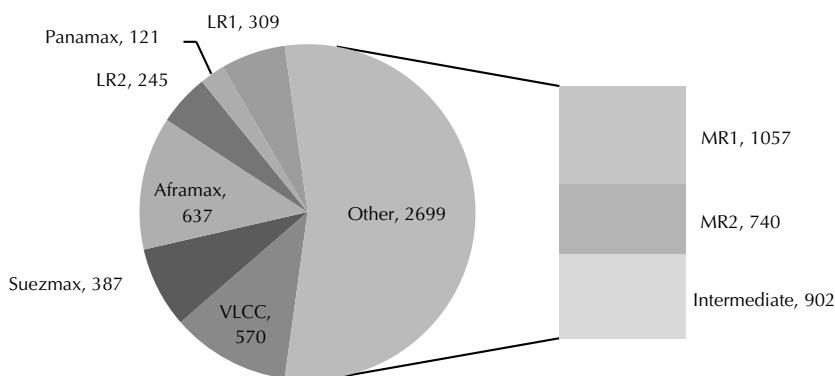
Tankers can operate under a number of different arrangements, known generally as

- bareboat charter (demise<sup>4</sup> charter);
- time charter;
- spot charter (voyage charter/trip charter); and
- contract of affreightment (CoA).

The contracts differ with respect to a number of provisions, including the way the freight is expressed (currency units<sup>5</sup> per metric ton or currency units per day), the allocation of costs between charterer and the ship owner, operational control over the ship, the level of details regarding the loading and discharge ports, the route, etc.<sup>6</sup> The costs include:<sup>7</sup>

- capital cost (the costs of financing the ship through a combination of debt and equity investments);
- operating costs (manpower costs, maintenance, insurance, management, registration);

**Figure 16.1** Size of tanker fleet (November 2011)



Source: Henry Curra, ACM Shipping, "Tanker Market Overview & Outlook," FFABA Tanker Derivatives Forum 2011.

- voyage costs (bunker fuel, port and pilotage charges, canal dues); and
- cargo handling costs (loading, stowage,<sup>8</sup> lightering,<sup>9</sup> discharging).

Many energy-trading organisations actively engage in arranging shipments of crude and refined products. It is critical to understand who is responsible for different expenses and who has a legal liability in case of an accident or environmental damage. Given that an accident may happen in a foreign country, with a very different legal system, securing support of a lawyer with a good understanding of maritime law is critical.

A bareboat contract is the equivalent of a financial lease for a car. The charterer has a full operational control of a tanker and is responsible for all the costs, with exception of the capital costs. A tanker is carried on the balance sheet of the owner, who is effectively a provider of financing. Similar arrangements are used often by the airlines for planes.

Under a time charter contract, a charterer takes over operational control of a vessel for a specified period of time. This arrangement is very similar to a car rental. Freight rates are usually expressed in currency units per day, with the charterer being responsible for the voyage costs. For example, the charterer is responsible for bunker fuel, in the same way a person renting a car pays for gasoline. The owner commits to providing a vessel with certain operational characteristics (size, speed, condition, etc). The charterer may use the ship for one or multiple trips, subject to constraints specified in the negotiated contract.

A spot contract may be negotiated as a voyage charter or a trip charter. A voyage charter specifies the destination and the loading port, the type and volume of cargo. The freight is specified usually in currency units per ton or as a lump sum. The ship owner is responsible for all the expense and for timely arrival in the loading port. Delays in unloading the cargo are handled through adjustment to the negotiated freight. This contract is equivalent to renting a taxi to get from a hotel to the airport. If the cab has to stop along the way and wait to pick-up additional passengers, an adjustment to the fare, based on additional trip time, will be made. Under a trip charter contract, a charterer hires a vessel for a specific trip. The freight is

specified in terms of currency units per day. The operational control of a ship rests with the ship owner, but the voyage expenses are covered by the charterer. The owner is not responsible for the delays and is shielded from the risk related to unexpected changes in expenses (such as the bunker fuel costs). The trip charter is equivalent to renting a car with a driver for a period of time.

A CoA is an arrangement used to move on a regular basis large volumes of a commodity over time, exceeding the size of a single vessel. A counterparty to such a contract has the option of using their own fleet or subcontracting other ships. The rates are expressed in currency units per metric ton; the ship owner has responsibility for all the costs. The contract benefits both sides. A ship owner has a steady and predictable revenue stream and the ability to optimise the utilisation of their fleet; the charterer has a guarantee of uninterrupted supplies and a fixed cost of transportation.

### **Worldscale**

Historically, the freights were quoted in dollars or pence per long ton for individual voyages.<sup>10</sup> In principle, determination of the freight for a specific trip would require long negotiations, given multiple loads and points of loading and discharge. The effort to streamline procedures for chartering ships was started during the World War II in order to manage shipments of military materials across the Atlantic. The last schedule of tanker voyage rates was issued by the British Ministry of Transport (MOT) for rates effective January 1, 1946, and by the United States Maritime Commission (USMC) for rates effective February 1, 1946.<sup>11</sup> The industry realised the benefits of having standardised schedules for the purpose of negotiating freight rates and developed a convention of quoting the rates as a percentage premium or discount to MOT or USMC. In the early 1950s, Shell Oil and BP approached the London Tanker Brokers' Panel (LTBP) and asked for development of the average freight rate assessment (AFRA), effectively a large scale of rates. The tankers were classified as general purpose (25,000 tons dwt), medium range (25–45,000 tons) and long range. The system developed in London became known as Intascale; a competing system in New York was called American Tanker Rate Schedule (ATRS). In 1969, Intascale was merged with ATRS in New York, creating the Worldwide Nominal Tanker Freight Scale

(Worldscale).<sup>12</sup> Worldscale was revised on January 1, 1989 (New Worldscale) to reflect much higher tanker rates.

Under Worldscale, a tanker owner earns roughly the same amount per day after costs are deducted. The freight rates are derived using the following formula:

$$\text{Daily hire} = \frac{\text{Voyage revenue} - \text{Voyage costs}}{\text{Voyage elapsed time}}$$

Voyage revenue is based on the rate the owner of a tanker receives (in US\$/MT) times the cargo tonnage (in MT). Voyage costs are expenses, such as port and fuel costs, canal charges, etc. Voyage time is the time spent to carry the cargo to the discharge port and during the ballast transit back to the load port. In other words:

$$\text{Daily hire} = \frac{\text{WS100} \left( \frac{\text{US\$}}{\text{MT}} \right) \times \text{Cargo quantity (MT)} - [\text{Port + Fuel + Canal charges}]}{\text{Round trip miles / Speed} / 24}$$

The WS100 rate applies to a standard vessel, defined as follows:

- total capacity: 75,000 MT;
- average service speed: 14.5 knots;
- in-transit bunker consumption: 55 MT/day;
- other bunker consumption: 100 MT/voyage;
- bunker grade: 380 cst;
- fixed hire element: US\$12,000/day; and
- port time: four days.

The WS100 is calculated subject to additional assumptions related to bunker fuel costs, port charges, etc. The annual Worldscale rate schedule lists over 60,000 reference rates quoted in US\$/MT. The rates are available for different port combinations. Spot rates are quoted as a percentage premium or discount of the published rate or WS100 rate. A charter based on WS90 (for example) means that 90% of the Worldscale flat rate for a specific voyage will be paid by the charterer.

Freight rates are highly variable, as shown in Figure 16.2 and Figure 16.3. The rates are very sensitive to macroeconomic conditions, which translate into trade flows of crude and refined products.

High volatility creates an obvious need for risk management instruments, covered in Chapter 18.

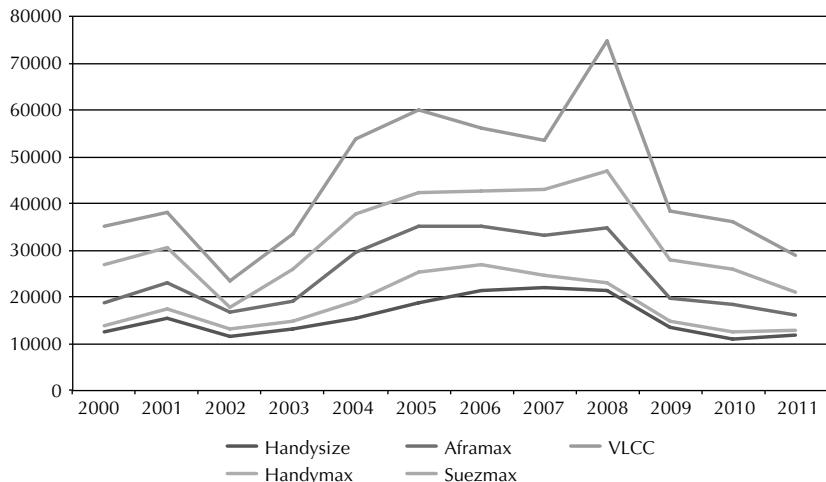
## OTHER MODES OF TRANSPORTATION AND PIPELINE OPERATIONS

### US oil and product pipelines

In addition to product tankers, there are several options available to refinery operators to deliver their products to the end users or distribution centres over land. Distribution of the US refined products is heavily dependent on the pipeline mode of transportation. This is illustrated in Figure 16.4 and Figure 16.5, which present statistics for the US for 1990–2009.<sup>14</sup> Figure 16.4 illustrates the evolution of crude transportation over that time period, while Figure 16.5 depicts the evolution of crude and refined products transportation. Pipelines accounted for 79.8% of crude transportation in 2009 and 70.2% of petroleum transportation.

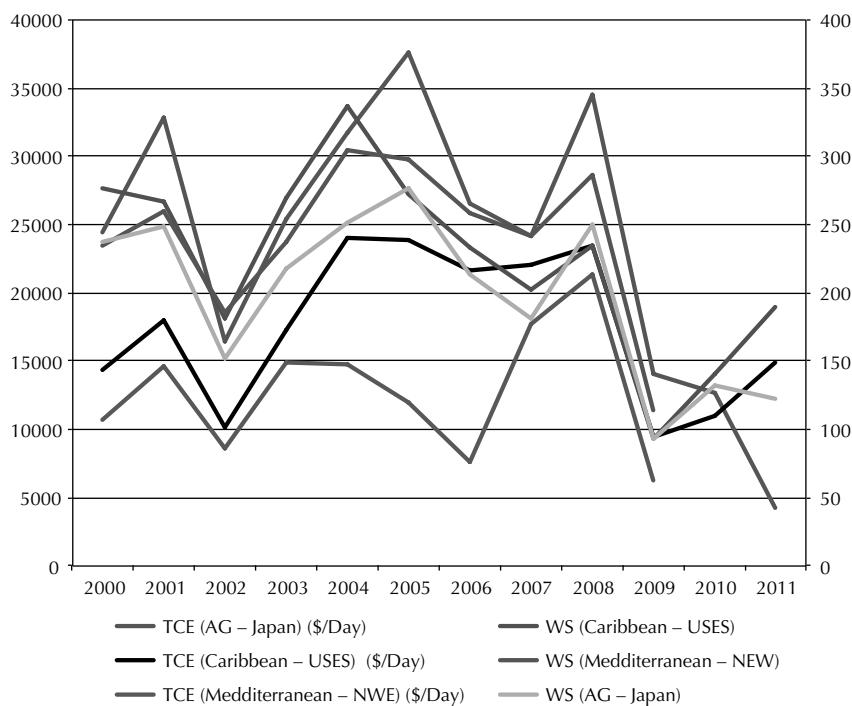
The product pipelines in the US are classified either as trunk pipelines or delivering pipelines. The trunk pipelines connect major production and consumption centres and represent the backbone of the system. The delivering pipelines are shorter and have multiple

**Figure 16.2** Oil tanker one-year time charter rates: 2000–11 (US\$/day, period averages)<sup>13</sup>



Source: [http://media.corporate-ir.net/media\\_files/irol/19/190546/2010NewLeadHoldingsLtd.pdf](http://media.corporate-ir.net/media_files/irol/19/190546/2010NewLeadHoldingsLtd.pdf)

**Figure 16.3** Product tanker spot charter rates: 2000–11 (US\$/day, period averages)

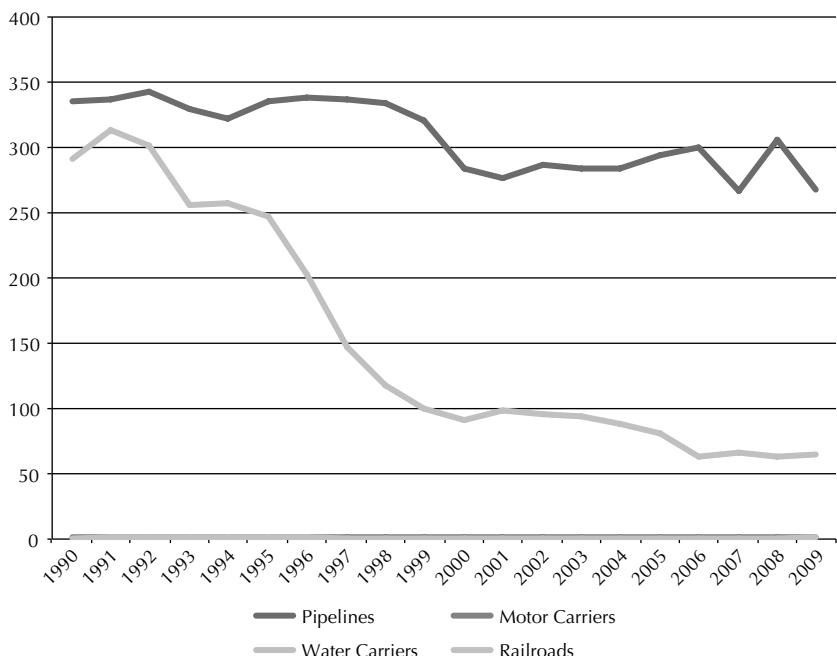


Source: [http://media.corporate-ir.net/media\\_files/irol/19/190546/2010NewLeadHoldingsLtd.pdf](http://media.corporate-ir.net/media_files/irol/19/190546/2010NewLeadHoldingsLtd.pdf)

Notes: AG – Arab Gulf; NEW – north western Europe; USES – US eastern seaboard; World Scale (WS) right axis; 2011 – March

extensions reaching to different locations. An example of a trunk pipeline is the Colonial Pipeline, and an example of a delivery pipeline is the Buckeye Pipeline Company, which operates in the middle Atlantic and upper Midwest regions.

It is convenient to discuss pipeline operations in the context of Petroleum Administration Defense Districts (PADDs), created during the World War II to implement rationing and production of oil and refined products. This system has been used to collect statistics about the oil products. PADD I (East Coast) has a deficit of refining capacity and is heavily dependent on the shipment of gasoline and heating oil either from the producing regions (PADD III) or imports from Europe and other regions. Two trunk

**Figure 16.4** Crude oil US transportation (billions of ton-miles)

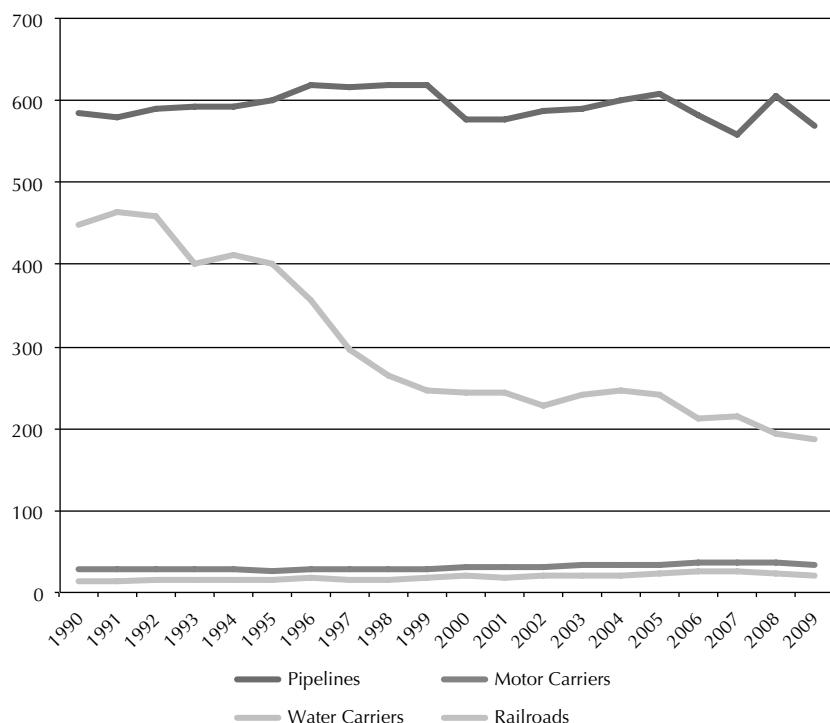
Source: [http://www.aopl.org/pdf/AOPL\\_Shift\\_Report\\_Press\\_Release\\_Feb\\_7\\_20121.pdf](http://www.aopl.org/pdf/AOPL_Shift_Report_Press_Release_Feb_7_20121.pdf)

pipelines (Colonial Pipeline and Plantation Pipeline) connect the Gulf Coast to the New York and Washington, DC areas, respectively.

PADD II (Midwest) is more balanced from the point of view of indigenous oil production, refining capacity and final demand. The deliveries of distillates are shipped from the Gulf Coast and PADD I, through which imported products transit westward. The main trunk pipelines serving this area are TE Product Pipeline (TEPPCO) and Explorer Pipeline. The delivery pipelines, clustered primarily around the Chicago area, include Williams, Equilon, Phillips, Citgo, Marathon Ashland, Buckeye, Wolverine, Kaneb and Conoco.

PADD III (Gulf Coast) contains most of the US refining capacity and is an origination point of two trunk pipelines mentioned above.

**Figure 16.5** Crude oil and refined products US transportation (billions of ton-miles)



Source: [http://www.aopl.org/pdf/AOPL\\_Shift\\_Report\\_Press\\_Release\\_Feb\\_7\\_20121.pdf](http://www.aopl.org/pdf/AOPL_Shift_Report_Press_Release_Feb_7_20121.pdf)

PADD IV (Rocky Mountain) is heavily dependent on diesel fuel (given the long distances that trucks have to drive) that is produced in local refineries (Denver area, Billings, MT and Casper, WY). The region is served by multiple pipelines crossing it or contained within its perimeter. The deliveries pipelines include Yellowstone, Cenex, Chevron, Phillips, Chase, Conoco, WYCO and Sinclair.

PADD V (West Coast) does not get much respect in the oil trading community.<sup>15</sup> It is self-contained and relatively isolated, as the Rockies create an obstacle to the construction of pipelines from other regions. The pipelines serving this region include the Calnev Pipeline, the Kinder Morgan Pipeline, Olympic, Chevron and Yellowstone.

The pipelines operate either in the batch or fungible mode. For the

batch mode, the identity of the products (ie, the identity of the shipper) is tracked during transportation and the same molecules that were injected into the pipe are delivered to the customer.<sup>16</sup> In the fungible mode, the molecules that correspond to the same product specification are comingled and the customer receives the desired product, but not the original material.<sup>17</sup> The fungible mode tends to be more efficient, as it gives a pipeline more flexibility to manage its operations. The batch mode may require suboptimal utilisation of tanks, as the molecules that belong to different customers have to be stored separately. A pipeline may find a compromise by using smaller tanks, which are associated with higher maintenance and labor costs.

### Pipeline operations and market disruptions

One of the industry developments of the last 20 years is the proliferation of many different grades and types of products. A pipeline is capable of carrying multiple products. This is accomplished by scheduling sequential shipments of different products. For example, a 25,000-barrel batch of a product may take 40–50 miles of a 10-inch pipeline.<sup>18</sup> The product batches may be separated by pigs, but in most pipelines the batches are butted against each other. The products will necessarily mix to some extent and the mixed hydrocarbons are called *transmix* in the industry jargon. Transmix has to be sent for re-refining at destination or the products have to be separated in centrifuges. The interfaces between different grades of the same product (the compatible interfaces) are used for the blending of a lower grade (for example, low-octane and high-octane gasoline mix are used to blend lower quality gasoline). The volume of transmix can be reduced by sending only one product (for example, gasoline) over a number of days (cycles), switching to a different product after one cycle is completed. Cycles typically take 6–10 days.<sup>19</sup>

It may seem that all these technical details are of minor importance to a trader. It is, however, critically important to understand the basics of pipeline operations. We can use an example of the Colonial Pipeline. It usually takes from 14 to 24 days for a batch to travel from Houston, Texas, to New York Harbor, with 18.5 days being the average time. After a hurricane or pipeline outage, a shortage of products may develop at the downstream locations, such as Atlanta, Washington DC and New York. This is an example of what

happened in certain US regional markets following Hurricane Katrina.<sup>20</sup>

AAA Mid-Atlantic spokesman John B. Townsend II said average gas prices in the Washington area jumped 16 cents yesterday, to a new high of US\$2.89, as more stations began to feel the pinch of dwindling supplies. In Virginia, prices rose 20 cents, to US\$2.82, Townsend said, while in Maryland prices climbed 19 cents, to US\$2.92. Prices in the District rose 10 cents, to US\$2.73, he said. Although a growing number of gas stations closed in the area yesterday, the Washington region remains in relatively better shape than many other places, such as North Carolina, Georgia and Louisiana, where customers have had to wait in line for as long as two hours. Officials urged travelers to conserve gas this weekend and tried to calm residents with news that the hurricane-damaged pipelines and refineries are slowly getting back up and running.

If a pipeline operates in a batch mode, the disruptions may be exacerbated.<sup>21</sup> Some terminals along the pipeline route may have to be supplied first and this may cause congestion, especially if a terminal upstream has a limited off-take capacity and the terminals downstream have to wait.

In Europe, the transportation of oil and refined products relies on a number of options, including pipelines, coastal short-range tankers, barges that use extensive networks of navigable rivers and channels, and road tankers. Refineries located along the coast can use product tankers, which typically range up to 10,000 tons for coastal vessels and up to 40,000 tons for tankers that travel over longer distances. River barges are typically designed to carry between 1,000 and 3,000 tons and operate on navigable rivers and canals. The oil use of an average car varies from 40 to 100 tons, and an average length of a train is about 20 cars (in Europe). Rail shipments are used often for products such as LPG, gasoline and diesel, and dominate in the case of heavier products with high viscosity, such as bitumen and heavy fuel oil, which are not adapted for pipeline transportation. Table 16.3 summarises the alternatives and unit costs available to the European shippers. The pipelines in general offer the lowest unit cost, but require high initial investments and limit flexibility with respect to the choice of receipt and delivery points. The pipelines allow for transportation of products with relatively low viscosity.

## THE IMPORTANCE OF OIL INVENTORIES

The inventory data is critical to any commodity market, and the oil industry is no exception. In the case of crude and refined products, inventories play three principal roles:

- ❑ they serve as buffers between certain assets and the rest of the system to guarantee their smooth operations; they are, for example, placed at the starting and ending points of a pipeline, and at a refinery; pipelines and refineries operate in continuous mode, the shipments (by tanker or rail) happen in discrete, lumpy steps;
- ❑ seasonality of demand for certain refined products and potential for supply disruptions (due to operational or geopolitical reasons) require a reserve to avoid potential shortages and propagation of shocks from the oil industry to the rest of the economy; and
- ❑ they are increasingly used by financial players for inter-temporal arbitrage; under the conditions of a strong contango, the arbitrage involves buying crude oil spot, storing it (often using decommissioned tankers, ie, floating inventory) and selling crude forward.

There are several similarities between natural gas and oil/refined products inventories, but there are also some important differences. The most important similarities are related to the following factors explaining the demand for inventories:

- ❑ technological constraints of the industry infrastructure;

**Table 16.3** European transportation options

Transport Mode	Costs Eurocents/ tonne/km	Size of Consignment (tonnes)
Product Pipeline	8–10	Continuous
Coastal Tanker (100–300 km)	10–18	Up to 10,000
Rail (20–40 tank cars)	16–25	1,200 to 2,000
Barge (100–300 km)	27–30	1,000 to 3,000
Road Tanker	35–80	18 to 44

*Source:* "Survey of the Competitive Aspects of Oil and Oil Product Markets in the EU," A report to Directorate-General Energy and Transport of the European Commission December 2009, Econ Pöyry AB.

- seasonality of demand;
- protection against supply disruptions; and
- speculation.

As in the case of natural gas, the oil inventories are required to support proper functioning of the physical infrastructure of the industry. Cushion gas in a storage facility is required to provide necessary internal pressurisation to extract working gas. The oil industry equivalent of cushion natural gas are the volumes of liquids in storage tanks that cannot be removed (*tank bottoms*), pipeline fill, in-transit inventory and working inventory. These volumes are required to support the routine operations of physical assets and cannot be used to satisfy production needs and end-user demand. The physical infrastructure of the oil industry is much more complicated and diversified than that of the natural gas business. Different parts of the infrastructure (production rigs, pipelines, tanks, tankers, port and rail terminals and distribution centres) operate at different speeds, and inventories are required as a buffer at the points of interface between different parts of the system.

Another reason to maintain inventories of oil/refined products is related to inevitable interruptions in volumetric flows due to accidents, natural disasters (such as hurricanes or floods) and man-made problems, such as strikes or military conflicts. Industrial action in France in 2010 demonstrated vividly how modern economies depend on the smooth supply of critical liquid fuels and how quickly a supply shock propagates across the system.<sup>22</sup>

Speculation plays an important factor in determination of the inventory level. A strong contango supports accumulation of inventories and selling forward stored commodities. The spot-forward arbitrage (explained in Chapter 4) has a simple mechanics. A speculator buys oil/refined products in a spot market, finds a storage facility and sells the commodity forward, locking a profit dependent on the cost-of-carry (the financing charges, storage costs, insurance, etc). The critical factor in this trade is an ability to synchronise the physical and financial operations and an extensive network of industry contacts to arrange access to the physical assets when the opportunities materialise.<sup>23</sup> Few global players have the ability to scale-up this operation to a profitable size and coordinate all the physical operations involved in the process. Oil majors and a few big

financial institutions have a major advantage in this game. The advantages include access to financing on a very short notice and favourable terms, direct control of physical assets required to engage in these strategies (either through direct ownership or contractual arrangements), physical presence in main energy centres and qualified personnel with a portfolio of necessary skills (legal, financial, logistical).

One important difference between the motivation behind inventory investments in natural gas and liquids for the US is geopolitical risk. Natural gas for the US is produced in North America and the industry is insulated from the risk related to international conflicts and supply disruptions due to potential hostile actions by other nations. In the case of petroleum, the US and other countries counter this threat by creating special reserves of crude oil, which are financed by governments. In the US, the SPR was created in 1975 under the Energy Policy and Conservation Act. As of 2010, the SPR includes the following fields.<sup>24</sup>

- ❑ “Bryan Mound holds 254 MMB in 20 caverns: 78 MMB sweet and 176 MMB sour.
- ❑ Big Hill: holds 170.1 MMB in 14 caverns: 73 MMB sweet and 98 MMB sour.
- ❑ West Hackberry: holds 228.2 MMB in 22 caverns: 120 MMB sweet and 108 MMB sour.
- ❑ Bayou Choctaw: holds 73.2 MMB in six caverns: 22 MMB sweet and 52 MMB sour.”

The capacity of oil in the SPR storage was equal to 727 million barrels (with 695.9 million barrels in storage as of August 3, 2012), about 50% of the US primary oil inventories (as of end-October, 2010). According to the Department of Energy (DOE), SPR provides 75 days of import protection (based on 9.7 million barrels a day for 2009 net imports),<sup>25</sup> with maximum drawdown capability of 4.4 million barrels a day (ie, much less than daily average imports). Oil from SPR can enter the US production system 12 days following a President’s decision.<sup>26</sup>

The quality of inventory data is much inferior in comparison with the natural gas markets, for reasons which will be discussed below.

There are two sources of the US inventory information: American Petroleum Institute (API) and the EIA.

API weekly inventory data is released every Tuesday at 10:30 Eastern Time. The statistical information available from API for subscribers includes:<sup>27</sup>

- ❑ *Weekly Statistical Bulletin*: refinery operations and inventories for refined products (including reformulated and other conventional motor gasoline, kerosene jet fuel, distillate (by sulfur content) and residual fuel oil);
- ❑ *Basic Petroleum Data Book*: historical data on energy, reserves, exploration and drilling, production, finance, prices, demand, refining, imports, exports, offshore transportation, natural gas and the OPEC activities;
- ❑ *Monthly Statistical Report*: commentary on the market conditions, year-end supply/demand estimates (December issue), quarterly estimates;
- ❑ *Quarterly Well Completion Report*: report on drilling activity, estimates of the total number of wells and footage drilled;
- ❑ sales of natural gas liquids and liquefied refinery gases;
- ❑ imports and exports of crude oil and petroleum products;
- ❑ inventories of natural gas liquids and liquefied refinery gases; and
- ❑ Joint Association Survey (JAS) on drilling costs.

*EIA Weekly* details US stocks of crude oil, petroleum products and the level of the strategic petroleum reserve, released at 10:30 am Eastern Time on Wednesdays. For some weeks starting with a public holiday, the release is delayed by one day. Detailed information for the current week becomes available on the EIA website later on the same day as "This Week in Petroleum". The week is defined as the time period beginning at 7:01 am each Friday, and ending at 7:00 am the following Friday.

The supply data is collected by the Petroleum Division (PD) in the Office of Oil and Gas (OOG) of the EIA, and the results are made available in a number of publications in addition to *This Week in Petroleum*, including:

- ❑ *Weekly Petroleum Status Report* (WPSR);

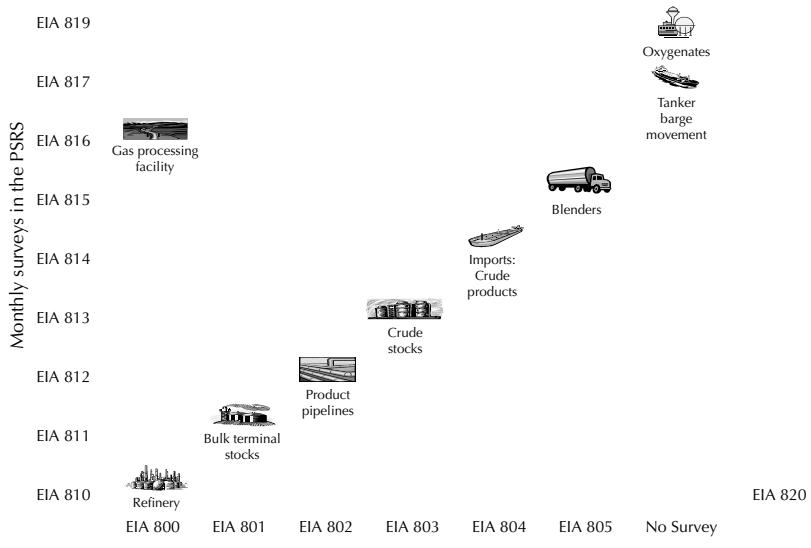
- Petroleum Supply Monthly (PSM); and
- Petroleum Supply Annual (PSA).

The data is collected through a number of different EIA forms, illustrated graphically in Figure 16.6, using a graph based on an EIA presentation.

Each issue of *This Week in Petroleum* contains a wealth of statistical data, including downloadable Excel spreadsheets. The time series (available for PADD regions and often by different grades) include information about crude spot price (in the US and worldwide), crude stocks (including days of supply), production, imports and demand. For distillates and motor gasoline, the information also includes retail prices (no prices for locations outside the US).

One useful way of looking at the inventory data is as information about days of supply – ie, the number of days for which consumption could continue at a normal level relying exclusively on accumulated stocks. The trend over the last 30 years, both for crude

**Figure 16.6** Weekly surveys in the Petroleum Supply Reporting System (PSRS)



Source: Based on Tammy G. Heppner and Matthew M. Breslin, 2009, "Accuracy of petroleum supply data," *Petroleum Supply Monthly*, February, Energy Information Administration.

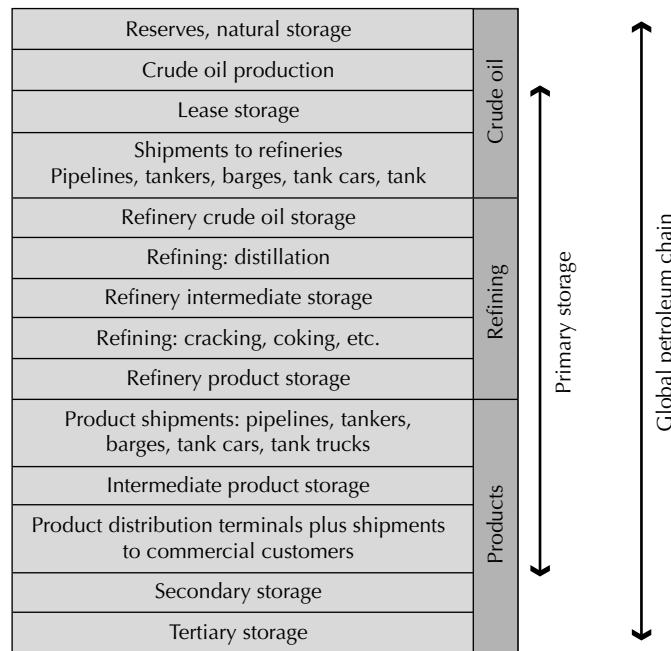
and refined products, is unmistakable: the industry is relying on smaller levels of inventories relative to usage. The trend has been reversed to some extent in the early years of the first decade of this century, but inventories of crude and products (measured in days of supply) are below the levels of the 1980s and the 1990s.

From the point of view of a trader, this trend creates a potential for periodic market shocks in instances of unexpected outages and supply interruptions. In order to capitalise on such opportunities, a trader must understand the nature of inventories and the factors explaining their dynamics.

### Inventory classification

Inventories are classified as primary, secondary and tertiary, with the two first categories more important for trading decisions. A picture is worth a thousand words, and this is roughly what we can

**Figure 16.7** Primary, secondary and tertiary storage graph



Source: "Observations On Petroleum Product Supply," The National Petroleum Council, December 2004.

save with Figure 16.7 – primary storage is well explained in the graph. Secondary storage includes volumes held by the local distribution terminals and commercial inventories. Tertiary storage includes volumes held by end users, usually at the vehicle tanks. The inventories of the end users are very important, as anybody who witnessed shortages following a stampede to refill car tanks prior to a hurricane. The secondary and tertiary storage levels are not included in the official inventory reports. EIA data collection captures inventory data for crude oil and petroleum products held in storage at.<sup>28</sup>

- refineries;
- pipelines;
- tank farms;
- bulk terminals that can store at least 50,000 barrels of petroleum products;
- crude oil in transit by tanker from the terminus of the Trans Alaska Pipeline in Valdez, Alaska, to other US ports; and
- oil in the SPR.

Inventory reports contains distortions that both underestimate and overestimate available supplies. Most installations, such as pipelines and storage tanks, cannot be completely emptied as their normal operations depend either on maintaining minimum volumes to keep the pressure at operable levels or because the tank bottoms cannot be effectively accessed. Supplies of crude and refined products in transit to the US and awaiting shipments in foreign ports may become available at short notice and, depending on the number of days required for one voyage, have to be at least partially counted towards the volumes that can relieve market pressures.

The factors behind falling days of supply should also be carefully assessed by a trader, in order to distinguish between normal fluctuations and stresses accumulating in the system. The decreasing number of refineries, combined with *capacity creep*, reduces the normal level of inventories any facility has to maintain, due to economies of scale.<sup>29</sup> The advances of information technology and management science, combined with improvements of the transportation infrastructure, reduce the optimal level of inventories (a phenomenon known as just-in-time inventories). The growth of the

derivatives markets creates an alternative to spot purchases and storage. This strategy of buying forward is especially attractive under the condition of market backwardation, with forward prices below the current spot prices. Falling production of oil in the US reduces inventories that would have been held in otherwise decommissioned pipelines and tank farms. The opposite happens when the oil output is increasing (as was the case at the time of writing).

Another limitation of the usefulness of the US inventory data is poor quality of data. The existing system of data collection is obsolete and requires a major overhaul.<sup>30</sup> Until this can happen, traders and analysts have to rely on approaches that combine quantitative systems with intuition and the collection of anecdotal data. The potential rewards may be significant as inventory data surprises occasionally result in periods of increased price volatility.<sup>31</sup>

Outside the US, weekly inventory data is provided by Japan. The IEA compiles and distributes monthly inventory data for the OECD countries. Saudi Arabia and a few other countries produce inventory data through the JODI framework. One type of data critical to the analysis of oil market trends is not directly available: Chinese demand, and commercial and strategic inventories. Estimates are available from different consulting firms and companies like Platts, which make available to their customers what is known as Chinese apparent or implied demand. The calculations of the country's oil demand are based on official data on refiners' crude throughput and net oil product imports. Refinery runs are available from China's National Bureau of Statistics (NBS), imports from China's Customs Statistics (CCS).

## CONCLUSIONS

As in the case of natural gas, the smooth functioning of the oil market is dependent on a very complex transportation and storage infrastructure. This is an important and also an extremely vulnerable link in the value chain of the energy industry. The need for different modes of transportation arises from concentration of production and processing facilities in locations that are usually removed from consuming centres. Vulnerability arises from dependence on a small number of choke points, such as the Strait of Hormuz, the Straits of Malacca and the Suez and Panama canals.

Oil transportation is associated with significant ecological risks, as proved by several high-visibility accidents that have been burned into the collective consciousness of billions of people. Inventory dynamics is important not only because storage creates a necessary buffer between consumption and production; as with natural gas storage, crude and products inventories provide a window into this industry, offering information about unfolding supply and demand trends.

- 1 DWT is the weight a ship is can safely carry. It includes cargo, fuel, fresh water, ballast water, provisions, passengers and crew (although, in this last case, the term deadweight may not be appropriate).
- 2 <http://www.drewry.co.uk/>.
- 3 [http://media.corporate-ir.net/media\\_files/irol/19/190546/2010NewLeadHoldingsLtd.pdf](http://media.corporate-ir.net/media_files/irol/19/190546/2010NewLeadHoldingsLtd.pdf).
- 4 The word "demise" denotes transfer of property or liability in the legal language.
- 5 The currency used in most contracts is US dollars.
- 6 See Amir H. Alizadeh and Nikos K. Nomikos, 2009, *Shipping Derivatives and Risk Management* (Basingstoke, England: Palgrave Macmillan), for more details.
- 7 As above.
- 8 Stowage denotes the operation of loading the ship in the most efficient and safe way and requires, in many cases, considerable skills.
- 9 Lightering is the process of transporting cargo from a ship to shore using smaller (lighter) vessels.
- 10 See "Worldscale: A Primer. Worldwide Tanker Nominal Freight Scale," presentation by McQuilling Services at The Tanker Derivatives Forum, New York March 2008. The rest of this section is also based on this document.
- 11 See above.
- 12 See also <http://www.worldscale.co.uk/>.
- 13 Based on Drewry Research quoted in the New Lead Holdings filing.
- 14 [http://www.aopl.org/pdf/AOPL\\_Shift\\_Report\\_Press\\_Release\\_Feb\\_7\\_20121.pdf](http://www.aopl.org/pdf/AOPL_Shift_Report_Press_Release_Feb_7_20121.pdf).
- 15 The traders analysing the weekly oil and products inventory numbers typically ignore the information for PADD V, as it has a very limited impact on the rest of the US market.
- 16 For example, on the Colonial Pipeline Mainline batches vary from 75,000 to 3,200,000 barrels. The batch of 75,000 barrels may be made up of three 25,000-barrel fungible batches. Batches delivered on smaller sublines vary from 2,500 barrels minimum up to 350,000 barrels. See [http://www.colpipe.com/ab\\_faq.asp](http://www.colpipe.com/ab_faq.asp).
- 17 An example of a fungible mode pipeline is the Buckeye Pipeline.
- 18 One mile of a 10-inch pipeline has a volume of 683 barrels.
- 19 One serious complication is recent introduction of ultra low sulphur diesel (ULSD). Transporting regular diesel and ULSD may result in contamination of the latter product and economic losses.
- 20 Sara Kehaulani Goo and Ylan Q. Mui, 2005, "Rumors send panicked residents to the pump," *Washington Post*, September 3 (<http://www.washingtonpost.com/wp-dyn/content/article/2005/09/02/AR2005090202662.html>).
- 21 See Cheryl J. Trench, 2001, "How pipelines make the oil market work – their networks, operation and regulation," Allegro Energy Group, for more technical details.
- 22 Matthew Saltmarsh, 2010, "French strikers block refineries," *The New York Times*, October 14.
- 23 Weak economy and the expectations of industrial rebound contributed to high levels of carry trades in oil in 2008–09. "The Energy Department said in May 2009 that as much as 130

million barrels were floating on ships around the world. The strategy helped boost BP's profit by US\$500 million in the first quarter of 2009, the company said. The TI Europe, the tanker hired by ConocoPhillips for storage in the Gulf, has a capacity of three million barrels, and ConocoPhillips is paying US\$41,000 a day, according to RS Platou's website." See Asjlyn Loder, 2010, "Oil contango doubles as traders bet on 2011 recovery," *Business Week*, September 8 (<http://www.businessweek.com/news/2010-09-08/oil-contango-doubles-as-traders-bet-on-2011-recovery.html>).

- 24 <http://www.fossil.energy.gov/programs/reserves/spr/spr-facts.html>.
- 25 The IEA recommends 90 days of import protection. It is important to recognise that in addition to SPR, the US economy will have the benefits of private inventories to rely on in the case of an emergency.
- 26 Oil has been released from the SPR on several occasions in the past (Hurricane Katrina, Desert Shield/Desert Storm), and through exchanges with US businesses. Exchange is usually a form of a loan.
- 27 [https://accessapi.api.org/accessapi/docs/API\\_DATA\\_BROCHURE\\_2010.pdf](https://accessapi.api.org/accessapi/docs/API_DATA_BROCHURE_2010.pdf).
- 28 A useful discussion of inventory trends can be found in US Government Accountability Office, 2007, "Energy markets. Increasing globalization of petroleum products markets, tightening refining demand and supply balance, and other trends have implications for US energy supply, prices, and price volatility," report to Congressional requesters, December.
- 29 In 1980, there were over 300 refineries in the US, in 2006 this number had dropped to about 150.
- 30 "The system was written decades ago and tinkered with over time," said Stephen Harvey, director of the EIA's office of oil and gas, which puts out the weekly data. "The question we're asking now is [whether] it's better to stop tinkering" and overhaul the entire system. See Brian Baskin, 2010, "US delays plans to improve oil-inventory surveys," *Wall Street Journal*, July 13.
- 31 Angi Rösch and Harald Schmidbauer, 2010, "Effects of weekly inventory data releases on crude oil spot prices," working paper, May 9. According to the authors, the "findings suggest that there is a pronounced impact of the market's perception of deviation on the behaviour of daily WTI price changes in the period from 2005 through 2009. There is an asymmetry w.r.t. the direction of deviation. An unexpected surplus of inventories implies a much higher volatility of price changes than if the forecast was undercut."

## *Oil Pricing*

Any casual observer of the energy markets living in the US associates the price of crude with the price of the prompt WTI contract traded on Nymex. The reality of crude pricing, given the multiplicity of grades and locations, is quite complex and poorly understood even by many practitioners in the energy industry. The world oil markets have evolved over the last 50 years towards increased transparency and efficiency of price discovery. Having said that, they have remained quite opaque and this condition gives a significant advantage to experienced traders operating in organisations with access to information about inventories, production trends, transportation flows, conditions of the physical infrastructure for the production and processing of crude oil, and other critical industry factors.

We will start this chapter with a review of the different ways of pricing oil in the past. The oil market today is based on a number of benchmarks (Brent, WTI, Oman/Dubai) that everybody complains about but still uses in the absence of a better alternative. The word “benchmark” may be somewhat misleading here, because in practice we are not dealing with a single easily identifiable price, but rather with a system of related and interacting physical prices and financial instruments, conventions developed over a long period of time and learned only through immersion in the market. This chapter may be the most important in this book, given how important oil is in the world economy, but it was also the most difficult to write, given how complicated the current oil pricing regime is.

### **OIL PRICING REGIMES**

As in other energy markets, history is key to an understanding of current conditions, and we start with a brief review of the history of oil pricing regimes. The author of this book was exposed as a student

to the concept of the *posted price*<sup>1</sup> of oil set by a few majors dominating the industry (outside the Soviet bloc).<sup>2</sup> The majors (also known as the *seven sisters*)<sup>3</sup> controlled production in the developing countries through concessions and joint ownership of exploration and production companies. The system of posted prices<sup>4</sup> was used for the calculation of royalties and taxes paid to the host governments. The majors were vertically integrated and relied on a system of transfer prices when oil was moved across borders to processing centres in Europe and the US, and between different units of the same corporation. The transfer prices were designed to minimise worldwide tax liabilities and, like posted prices, had no true free-market significance. Bilateral transactions between the majors were based on negotiated prices that constituted commercial secrets and were not available as price signals to other market participants.

The posted prices were used to calculate *ad valorem* royalties and taxes on notional profits. The formula to calculate the tax was given as:<sup>5</sup>

$$\text{Royalties} + \text{Taxes} = (P - C - P \times R) \times T + P \times R$$

where  $P$  is the posted price,  $C$  is the notional cost of production,  $R$  is the royalty rate and  $T$  is the profit tax rate.

Another concept closely associated with the posted price was the tax-paid cost:

$$Y = C + \text{Tax}$$

The posted prices reflected the balance of power and bargaining skills of the host countries and the international oil companies, and were ineffective as a price signal outside a vertically integrated industry.

Several different factors slowly eroded this system. Independent oil companies started to compete with the majors for concessions in the producing countries, other sources of oil (the Soviet Union) became available and, most importantly, the demand for oil increased at a high rate following several decades of economic prosperity in the West. Demand growth also increased the bargaining power of the OPEC countries.<sup>6</sup>

In September 1970, Libya negotiated an increase in the posted price of 30 cents, retroactive to 1965. Similar demands were successfully made by other OPEC countries at negotiations conducted in

Tehran in February 1971, with the increases in the posted price ranging from 50% to 55%. The agreement was negotiated for a period of five years, until the end of 1975.

Price negotiations were reopened in September 1973, as the bargaining position of the OPEC countries was improving with a booming world economy and growing demand for oil. However, the negotiations did not succeed and OPEC took recourse to unilateral action by raising prices and cutting production. After several such increases, the posted price reached the level of US\$11.651 for Arabian Light, a development remembered ever since as the *first oil shock*. The unilateral price increases were followed by demands for equity participation, with the Gulf countries winning in 1972 an initially 25% equity stake, later increased to 51% in 1983. This was a crucial event from the point of view of market evolution. The equity participation gave the OPEC governments an oil supply that had to be marketed to other parties. To begin with, the transactions were based on a variety of different prices, including posted prices in place, buyback prices and government selling prices (GSP), also called official selling prices. After 1975, the preferred pricing mechanism was based on GSPs defined in terms of a reference price plus a quality/location differential. The preferred reference price was 34° API Arab Light crude.

OPEC engaged in largely unsuccessful efforts to manage the GSPs set by different member countries by establishing a reference price managed by the Conference of OPEC Oil Ministers. The GSPs were defined through differentials (positive or negative) to the reference price. A very soft market for oil in the 1980s created incentives for the individual member countries to discount the oil they were selling, undermining the very principle of a managed market.

The equity participation affected also the business model of the majors. They lost control over part of the production stream of crude oil, while maintaining the midstream and downstream assets. They had no choice but to diversify sources of supply through market-based arrangements, further contributing to the growth of the open market for oil. The expansion of production from the non-OPEC countries (in the late 1970s and early 1980s: North Sea, the Soviet Union, Mexico) intensified competition in the spot markets, leading to oil glut and sharply lower oil prices in the 1980s. The expansion of non-OPEC production led eventually to an unravelling of the system

---

of administered prices. The foundations of the system were further weakened by discord inside OPEC and conflicts between Saudi Arabia and Gulf countries interested in lower prices and market share, and more hawkish members of the organisation.<sup>7</sup>

The alternative pricing system was introduced in 1986 by Saudi Arabia through adoption of a netback price. Netback is a pricing system for crude (used as well for other commodities) which ties the prices of crude to the prices of refined products adjusted for transportation and processing costs back to the source. The attraction of this system is that a processor has a guaranteed profit margin, even if prices collapse. There are, however, good reasons to believe that these contracts contributed to a significant degree to a drop in oil prices in the late 1980s,<sup>8</sup> which may explain why the experiment with netback pricing was relatively short lived.

The next phase in the evolution of the oil markets was associated with adoption of the market-based system, still in place today, with the reference price (referred to as a marker or benchmark) taken from the market. The market in this case means a system of related spot, forward, options and futures prices related to crude oil of given quality, produced at a specific location. The prices of oil at other locations and of different grades are calculated by adding a negotiated differential to the benchmark price.

This system, although logical and based on solid economic principles, is under stress and its evolution should be monitored closely by all market participants. The cracks in the edifice are a result of two fundamental problems:

- ❑ the output of crudes used as benchmarks is falling, undermining their viability as reference points; the small volumes make the market vulnerable to manipulation and short squeezes; and
- ❑ the traditional benchmarks represent quality grades which are increasingly unrepresentative of the crudes produced worldwide; the structure of production is shifting towards more sour and heavy crudes, whereas the popular benchmarks tend to be rather light and sweet.

As we will see, the consequence of these two developments is that the price benchmarks increasingly represent local conditions of

shrinking pockets of production, with local infrastructure conditions increasingly dominating pricing.

## BRENT CRUDE

Brent crude is the most important price benchmark for the biggest commodity market in the world. It is also a good topic to discuss in a book on the energy commodity markets, as it provides a good illustration of all the important issues one has to think about to understand the process of price formation, interactions between the physical and financial transactions and different types of risks:

- ❑ the importance of physical infrastructure and market conventions;
- ❑ different interests and motivation of financial and physical players;
- ❑ the importance of regulation and tax laws; and
- ❑ the dangers of manipulation.

The Brent complex is a good illustration of the old truth that the market is a product of human action but not human design.<sup>9</sup> The growth of Brent was driven by a multiplicity of factors that will be covered in details below. The Brent market is not only very complex, but it also continues to evolve – making the task of explaining it quite difficult.<sup>10</sup>

In a narrow sense, Brent is oil produced from the Brent field in the North Sea.<sup>11</sup> It is a crude of high quality, with an API of around 38.3 (for the “original” Brent) and a sulphur content of 0.37%.<sup>12</sup> Standard Brent contracts represent 500,000 bbls (historically, a typical tanker, although now a typical cargo is 600,000 bbls), with the port of loading being Sullom Voe in the Shetland Islands, operated by Shell.<sup>13</sup> Market participants can trade Brent oil forward specifying only the delivery month, without the precise loading date (open contract). Producers provide to the terminal operator indications how much they expect to produce in a given month. Once the terminal operator conveys the loading schedule for a given month to the seller, the latter is obligated to nominate a cargo of physical Brent to a buyer with a three-day lay-day<sup>14</sup> loading date with a minimum 25-day notice (15 days prior to 2002, and 21 days prior to January 6, 2012). The buyer is obligated to charter a vessel available for loading

during this time. The volumetric optionality in the contract is 1% (5% prior to 2002).<sup>15</sup> The minimum notice period means nominations have to begin 25 days before the beginning of a calendar month and be completed 28 (25 + 3) days before the end of the month.

Once the loading range has been established, the cargo becomes tradable as “dated Brent” – ie, a tanker with a specific date assigned to it. An active forward market for dated Brent has developed, with the actual physical cargoes being traded at a differential to a forward price or the price of dated Brent. As in many other energy markets, price discovery takes place primarily in the forward markets, with the spot price, if one exists at all, being defined with respect to the forward benchmark. An alternative to selling a cargo forward is to use it internally in one of the refineries operated by the producer, or to use it for an EFP (more about the EFPs below).

The buyer who is selected by a producer may choose to accept the cargo<sup>16</sup> or may not need oil at all (it may be a financial intermediary, hedging or speculating), and may seek to re-nominate a given cargo to a counterparty (another player who happened to buy a cargo for a given month from them). This practice resulted in the past in a “daisy chain” of transactions in which market participants pass a hot potato until it lands in the hands of a more or less consenting adult. The game has to stop eventually when the final deadline is reached, at 5 pm on one fateful day. The event of receiving a nomination at the last

**Table 17.1** Brent Blend summary assay

<b>Crude</b>	Brent Blend
<b>Location</b>	Northern North Sea
<b>Load terminal</b>	Sullom Voe
<b>Parcel sizes</b>	Up to VLCC
<b>API°</b>	37.5
<b>Sulphur % wt</b>	0.40
<b>Acidity mg KOH/gm</b>	0.01
<b>Distillation yields (%wt)</b>	
<b>C1 to C4</b>	2.4
<b>Naphtha (C5 to 149° C)</b>	19.1
<b>Kerosene (149° C to 232° C)</b>	14.2
<b>Gas oil (232° C to 342° C)</b>	20.9
<b>Atmospheric residue (342° C+)</b>	43.4
<b>Ni ppm wt</b>	1
<b>Va ppm wt</b>	6

Source: <http://www.bp.com/extendedsectiongenericarticle.do?categoryId=16002743&contentId=7020303>

minute prior to the expiration is known in in the industry folklore as being “five o’clocked.” Some traders used to manage this risk through the sophisticated practice of disconnecting their telephones and telexes, but modern communications tools make it increasingly difficult to dodge the bullet.<sup>17</sup>

Once the game of musical chairs stops, the forward (dry, or paper, in the industry jargon) contract mutates into a wet (physically deliverable) contract. This triggers the following actions the ultimate buyer has to take:<sup>18</sup>

- to nominate whether the cargo will be “min” (for example, 594,000 bbls) or “max” (for example, 606,000 bbls) for a 600,000 bbls tanker; and
- sell the cargo on the “spot” (dated Brent or Forties or Oseberg) market; or
- nominate a vessel to collect (lift) the cargo within the lay-day; and
- pay for the cargo in full 30 days after the bill of lading date.

The payment (as well as the physical operations) can be avoided if two counterparties have offsetting positions. In this case, the dry book-out may take place. The transactions are netted and the difference in pricing (if any) is settled in cash, based on prevailing market prices.

The initial growth of the Brent market was driven by tax considerations, illustrating how critical it is to understand the tax situation and strategies of a counterparty. Tax considerations may make the other side more flexible with respect to the many provisions of a contract (for example, tax savings can induce a trader to make price concessions). In the 1980s, integrated UK-based oil companies had to use the BNOC<sup>19</sup> price, set by the government, for oil transferred from one part of the company to another (from upstream to downstream). However, if oil was sold and then repurchased in the open market, one could use the transaction price (most of the time, this was lower than the BNOC price) for tax calculations. This explains the emergence of the daisy chain of transactions. Once the market took off, it survived the eventual abolishment of BNOC.

Brent became a dominant price benchmark in the 1980s (it is now used for pricing about 65–70% of world oil output)<sup>20</sup> and still

continues in this role, although not without challenges. Output from the Brent field equal to about 850,000 bbls a day in the 1990s dropped to about 400,000 bbls a day in 2004, less than one tanker a day.<sup>21</sup> It became relatively easy to obtain control of the physical supply, and the market evolved into a condition where the absence of a short squeeze in a given month would become newsworthy. The first action was taken in 1991, when original Brent was comingled with Ninian crude to create what was called Brent Blend. In 2002, the falling production of Brent oil prompted Platts to redefine it and expand its specification by adding two additional fields: Forties and Oseberg,<sup>22</sup> with physical characteristics close to those of Brent, and creating a new standard: BFO. Brent and Oseberg outputs in 2004 were roughly equal; the Forties field produced 825,000 bbls/day. In 2006, another field was included in the Forties: Buzzard.<sup>23</sup> In 2007, the BFO standard was expanded by including Ekofisk crude<sup>24</sup> (for price assessment) and became known as BFOE. Crude oil from the Buzzard field consists of medium sour crude (32.6° API, 1.4% sulphur), with minor wax and asphaltenes.<sup>25</sup> The modification of the definition of Brent could not change the physical realities of the North Sea production: the decline of output continued and, in 2008, BFO production was about 18 million barrels a month, or roughly one cargo a day.

The important role of Brent is supported by the Brent futures contract traded in London on the ICE Futures Exchange (previously IPE), launched on June 23, 1988. The contract specifications are shown below.<sup>26</sup>

### Brent complex

The instruments and transactions mentioned above are closely connected and represent what can be called a “Brent complex.”

- ❑ *Dated Brent*: a market for cargoes of Brent, Oseberg, Forties, Ekofisk crudes with confirmed loading range in the next 25 days.<sup>27</sup>
- ❑ *Forward Brent* (25-day BFOE). This is a forward market for future cargoes of the four crudes underlying the Brent complex. The exact loading dates or grades are not specified at inception. This contract is referred to sometimes as cash or paper BFOE. This is somewhat misleading because the term “cash” is usually used with respect to spot transactions.

- ❑ *Brent futures contracts* traded on ICE Futures in London, Nymex and DME.
- ❑ *OTC Brent-related derivative contracts* (options, swaps, swaptions) that settle based on the prices mentioned above (futures, forwards, dated Brent). The most important derivative is the contract-for-difference (CFD) and related dated-to-frontline (DFL) contract.

We shall review these components below in more detail.

### **Dated Brent**

The dated Brent assessment by Platts<sup>28</sup> considers the prices of physical Brent, Forties, Oseberg and Ekofisk crudes loading within a window of 10–25 days (10–27 days on Friday). The assessments are based on outright prices of consummated transactions (FOB terminal and ship-to-ship transfers, STS), as well as bids and offers. Dated Brent comes closest to what may be considered a “spot” price, although it is important to recognise that there is an element of “forwardness” in it, and that it is a price reflecting activity over a specified time window, not on a single day.

### **25-day BFOE**

25-day BFOE is a forward contract with three monthly assessments available from Platts. The convention adopted by Platts was to price the cash BFOE contract using the most competitive – ie, the cheapest – grade (not on an average of four crudes). Historically, this was Brent crude but the addition of the Buzzard field, with its relatively high-sulphur content, changed this pattern. Forties crude tends to be the cheapest grade. The last assessment for front month cash BFOE is published by Platts on the last business day of the preceding calendar month.

At the time of writing, a development in the Brent market has demonstrated both the importance of understanding the minutiae of construction of oil price benchmarks and of the incessant tracking of the developments across different markets and connecting the dots. The price of Brent has received unexpected support from the South Korean oil importers taking advantage of tax waiver under the EU-South Korea free trade agreement signed in 2011. The waiver eliminates a 3% tax and effectively offsets transportation cost from

Europe to the Far East.<sup>29</sup> The Korean importers increased purchases of the Forties crude, the cheapest oil in the Brent complex. Price was just one reason. Forties oil has a high level of sulphur (compared to other North Sea crudes), and can be processed in the Korean refineries outfitted for this type of crude. As explained above, Brent price is set by the cheapest component in the basket. Increased Korean imports moved this floor level up: a rising tide lifts all the boats. There were other impacts as well; as explained by Javier Blas:<sup>30</sup>

Oil traders said the Korean buying was not only inflating Brent prices, but also putting pressure on the price of crude for immediate delivery, keeping it at a premium over forward-dated contracts, a condition known as backwardation. The backwardated market generally hurts the profitability of the refining sector.

### **Contract for difference**

Brent Contract for Difference (CFD) is a fixed for floating swap, with the underlying being the price differential between dated Brent and forward price. The contract is cash-settled, based on Platts's assessments of dated Brent and a forward price specified in the swap contract. In other words, Brent CFD is a transaction under which one party pays a floating calendar spread and receives a fixed differential. This is an OTC contract, and terms may vary from transaction to transaction, but most counterparties follow certain conventions, which make it possible to explain it in general terms. The Brent CFD emerged in 1988 and became quite important in 1992.<sup>31</sup> The contract has been developed to facilitate hedging of the basis risk between dated Brent and the forward/futures Brent prices. The contracts are settled in cash, based on the prevailing market at the expiration of the contract. As in the case of any basis swap, the details of the contract include:

- notional volume;
- definition of the differential;
- fixed differential; and
- assessment window.

Notional volumes are not reported, but the conventions to this market are likely to make market participants gravitate towards deal sizes corresponding to a typical tanker size – ie, 500,000–600,000

barrels. The transactions happen in incremental clips of 100 lots (ie, 100,000 barrels).

The differential is defined as the difference between dated Brent and forward Brent. It is important to recognise that dated Brent referenced in a CFD contract will be assessed over a future time window. For example, on June 1, 2012 (Friday), we would have CFDs quoted by Platts for the week of June 4–June 8, June 11–June 15, etc, up to eight weeks into the future ahead of the current day.<sup>32</sup> The definition of the differential requires specification of the assessment window – ie, the time period during which the average realised market differential is calculated – and the forward contract against which the spread is calculated (for example, the prompt or the second available contract).<sup>33</sup> For example, a typical five-day (usually Monday–Friday or, less frequently, Wednesday–Tuesday) assessment window means that the reported prices of dated Brent will be averaged, along with calculation of the average forward price for the specific month. The assessment window may be, of course, defined in any reasonable way as long as the counterparties agree and may be two to four days long, or may cover the entire calendar month or two.

At the inception of a CFD contract, dated Brent prices used for settlement are not known. In our example above, for the first CFD the prices of dated Brent over the week of June 4–June 8, will be averaged, along with the prices of paper BFOE, and the averages will be used for the final cash settlement. This is the reason why CFDs are sometimes defined in terms of forward dated Brent, where “forward” means dated Brent averaged over some future time window. To summarise:

- ❑ Differential = Forward-dated Brent – Second month forward Brent; or
- ❑ Forward dated Brent = Differential + Second month forward Brent.

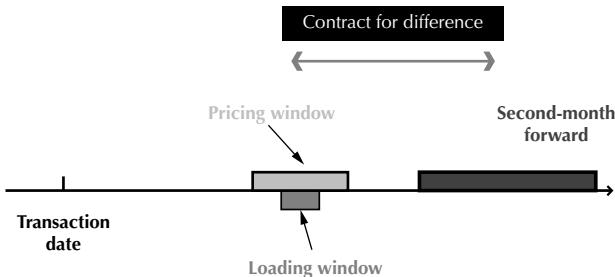
As we can see, the price levels are often derived from the spreads and forward prices, not the other way around. Most humans are conditioned to derive spreads from the absolute levels, and not vice versa, and this carries to the design of the software systems, which usually treat flat prices as the primary data and the spreads as the derived values.

For a CFD contract, both bid and offer prices are quoted. The quoted price may be negative if the market is in contango. Contango means that forward dated Brent < second month forward Brent, and this translates into a negative differential (spread). If the spread changes from minus US\$2.0 to minus US\$1.0, this means that the spread is increasing! As we mentioned in the chapter on swaps, the buyer of a swap pays the fixed price. In the case of a CFD, a buyer of this contract would pay the fixed differential and receive the floating differential at settlement. For illustration, suppose that a trader hedges the purchase of a physical cargo buying a CFD contract quoted at -US\$1.02/-0.98 per bbl. They pay *minus* US\$0.98 (the offer), but this is the same as receiving US\$0.98 /bbl.<sup>34</sup> If the spread at settlement is -US\$0.75/bbl, they will receive this amount, but receiving a negative amount is the same as paying US\$0.75/bbl. The trader in this case nets 23 cents/bbl and comes out ahead.

The importance of the Brent CFD can be explained by the pricing conventions used in the international oil markets. Many crudes traded in what can be defined as the spot market are priced at a differential to dated Brent. The differential to dated Brent<sup>35</sup> is fixed at the inception of a transaction, but the dated Brent itself will become known at the time of the loading of a cargo. Such arrangements represent a form of risk management. The time gap between a transaction date and loading of the cargo may be 10–25 days long.<sup>36</sup> Delaying final pricing to the time of the loading reduces price risk.

The following example explains the use of CFD as a hedging tool. Suppose a trader sells Nigerian Qua Ibo cargo priced in the November 2–6 week, at a differential to forward dated Brent. The producer will make less money if forward dated Brent for that week decreases. A hedge involves selling a CFD contract at a fixed differential to the cash contract. This is equivalent to receiving quoted fixed differential and paying floating. The hedger also sells Brent forward (cash BFOE contract for January delivery). Suppose that the forward is sold at US\$100/bbl and the CFD was quoted at US\$-1.02/-0.98. This translates into a forward dated price of US\$98.98. Suppose the final outcomes for the settlement week are (these are averages):

- dated Brent: US\$95/bbl;
- spread: -US\$0.75/bbl; and
- BFOE January forward: US\$95.75/bbl.

**Figure 17.1** Hedging sale of a cargo<sup>37</sup>

The hedger receives the quoted CFD differential (locked-in at inception) equal to -US\$1.02. This means that they pay US\$1.02/bbl. They have to pay the negative US\$0.75/bbl differential, which translates into receiving US\$0.75/bbl. The net result is that the hedger pays 27 cents per barrel. Liquidation of their position in the BFOE forward contract yields US\$4.25/bbl (the contract was sold at US\$100/bbl and bought back at US\$95.75/bbl). The hedges net US\$3.98/bbl (US\$4.25–0.27). The hedger receives the final price of US\$98.98/bbl (US\$95 + US\$3.98), the price they locked in at inception.

Figure 17.1 helps to explain the mechanics of this process. It illustrates the time lag between a transaction date and the loading of a cargo. Pricing is based on dated Brent quotes averaged over the pricing window, which is usually longer than the loading window. Uncertainty over the level of dated Brent prices during the pricing window can be removed by transacting in the second available forward Brent contract and in the CFD. For example, a refiner, exposed to higher Brent prices, would buy the second month forward and the CFD, corresponding to the pricing window.

One development in the energy markets has been a very rapid proliferation of different financial contracts that are related to spreads. In the context of the Brent market, the example is the CME dated to frontline Brent swap futures (ICE offers an analogous contract). The highlights of the contract (symbol FY) specification include the following features.<sup>38</sup>

- ❑ Settlement type: financial.
- ❑ Contract unit: 1,000 barrels.
- ❑ The floating price: The floating price for each contract month is

the arithmetic average of the mid-point between the high and low quotations from the Platts Crude Oil Marketwire for dated Brent minus the Brent crude oil (ICE) futures first nearby contract settlement price for each business day that are both determined during the contract month (using common pricing), except as set forth in Section (B) [...]. (B) The settlement price of the first nearby contract.

- ❑ Termination of trading: Trading shall cease on the last business day of the contract month will be used except on the last day of trading for the expiring Brent Crude Oil Futures contract when the settlement price of the second nearby Brent contract will be used.

Another term used in the market is the so-called North Sea strip. This is an average of prices of physical crude delivered in the North Sea 10–25 days forward, with the prices derived using the CFD quotes. Other price strips based on Brent and provided by Platts are (with day ranges given in the parentheses): Mediterranean Strip (13–28), BTC (Baku-Tbilisi-Ceyhan) Strip (13–33), West African Strip (18–48), Angolan Strip (15–45), Canadian Strip (31–45).

### **Other Brent-related contracts**

It is widely recognised that about 65% of world oils are priced based on Brent crude.<sup>39</sup> Platts makes three forward assessments for 25-day<sup>40</sup> cash (paper) BFOE, which reflects prices of entire or partial cargoes for physical delivery scheduled within the month of the contract.<sup>41</sup> A prompt forward contract expires on the fifth day of a month but Platts continues to provide assessments for the expired contract through to the end of the calendar month. The price information collected includes executed deals, bid and offers, Brent spreads and CFDs.

The important role of Brent in the world oil market is supported by the Brent futures contract traded in London on the ICE Futures Exchange. The highlights of the contract (ticker B) specifications are shown below.<sup>42</sup>

- ❑ Description: The ICE Brent crude futures contract is a deliverable contract based on EFP delivery with an option to cash-settle.
- ❑ Expiration date: Trading shall cease at the end of the designated

settlement period on the business day (a trading day which is not a public holiday in England and Wales) immediately preceding:

- either the 15th day before the first day of the contract month, if such 15th day is a business day; or
- if such 15th day is not a business day, the next preceding business day.

- Contract size: 1,000 barrels (42,000 US gallons)
- Settlement price: The weighted average price of trades during a two-minute settlement period from 19:28:00, London time.
- Trading methods: Electronic futures, EFP, EFS and block trades are available for this contract.
- The ICE Brent crude futures contract is a deliverable contract based on EFP delivery with an option to cash-settle, ie, the ICE Brent Index price for the day following the last trading day of the futures contract.

As we can see from the contract specifications, the underlying is a short-term forward contract for future delivery Brent crude, given the option to settle in cash against the ICE Brent Index. The definition of the index reads as follows:<sup>43</sup>

The index represents the average price of trading in the 25-day BFOE market in the relevant delivery month as reported and confirmed by the industry media. Only published cargo size (600,000 barrels) trades and assessments are taken into consideration.

### Calculation

The index is calculated as an average of the following elements:

1. A weighted average of first month cargo trades in the 25-day BFOE market.
2. A weighted average of second month cargo trades in the 25-day BFOE market plus a straight average of the spread trades between the first and second months.
3. A straight average of designated assessments published in media reports.

The alternative to a cash settlement is an EFP transaction. The option to use this settlement establishes a link between physical and financial markets, a very important and often under-appreciated component of the Brent complex.

The importance of the Brent futures contract was elevated in July 2000, when Saudi Aramco decided to use the so-called *Bwave* for pricing shipments of its crudes to Europe (followed in six months by

Kuwait, and later by Iran).<sup>44</sup> This decision was dictated by frustration with manipulation of the prices of dated Brent, the problem that became quite serious in the early 2000s before the BFO standard was introduced. Bwave is the volume weighted average of all the transactions during a given day.

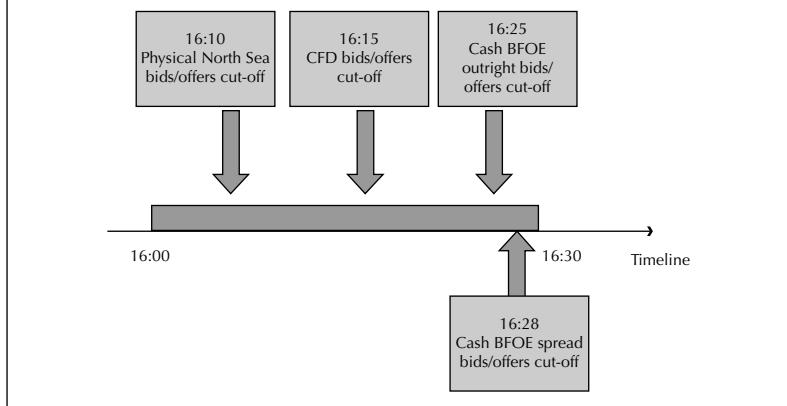
Recent revision of the pricing window from 21 to 25 days complicated the design of the ICE Futures Brent contact. As explained by ISDA:<sup>45</sup>

Platts implemented a methodology change for the forward BFOE market from a 21 day to ICE which was constructed to reflect a 15 day BFOE forward basis. These are physically delivered contracts and the 10 day difference has resulted in up to 40% of the BFOE assessed programme not being available to be traded on the ICE Brent futures expiry day. This discrepancy between the original futures construction and expiry with the BFOE forward increases the potential for price distortions at expiry. To address this ICE has launched a new ICE Brent contract to run in parallel to the current one with the first delivery date being December 2012.

### Platts E-Window

European crude assessments are made by Platts using a platform known as E-Window, which was developed in collaboration with the ICE. Platts' reporters receive bids and offers from market participants through electronic messages and enter the information into the system. The same data are displayed on Platts wire system

**Figure 17.2** North Sea price assessments process (Platts E-Window, London Time)



(PAG003). Major oil traders can enter bids and offers directly from their computers using screens that resemble closely other ICE screens. They also have the option to act on their bids and offers and clear their transactions through ICE if they choose to. Orderly price discovery requires that a well-defined procedure for submitting information is followed (as shown in Figure 17.2).

The price assessment process for cash BFOE, WTI and Mars, as well as other spot crude prices, is based on a system known as market-on-close (MOC). This system captures a prevailing value of different crudes at a specific point in time. This approach corresponds to the solutions adopted by the exchanges, which typically establish settlements prices using averages of the last few minutes of trading.

## WTI

The West Texas Intermediate futures contract is traded on Nymex, with an option to make physical delivery in Cushing, Oklahoma. The Nymex WTI contract is the most successful Nymex contract in history and probably the most closely watched and widely commented commodity price. The specifications of the contract are shown below.<sup>46</sup>

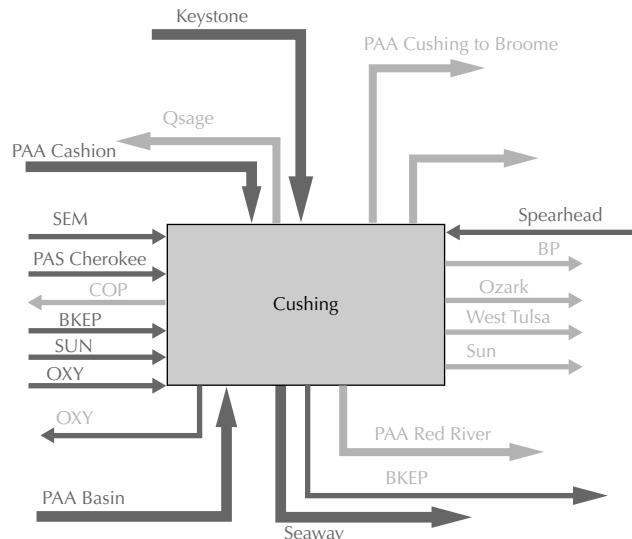
WTI is a high-quality crude produced in the US, light and low-sulphur, with an API gravity of 39.6 degrees and about 0.24 percent of sulphur.<sup>47</sup> Given its high quality, WTI historically traded at a premium of US\$1–2 per barrel with respect to Brent and US\$5–6 over the OPEC basket.<sup>48</sup> The historical relationship to Brent crude changed a few years ago and became quite unstable, as will be explained below. The viability of WTI as a universal price benchmark has been frequently questioned. Given that, for most people, WTI is the “price” of oil, most closely followed and reported on TV screens and in the newspapers, this is a very significant development.

The evolution of WTI spot and forward prices cannot be fully understood without analysing the details of volumetric flows, refinery operations and demand at different locations surrounding Cushing. One has to understand also the changes in the configuration of physical assets used to produce, transport and process oil in this area, and the production trends. Fortunately, a very incisive study<sup>49</sup> on this topic, prepared by the consulting firm Purvin & Gertz, is

**Table 17.2** Estimated Cushing Storage Tank Capacity (millions of barrels), 2012

Owner/Operator	Capacity
Plains All American	19
Enbridge 16.4 Adding	16.4
Magellan	12
Blueknight (Vitol)	6.8
Rose Rock	5
Gavilon	4
Enterprise	3.1
Parnon	3
Deeprock	1.8
CVR Energy	1
Phillips 66	0.8
Sunoco	0.4
Occidental (Centurion)	0.3
Total	73.6

Source: [http://articles.chicagotribune.com/2012-05-08/news/sns-rt-oil-storage-cushing1e8g2kva-20120508\\_1\\_cushing-plans-crude](http://articles.chicagotribune.com/2012-05-08/news/sns-rt-oil-storage-cushing1e8g2kva-20120508_1_cushing-plans-crude)

**Figure 17.3** Oil pipelines configuration around Cushing, OK

Source: Based on Plains All American Pipeline, 2011, "PAA: A core holdings through the cycles," analyst meeting, June 9

available to the industry. Although this document is somewhat dated, it still remains the best source on this topic. The section below is based to a large extent on this document.

Cushing itself is a small town attached to a cluster of oil tank farms (with no disrespect intended to the citizens of Cushing). The most recent information about its shell storage capacity<sup>50</sup> is shown in Table 17.2.

The working crude storage capacity at Cushing is about 64 million barrels and its utilisation is of critical importance to the industry. This information is available from Genscape, based on aerial reconnaissance photos.<sup>50</sup> Cushing is connected to other important oil industry centres, such as refinery clusters, storage centres, ports and producing regions, through a number of pipelines that can be depicted with some simplifications (see Figure 17.3). Some of these pipelines deserve a few comments, given their importance.

*Basin Pipeline*, owned by Plains All American Pipeline (87%), connects West Texas and southern New Mexico to Cushing. The 519-mile pipeline has a capacity of between 144,000 barrels per day and 400,000 barrels per day along different segments.<sup>52</sup>

*Centurion Pipeline*, owned by Occidental Petroleum Corporation, is a cluster of assets combining gathering systems, a common carrier pipeline and storage tanks, extending from southeast New Mexico to Cushing. Throughput capacity is about 350,000 barrels per day, with five million barrels of storage capability. The length of the entire system is 2,750 miles.

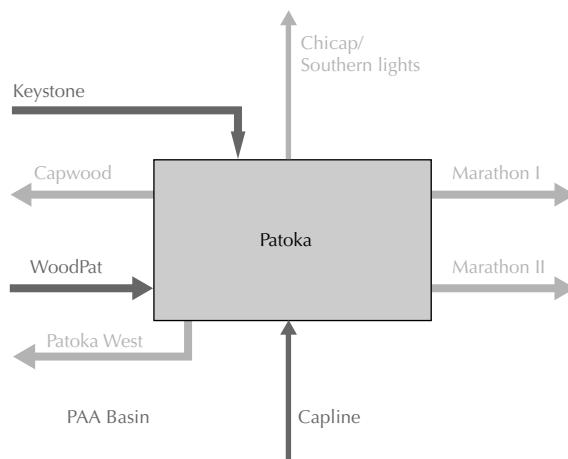
*Spearhead pipeline*, owned by Enbridge, is a 22–24-inch diameter system of 650 miles, with a daily capacity of 190,000 barrels, extending from Chicago to Cushing.<sup>53</sup> The pipeline flew oil initially from south to north, but Enbridge reversed the direction of the flow in 2006 after acquiring the asset. This decision enabled deliveries of Canadian crudes (through Spearhead and connecting Canadian pipelines) into Cushing, contributing largely to the current market conditions around Cushing.

*Seaway pipeline* was owned by a partnership of Enterprise Products Partners (EPD) and ConocoPhillips (COP), and operated by EPD. The pipeline extends from Freeport, TX, to Cushing, with a length of

669 miles and a capacity of 285,000 and 350,000 barrels per day (depending on crude grade and viscosity). Over the last few years, market observers have debated whether the flow direction on this pipeline should be reversed to relieve congestion at Cushing. The decision to reverse the flow direction on Seaway was finally announced in November 2011, when ConocoPhillips sold its 50% stake in the pipeline for US\$1.1 billion to Enbridge. Enbridge and its partner Enterprise Products Partners completed the first stage of the reversal operation in May 2012, with the initial pipeline capacity of 150,000 bbl/day (a 12-day trip between Cushing and Freeport). The capacity will be eventually increased to 400,000 barrels in 2013 and 850,000 barrels a day by mid-2014 (a five-day trip).<sup>54</sup>

Other critical centres than Cushing include the Louisiana Offshore Oil Port (LOOP), Patoka and Midland. LOOP, owned by Marathon Pipe Line, Murphy Oil Corporation and Shell Oil, is a deepwater port facility located near Port Fourchon, LA. It is the only port in the US capable of offloading ULCC and VLCC.<sup>55</sup> Oil is offloaded through hoses connected to a single point mooring (SPM) base and transported to the terminal through an underwater pipeline. Crude is stored in underground storage facilities (the Clovelly Dome Storage

**Figure 17.4** Pipeline connections to Patoka



Source: Based on Plains All American Pipeline, 2011, "PAA: A core holdings through the cycles," analyst meeting, Houston, June 9

Terminal) and in above-ground tanks. Patoka, Illinois, is a town with a cluster of refineries and multiple pipeline connections (see Figure 17.4).

A critical concept required to understand the flows of oil between different points is *parity pricing*, the level of prices of different crudes which results in the same refining margin. Refining margins depend on the prices of competing crudes and refined products, crude qualities, transportation costs and refinery configurations. The differences in margins will trigger the flows of crudes between different locations until parity is restored. A refinery operator and a crude trader have to monitor constantly shifting parity relationships and position themselves in anticipation of the expected changes (the shifting parity relationships for Cushing and other critical refining centres are covered extensively in the study by Purvin & Gertz).

Three important factors contributed to evolving price dynamics in Cushing. One of them was the increasing availability of Canadian crudes, made possible, among other developments, by reversal of the flows of oil on the Cushing–Chicago pipeline, acquired by Enbridge and renamed the Spearhead Pipeline.<sup>56</sup> The second development was a periodic shortage of storage capacity at Cushing that contributed to the steeper contango of WTI prices. Shortage of storage creates a downward pressure on spot prices as the opportunities for storage arbitrage become restricted and crudes are dumped into the open market. The third factor was limited pipeline capacity to flow crude from the Cushing area to the refineries located at the US Gulf Coast (USGC). The combination of these factors, in conjunction with prevailing transportation rates, “created the most volatile market experienced with respect to price relationships for WTI versus other benchmarks like LLS [Light Louisiana Sweet] and Brent.”<sup>57</sup>

Another development was the drop in the North Sea oil output discussed earlier. Scarcity of Brent had a profound impact on the relative prices of WTI and Brent. The report by Purvin & Gertz summarises this succinctly:<sup>58</sup>

It is important to note here that as North Sea production declined, the proportion of these crudes remaining to be consumed in Europe grew and the availability for export to the US dropped sharply to almost nothing through 2008. In addition, most of what was imported came to the East Coast with its logistics advantage relative to the USGC. This trend resulted in an arbitrage delink between the

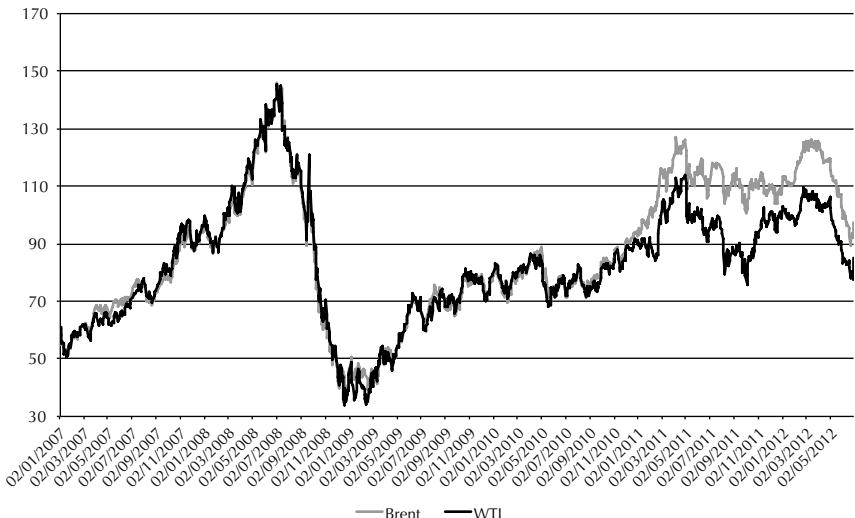
North Sea prices and the USGC prices for a good portion of the time during this period.

The discussion above illustrates the importance of understanding the physical energy infrastructure in making trading decisions, one of the central themes of this book. In 2007, the relationship between WTI and Brent reversed from the usual US\$1.00–1.50 premium of WTI over Brent (roughly the cost of moving Brent to the USGC refineries) to a discount which was sometimes quite significant. The author received at this time a number of calls from friends in different financial institutions asking for advice. They were swamped with proposals from different aspiring speculators who saw the WTI/Brent price reversal as a great trade idea and were ready to put on bets of epic proportions as long as somebody was willing to fund them. The bet was on the spread reverting to original levels, creating “the opportunity of the lifetime.” As Shakespeare wrote:<sup>59</sup>

There is a tide in the affairs of men.  
Which, taken at the flood, leads on to fortune;  
Omitted, all the voyage of their life  
Is bound in shallows and in miseries.  
On such a full sea are we now afloat,  
And we must take the current when it serves,  
Or lose our ventures.

Such strategies represent bets on the reversal of relative prices to historical relationships following a shock to the market, and are very popular in the fixed income and foreign exchange markets. Our advice was consistently to look for the structural changes in the industry that could invalidate historical regularities. Betting on price adjustments following a displacement without investigating the shifting fundamentals is akin to using astrology in trading. The problem was that no tankers loaded with Brent crude were making trips to the US Gulf Coast, and this was not helping the speculators with “the voyage of their lives.” As far the author knows, many such trades were eventually executed, sometimes to the chagrin of investors. The saving grace was the volatility of the WTI/Brent ratio, which created opportunities for a graceful exit.

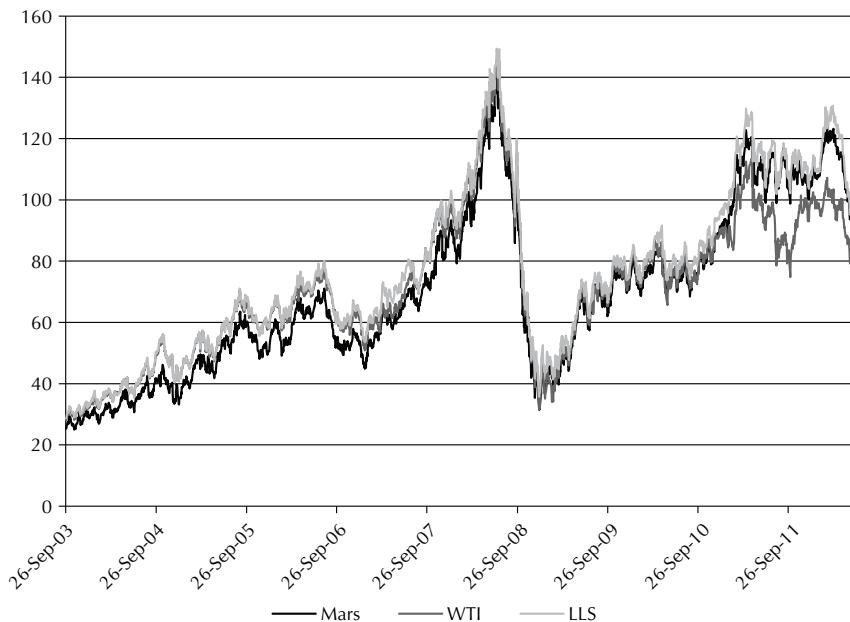
The price fluctuations of the Brent and WTI prices over the relevant period are shown in Figure 17.5. At the time of writing (mid-2012), the tide has turned and we can detect a growing number of trades betting

**Figure 17.5 WTI divergence from Brent**

Source: EIA for WTI, BP for Brent

on a gradual narrowing of the Brent/WTI spread, perhaps back to historical levels, perhaps to about US\$5 premium of Brent over WTI. Opinions differ. It is an obvious indication that it is time to move on to other bets. Figure 17.6 illustrates that the decoupling of WTI from Brent extended to other grades of crude. We can see that WTI dropped below the prices of LLS and Mars crudes, the grades with unrestricted access to the Gulf Coast refineries and representing, therefore, higher value. It is also a good illustration of the WTI dilemma. A refiner using WTI to hedge a refinery processing crudes similar to Mars or LLS would have serious worries about hedge efficiency.

Historically, unusual price relationship with respect to Brent led many market observers to conclude that WTI was a broken benchmark, with obvious and expected protestations from Nymex, the exchange most exposed to an impairment of WTI credibility.<sup>60</sup> It is obvious that the dynamics of WTI prices was harmful to the interests of the oil-producing countries using this benchmark for pricing their oil exports. In autumn 2009, Saudi Aramco made a decision to abandon WTI as a price marker as of January 1, 2010, opting instead to use the Argus Sour Crude Index (ASCI), introduced in May 2005.<sup>61</sup> The exchanges reacted swiftly, offering futures contracts based on

**Figure 17.6** WTI versus LLS and Mars

Source: © 2012 by Platts, a division of The McGraw-Hill Companies, Inc.

the new indexes. ICE Futures announced on November 11, 2009, the introduction of two cash-settled futures contracts based on the ASCI and cleared by ICE ClearEurope. The two futures contracts are the ICE ASCI Futures (an outright contract), and the ICE ASCI, a contract on the spread between the ASCI and WTI prices. The futures contracts, listed by ICE Futures Europe, became available on December 7. This step was preceded by introduction of two OTC-related ASCI contracts, which became available November 13, 2009. NYMEX also offers a futures contract settling on the ASCI, as well as a number of contracts on ASCI related spreads.

### New Pipeline Projects

Congestion at and around Cushing prompted a number of energy companies to explore opportunities for construction of the new, and expansion of the existing pipelines, terminals and storage tanks. Completion even of a subset of proposed pipelines is likely to change dramatically the price dynamics of WTI and its spread to Brent by

2013. At the time of writing, several major projects are either in the initial stages of construction or regulatory approvals.

A joint venture was formed between EPD and Energy Transfer Partners (ETP) to explore a 584-mile pipeline from Cushing to Houston. This project has been cancelled but deserves a mention. Many deals in this business are often revived.

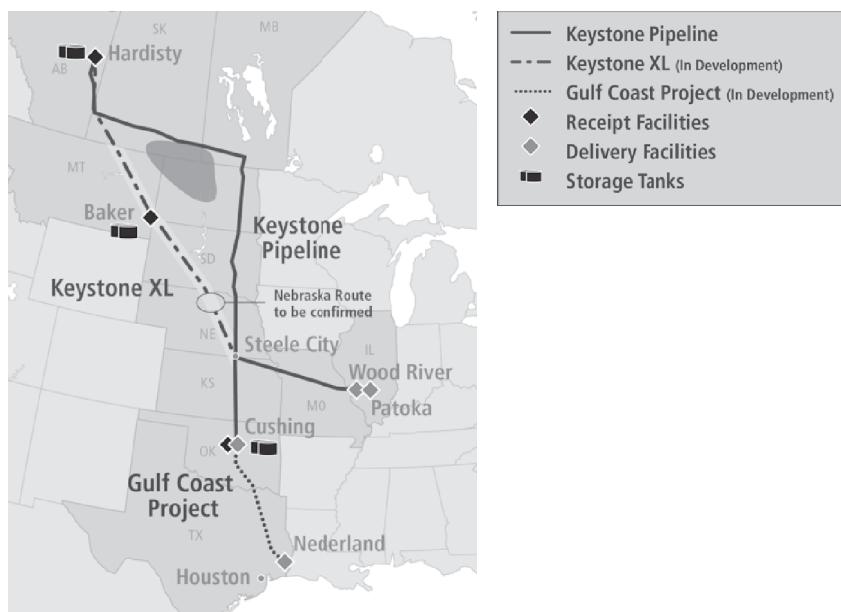
Transcanada continues work on its Keystone Pipeline, a 3,460-kilometre, 590,000 barrels per day project connecting Hardisty, Alberta, to the US markets (Wood River and Patoka in Illinois (completed in 2010), and Cushing (completed in 2011)). The pipeline, through the Dakotas and Nebraska, forks in Steele City, Nebraska, into two branches, one to Illinois and the other to Oklahoma.

The proposed expansion of the Keystone Pipeline (Keystone XL) is highly controversial and may be delayed beyond its initial completion date of 2013. The new 2,673-kilometre (1661-mile), 36-inch pipeline would extend from Hardisty to Nederland, Texas (near Port Arthur), and would incorporate part of existing pipeline (from Steele City to Cushing). The initial capacity is 500,000 barrels per day, with potential expansion to 1,000,000 barrels. The pipeline requires, among other regulatory permits, the approval of the Department of State, as it will cross the US national border. The controversy is related to the quality of crude the expanded pipeline would carry (syncrude from the tar sands in Alberta, with a high content of heavy metals) and as it travels through ecologically sensitive areas, including the Ogallala aquifer in Nebraska.

At the time of writing, the permission for the southern leg of the project (Cushing to the Gulf Coast) was issued, with an expected completion date of 2013. Readers should follow developments closely, not only because of the importance of congestion around Cushing, but also because the outcome of this controversial project is the weather vane for US energy policy.

Other pipeline projects currently under consideration include:<sup>62</sup>

- ❑ Enbridge Energy Company is considering the Flanagan South Pipeline Project, a 600-mile, 585,000 barrels per day 36-inch pipeline connecting Flanagan, Ill, and Cushing, parallel to the existing Spearhead pipeline.
- ❑ Bakken Marketlink, Transcanada: a 100,000 barrels per day of pipeline from Baker, Montana, that would link the northern US

**Figure 17.7 Keystone XL Proposed Route**

Source: Transcanada

portion of the Bakken field with the Keystone XL line, creating a connection to Cushing.

- ❑ Enbridge is also planning the 24-inch, 350,000 barrels per day Monarch Pipeline from Cushing to Houston. The status of the project is unclear at the time of writing.
- ❑ Oneok Partners' Bakken Crude Express (200,000 barrels a day) would connect Bakken to Cushing (possibly by 2015). It would connect to the Overland Pass Pipeline, that could carry some of the oil to Opal, WY.

Given the uncertainty regarding the future of some of these projects, we can expect that the WTI–Brent spread will be highly volatile.

### WTI CME futures contract

The highlights of the Nymex WTI futures contract (symbol CL):<sup>63</sup>

- ❑ Termination of trading: Trading in the current delivery month shall cease on the third business day prior to the 25th calendar day of the month preceding the delivery month.<sup>64</sup>

- Contract unit: 1,000 barrels
- Settlement type: Physical
- Delivery: (A) Delivery shall be made FOB at any pipeline or storage facility in Cushing, Oklahoma, with pipeline access to TEPPCO, Cushing storage or Equilon Pipeline Company's Cushing storage.<sup>65</sup>
- Settlement price: The first six contract months in Nymex WTI crude oil futures (CL), [...] are settled by CME Group staff based solely upon trading activity on CME Globex between 14:28:00 and 14:30:00 Eastern Time (ET).<sup>66</sup> On the day of expiration, the expiring month will settle based on the volume-weighted average price (VWAP) of the outright CME Globex trades executed between 14:00:00 and 14:30:00 ET, and the second month will settle based on the VWAP of the outright CME Globex trades executed between 14:28:00 and 14:30:00 ET.<sup>67</sup> (See the documents in this footnote for additional important details).
- Venue: CME Globex, CME ClearPort, Open Outcry (New York).

### **Physical WTI trading**

Physical transactions related to WTI evolve around two types of arrangements known as P+ and CMA. Due to space limitations, the discussion of the P+ and CMA markets is a bit simplified. A detailed discussion of both markets can be found in recent posts by Sandy Felden at <http://www.rbnenergy.com/>.

#### *P+ price*

A posted price concept is quite important in the energy oil markets. This is a price at which various refiners and gatherers purchase physical domestic crudes. In the old days, the bulletin stating the current price level would be nailed to the refinery gates (hence the term "posted price"). There are different postings for different gathering points. The postings for the Cushing location are usually specified at a discount to Nymex WTI prices in order to compensate the gatherer for transportation costs.

The Platts document offers the following summary of the pricing rules:<sup>68</sup>

*P-Plus WTI:* The assessment reflects the price of WTI sold into Cushing on the basis of "postings plus." P-plus deals are invoiced at a later date on the basis of a differential to an average of one or more

crude oil postings. For example, a deal done at P-plus 75 cts would be invoiced at 75 cts more than the previously agreed-upon postings basis.

*WTI Calendar Delta:* The assessment reflects the price of WTI crude oil sold into Cushing/Oklahoma on the basis of a delta versus a monthly WTI average. WTI Calendar Delta deals are invoiced at a later date: For instance, March WTI calendar delta transactions would be based on the average of the Nymex WTI front-month during March, plus or minus a delta, and then versus cash front-month WTI after the Nymex WTI front-month expiry. The delta fluctuates with first/second and first/third month WTI spreads, and with bids/offers in the market. The Platts WTI Calendar Delta assessment reflects where the delta is traded and/or talked in the market.<sup>69</sup>

A premium over the posting price, which may be paid for barrels delivered at the gathering point, is referred to as the *posting bonus* or *P-plus* price.

The following relationship holds:

$$\text{Price of delivered physical} = (\text{Posting price}) + (\text{Posting bonus})$$

During the delivery month, adjustments in the daily posting price tend to follow the Nymex prompt WTI contract, although posting prices are adjusted only in US\$0.25 increments. The difference between the prompt Nymex and the WTI posting is called the *posting basis*.

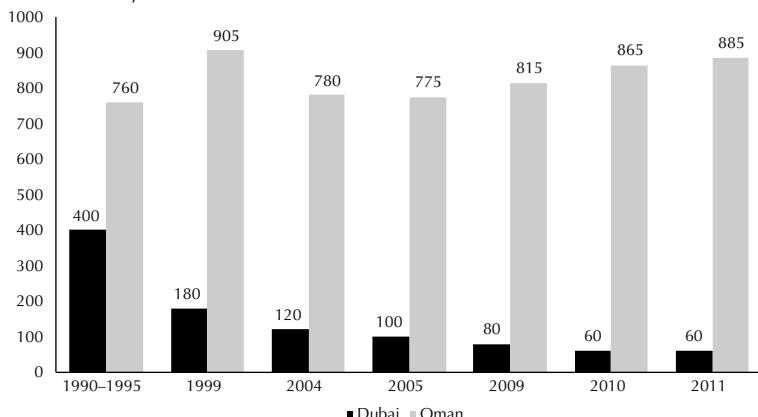
So, by definition of the posting basis, we have:

$$\text{WTI prompt contract} = \text{Posting price} + \text{Posting bonus} + \text{Posting basis}$$

Historically, the industry was relying on the Koch posted price for Cushing. This price was discontinued a few years ago, and physical traders switched to the ConocoPhillips posted price.

A Calendar Merc Average (CMA) transaction is a link between an expiring WTI contract and monthly physical markets. The cash window at Cushing is a three-day period between the expiration of the prompt WTI futures contract and the beginning of the delivery month. The transactions in this time period are used to adjust positions, make delivery arrangements with customers, storage facilities and pipelines. Prices are based on an average of the WTI settlement prices of the prompt contract during the delivery month plus/minus a negotiated differential. Trading during the cash window provides information about tightness of the physical market around Cushing and is an important source of market information. CMA

**Figure 17.8** Dubai and Oman crude production estimates (thousand barrels a day)



Source: Thomas Leaver, 2012, "Continuing evolution in East of Suez markets: The growing case for benchmark change," MPCG, Bahrain, May 7

contracts are traded electronically on a specialised platform called Houston Street,<sup>70</sup> through voice brokers and directly between the counterparties.

### OMAN–DUBAI

The third price benchmark, Oman–Dubai crude, used by the Persian Gulf countries and the Asian importers, has a unique set of issues associated with its design and applications. The volume of priced crudes based on a combination of futures prices and price assessments at these two locations exceeds 10 million barrels a day.<sup>71</sup> The preference for these two benchmarks can be explained by quality issues. Persian Gulf crudes are more sour than both WTI and Brent, and the reliance on price benchmarks representing oil of higher quality may lead to mispricing and mishedging. What makes Oman–Dubai comparable to Brent and WTI is falling physical production in the case of the Dubai crude, as illustrated in Figure 17.8. Two pricing services, Platts and Argus, attack this problem in significantly different ways, creating a potential for arbitrage transactions.

Platts started Dubai and Oman spot price assessments in January 1984.<sup>72</sup> Falling output of Dubai crude affected the performance of the index and Platts decided to address the problem by allowing the delivery of Oman crude, with stable production levels, into Dubai

purchase contracts as of November 16, 2001. This decision revived the index but did not address another market concern: the standard size of Dubai–Oman contract corresponding to the standard cargo size of 500,000 barrels.<sup>73</sup> Many end users had exposure that did not correspond to the multiples of this volume, and the large volume transactions could not be used effectively for hedging. Platts addressed this problem in a creative way by advancing the concept of *partials*, smaller size transactions (25,000 barrels) settled in cash,<sup>74</sup> which could be converted into full tankers if a sufficient number of contracts for the same delivery date was traded between two parties (19 contracts totalling 475,000 barrels were required for conversion into a physical cargo, given a loading tolerance).<sup>75</sup> Partials were accepted by the market, leading to an increase in the number of participants, including non-equity players. The bid–offer spreads were also reduced. However the volumes of these contracts remain rather low.

For example, as reported by the Dubai Mercantile Exchange (DME),<sup>76</sup> no Oman partials traded on 93% of trading days during the period September 1–December 31, 2010. The Dubai market for partials is more liquid (in terms of higher volumes) but is concentrated on (in terms of the number of participants) both sides (ie, buy and sell). Given that the Oman partials are assessed based on the Dubai benchmarks, Oman assessments can be influenced by a small number of players.

Argus uses an alternative pricing solution to address the low liquidity in the market for Dubai crude. Their approach relies on the financial instrument that links Brent and Dubai prices: the Exchange for Swaps (EFS) contract, which reflects the difference between the two markets. This instrument allows the market participant to hedge their positions in Persian Gulf crudes with Brent, covering the basis risk at the same time.<sup>77</sup>

The DME was launched in 2005, through as a joint venture between Tatweer (a member company and subsidiary of Dubai Holdings) and Nymex. DME has a number of stakeholders: a subsidiary of Dubai Holding owns 9% of the company, Oman Investment Fund has 29%. CME Group holds a stake equal of 50% and strategic investors (including Goldman Sachs, JP Morgan, Morgan Stanley, Shell, Vitol and Concord Energy) account for another 12%. The first futures contract on the DME was initiated in 2007. The DME now offers three contracts:

- 
- DME Oman Crude Oil Futures Contract (OQD)
  - DME Oman Crude Oil European Style Option Contract (OQE)
  - DME Oman Crude Oil Financial Contract (ZGD)

The OQD contract is used for calculation of the official selling price (OSP) for both Oman (since June 2007) and Dubai (since June 2009). The Sultanate of Oman sets its OSPs on a forward-pricing basis using the monthly average of the DME Oman daily settlements. The Dubai Department of Petroleum Affairs (DPA) uses a monthly average of the futures prices adjusted by a monthly differential.

In February 2009, the DME switched to the CME Globex platform, on which two other benchmarks are also traded (WTI, Brent). In December 2010, CME launched a number of DME-related contracts:

- DME Oman Crude Oil Swap Futures;
- DME Oman Crude Oil BALMO Swap Futures;
- ICE Brent (Singapore Marker) versus DME Oman Crude Oil Swap Futures;
- ICE Brent versus DME Oman Crude Oil Swap Futures;
- DME Oman Crude Oil Average Price Option;
- DME Oman Crude Oil versus Dubai (Platts) Swap Futures;
- Singapore MOGAS 92 Unleaded (Platts) versus DME Oman Crude Oil Swap Futures; and
- Singapore Gasoil (Platts) versus DME Oman Crude Oil Swap Futures.

As mentioned above, Persian Gulf crudes going to Asia are priced mostly on Platts indexes. The OSP of the Kingdom of Saudi Arabia, Kuwait, Iraq and Iran is given by:

$$OSP_M = \frac{PlattsOman_{M+2} + PlattsDubai_{M+2}}{2} + \text{Adjustment factor where } M \text{ stands for month}$$

Falling liquidity in the partials market may increase the probability that the DME crude futures contract will be elevated to the level of regional benchmark. In April 2012, the daily volumes of this contract jumped several times over the 10,000 mark, a critical level for its credibility.<sup>78</sup> There are, however, several reasons that may slow down recognition of this contract as a pricing benchmark.<sup>79</sup> The

contract, the biggest physically settled oil futures contract, is seen more as a conduit to local supply than a price discovery mechanism. The exporters are conservative and reluctant to switch to a different pricing mechanism from the traditional system based on the Platts assessments. Finally, a system of derivative transactions linking the Dubai market to Brent developed over time, creating alternative hedging solutions. Specifically, Brent/Dubai EFS transactions and Dubai calendar swaps, traded actively in the OTC markets, allow traders to transform Dubai price exposure into Brent exposure.

## CONCLUSIONS

The oil pricing regime in place since the late 1980s has demonstrated both flexibility and resilience in spite of many trends eroding its foundations: shifts in geographical distribution of production, divergence of quality of crude oil produced worldwide compared to the grades behind the benchmarks, dwindling flows of Brent and related crudes, logistical challenges in the North Sea and the US, emergence of other market participants such as big Chinese oil companies and national oil companies (NOCs). One of the reasons for the durability of this system was the absence of viable alternatives. Inertia and the difficulty of a synchronised transition of the entire industry to a new pricing framework is an additional explanation. At some point, however, a change will take place. One potential solution is the development of a system based on an index representing an average of multiple grades. The BFOE standard may be an early precursor of such a solution.

- 1 Based to a large extent on Bassam Fatouh, "The origins and evolution of the current international oil pricing system," in Robert Mabro (ed), 2006, *Oil in the 21st Century: Issues, Challenges and Opportunities* (Oxford, England: Oxford University Press).
- 2 In 1950, the majors controlled 85% of global crude oil production outside North America and the Soviet Bloc.
- 3 The term *seven sisters* (sette sorelle) was coined by Enrico Mattei, the head of Italian company Ente Nazionale Idrocarburi, after his efforts to join the so-called Iranian Consortium were ignored. They were: Standard Oil of New Jersey (a predecessor of Exxon Mobil), Royal Dutch Shell, Anglo-Persian Oil Company (a predecessor of BP), Standard Oil Company of New York (a predecessor of Mobil, and by extension Exxon Mobil), Standard Oil of California (a predecessor of Chevron), Gulf Oil (which became part of Chevron), Texaco (merged with Chevron in 2001). The group of original sisters shrank over time, often through same-sex marriages.
- 4 Posted prices are made public by sellers and buyers that declare their intention to transact at certain levels for a period of time. In the early days of the industry, the price lists were liter-

- ally nailed to the gates of a refinery. The concept of a posted price survives in an obscure corner of the US crude markets as the so-called P+ price (as explained later).
- 5 Robert Mabro, 2005, "The international oil price regime. Origins, rationale and assessment," *The Journal of Energy Literature*, (XII)1, June, pp 3–20.
- 6 The history of OPEC and oil price shocks is covered in a number of books, including Dan Yergin, 1991, *The Prize: The Epic Quest for Oil, Money, & Power* (New York, NY: Simon & Schuster); Eric Laurent, 2006, *La Face Cachée du Pétrole* (Paris, France: Plon). The efforts to establish informal consultations between the oil-producing countries go back to the late 1940s, following the first initiatives undertaken by Venezuela. The creation of OPEC during the Baghdad meeting (September 10–14, 1960) was the result of the efforts by Venezuelan Energy and Mines minister Juan Pablo Pérez Alfonzo and the Saudi Arabian Energy and Mines minister Abdullah al-Tariki. The secretariat of OPEC was initially located in Geneva, but following the denial of diplomatic status by the Swiss authorities, it relocated to Vienna in 1965. During the first 10+ years of its existence, OPEC did not receive much attention, but this oversight by the world powers has been more than compensated in subsequent decades.
- 7 Discords inside OPEC lead to the emergence of a two-tiered price structure in the late 1970s, with a one price marker set by Saudi Arabia and a competing price marker established by some other countries.
- 8 "The netback contract, with its few merits and small defects, is now associated with a disastrous event in the history of the oil industry. Its prospects as a method, among alternatives, for disposing of crude oil may suffer from the bad memories of this unlucky association." See Robert Mabro, 1987, *Netback Pricing and the Oil Price Collapse of 1986* (Oxford, England: Oxford Institute for Energy Studies).
- 9 The statement was made famous by Friedrich Hayek, but was coined by Adam Ferguson in 1767 in "An essay on the history of civil science." "Every step and every movement of the multitude, even in what are termed enlightened ages, are made with equal blindness to the future; and nations stumble upon establishments, which are indeed the result of human action, but not the execution of any human design."
- 10 Paul Horsnell of the Oxford Institute for Energy Studies remarked that "a major feature of the Brent market is that it works extremely well as long as one does not think about it too hard."
- 11 "In 1971 the Brent Field was discovered by Shell/Esso and tested in 1972 with 1.8 billion barrels of recoverable oil; nine major Brent sandstone fields were discovered by the end of 1973 (Brenn *et al.* 1990). In 1980 the northern North Sea (overwhelmingly comprising fields with Brent Group reservoirs) was ranked as the 13th largest petroleum province in the world, containing 1.6% of produced and recoverable oil equivalent reserves (Ivanhoe 1980). By 1988, discovered Brent hydrocarbons comprised some 49% of the UK's recoverable reserves, totaling 22.5 billion barrels of oil equivalent." See A. C. Morton, R. S. Haszeldine, M. R. Giles and S. Brown, 1992, "Geology of the Brent Group: Introduction," *Geological Society, London, Special Publications*, 61; p 1–2.
- 12 [http://tonto.eia.doe.gov/ask/crude\\_types1.html](http://tonto.eia.doe.gov/ask/crude_types1.html).
- 13 Today the typical size of a Brent cargo (and the standard volume of a dated Brent contract) is 600,000 barrels.
- 14 Lay-days is the time a party in a charter contract is allowed for taking in and discharging cargo.
- 15 Shell amended the 1990 "Agreement for the sale of Brent Blend crude oil on 15-day terms" in April 2002. With the cargo size of 500,000 bbls, the buyer had the right to nominate an actual volume of 475,000–525,000 bbl. The change reduced this interval to 495,000–505,000 bbl.
- 16 Accepting a cargo is called *lifting* a cargo.
- 17 This practise ("four o'clocked" now) is much less frequent today and is frowned upon in a polite society.

- 18 Based on <http://www.barrettwells.co.uk/21daybfos.html>.
- 19 The British National Oil Corporation formed in 1975, eventually privatised and bought by BP (also known as Britoil).
- 20 This claim is contested by the CME (see Robert Levin, 2009, "WTI: Understanding the global crude oil benchmark," March 12 (<http://www.cmegroup.com/trading/energy/files/WTI-Global-Benchmark.pdf>); "Why the 'two-thirds' myth for oil price referencing doesn't add up" (<http://archive.opnmkts.com/energy/why-the-%E2%80%9Ctwo-thirds%E2%80%9D-myth-for-oil-price-referencing-doesn%E2%80%99t-add-up/>)).
- 21 Stella Farrington, "Blending the rules," *Energy Risk*, January 2005
- 22 Physical Brent has been redefined for the first time in 1991, when Ninian crude was assigned to the Brent stream (to counter falling Brent volumes). The term Brent Blend was coined at this time. Crude from the Forties is delivered to Hound Point terminal in Scotland located in the Firth of Forth; Oseberg crude goes to Sture in Norway (near Bergen).
- 23 The API of the oil from Buzzard field is about 32, compared to 45 at the Forties. This is not necessarily bad news, as addition of Buzzard makes the modified Brent standard resemble closer the average quality of oils produced worldwide.
- 24 Ekofisk crude is delivered to the Teeside terminal.
- 25 "The latest BP assay sampled in April 2007 shows Forties Blend to have an API of 41.8° and a sulphur level of 0.49%. This is with a Buzzard composition in the Blend of approximately 22%." See <http://www.offshore-technology.com/projects/buzzard/>.
- 26 <https://www.theice.com/productguide/ProductDetails.shtml?specId=219>.
- 27 On January 6, 2012, the 21-day time window for Brent cargoes was extended to 25 days at the initiative of Platts.
- 28 <http://www.platts.com/IM.Platts.Content/methodologyreferences/methodologyspecs/crudeoilspecs>.
- 29 Javier Blas, 2012, "Korean buying spree boosts Brent price," *Financial Times*, June 7.
- 30 Ibid.
- 31 Fernando Barrera-Rey and Adam Seymour, 1996, "The Brent contract for differences (Cfd). A study of an oil trading instrument, its market and its influence on the behavior of oil prices," Oxford Institute for Energy Studies.
- 32 As explained by Platts, "the first weekly balance is on a forward week basis on Thursday and Friday, and becomes a balance week quotation between Monday and Wednesday. It is rolled forward every Thursday." See <http://www.platts.com/IM.Platts.Content/methodologyreferences/methodologyspecs/crudeoilspecs.pdf>.
- 33 The market convention is to use the second available forward.
- 34 As in "This ain't no pipe", a double negative in logic always means a positive statement, or "minus minus" in arithmetic translates into a positive number.
- 35 Dated Brent is referred to as Dated North Sea Light (Platts) or Argus North Sea dated.
- 36 Platts proposed in mid-2011 to expand this window to 25 days. The change was implemented in 2012. The rationale behind it was increasing the number of cargos covered in the price assessments.
- 37 The Chicap Pipeline is operated by BP, Southern Light by Enbridge (this pipeline carries diluents to Alberta).
- 38 [http://www.cmegroup.com/trading/energy/crude-oil/dated-to-frontline-brent-crude-oil-swap-futures\\_contract\\_specifications.html](http://www.cmegroup.com/trading/energy/crude-oil/dated-to-frontline-brent-crude-oil-swap-futures_contract_specifications.html).
- 39 Specific percentages are contested but the time to dethrone the Brent benchmark has not come yet.
- 40 As explained above, the buyers are notified about the three-day loading window 25 days ahead of delivery.
- 41 Platts, 2009, "Methodology and specifications guide," *Crude Oil*, December (<http://www.platts.com/IM.Platts.Content/MethodologyReferences/MethodologySpecs/crudeoilspecs.pdf>).

- 42 <https://www.theice.com/productguide/ProductSpec.shtml?specId=219>.
- 43 [https://www.theice.com/publicdocs/futures/ICE\\_Futures\\_Europe\\_Brent\\_Index.pdf](https://www.theice.com/publicdocs/futures/ICE_Futures_Europe_Brent_Index.pdf).
- 44 "Saudi switch to Brent futures pricing gets mixed reviews," *Reuters*, March 30, 2001.
- 45 "IOSCO consultation paper – Oil Price Reporting Agencies," ISDA, March 2012.
- 46 [http://www.cmegroup.com/trading/energy/crude-oil/light-sweet-crude\\_contract-specifications.html](http://www.cmegroup.com/trading/energy/crude-oil/light-sweet-crude_contract-specifications.html).
- 47 "Pricing differences among various types of crude oil" ([http://tonto.eia.doe.gov/ask/crude\\_types1.html](http://tonto.eia.doe.gov/ask/crude_types1.html)).
- 48 The OPEC Reference Basket, as defined on June 15, 2005, consists of a weighted average of the following blends of oil: Saharan Blend (Algeria), Minas (Indonesia), Iran Heavy, Basra Light (Iraq), Kuwait Export, Es Sider (Libya), Bonny Light (Nigeria), Qatar Marine, Arab Light (Saudi Arabia), Murban (UAE), BCF 17 (Venezuela).
- 49 K. D. Miller, M. T. Chevalier and J. Leavens, 2010, "The role of WTI as a crude oil benchmark prepared for CME," *Purvin & Gertz*, January.
- 50 Shell storage capacity is the maximum volume that can be stored at any point in time. The capacity available to the industry is typically much smaller because some storage is lost to heel volumes (ie, the volumes that cannot be removed due to operational constraints) and the volumes required for blending and for other technical reasons.
- 51 See Gregory Meyer, 2010, "Oil traders look to James Bond-style data collection," *Financial Times*, March. Most oil and product tanks have floating roofs, installed to avoid accumulation of vapours above the surface of liquids stored in the tank, and reduce the risk of explosions. The position of a roof can be determined from a picture taken from above. "Hired by Genscape, an energy industry data provider, the helicopter pilot brings back color and infrared images to help peer at – and inside – 322 tanks around Cushing. Analysts then examine the images to discern how much oil is parked there."
- 52 The pipeline supports multiple flows of oil, as explained in the company financial filings: "The Basin system consists of four primary movements of crude oil: (i) barrels that are shipped from Jal, New Mexico, to the West Texas markets of Wink and Midland; (ii) barrels that are shipped from Midland to connecting carriers at Colorado City; (iii) barrels that are shipped from Midland and Colorado City to connecting carriers at either Wichita Falls or Cushing; and (iv) foreign and Gulf of Mexico barrels that are delivered into Basin at Wichita Falls and delivered to connecting carriers at Cushing." <http://www.b2i.us/profiles/investor/secxml.asp?Bzid=789&fg=1>.
- 53 <http://www.enbridgeus.com/Main.aspx?id=230&tmi=230&tmt=1>.
- 54 <http://www.businessweek.com/articles/2012-05-16/unlocking-the-crude-oil-bottleneck-at-cushing#p2>.
- 55 <https://www.loopllc.com/>.
- 56 "The first continuous delivery of western Canadian crude oil was initiated in early March 2006 through Spearhead Pipeline to Cushing, Okla. The 650-mile, 22- and 24-inch diameter pipeline historically operated in south-to-north service, but Enbridge bought the pipeline and reversed its flow to provide Canadian crude oil producers and shippers with access to markets in the Mid-Continent and southern United States. Enbridge recently expanded the Spearhead Pipeline, increasing the average annual capacity from 125,000 barrels per day (bpd) to 190,000 bpd." (See <http://www.enbridgeus.com/Main.aspx?id=230&tmi=230&tmt=1>).
- 57 Purvin & Gertz, p 45.
- 58 Ibid, p 112.
- 59 William Shakespeare, *Julius Caesar*, Act 4, Scene 3.
- 60 "US West Texas Intermediate crude futures traded on the Nymex dropped to a record US\$11-per-barrel discount to Brent last week amid concerns about brimming stocks at the Cushing, Oklahoma, storage hub, leading some analysts to question whether WTI remained a useful benchmark for the global oil market. Defending the contract as 'the most transparent benchmark crude,' Nymex researchers said they believe there is at least five million

barrels of operable storage space left at Cushing and denied there were any logistical problems looming there that would force prices lower." Joshua Schneyer, 2009, "Nymex defends WTI, says Cushing storage ample" (<http://uk.reuters.com/article/idUKN205732792009120>).

- 61 This index is based on three US produced blends: Mars, Poseidon and Southern Green Canyon. See Izabella Kaminska, 2009, "Saudi Aramco's WTI snub," *FT Alphaville*,. As reported in this post, "Saudi Aramco will publish a monthly price differential to a month's average of the daily ASCI price published by Argus. Argus prices are already extensively used in the US midcontinent and Gulf coast crude markets to price long-term supply contracts."
- 62 A word of caution: the situation is fluid: some projects can be delayed or cancelled and other projects can be added to the list.
- 63 [http://www.cmegroup.com/trading/energy/crude-oil/light-sweet-crude\\_contract\\_specifications.html](http://www.cmegroup.com/trading/energy/crude-oil/light-sweet-crude_contract_specifications.html).
- 64 Some details have been omitted. Deliverable crude streams include West Texas Intermediate Low Sweet Mix (Scurry Snyder), New Mexican Sweet North Texas Sweet, Oklahoma Sweet, South Texas Sweet, Brent Blend Bonny Light, Qua Iboe, Oseberg Blend and Cusiana. For more on grade and quality specifications, see <http://www.cmegroup.com/rulebook/NYMEX/2/200.pdf>.
- 65 Some details have been omitted.
- 66 [http://www.cmegroup.com/trading/energy/files/NYMEX\\_Energy\\_Futures\\_Daily\\_Settlement\\_Procedure.pdf](http://www.cmegroup.com/trading/energy/files/NYMEX_Energy_Futures_Daily_Settlement_Procedure.pdf).
- 67 [http://www.cmegroup.com/trading/energy/files/NYMEX\\_Energy\\_Futures\\_Final\\_Settlement\\_Procedure.pdf](http://www.cmegroup.com/trading/energy/files/NYMEX_Energy_Futures_Final_Settlement_Procedure.pdf).
- 68 Platts. 2010, "Methodology and specifications guide," *Crude Oil*, October.
- 69 Platts considers postings of the following companies Plains, Sunoco, Shell, Murphy and ConocoPhillips.
- 70 <http://www.houstonstreet.com/>. The platform supports trading of many physical crudes, including: St.James, Teppco Cushing, Guernsey, Empire, Houma, Clovelly, Teppco Midland, ExxonMobil Midland, Patoka, and Equilon Basin Pipeline, Wood River, Guernsey Express, Mokena Enbridge. Grades traded include HLS, LLS, Bonito Sour, Eugene Island, Poseidon, Mars, WTS, WTI, Foreign Sweet, Foreign Sour, Domestic Sweet, Hoover-Diana and Wyoming Sweet. It also supports Canadian grades including: Lloyd Blend, Lloyd Wainwright, Bow River and Mixed Sweet.
- 71 Adrian Binks, 2005, "Middle East crude pricing: The Dubai debate," *Middle East Economic Survey*, (XLVIII)7, February 14. According to the DME, 0.9 million barrels a day are priced based on the exchange futures prices versus 13.1 million based on Platts assessments. See "DME/CME Forum," March 27, 2012, Dubai.
- 72 "Platts, 2004, Review of the partial Dubai and Oman price discovery mechanism."
- 73 The operators of the terminals did not allow for loading cargos smaller than 200,000 barrels, with both countries pricing their exports on the FBO basis for standard size tankers.
- 74 The partials are traded during a 30-minute trading window between 5:00 pm and 5:30 pm Singapore time. Platts moved the Dubai-Oman assessment window from New York to Singapore on June 3, 2002.
- 75 Partials were used in the 1980s in the Brent and WTI markets, but the development of the futures markets pre-empted the need for such arrangements. Brent partials are still used by the traders in Asia to trade Brent-Dubai spreads.
- 76 Thomas Leaver, 2011, "DME and the role of oil price benchmarks," DIFC Oil Trade and Finance Conference, Dubai, March 28.
- 77 "Argus assesses the price of Dubai in the same way that the market assesses the value of Dubai. Companies that trade the Dubai market convert their Dubai price exposure into a Brent price exposure. They do this because Brent price exposure is easier to manage. The underlying traded volumes of Dubai swaps and the associated ICE Brent contract provide a

more reliable and transparent basis for assessing the price of Dubai than any other methodological approach. *Argus* begins its assessment of the Dubai price by reporting the active market in Dubai swaps based in Singapore. Dubai swaps are traded as an exchange of futures for swaps (EFS). *Argus* identifies and reports the prevailing EFS price, expressed as a differential to the ICE Brent price. This EFS value is subtracted from the prevailing price of Brent on the ICE at 4:30 pm Singapore time to establish a fixed price for Dubai swaps. *Argus* identifies the EFS value for the first three months forward. Applying these differentials to the first three months of the ICE Brent price provides *Argus* with absolute values for the first three months of the Dubai swaps market.<sup>78</sup> See “*Argus crude: Methodology and specifications guide*” ([http://web04.us.argusmedia.com/ArgusStaticContent//Meth/crude\\_meth\\_latest.pdf](http://web04.us.argusmedia.com/ArgusStaticContent//Meth/crude_meth_latest.pdf)).

- 78 Javier Blas, 2012, “Rise of the Oman crude benchmark,” *Financial Times*, May 9.
- 79 Bassam Fattouh, 2012, “The Dubai benchmark and its role in the international oil pricing system,” *Oxford Energy Comment*, March.



## *Transactions in the Oil Markets*

As was explained in Chapter 4, forward and futures can be used for hedging, speculation or arbitrage. This chapter will review some of the most popular structured transactions used in the oil and refined product markets, combining different derivative instruments covered earlier in the book (Chapters 4 and 5). Many of these structures have been designed to offer the hedging companies more flexibility in structuring their risk-mitigation strategies. This can be accomplished in a number of ways:

- ❑ combining in one structure instruments of different types (for example, swaps and options or options of different types) offering the buyer ability to execute the hedge in one step, instead of building the hedging portfolio step by step from atomic components);
- ❑ combining in one structure the same instruments (for example, forwards) related to different underlying physical commodities; and
- ❑ combining different instruments and different underlying commodities.

The techniques used to structure hedging instruments that will be discussed apply to different markets. The principles are the same, although specific conventions will vary from commodity to commodity.

The ability to use more complex strategies offers a number of benefits to end users. To be able to transact in one step and rely on the financial engineering skills of hedge providers offers time and cost savings to end users. Accounting rules push end users towards hedging strategies that match their underlying exposures and which

require advanced customisation. The disadvantage is that many structured transactions may be difficult to value for all but the most sophisticated market participants. A complex transaction may be a convenient opportunity for locking up the end user in a transaction that is difficult to unwind (or can be unwound at a considerable cost), and has significant profit margins accruing to the provider that are buried in a convoluted structure. The final decision regarding whether to rely on more complex structures or use plain vanilla hedges is a difficult one, and requires significant investment in risk management skills.

We will start with a review of oil and refined product market participants and a discussion of their main objectives, as well as the constraints they face. This will be followed by a description of strategies such as collars, 3:2:1 crack spreads and participating swaps. The chapter will end with a brief description of the market for freight derivatives.

### **CRUDE AND REFINED PRODUCTS: MARKET PARTICIPANTS AND TRANSACTIONS**

Unlike the North American natural gas market, which is a practically closed system (except for very insignificant LNG flows), oil and refined products are traded worldwide. This global character of this market creates a number of unique challenges related to the pricing of complex trades, collecting fundamental information and managing the positions in a trading book, given that the underlying physical commodities and financial instruments are traded continuously around the globe. The market participants include a number of different categories of companies and financial institutions with varying objectives, including:

- producers of crude oil;
- refiners;
- consumers and end users of refined products;
- commercial and investment banks; and
- investors:
  - hedge funds;
  - index funds; and
  - institutional investors.

Producers of crude oil participate in the market as sellers of their volumetric flows and as hedgers seeking to eliminate downside risk. Some big integrated oil companies follow a corporate policy of riding the crude oil price cycles without hedging, based on a number of considerations, such as the following.

- ❑ The conviction that the market for oil and refined products-related derivatives is not sufficiently deep and liquid to provide insurance for their entire production stream. This may certainly be true for some of the largest oil companies.
- ❑ The investments in oil fields and fixed assets create exposures that have a horizon as long as 30 years, and the markets do not offer instruments with corresponding maturities.
- ❑ The economic exposure of an integrated oil company (ie, company with operation both upstream and downstream) is with respect to crack spreads (ie, the price differentials between refined products – baskets of refined products – and different types of crude), and the ability to hedge this exposure may be limited as the markets related to different legs of the spread have different depth and liquidity. Vertical integration provides a level of protection against price fluctuations as adverse price movements on one side of the crack spread can be partially or completely offset by a compensating change in prices on the other side. The use of transfer prices between different units of the same company located in different jurisdictions, and complex tax issues, create additional impediments to active hedging. As an executive of a big, integrated oil company once told the author: “Our problem is in which part of the company we are going to book the profits.”
- ❑ Some big integrated oil companies believe that they have sufficiently strong balance sheets to allow them to ride the market. Incurring incremental costs to hedge their exposures, sometimes with the counterparties that may represent lower credit quality, is not necessarily a wise strategy.
- ❑ Some big oil producers take a different view and operate massive trading units, covering not only exposures related to their operations in the crude, natural gas and products markets, but also offering services to other market participants across the entire spectrum of the physical and derivative energy markets. Such

trading units are treated as profit centres, and they make a significant contribution to the overall bottom line. The sophistication of their trading units matches, and sometimes exceeds, the competence levels of their counterparts in the financial institutions.

Oil producers who engage in hedging typically cover their exposures over a horizon of two or three years, although longer-term products in the markets with sufficient liquidity are not unheard of. The volumes typically decline over time reflecting a generally sound strategy of leaving certain volumes uncovered. Some producers are active hedgers, adjusting their exposures frequently and unwinding portions of the forward positions as they move closer to expiration.<sup>1</sup> Opportunities for large-scale transactions occur when producers engage in periodic adjustments to the hedging programmes (often close to the end of calendar years) and when large-scale transactions related to asset acquisitions and mergers take place. In the latter case, the existing hedges often have to be unwound and new positions put in place, creating an opportunity not only for those traders who were fortunate enough to win the business, but also for those who were able to correctly anticipate the market moves associated with a big transaction.

The end users hedge their price exposures related to the buying of feedstocks to industrial processes or the refined products they use as a source of heat or transportation fuel. As in the case of producers, the end users typically put on hedges for two or three years out, extending the hedges up to five years with sufficient market liquidity. Some marketing companies, serving retail consumers of heating oil and propane, hedge extensively the fixed price programmes offered to their customers.

Both the end users and producers have a strong preference for hedging in the OTC markets, and rely extensively on Asian swaps and options.<sup>2</sup> Asian swaps and options settle on an average price calculated over a specified time window, using publically available market prices. The users of hedging instruments believe that such structures correspond better to the profiles of the underlying exposures. Another frequently mentioned rationale for such options is protection against market manipulation and random price shocks in relatively shallow markets.

In the US, the most popular structures use daily settlement prices of the Nymex WTI contract for calculations of the average price,

although an OTC contract can be structured in any way. One complication in pricing Asian swaps is that the WTI futures contract matures during a calendar month. This means that the average monthly price used in a swap settlement references the prices of two futures contracts. This creates a set of unique modelling challenges in pricing swaps and options. In the case of options, the prices and volatilities of two separate futures contracts have to be used in the pricing model, and the extent of co-movement between two prices has to be captured.

Over the last few years, there has been a pronounced trend towards greater participation of investors in the oil markets. In addition to commodity trading advisors (CTAs) investing on behalf of their clients, many hedge funds have moved into this space. Hedge funds operate along the entire forward price curve, with a pronounced interest in long-dated options and variance swaps. These strategies reflect the expectations of potential disruptions in the oil markets related to geopolitical events and, potentially, oil production failing to keep up with growing demand. Hedge funds often behave opportunistically, entering and exiting the market at a very short notice.

Institutional investors (such as pension funds, university endowments, mutual funds) and, in some cases, sophisticated individual investors see the oil and refined product markets as a diversification strategy or a hedge against risks related to geopolitical events, inflation and spiking interest rates. Growing imbalances in the international economy create potential risks that cannot be managed with traditional financial instruments. The institutional investors have a longer horizon and do not engage in short-term transactions (except for regular rollovers of futures positions). Some investors participate in the energy commodity markets through the exchange traded funds (ETFs). USO is an example of such a fund.

Both producers and end users have shown an increasing interest in using option-based strategies. This preference may be explained by a number of factors, primarily the desire to eliminate the uncertainty related to collateralisation and margining of the swap and forward/futures positions. Hedging by buying calls and puts is associated with a cash outflow at inception, but the cost of strategy is known, with no risk of unexpected and costly calls for additional collateral, due to price movements or certain credit events.<sup>3</sup> Some producers and end-users use an option strategy known as a fence (or

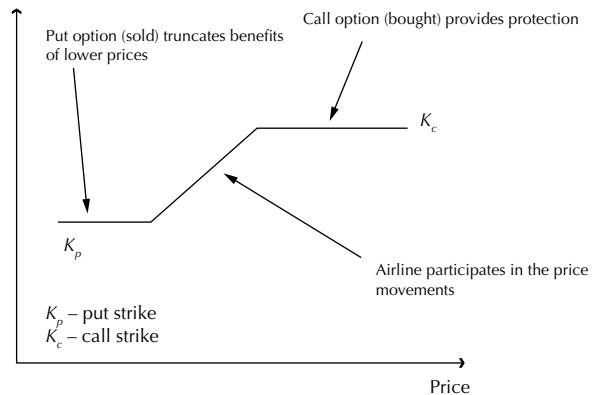
collar), which requires a closer scrutiny. These structures may be quite risky.

## EXAMPLES OF TRANSACTIONS

### Collars

A collar is a strategy consisting of buying an option as a hedge and selling at the same time another option in order to finance (fully or partially) the purchased option. This structure has been covered in Chapter 5 but we shall repeat a few details for the reader's convenience. For example, a producer may buy a put (a floor) as a protection against a price drop, selling at the same time a call with a strike greater than the strike price of a put in order to generate positive cashflow, which helps to reduce the cost of their hedging strategy. The logic behind it is that a producer is willing to give up part of the upside in order to reduce (or completely offset) cash outflows related to hedging. Carefully calibrated strike prices may make the cost of buying puts equal to the proceeds from selling calls, making the transaction cash-neutral (at inception). Such structures are called *costless collars*, although, as we shall see, these are often quite expensive. The confusion is related to the distinction between the cash outlay at inception and the ultimate cost of the strategy. An end user may buy a call, selling a put with a lower strike price (see Figure 18.1), following the logic analogous to the motivation of a producer: giving up part of the upside (in this case, the upside comes from lower prices), to reduce the cost of hedging. Some collar structures are more complicated. For example, an end user (such as an airline) may buy a call and sell a put (with WTI as the underlying), the strike of a call ( $K_c$ ) being greater than the strike of the put ( $K_p$ ). In addition, the end user may sell a call with a strike  $K_{c1} > K_c$ . In other words, the hedger buys a call (strike  $K_c$ ) and sells two options (a put and a call with strike  $K_{c1}$ ). The logic behind this strategy, called a three-way collar, is that the end user wants to generate additional positive cashflow by selling the call with a strike  $K_{c1}$ , hoping that the prices of crude above the level of the second call strike represent a low probability scenario. They are willing to take the risk of an extreme spike in crude prices in return for a reduction in the cost of the implementation of their hedging strategy.

The costless collars may be a very expensive way of buying protection. To use an example of a producer, let us assume that they buy a

**Figure 18.1** An option collar

collar for a period of three years, by buying a strip of puts (put options with varying maturity dates) and selling a strip of calls, with corresponding maturity dates, and the strikes calibrated to reduce to zero the initial cash cost. The risk the hedger is taking is a sudden upward move in the forward price curve of the underlying, reducing the value of the portfolio of puts (a long position) and increasing the value of a short position in calls. Depending on the negotiated credit arrangements with a hedge counterparty, the end user may have to post additional collateral or may draw on an automatic credit line negotiated with the provider of hedges.<sup>4</sup> In the latter case, the provider of hedging instruments waives the right to collateral and grants a "loan" secured with the counterparty's assets.<sup>5</sup> The hedge providers see it as "good way" risk, as the producer's assets securing the loan increase in value (the price moved significantly up in this example). The hedgers are happy because they keep the cash, which they can use for general corporate objectives.<sup>6</sup> The picture is, however, more complicated. One can argue that by committing the assets as security behind the "loans", the hedging company loses the flexibility to use the assets to secure financing on better terms. An alternative is to borrow directly and use the funds to post collateral. The advantage of relying on the arrangement with a hedge provider described above is predictability of outcomes and availability of financing at short notice, without unpleasant surprises. This has to be weighed against reduced flexibility in managing the balance sheet

and the potential danger of becoming a captive customer of a financial institution providing hedging solutions.

The example of potential dangers of costless collars is provided by the liquidity crisis experienced by Ashanti Goldfields – described in a paper by Christopher Gilbert,<sup>7</sup> which should be a mandatory reading for any risk manager in the commodity space. This event is related to the gold market, but the lessons one can learn from this saga are universal. From this example, we can see that it is critical to assess not only the mark-to-market implications of hedges but also the cashflow/collateral consequences. A recurring story in the energy industry (and in the commodity business in general) is a company putting on hedges only to be crushed by margin calls due to a shift of the entire forward price curve. The hedges are liquidated at a loss, then the price shift is reversed and a company is facing a double punishment. It has recognised trading losses and has no hedges in place.

The collar structures are sometimes associated with an option to extend. Such options are granted by the end users to hedge providers, sometimes for periods as long as 10 years. This strategy lowers the overall cost of establishing a collar hedge but may put the end user in a very unfavourable position for a very long time. The options to extend are sometimes used as a way to roll forward the losses related to established hedges. A typical sequence of events unfolds as follows: the end user incurs losses on a costless collar due to a significant market move and decides to avoid recognition of a loss by extending the maturity of the transaction, and also selling additional options to the hedge counterparty. This strategy is more likely to be used by end users domiciled in countries with lax accounting rules which do not require reporting the mark-to-market value of the entire transaction. This may offer an incentive to hedge providers to drag a client into a long-term, high-profit-margin situation on less than favourable terms, using complex options most end users cannot value or risk manage correctly. In the long run, such transactions tend to be rather myopic and of questionable wisdom for both sides. The end user may eventually default or seek legal redress. The governments may even become involved on occasion. One general conclusion is that an end user of derivatives should never use a hedge instrument that cannot be valued internally or through a reliable outside consultant. When confronting a hedge strategy that cannot be independently assessed, the correct reaction is to reach for garlic and crucifixes.

Other option structures related to collars and three-way collars<sup>8</sup> are four-way collars. A four-way collar is a straightforward extension of a three-way collar. For example, an airline may enter into a collar transaction involving two calls (long and short) and two puts (short and long).

A good example of the challenges of hedging end-user exposures related to refined products is the case of jet fuel. The challenges of hedging jet fuel are related to:

- ❑ the proxy nature of the hedges; and
- ❑ the difficulty of hedging the exposure in the context of the business model of a company.

The fundamental problem with hedging corporate exposures is that one does not seek protection by piecemeal hedging of different risks in isolation, but the objective is rather to protect the bottom line – and sometimes hedging may be a cure worse than the disease when this is forgotten. Fortunately, there is an extensive literature about airline hedging strategies.

One example of losses related to the use of the collar structures can be found in the experience of one US airline. As reported in the press:<sup>9</sup>

UAL/United Airlines said Wednesday its third-quarter earnings will include US\$544 million in losses from fuel hedging contracts. That's the two-edged sword of fuel hedges and fuel expenses. United will benefit from falling fuel prices going forward. But as its hedging contracts are "marked to market," it has to record the changes in the market value of those contracts as of Sept. 30. Broadly speaking for most airlines' hedging activities, the existing hedges go up in value when fuel prices rise. When fuel prices go down, the hedges go down in value. In a filing with the Securities and Exchange Commission, UAL estimated that it has US\$472 million in unrealised mark-to-market losses, and US\$72 million in realised losses. It also netted out to US\$8 million in gains on settled contracts.

Information provided in the SEC filing, referring to the happier days of the second quarter of 2008, throws additional light on this strategy (from UAL's Q2 2008 10-Q):<sup>10</sup>

*Aircraft Fuel Hedges.* [...] In order for the Company to obtain more favorable terms for a portion of its hedge positions, the Company entered into collars with additional features. These hedge positions include extendable collars referred to above and collars that include

twice the amount of put volume as call volume. [...] In addition, for the six months ended June 30, 2008 and 2007, the Company recognised net hedge gains of US\$279 million and US\$39 million, respectively [...] As of June 30, 2008, fuel expense and nonoperating income included US\$237 million and US\$21 million, respectively, of unrealised mark-to-market gains for contracts settling after June 30, 2008.

The 10-Q document points out to the potentially risky strategy of using extendable collars and using a larger number of put options than call options (or vice versa). Large price movements can make calls (providing protection) worthless and trigger significant losses on the puts. If the losses have to be collateralised with cash or marketable securities, a liquidity crisis may suddenly materialise out of nowhere.

### **3:2:1 crack spread**

The 3:2:1 crack spread is a trading position that contains three barrels of crude oil, two barrels of gasoline and one barrel of heating oil. A long position in crude implies a short position in the products, and vice versa. The most convenient platform for putting on this position for US-based market participants is Nymex,<sup>11</sup> and in this case the WTI, #2 Heating Oil and RBOB Gasoline contracts described previously would have been used. The logic behind using this position is that many US-based refineries are configured to produce double the volume of gasoline compared to the volume of distillates, such as heating oil and diesel (that happen to be very close in terms of their physical and chemical characteristics). A refiner who wishes to protect their processing margin would establish a position of three contracts of WTI (long), two barrels of gasoline and one barrel of heating oil (short). Using the prices as of September 3, 2010, the gross margin implied by this position is given by the calculation shown in Table 18.1, and is equal to US\$17.50 per barrel.

At the expiration of the contracts (or close to the contract expiration), the refiner will close the futures positions and transact in the spot markets or short-term forward markets closer matching the conditions of the local market where the refiner operates. The reason for this strategy is that at the point in time the initial futures hedge was created, the liquidity or the bid–offer spreads in the location-specific forward markets may have been too low and unattractive.

The 3:2:1 hedge can be characterised at best as a dirty hedge, with

a number of potential risks that have to be identified *ex ante* and closely monitored over the life of the transaction. The potential risks include the following.

- ❑ *Basis risk.* This risk is related to the potential discrepancy between the hedging instruments and the true economic exposures of a refinery. A refinery may be located on the Gulf Coast or in the Midwest and transacting in the local spot or forward markets. The basis risks typically exist across the geographical and quality characteristics of the hedged products. The prices of gasoline sold at a local spot market may not follow exactly the prices of gasoline traded on Nymex due to regional economic trends and the regional quality specifications. A refinery may rely on the use of crude of different quality than WTI, and the disconnect between the two prices may be amplified by the decoupling of WTI Cushing prices from the prices of other crudes. A refinery has two choices to address the basis risk, in addition to the obvious solution of doing nothing. One is to rely to a greater extent on the trading skills and experience of the managers implementing the hedging programme, and their ability to adjust the positions and improving timing of the transactions. This requires, in turn, an investment in the risk management system, monitoring of traders to avoid excessive risk taking and the careful balancing of conflicting objectives. An alternative strategy is based on the construction of a more complex hedge, as illustrated below. A more structured hedge can reduce the basis risk but will require trading across the futures and OTC markets, potentially at higher transaction costs, and with additional investment in trading and risk management infrastructure.
- ❑ *Volumetric risk.* One potential risk in any hedging strategy is an outage of a physical asset that reduces or completely destroys the underlying exposures. A refinery outage may require closing the hedges at a loss, without the ability to realise an offsetting gain through refinery operations. The losses may be compounded by the market impact of the outage. Local prices of refined products increase (as the supply drops) and the price of crude may go down (as the local demand is curtailed due to the outage). Given that the refiner's hedge is being short products and long crude, this is not the best outcome.

❑ *Refinery configuration risk.* The 3:2:1 hedge effectively defines a virtual, simplified refinery, but it is unlikely to match the true underlying exposures of a specific asset. Each refinery is unique and the basket of refined products that is produced is determined by the availability of local crudes and the ability to sell refined products in different geographical markets, and the availability of different processing units beyond simple distillation units. A manager with a non-standard refinery configuration can use a wider basket of hedge instruments or modify the proportions in the 3:2:1 basket. One frequently used structure is 5:3:2 hedge (see Table 18.2), or a 2:1:1 basket. This latter hedge would be preferable to a refinery with a lower relative gasoline yield. The choice of the crack-spread proportions is also dictated by the price levels and often reflects seasonality of prices.

The traditional 3:2:1 hedge worked relatively well for the refineries based on the US Gulf Coast for many years, after the introduction of Nymex contracts underlying the crack spread. These refineries relied to a large extent on light sour crude inputs (less expensive than WTI) and sold refined products loco or FOB Gulf Coast locations (at the prices lower than the New York Harbor, referenced in the heating oil and gasoline contracts). The 3:2:1 hedge would overstate the product

**Table 18.1** 3:2:1 Hedge

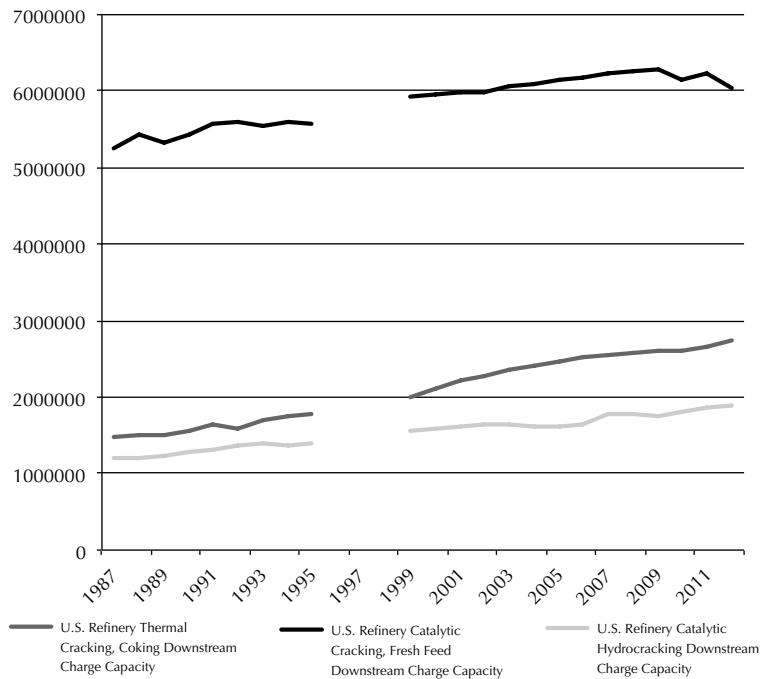
Product	Price	Unit	Barrels	Market value
WTI	77.49	US\$/bbl	3	(US\$232.47)
Gasoline	1.922	US\$/gallon	-2	US\$161.45
Heating oil	2.1076	US\$/gallon	-1	US\$88.52
				US\$17.50

**Table 18.2** 5:3:2 Hedge

Product	Price	Unit	Barrels	Market value
WTI	77.49	US\$/bbl	5	(US\$387.45)
Gasoline	1.922	US\$/gallon	-3	US\$242.17
Heating oil	2.1076	US\$/gallon	-2	US\$177.04
				US\$31.76

*Note:* Gasoline and heating oil are priced in USD/gallon  
Prices of gasoline and heating oil have to be multiplied by 42 (a barrel contains 42 gallons)

**Figure 18.2** US refinery downstream charge capacity (1987–2012, barrels per stream day, as of January 1)



Source: U.S. Energy Information Administration

revenue, but would overstate the cost of crude as well, helping to match the refining margin more closely. Overhaul of the US refineries over the last few decades is reflected in the growing share of the coking refineries. Additional investments in coking units was necessitated by the switch to lower quality crudes, such as the Maya and Venezuelan grades, as the sources of better quality crudes began to dry up. The changes in the configuration of the US and Gulf Coast refineries are reflected in Figures 18.2 and 18.3, which illustrate the evolution of refinery downstream charge capacity for the US and PADD III.

Refinery downstream charge capacity is defined as:<sup>12</sup>

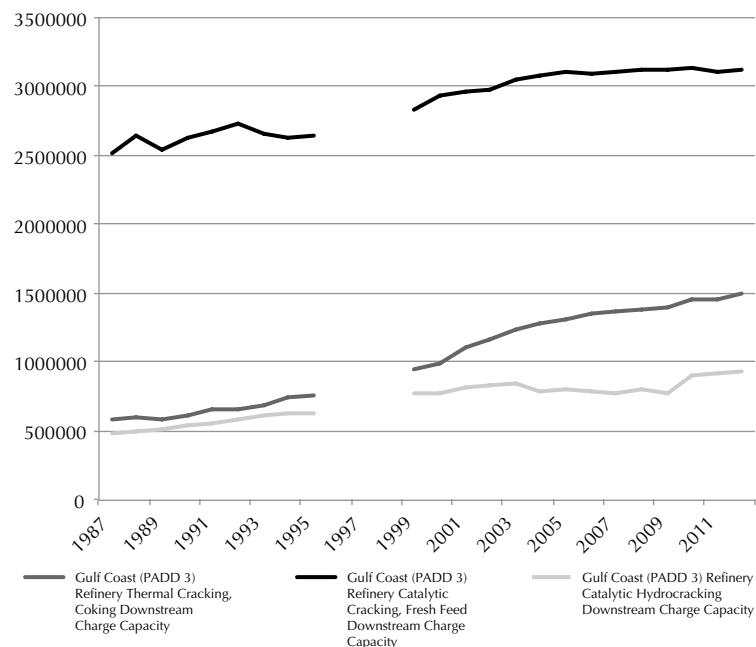
The amount of input that a distillation facility can process under usual operating conditions. The amount is expressed in terms of capacity during a 24-hour period and reduces the maximum processing capability of all units at the facility under continuous

operation [...] to account for the following limitations that may delay, interrupt, or slow down production:

- the capability of downstream facilities to absorb the output of crude oil processing facilities of a given refinery. No reduction is made when a planned distribution of intermediate streams through other than downstream facilities is part of a refinery's normal operation;
- the types and grades of inputs to be processed;
- the environmental constraints associated with refinery operations;
- the reduction of capacity for scheduled downtime due to such conditions as routine inspection, maintenance, repairs, and turn-around; and
- the reduction of capacity for unscheduled downtime due to such conditions as mechanical problems, repairs, and slowdowns.

Changes to refinery configuration undermine the hedging efficiency of the traditional 3:2:1 crack spread. An example of a basket used for hedging a US Gulf Coast coking refinery is borrowed from a presen-

**Figure 18.3** Refinery downstream charge capacity (barrels per calendar day, PADD III)



Source: U.S. Energy Information Administration

tation by Baker & O'Brien, Inc.<sup>13</sup> The basket they found preferable to the traditional 3:2:1 hedges is composed of:

- 0.5 × Regular Unleaded Conventional Gasoline, plus
- 0.5 × 0.2% Sulfur No. 2 Heating Oil, minus
- 0.33 × WTS–Midland, minus
- 0.33 × Dated Brent, minus
- 0.33 × 3.0% Sulfur No. 6 Fuel Oil

The reader is encouraged to take a closer look at the presentation from which this example has been taken.

### **Participating swaps**

A swap locks in the buyer (the payer of fixed price) into a predetermined price. This may be a sub-optimal strategy from the point of view of the end user, should the prices drop. A hedging company would be paying above-market prices (or selling below market prices), being vulnerable to the competitors who did not hedge or hedged at more favourable levels. One solution is to enter into a swap that gives the buyer a credit in case a floating price drops below a fixed price. The source of the credit is the put option embedded in the swap. The buyer acquires this option by paying a fixed price that is somewhat higher than the price of the plain vanilla swap (with no embedded option). This option is paid in installments, every time the periodic fixed-price payment is made.

The details of such transaction can be illustrated with the following example:<sup>14</sup>

Description	Financial Crude Oil 50% Participation Swap Contract
Contract Maturity	One year
Fixed Price Payer	The End User
Floating Price Payer	Sempra Energy Trading Corp.
Settlement Type	Financial
Settlement Dates	Five business days after the last day of the Pricing Period, for a total of 12 Settlement Dates
Pricing Periods	Each full calendar month, corresponding to the Settlement Date at the end of the month

Reference Quantity	100,000 barrels per month
Reference Price	The daily Official Settlement price of the prompt NYMEX WTI Futures Contract in \$/bbl
Floating Price	The arithmetic average of the Reference Price during the Pricing Period rounded to the nearest \$0.01/bbl
Floating Payment	Floating Price * Reference Quantity
Fixed Price	\$23.00 per barrel
Fixed Payment	(Fixed Price * Reference Quantity) – Participation Credit
Participation Credit	No credit is due unless the floating price is lower than the fixed price. If due, the credit is calculated as: (Fixed Price – Floating Price) * References Quantity * 0.50
Payment Date	Five business days following the Settlement Date
Documentation	Sempra Energy Trading Corp. Standard Swap Agreement
Credit Arrangements	To be determined

Source: Sempra Energy Trading

This example requires a few comments. First, this is an example of an Asian swap, with the floating price determined over the course of the month as an average of the daily settlement prices of the Nymex WTI futures contract. The complication is that the WTI contract expires during the course of the month. This somewhat complicates both swap settlements and valuation. The back office has to pay attention to the cashflow settlement dates of the swap. The valuation of the swap requires using a composite forward price for a given calendar month, calculated as the weighted average of two forward prices, with the weights being the number of days. For example, during the month of November, the futures contract for December delivery expires after about 20 days, and the second available contract (for January delivery) becomes prompt. In valuation of the swap, the forward price (the best available guess of the settlement levels of the floating price) is calculated as weighted average of the December and January futures prices (at the time of valuation) weighted by the number of trading days when each contract is prompt.

In case the floating price calculated monthly following the rules described above is below the fixed price of the swap, the hedging company receives a credit equal to 50% of the difference. Receiving

50% of this difference on the reference volume is equivalent to receiving the full difference on 50% of the volume. The customer effectively bought a put option on 50% of the reference (contractual) volume, with the strike equal to the fixed price. The premium for this option is not paid at inception of the contract, but in installments as the swap settles. The premium is distributed over time and added to the fixed price of the swap, which is as a result bigger than the fixed price of the plain vanilla swap.

The calculation of the adjusted swap price is carried out iteratively. At the starting point, the fixed price of the swap, adjusted for the option premium, which is also the strike price of the option, is unknown. One has to start with a guess of the fixed price, and then iterate adjusting this estimate until the fixed price of the swap covers exactly the cost of the option and the fixed price of a corresponding swap with no embedded option.

## FREIGHT RISK MANAGEMENT

The owners of tankers, as well as shippers, can use a vibrant and rapidly growing market for shipping derivatives to hedge market risk, with the added benefit of controlling credit risk. Historically, price and credit risk were controlled through a careful selection of counterparties and through the choice of contractual arrangements of longer duration (time charters and CoAs). An alternative tool for freight risk management was offered in 1985 when the Baltic Exchange<sup>15</sup> introduced the Baltic Freight Index (BFI). The BFI was a basket of 13 voyage routes, ranging from 14,000 mt of fertiliser up to 120,000 mt of coal (no time charter routes), which underwent over time a number of modifications.<sup>16</sup> The BFI (later renamed as the Baltic Panamax Index (BPI)) became a settlement benchmark for the Baltic International Freight Futures Exchange (Biffex) cash-settled futures contract. The design of this contract had a number of flaws, including a significant basis risk. A basket of freight contracts may not reflect correctly the evolution of rates in a specific market segment.<sup>17</sup> Frequent revisions to the index represented an effort to chase a better design in the presence of a fundamental problem. Dwindling volumes led the London International Financial Futures Exchange (Liffe), which took over Biffex in 1991, to suspend trading of this contract (in 2001). However, the need for publicly available price benchmarks for contract valuation and risk management did

not go away, and the Baltic Exchange continued to develop and publish indexes for different types of ships and cargos. Tanker market reporting started in 1998.

The Baltic Exchange tanker indexes over time morphed into two indexes: one for dirty and one for clean tankers (as much as the latter is an obvious oxymoron). Dirty tankers transport crude, and possibly heavy, fuel oil; clean tankers are used for gasoline and jet fuel. The indexes are based both on daily Worldscale and non-Worldscale assessments of international tanker routes and a selection of basket and individual time charter equivalents (TCEs).<sup>18</sup>

The Baltic Clean Tanker Index (BCTI) includes the following routes:<sup>19</sup>

<b>Route</b>	<b>Route description</b>	<b>Size mt</b>	<b>Following indicative routes from the basis</b>
TC1	75,000 mt Middle East Gulf – Japan	75,000	Ras Tanura to Yokohama
TC2_37	37,000 mt Continent to USAC	37,000	Rotterdam to New York
TC3_38	38,000 mt Caribbean – USAC	38,000	Aruba to New York
TC4	30,000 mt Singapore to Japan	30,000	Singapore to Chiba
TC5	55,000 mt Middle East to Japan	55,000	Ras Tanura to Yokohama
TC6	30,000 mt Algeria/Euromed	30,000	Algeria/Euromed

The details of the routes are shown below.

---

#### **Route: TC1**

75,000 mt, CPP/UNL Naphtha Condensate, Middle East Gulf to Japan. Ras Tanura to Yokohama with lay-days/cancelling 30/35 days in advance. Maximum age 12 years.  
Weighting: 0%

---

#### **Route: TC2\_37**

37,000 mt, CPP/UNL Continent to USAC. Rotterdam to New York with lay-days/cancelling 10/14 days in advance. Maximum age 15 years.  
Weighting: 0%

---

#### **Route: TC3\_38**

38,000 mt, CPP/UNL Caribbean to USAC. Aruba to New York with lay-days/cancelling 6/10 days in advance. Maximum age 20 years. Assessment

basis – Oil Pollution Act premium paid.  
Weighting: 0%

---

**Route: TC4**


---

30,000 mt, CPP/UNL Singapore to Japan. Singapore to Chiba with lay-days/cancelling 7/14 days in advance. Maximum 15 years.  
Weighting: 0%

---

**Route: TC5**


---

55,000 mt, CPP/UNL naphtha condensate, Middle East/Japan. Ras Tanura to Yokohama with lay-days cancelling 30/35 days in advance. Maximum age 15 years.  
Weighting: 0%

---

**Route: TC6**


---

30,000 mt CPP/UNL Algeria/Euromed Skikda/Lavera with lay-days cancelling 7/14 days in advance. Max age: 15 years.  
Weighting: 0%

---

The Baltic Dirty Tanker Index (BDTI) includes the following routes:

<b>Route</b>	<b>Route description</b>	<b>Size mt</b>	<b>Following indicative routes from the basis</b>
TD1	280,000 mt ME Gulf to US Gulf	280,000	Ras Tanura to LOOP
TD2	260,000 mt ME Gulf to Singapore	260,000	Ras Tanura to Singapore
TD3	250,000 mt ME Gulf to Japan	250,000	Ras Tanura to Chiba
TD4	260,000 mt W Africa to US Gulf	260,000	O.S Bonny to Loop
TD5	130,000 mt W Africa to USAC	130,000	O.S Bonny to Philadelphia
TD6	135,000 mt Black Sea / Med	135,000	Novorossiysk to Augusta
TD7	80,000 mt North Sea to Cont	80,000	Sullom Voe to Wilhelmshaven
TD8	80,000 mt Kuwait – Singapore (Crude/DPP Heat 135F)	80,000	Mena al Ahmadi to Singapore
TD9	70,000 mt Caribs to US Gulf	70,000	Puerto la Cruz to Corpus Christi

TD10D	50,000 mt Caribs to USCA – Double hull vessel	50,000	Aruba to New York
TD11	80,000 mt Cross Med	80,000	Banias to Lavera
TD12	55,000 mt ARA to US Gulf	55,000	Antwerp to Houston
TD14	Trial – 80,000 mt SE Asia to EC Australia	80,000	Seria to Sydney
TD15	Trial – 260,000 mt West Africa to China	260,000	Zaffiro and Bonny to Ningbo
TD16	Trial – 30,000 mt Black Sea to Mediterranean	30,000	Odessa to Augusta

The details of the routes are shown below:

---

#### **Route: TD1**

280,000 mt, Middle East Gulf to US Gulf. Ras Tanura to LOOP with lay-days cancelling 20/30 in advance. Maximum age 20 years.  
Weighting: 0%

---

#### **Route: TD2**

260,000 mt, Middle East Gulf to Singapore. Ras Tanura to Singapore with lay-days/cancelling 20/30 days in advance. Maximum age 20 years.  
Weighting: 0%

---

#### **Route: TD3**

250,000 mt, Middle East Gulf to Japan. Ras Tanura to Chiba with lay-days/cancelling 30/40 days in advance. Maximum age 15 years.  
Weighting: 0%

---

#### **Route: TD4**

260,000 mt, West Africa to US Gulf. Off Shore Bonny to LOOP with lay-days/cancelling 15/25 days in advance. Maximum age 20 years.  
Weighting: 0%

---

#### **Route: TD5**

130,000 mt, West Africa to USAC. Off Shore Bonny to Philadelphia with lay-days/cancelling 15/25 days in advance. Maximum age 20 years.  
Weighting: 0%

---

#### **Route: TD6**

135,000 mt, Black Sea/Mediterranean. Novorossiysk to Augusta with lay-days/cancelling 10/15 days in advance. Maximum age 20 years.  
Weighting: 0%

**Route: TD7**

---

80,000 mt, North Sea to Continent. Sullom Voe to Wilhelmshaven, with lay-days/cancelling 7/14 days in advance. Maximum age 20 years.  
Weighting: 0%

---

**Route: TD8**

---

80,000 mt, Crude and/or DPP Heat 135F, Kuwait to Singapore. Mena al Ahmadi/Singapore with lay-days/cancelling 20/25 days in advance.  
Maximum age 20 years.  
Weighting: 0%

---

**Route: TD9**

---

70,000 mt, Caribbean to US Gulf. Puerto La Cruz to Corpus Christi with lay-days/cancelling 7/14 days in advance. Maximum age 20 years.  
Assessment basis – Oil Pollution Act premium paid.  
Weighting: 0%

---

**Route: TD10D**

---

50,000 mt, fuel oil, Caribbean to USAC. Aruba to New York with lay-days/cancelling 7/14 days in advance. Double hull vessel. Maximum age 20 years.  
Weighting: 0%

---

**Route: TD11**

---

80,000 mt, cross Mediterranean/Banias to Lavera with lay-days/cancelling 10/15 days in advance. Maximum age 20 years.  
Weighting: 0%

---

**Route: TD12**

---

55,000 mt, fuel oil, Amsterdam–Rotterdam–Antwerp range to US Gulf. Antwerp to Houston with lay-days cancelling 15/20 days in advance.  
Double-hulled vessels.  
Weighting: 0%

---

**Route: TD14**

---

80,000 mt, no heat crude, SE Asia to EC Australia, Seria to Sydney with lay-days/cancelling 21/25 days in advance. Double hull and maximum age 15 years old.  
Weighting: 0%

---

**Route: TD15**

---

260,000 mt, no heat crude, West Africa to China. Zafiro and Bonny to Ningbo with lay-days cancelling 20/30 days in advance. Double hull and maximum age 20 years.  
Weighting: 0%

---

**Route: TD16**

---

30,000 mt fuel oil heat 135 F, Black Sea to Mediterranean. Odessa to Augusta with lay-days/cancelling 8/14 days in advance. Double hull and maximum age 20 years.  
Weighting: 0%

---

Another index published by the Baltic Exchange covers certain Asian routes. The BITR-Asia index is published in Singapore at 16.00 local time and includes the following routes:

Route 4 30,000 mt, CPP/UNL<sup>20</sup> Singapore–Japan  
Route 7 30,000 mt, Singapore–East Coast Australia,  
Route 10 40,000 mt, South Korea–NOPAC<sup>21</sup> West Coast; and  
Route 11 40,000 mt, CPP South Korea–Singapore.

The BLPG index covers 44,000 mt, 5%,<sup>22</sup> 1–2 grades fully refrigerated Liquid Petroleum Gas, Ras Tanura to Chiba, laydays 10/40 days in advance. Laytime 96 hours total. Maximum tanker age is 20 years.

Spot freight indexes are compiled by a panel of independent shipbrokers<sup>23</sup> (ie, entities which do not take positions but facilitate transactions), using information on prevailing rates, on-going negotiations and recently concluded deals. As in the case of all PRAs, judgement and experience play a significant role in arriving at the price assessments. The submissions by the panellists are subject to audit by the Baltic Exchange.

The futures market was gradually replaced by forward freight agreements (FFAs), introduced in 1992 by Clarkson Wolff and traded for the first time in October of that year. FFAs are forward contracts that are cash-settled against an assessment of freight rates that are route specific. The assessments are provided daily by panellists associated with the Baltic Exchange (and a few assessments are provided by Platts).<sup>24</sup> These contracts were initially traded OTC as bilateral arrangements facilitated by specialised brokers. Starting in 2002, a new trend established itself, with a hybrid approach that combined clearing with screen trading, which takes place with the help of brokers. This trend was initiated by NOS (The Norwegian Futures and Options Clearing House, founded in 1987) and Oslo-based Imarex (International Maritime Exchange founded in 1999). The partnership offered a service consisting of electronic screens, facilitating the trading of shipping derivatives in small units called lots

(one day in a time charter or 1,000 tons of cargo) and clearing services for the executed trades. This arrangement gave market participants the ability to customise hedges to reflect their specific exposures.

The offering of Imarex-listed shipping derivatives includes:<sup>25</sup>

	Underlying product	FFA	Options
Dirty	Aframax: TD7, TD8, TD9, TD11, TD17	X	X
	Suezmax: TD5	X	X
	VLCC: TD3, TDS TCE, TD3USD	X	X
	MR: TD16	X	X
Clean	MR: TC2, TC2USD, TC4, TC11	X	X
	MR: TC6	X	X
	LRI: TC5	X	X
Oil products	Underlying product	FFA	
Fuel oil	Rotterdam 3.5% FOB	X	
	North West Europe 1% FOB	X	
	Singapore 180 CST FOB	X	
	Singapore 380 CST FOB	X	
	US Gulf no. 63% sulfur FOB	X	

Listed tanker voyage FFA products (Worldscale):

Underlying	Index	Index provider	Closing price provider
TD7, Aframax, North Sea–Continent, 80,000 mt	Baltic Exchange	Imarex	
TD 9, Aframax, Carib–USG, 70,000 mt	Baltic Exchange	Imarex	
TD 5, Suezmax, West Africa–USAC, 130,000 mt	Baltic Exchange	Imarex	
TD 3, VLCC, AG–East, 260,000 mt	Baltic Exchange	Imarex	
TD8, Aframax, Kuwait–Singapore, 80,000 mt	Baltic Exchange	Imarex	
TD17, Aframax, Baltic Sea–Continent, 100,000 mt	Baltic Exchange	Imarex	
TD11, Aframax, Cross–Med, 80,000mt	Baltic Exchange	Imarex	
TD16, MR, Black Sea–Mediterranean, 30,000mt	Baltic Exchange	Imarex	

TC4, MR, Singapore–Japan, 30,000 mt	Platts	Imarex
TC2, MR, Continent–USAC, 37,000 mt	Baltic Exchange	Imarex
TC5, LR 1, AG–Japan, 55,000 mt	Platts	Imarex
TC6, MR, Algeria–Euromed, 30,000 mt	Baltic Exchange	Imarex
TC11, South Korea–Singapore, 40,000 mt	Baltic Exchange	Imarex
TD18, MR, Baltic Sea–Continent, 30,000 mt	Baltic Exchange	Imarex
<b>Flat rates</b>	As published by the Worldscale Association (London) and the Worldscale Association (NY).	
<b>Price quotation</b>	Worldscale points	
<b>Minimum price fluctuation</b>	0.25 Worldscale point	
<b>Contract value</b>	#Lots x Lot size x Worldscale flatrate x (Worldscale points/100)  (The Worldscale flatrate applicable for each index day in the delivery period)	
<b>Delivery period</b>	<b>Month:</b> First index day of the month to last index day of the month.  <b>Quarter:</b> First index day of the quarter to last index day of the quarter. A quarter contract will be split equally into three-month contracts on the trading day and settled as month contracts.  <b>Year:</b> First index day of the year to last index day of the year. A year contract is split into equally into 12-month contracts on the trading day and settled as month contracts.	
<b>Final settlement day</b>	Last settlement day in the delivery period.	
<b>Settlement price</b>	The arithmetic average of the spot prices for the relevant underlying product over the number of index days in the delivery period.	
<b>Lot size</b>	1 lot = 1,000 mt	
<b>Minimum lots per contract</b>	0.01 lot in all products	

Listed tanker voyage FFA products (US\$/mt):

<b>Underlying</b>	<b>Index</b>	<b>Index provider</b>	<b>Closing price provider</b>
	TC2US\$, MR Continent-USAC, 37,000 mt	Calc by NOS	Imarex
	TD3US\$, VLCC AG-East, 260,000 mt	Calc by NOS	Imarex
<b>Flat Rates</b>	As published by the Worldscale Association (London) Limited and the Worldscale Association (NY)		
<b>Price quotation</b>	US\$/mt		
<b>Minimum price fluctuation</b>	0.0001 US\$		
<b>Contract value</b>	#Lots x Lot size x Price		
<b>Delivery period</b>	<b>Month:</b> First index day of the month to last index day of the month. <b>Quarter:</b> First index day of the quarter to last index day of the quarter. A quarter contract will be split equally into three-month contracts on the trading day and settled as month contracts. <b>Year:</b> First index day of the year to last index day of the year. A year contract is split into equally into 12-month contracts on the trading day and settled as month contracts.		
<b>Final settlement day</b>	Last settlement day in the delivery period.		
<b>Settlement price</b>	The arithmetic average of the spot prices for the relevant underlying product over the number of index days in the delivery period.		
<b>Lot size</b>	1 lot = 1,000 mt		
<b>Minimum lots per contract</b>	0.01 lot in all products		

Listed tanker time charter FFA products:

<b>Underlying</b>	<b>Index</b>	<b>Index provider</b>	<b>Closing price provider</b>
	TD3_TCE, VLCC AG-East, 260,000mt	Baltic Exchange	Imarex
<b>Flat Rates</b>	As published by the Worldscale Association (London) and the Worldscale Association (NY).		

<b>Price quotation</b>	US\$/day
<b>Minimum price fluctuation</b>	250 US\$/day
<b>Contract value</b>	#Lots x Lot size x Price
<b>Delivery period</b>	<b>Month:</b> First index day of the month to last index day of the month. <b>Quarter:</b> First index day of the quarter to last index day of the quarter. A quarter contract will be split equally into three-month contracts on the trading day and settled as month contracts. <b>Year:</b> First index day of the year to last index day of the year. A year contract is split into equally into 12-month contracts on the trading day and settled as month contracts.
<b>Final settlement day</b>	Last settlement day in the delivery period.
<b>Lot size</b>	1 lot = 1 day
<b>Minimum lots per contract</b>	0.01 lot in all products

---

The success of Imarex attracted a number of participants to this market, including London Clearing House (LCH), Nymex and the Singapore Global Exchange. LCH opened a platform for shipping derivatives business in September 2006. After initial problems, LCH catapulted to the leading position in this market.

As shown in Figure 18.4, the trend is towards the clearing of freight derivatives (as opposed to bilateral contracts). Figure 18.5 illustrates segmentation of the FFA market by end users. The numbers imply that the FFAs are used not only as a tool for hedging existing exposures (the ship owners and majors), but also for speculation.

## CONCLUSIONS

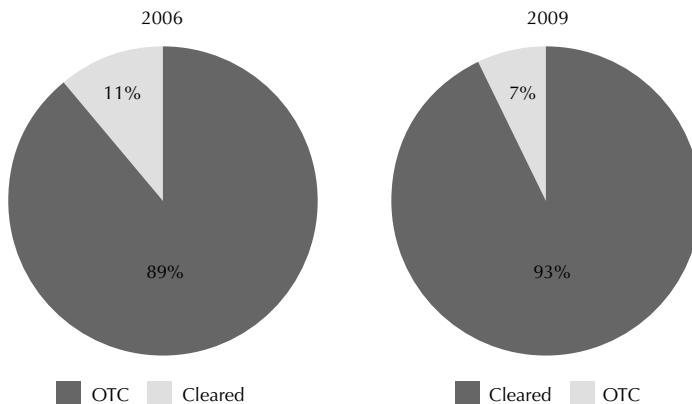
Transactions in the oil markets, as for commodity transactions in general, are undertaken for a variety of reasons, including:

- ❑ participation in the physical supply chain;
- ❑ hedging of future cashflows related to the sales and purchases of energy commodities;
- ❑ protection of processing and transportation spreads to optimise asset operations;

- protection of inventory values; and
- speculation.

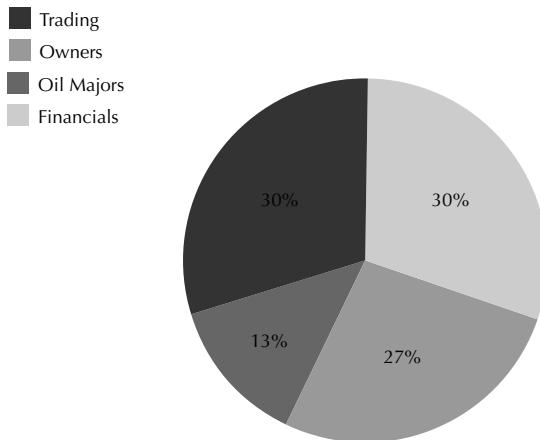
These objectives sometimes overlap and cannot always be neatly separated. What they have in common, however, is that they cannot

**Figure 18.4** Cleared versus OTC freight derivatives



Source: Stefan Albertijn, 2010, "Function and role in the freight derivatives market," Freight Market Information User Group, November 2

**Figure 18.5** Tanker FFA market – segmentation by participants



Source: Georgi Slavov, 2010, "FFA market dynamics and trading strategies," presentation to FFBA Shanghai Forum, September.

be carried out in a vacuum without a good understanding of the physical system behind them. A set of transactions related to the developments covered in this section of the book is a good example. A few years ago, when the prices of WTI and Brent crudes changed their historical relationship (Brent used to trade a dollar to two dollars under WTI), many traders put massive positions betting on the reversal of the market to historical norms. These bets were based on pure analysis of the historical price charts and, in many cases, were made by traders who could not find Cushing on the map, could not enumerate the pipelines crossing Cushing and did not understand the reasons why the cross-Atlantic Brent/WTI arbitrage was suddenly suspended. As Marx<sup>26</sup> used to say: "Learn from the mistakes of others. You can never live long enough to make them all yourself."

- 1 This may reflect the conviction that the price relationships in the forward markets are more stable than in the spot markets, and the quality of hedges may deteriorate closer to expiration.
- 2 The term *Asian option*, according to industry folklore, was coined by a team from Chase Bank working on a deal in Tokyo in the late 1980s. Average price structures were used historically in the oil markets, but at the time there were no accepted procedures for pricing such options. The Chase analysts came up with a model, so they called the average price options they used "Asian" for obvious reason.
- 3 For example, if a hedging company is downgraded, the negotiated credit agreements may require the posting of additional collateral.
- 4 As mentioned before, we shall cover this topic in detail in the section on risk management.
- 5 Chesapeake Energy Corporation (CHK) reported that a liability related to hedging activities of US\$6 billion was once secured with corporate assets worth US\$11 billion.
- 6 See Chapter 6 for additional information.
- 7 Christopher Gilbert, 2001, "Has the Ashanti Goldfields Loss Discredited Collar Hedges?" Vrije Universiteit, working paper.
- 8 Three-way collars (a call spread financed with the sale of a put) are called seagulls.
- 9 Terry Maxon, 2008, "United Airlines to record big losses from fuel hedging," *Dallas News*, September 17.
- 10 United Continental Holdings, 10-Q, Filed on July 24, 2008, p 31 (available from [www.unit-edcontinentalholdings.com](http://www.unit-edcontinentalholdings.com), Investor Resources section).
- 11 This strategy can be customised by relying on the OTC markets, using instruments which are better suited for hedging refineries located in a specific location (such as the US Gulf Coast or Midwest). The advantage of using Nymex is the ability to establish the position in one step. Nymex will treat a crack-spread transaction as a single order, with one margin applied against the entire basket.
- 12 See [http://www.eia.doe.gov/dnav/pet/TblDefs/pet\\_pnp\\_cap1\\_tbldef2.asp](http://www.eia.doe.gov/dnav/pet/TblDefs/pet_pnp_cap1_tbldef2.asp).
- 13 John B. O'Brien, Scott Jensen "A New Proxy for Coking Margins – Forget the Crack Spread," National Petrochemical & Refiners Association, Annual Meeting San Francisco, March 13–15, 2005, <http://www.bakerobrien.com/documents/Coking%20Spread%20from%20NPRA.pdf>
- 14 [http://www.sempracommodities.com/oil\\_endusers.asp](http://www.sempracommodities.com/oil_endusers.asp).

- 15 The Baltic Exchange started, as Lloyds of London, in the 18th century as an informal gathering of seamen and merchants meeting at the Virginia and Baltick coffee house in Threadneedle Street, London.
- 16 The modifications are documented in Baltic Exchange document “A History of the Baltic Indices” (<http://www.balticexchange.com/media/pdf/a%20history%20of%20baltic%20indices%20201011.pdf>).
- 17 The volume never met initial expectations. About 5,000 contracts were expected to be traded daily; the actual volume never exceeded 1,500 (see Jeremy Penn, “The role of the Baltic Exchange in the freight derivatives market,” in Manolis G. Kavussanos and Ilias D. Visvikis, 2011, *Theory and Practice of Shipping Freight Derivatives* (London: Risk Books)).
- 18 The TCE is calculated using a variable feed of bunker prices supplied by Argus Media.
- 19 [http://www.clarksonsecurities.com/bcti\\_routes.aspx](http://www.clarksonsecurities.com/bcti_routes.aspx).
- 20 Clean Petroleum Product, unleaded.
- 21 North Pacific.
- 22 Maximum 5% propene ( $C_3H_6$ ) content.
- 23 Current panelists are listed at <http://www.balticexchange.com/default.asp?action=article&ID=556>.
- 24 As reported by the Baltic Exchange, “The Baltic Forward Assessments (BFA) are an estimated mid-price of bids and offers for the dry and wet market based on submissions from brokers at 1730 (London). [...] Wet routes: Clean (TC2\_37, TC4, TC5, TC6), Dirty (TD3, TD5, TD7, TD11).” TC4 and TC5 assessments are provided by Platts (for historical and consistency reasons). These two routes are related closely to the naphtha markets. Consistent price and freight assessments facilitate netback calculations. See <http://www.balticexchange.com/default.asp?action=article&ID=5133>.
- 25 [http://www.imarex.com/products/contract-specifications/#dry\\_bulk\\_tc\\_baskets](http://www.imarex.com/products/contract-specifications/#dry_bulk_tc_baskets).
- 26 Groucho Marx, 1890–1977.



## Section 5

# **Electricity, Emissions and Coal**



## *Electricity: The Basics*

Each section of this book has started with a discussion of the physical properties of different energy commodities. However, this task is especially difficult for electricity. We cannot see it, we cannot smell it, we cannot touch it – and appeals to our sensory experience break down when we try to describe its movement. We can, of course, touch a conduit transmitting electricity, but this is not a recommended way to study the market. At the same time, it is difficult to think of any energy commodity for which an understanding of the technical aspects of production and transportation is more important in order to avoid nasty market surprises and bad deals. The challenge of explaining certain physical facts related to electricity is that it can be based either on a “trust me, this is the way it is” approach or would require a long dissertation involving complex mathematics. Complex means, in this context, both complicated and making use of complex numbers – ie, numbers containing real and imaginary parts. We have tried to choose a middle road and elaborate on some technicalities. Those who find the following discussion too sensational can skip directly to the section on units of measurement. This is important to anybody who trades electricity: this is how prices are quoted. It is, however, important to be aware of certain important facts, which are explained briefly below.

There is an important distinction between real power and reactive power, which add up to what is called *apparent power*, linked together by a Pythagorean-type formula:

$$(\text{apparent power})^2 = (\text{real power})^2 + (\text{reactive power})^2$$

Real power is what we pay for, with prices quoted in currency units per kWh (for our home electricity bills) or per megawatt-hour (MWh) when we trade. We are not paying directly for reactive power

measured in volt-ampere reactive (VAR) units, because we are consuming it half of the time and giving it back half of the time (in the case of pulsating alternating current) – for a net consumption of zero. It is possible to compare electric current to water distributed over local utility pipes. Voltage, related to reactive power that is critical for voltage support, can be compared to pressure in the pipe system. We are paying directly for water, but we would not get any without pressure, which is created by building water towers. The cost of building water towers and pumping water to the tanks has to be covered somehow – and is paid for by the ultimate consumers. The cost of reactive power has to be allocated to the users of electricity, but this is by no means an easy theoretical problem.

The second important fact is that electricity follows the path of least resistance. This means that one cannot control with prevailing equipment how electricity flows will be distributed over multiple connected systems. This is a serious problem from the point of view of system operators, who have to worry about unexpected and uncontrolled electricity flows undermining system reliability. It is also an important issue for power-trading desks, because the flexibility of arranging transactions at very short notice – a benefit to those who face sudden shortages or surpluses of electricity – translates into a headache for transmission system operators, who can see electricity sloshing unexpectedly through an area representing 50% of the US as a potential cause of system-wide outages.

In practice and in the technical literature, both the terms power markets and electricity markets are used as equivalents. A person paying attention to detail (or partial to splitting a hair) would argue that electricity is a term which is more general than power. There are certain markets that fall under the umbrella of electricity markets (ie, firm transmission rights, or capacity) that are not power markets in the narrow sense of the word. However, we shall use both terms without differentiating between them. Another useful distinction is the one between energy and power. The difference between the two can be best explained using an analogy between water and water flow. Water at the top of the water tower is stored energy. The rate at which energy flows is power. The same amount of water may be released slowly over time (through straw) or quickly (through a high-diameter pipe).

This chapter can be viewed as partly optional, and is recommended

primarily for fundamental and quantitative analysts. Having said that, it does help to understand the basic differences between direct and alternating current for a number of reasons. Our current electricity business developed through the competition between the two fundamental visions of the industry: one based on direct current (Edison) and the other on alternating current (Westinghouse). The chapter will begin with a restatement of a few facts from school physics, including Ohm's Law, Kirchoff's Voltage Law and Kirchoff's Current Law. The section on alternating current and power flow equations requires familiarity with complex numbers. One can, however, take a look at the numerical examples of power flows in a simple three-node network (Figure 19.5). This example is used practically in every textbook on electricity, and illustrates one basic fact of the transmission network: electricity flows where it wants to (subject to the laws of physics) and pays no attention to the contracts traders negotiate. The reader who is not interested in such technicalities can proceed directly to the section on units of measurement, which is indispensable.

## ELECTRICITY: DEFINITIONS

This section contains a basic summary of some introductory facts related to electricity. It is not an extensive exposition of the physics underlying electricity, but rather a compendium of definitions and terms which are required in order to understand the physical aspects of the power industry.

The unit of charge is one coulomb (C), or  $6.25 \times 10^{18}$  protons (positive elementary particles forming the nucleus of an atom). By definition, one proton has a charge of  $1.6 \times 10^{-19}$  C and an electron has a charge of  $-1.6 \times 10^{-19}$  C. Electric potential is the potential energy of a charge at a given location, relative to a neutral reference level (often called ground or an electrically neutral place), divided by the amount of the charge. The unit of potential is one volt (V), which is equal to one joule (J)/coulomb. The joule is a standard unit of energy. Voltage measures a difference of charge at two different points in a circuit. The voltage, like elevation, is a relative concept and thus can be measured relative to another agreed point.

Current is defined as a flow rate of charge. Using the imperfect analogy of a water pipe,<sup>1</sup> voltage can be compared to pressure (created by erecting a water tower) and current is the rate at which

water flows through the conduit. Current is denoted in most textbooks by  $i$  or ( $I$ ) and is measured in units called amperes ( $A$ ), where  $A = C/\text{sec}$ .

Ohm's Law relates current, voltage and resistance. Current through a conduit is proportional to voltage and inversely proportional to the resistance:

$$I = V/R \quad (19.1)$$

where resistance ( $R$ ) describes the opposition of the conductor to the flow of current. The resistance, an equivalent of mechanical friction for electricity, depends on the physical properties of the material and the area of the cross-section of the conductor (in the same way as a larger diameter pipe facilitates the flow of water) and  $l$  its length:

$$R = \frac{\rho l}{A} \quad (19.2)$$

where  $\rho$  is a resistance constant which is material-specific,  $l$  is the length and  $A$  is the cross-section of the area. The unit of resistance is the ohm ( $\Omega$ ), equal to one volt per ampere. Resistivity of different materials is measured in ohm metres. The inverse of resistance is called conductivity and is measured in units called mhos.

Power ( $P$ ) is defined as energy expended per unit of time and is measured in units called watts (joules per second). Power is equal to current multiplied by voltage drop:

$$P = IV \quad (19.3)$$

Given that voltage  $V$  can be expressed as  $I \times R$  (from Ohm's law), power is equal to:

$$P = I^2 R \quad (19.4)$$

Equation 19.4 is the most important formula as it explains the design of practically every electric system of reasonable size. According to the equation, power changes with the square of current but is linear in resistance. If a higher load is connected to the system, more power has to be transmitted. At a constant voltage, this means that more current has to flow, and this in turn increases resistive losses (as resistive heating is related to the square of the current).

The first power system developed in the US, by Thomas Edison, relied on direct current (DC) – ie, uni-directional current of constant

voltage across the system (the difference between direct and alternating current will be discussed below). This design, if it had become standard, would have run into insurmountable technical problems. Safety considerations require the use of relatively low voltages in local distribution systems (power companies prefer to keep their customers alive). Transmitting electricity at the same voltage across the entire grid would result in high losses that could be reduced either by using power lines with a higher cross-section (and are therefore heavier and more expensive) or building power plants closer to the clusters of load. The second solution would not only be very expensive, it would have environmental consequences. The obvious solution would be to generate electricity at relatively low voltage, transmit it at high voltages and reduce the voltage at the interconnection point to the local distribution system. This solution was viable at the dawn of the electricity era only for alternating current, using a device called a transformer.

An alternative approach is associated with Nikola Tesla, a Serbian engineer, and George Westinghouse, an American industrialist and inventor.<sup>2</sup> Their design, based on alternating current (AC), eventually won and DC was relegated to certain special applications (such as, for example, third rail systems).

Two fundamental theorems required to understand the physical aspects of the electricity business were formulated by the German physicist Gustav Kirchoff in the 19th century. The first law, known as the Kirchoff's Voltage Law (or Kirchoff's Loop Law), states that the algebraic sum of voltages around a closed loop is equal to zero. This law is a version of the principle of conservation of energy, whereby voltage is defined as the energy per unit charge. The total amount of energy gained per unit charge must equal the amount of energy lost per unit charge. This seems to be true as the conservation of energy states that energy cannot be created or destroyed; it can only be transformed from one form to another. The law is often formulated using alternative language: the voltage drops across the resistors add up to the voltage applied by the voltage source, or:

$$\sum_{i=1}^n V_i = 0 \quad (19.5)$$

where  $V_i$  stands for voltage at point  $i$ . The voltage drops are defined as voltage reduction between the source of voltage and load.

---

Figure 19.1 contains the example of this law for a simple circuit. The sum of voltages through the circuit is equal to zero ( $V_1 + V_2 + V_3 + V_4 = 0$ ).

The second law, known as the Kirchoff Current Law (or Kirchoff's Point Law), states that the sum of currents through any node of a circuit is zero:

$$\sum_{i=1}^n I_i = 0 \quad (19.6)$$

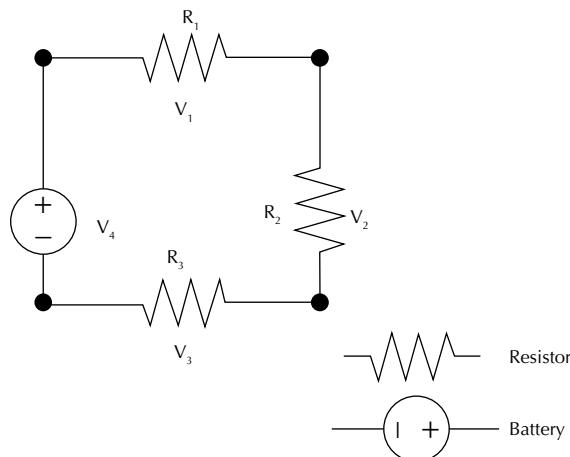
In Figure 19.2, we have currents  $I_2$  and  $I_3$  flowing into the junction, and currents  $I_1$  and  $I_4$  flowing out. According to the Kirchoff Current Law,  $I_2 + I_3 = I_1 + I_4$ . It is worth noting that the Kirchoff's laws assume constant current.

## ALTERNATING CURRENT: BASICS

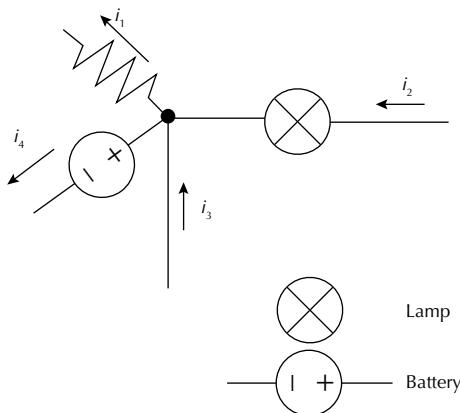
### Definitions

Practically all large-scale power systems use alternating current – ie, current with alternating and reversing polarity (direction) with frequency set typically to 60 hertz (US) or 50 hertz (Europe and most of the rest of the world). This section will introduce the terms and mathematical tools necessary to understand and model the integrated

**Figure 19.1** Kirchoff Voltage Law



Source: [http://en.wikipedia.org/wiki/Kirchhoff's\\_circuit\\_laws](http://en.wikipedia.org/wiki/Kirchhoff's_circuit_laws)

**Figure 19.2** Kirchoff Current Law

Source: <http://physics.about.com/od/electromagnetics/f/KirchhoffRule.htm>

generation and transmission systems based on alternating current. Limitations of the space do not allow us to explain all the concepts in detail, and this is better left to people with background in electrical engineering and science. A reader should accept the definitions and consult electrical engineering and physics books to develop a deeper insight into the field. The exposition requires some basic familiarity with complex numbers and the notations used in this field of mathematics. In addition, Panel 19.1 covers the bare essentials of complex numbers.

Alternating current can be described as the flow of electric charge changing direction (unlike direct current), with voltage oscillating periodically over time. The number of oscillations per unit of time is called frequency (the inverse of frequency is called a period). A natural choice for modelling current  $I(t)$  as an oscillating quantity is a sine function.<sup>3</sup>

$$I(t) = I_{\max} \sin(\omega t + \varphi_t) \quad (19.7)$$

where  $I_{\max}$  denotes the amplitude (the maximum value of current as it fluctuates),  $t$  is time (measured in seconds),  $\omega$  is angular frequency<sup>4</sup> and  $\varphi$  is phase shift (the angle by which the sine curve is shifted with respect to point zero).<sup>5</sup>

The choice of starting point is arbitrary. What is important, as will be explained below, is to measure how much voltage is leading (or

lagging) the current. In other words, voltage may peak (or bottom) before current peaks (bottoms). The use of an angle may be somewhat counterintuitive, but the system of measurement can be explained by observing that, if a complete oscillation takes 1/60 of a second (with 60 oscillations per second), this number can be represented as a fraction of the angle of 360° degrees, which is equal to  $360/60$  – ie, six degrees, or  $6 \times 2\pi / 360 = 0.10472$  radians.

Voltage, in turn, can be represented by the equation:

$$V(t) = V_{\max} \sin(\omega t + \varphi_V) \quad (19.8)$$

where  $\varphi_V$  represents voltage phase that does not have to coincide with the electric current phase.

Equations 19.7 and 19.8 provide the measures of instantaneous current and voltage. Measuring the magnitude of these quantities over longer time periods requires using an average value. This is done in practice by calculating the root mean square (RMS), which is defined for time interval  $(T_1, T_2)$  as:

$$rms = \sqrt{\frac{1}{T_2 - T_1} \int_{T_1}^{T_2} (I_{\max} \sin(\omega t))^2 dt} \quad (19.9)$$

$$rms = I_{\max} \sqrt{\frac{1}{T_2 - T_1} \int_{T_1}^{T_2} \frac{1 - \cos(2\omega t)}{2} dt} \quad (19.10)$$

$$rms = I_{\max} \sqrt{\frac{1}{T_2 - T_1} \left[ \frac{t}{2} - \frac{\sin(2\omega t)}{4\omega} \right]_{T_1}^{T_2}} \quad (19.11)^6$$

$$rms = I_{\max} \sqrt{\frac{1}{T_2 - T_1} \frac{T_2 - T_1}{2}} \quad (19.12)$$

$$rms = I_{\max} / \sqrt{2} \quad (19.13)$$

As one can see from equations 19.12 and 19.13, RMS is calculated as the square root of the average of the squared value of a function (current or voltage in our case). The reason for using RMS is that in practice we are interested more in average values, as opposed to instantaneous levels of current, voltage and power. The rationale for using the squared value is that a sine function fluctuates between 1

and  $-1$  and, therefore, taking a simple average over a number of full cycles would produce a zero value.

RMS may seem to be a very abstract concept, but it is something we encounter quite often and are very familiar with. All alternating electric power systems are rated in the RMS voltages and currents including our morning friends: the toaster and the coffee machine. A 120 V system really means that the RMS value for voltage is 120.

The equivalent of resistance in the case of alternating current is known as impedance, the measure of opposition of a circuit to the flow of current. Impedance is represented as a complex number, with reactance being its imaginary part ( $X$ ), and resistance ( $R$ ) its real part. Reactance is defined as either capacitive reactance or inductive reactance.<sup>7</sup> The combination of reactance and resistance is defined as impedance  $Z$ , given by a complex number

$$Z = R + iX \quad (19.14)$$

Inductive reactance is equal to the angular frequency multiplied by inductance ( $L$ ), which in turn is defined by the physical properties of an inductor:<sup>8</sup>

$$X_L = \omega L \quad (19.15)$$

Inductance is a link between voltage and current. In the case of direct current, voltage is given as a product of current ( $I$ ) and resistance ( $R$ ), or  $V = IR$ . In the case of alternating current, the corresponding relationship is:

$$V = L \frac{dI}{dt} \quad (19.16)$$

where  $dI/dt$  is the rate of change of current with respect to time. This equation explains why, after a sudden change in current (for example, due to a sudden interruption of transmission), a voltage spike may follow.

Capacitive reactance is given by the inverse of the product of the angular frequency and the capacitance ( $C$ ), which is in turn given by the physical properties of a device:

$$X_C = -\frac{1}{\omega C} \quad (19.17)$$

It is also possible to establish a relationship between current and voltage, using capacitance:

$$I = C \frac{dV}{dt} \quad (19.18)$$

The combined effect of inductors and capacitors is to change the phase shift between current and voltage; in general, capacitors cause the current to lead voltage, while inductors have the opposite effect.

Impedance  $Z$  (see equation 19.14) can be written in the polar form. The angle  $\varphi$  of the impedance corresponds to the phase shift between voltage and current caused by an electronic device.<sup>9</sup> Depending on whether the device is an inductor or a capacitor, the sign of  $\varphi$  will be either positive or negative.

The inverse  $1/Z$  of impedance is called admittance,  $Y$ , where:

$$Y = G + iB \quad (19.19)$$

where  $G$  (conductance) and  $B$  (susceptance) can be related to impedance, following the rules of division of the complex numbers.<sup>10</sup>

$$G = \frac{R}{Z^2}; B = -\frac{X}{Z^2} \quad (19.20)$$

## Power

Power is defined as energy spent per unit of time and is measured in watts.<sup>11</sup> A 100 W bulb turned on for an hour will consume 100 watt-hours of electricity. Watts measure the instantaneous power that has to be adjusted for the time power was being consumed. In order to determine the consumption of power, one has to multiply the rate at which energy is used by the time. This leads to the well-known units of kilowatt-hours and megawatt-hours used to quote prices in electricity trading.

Representation of power in terms of complex numbers power is given as the product of two complex numbers:

$$S = I^*V \quad (19.21)$$

where the asterisk denotes the conjugate of a complex number ( $I^*$  is a conjugate of  $I$ ). By convention, the voltage has a zero phase angle (by the choice of units and the starting point), a lagging current has negative phase angle.

At the intuitive level, voltage measures energy per unit charge, current is the flow rate of charge. The product of the two quantities

## PANEL 19.1 COMPLEX NUMBERS

Power flow equations rely on complex number arithmetic. As a reminder, a point in the Euclidean plane can be represented using its Cartesian coordinates  $(x, y)$ . The same point may be represented as a complex number,  $z$ , given by  $z = x + iy$ , where  $i$  is the square root of  $-1$ . The component  $x$  is called the real part, and  $y$  is called the imaginary part. An alternative representation of complex numbers uses the polar form, with  $z = r \times \cos(\theta) + ir \times \sin(\theta) = r \times e^{i\theta}$ , where  $r$  is the modulus of the number  $z$ , given by  $|z|^2 = zz^*$ . Graphically, the modulus is equal to the length of the straight line interval connecting points  $(0,0)$  and  $(x, y)$ . The angle  $\theta$  of a complex number is represented in electrical engineering by symbol  $\angle$ . The number  $z^* = x - iy$  is called the conjugate of the complex number  $z = x + iy$ . Figure 19.3 illustrates these concepts for an arbitrary number  $(x, y)$ .

There are several notations used in many textbooks, which may be confusing in practice, as there are no standard conventions. Instead of using  $i$  to denote the square root of minus 1, some authors use the letter  $j$ . The reason for this is that the letter  $i$  is also used to denote current. The expression

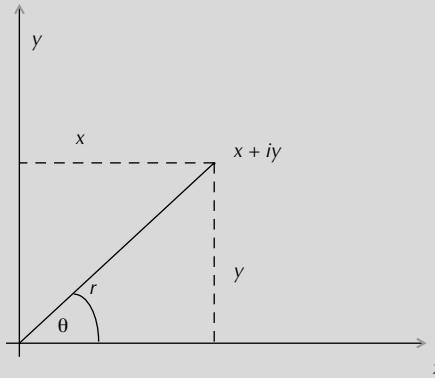
$$r \times e^{i\theta}$$

is sometimes written as

$$r \angle \theta$$

Also, multiplication of two complex numbers written in a polar form results in addition of exponents (angles in this case). A reader should consult an introductory textbook on complex analysis for more information.

**Figure 19.3** Complex number  $x + iy$



provides information about the amount of energy carried by the flow of electrons.<sup>12</sup> Both current and voltage denote instantaneous quantities – ie, levels at a specific point in time. What is more relevant is the level of power over the entire cycle of oscillating current and voltage. In the resistive case, average power can be calculated from the RMS quantities.

$$S_{\text{avg}} = I_{\text{rms}} \times V_{\text{rms}} \quad (19.22)$$

Once reactance is considered, equation 19.22 has to be modified by the cosine of the angle of the shift between voltage and current.

$$S_{\text{avg}} = \cos(\varphi) \times I_{\text{rms}} \times V_{\text{rms}} \quad (19.23)$$

Power can be represented as a point in the complex plane. Apparent power ( $S$ ) is represented by the interval of length  $S$ , at an angle  $\varphi$  with the real axis. This angle corresponds to the phase difference between the voltage and the current. The projection of the vector  $S$  on the real axis of length  $P$  is called real power, the projection  $Q$  on the imaginary axis is reactive power. The ratio of real to reactive power is the same as the ratio of resistance to reactance. Another important fact (introduced without detailed explanations) is that, under the conventions used in electrical engineering, inductive loads “consume” reactive power and capacitive loads supply it. Most loads are inductive; this explains the conventions used with respect to real and reactive power.

### Reactive power

Reactive power is a relatively unknown aspect of the electric industry in spite of its crucial importance to the operations of the integrated generation transmission grid.<sup>13</sup> Real power and reactive power concepts were analysed from the point of view of the mathematical representation of a natural phenomenon. This section covers the supply and pricing aspects of reactive power. Real power, measured in watts, is used to produce motion and heat and performs useful work. Reactive power, measured in VAR, is necessary to provide voltage support. Real power can be compared to water distributed over the network of a local utility. Reactive power can be compared to pressure that is necessary to push water through the pipes. Consumers of water have to incur the cost of water towers used to produce the pressure across the system. One of the chal-

lenges facing the designers of any power market is to create sufficient incentives to secure adequate supply of reactive power. Shortage of reactive power may cause, or contribute to, a blackout.

The components of an electrical system are characterised by inductance and capacitance. Capacitance is a source of reactive power, while inductance consumes reactive power. Generators and power lines may supply or consume reactive power, depending on the circumstances. Most residential, commercial and industrial loads are inductive, and therefore consume reactive power.

In the case of generators, a lagging power factor means that the generator is supplying reactive power (the generator is over-excited); a leading power factor means that it is consuming reactive power factor (is under-excited). The terms under- and over-excitation refer to a special device called an exciter, which is the source of direct current required to energise the magnetic field of the generator. Over-excitation means that more power has to be provided to the generator unit from a DC power source.

Transmission lines may either produce or consume reactive power. Each line is characterised by surge impedance loading, at which the impact of capacitance (production of reactive power) is equal to the impact of inductance (consumption of reactive power). The consumption of reactive power increases with the square of current. This means that, at the time of large power transfers, the losses of reactive power increase. One solution for this is to place the devices that supply reactive power closer to the load.

The relationship between current and consumption of reactive power explains the mechanism behind a blackout. Inadequate supply of reactive power causes voltage drops. This in turn causes an increase in current to maintain the same supply of power (as explained in the previous section,  $\text{Power} = \text{Current} \times \text{Voltage}$ ). The increase in current increases the consumption of reactive power, and this leads to further voltage drops. Dropping voltage may cause some generators to disconnect automatically in order to avoid permanent damage, and also line trips, which further reduce generation available to the loads. Dropping voltage may have a cascading effect causing the entire system to shut down.

In addition to generators and transmission lines, reactive power may be supplied or consumed by other devices, which can be classified as static or dynamic. Static devices (capacitors or inductors)

cannot adjust the level of reactive power if voltage is unchanged and reduce the reactive power production if voltage drops. Dynamic devices can adjust the output of reactive power when voltage drops. The dynamic devices include synchronous generators, synchronous condensers, flexible AC transmission systems (FACTS), static VAR compensators (SVC), static compensators (STATCOM) and dynamic VAR (D-VAR). These devices are distributed across the power system to provide reactive power (when loads increase) or to consume reactive power when its supply exceeds the system needs.

The evolving design of power systems increases the importance of reactive power. One example is provided by the growing significance of wind power. The older wind turbines were not producing reactive power at all, but were consuming it. Modern wind turbines are designed in a way that allows for adjustment of the power factor from 0.90 lagging to 0.95 leading. D-VAR control allows for the dynamic control of voltage. This is important because wind farms are typically located at remote locations at the end points of long transmission lines, where supply of reactive power from alternative sources may present a challenge.

### Power flow equations

An explanation of the concepts of the marginal location prices and the development of numerical examples illustrating how they are calculated requires the introduction of many concepts from both electrical engineering and optimisation theory. This and subsequent sections will be comparatively superficial, and should not be seen as the last word on the subject. The objective is to present certain critical definitions and rely more on intuition than an understanding of complex maths. An interested reader can consult a number of specialised textbooks on the subject.<sup>14</sup>

Power flow equations rely on the complex number arithmetic explained earlier.

The current flowing from a node  $i$  to node  $j$ ,  $I_{ij}$ , is given by the complex number defined as:

$$I_{ij} = \frac{V_i - V_j}{Z_{ij}} = Y_{ij}(V_i - V_j) \quad (19.24)$$

where  $V_i$  ( $V_j$ ) is a voltage at the node  $i$  ( $j$ ), represented as a complex number, and  $Z_{ij}$  ( $Y_{ij}$ ) are the impedance and admittance between the two nodes, respectively. Impedance is given by:

$$Z = Ze^{i\theta} \quad (19.25)$$

where the magnitude  $Z$  is the ratio of the voltage amplitude to the electric current amplitude, while the argument  $\theta$  gives the phase difference (phase angle) between voltage and current. The concept of impedance is the extension of resistance to the alternating current framework. Impedance is a complex number with the real part corresponding to resistance, and the imaginary part being the reactance. In the case of a direct current flow, impedance and resistance will be equivalent, as the phase angle is zero. Admittance is the reciprocal of impedance (hence equation 19.24), and can be written in two different ways, using either admittance or impedance.<sup>15</sup>

The power (given by a complex number) flowing from node  $i$  to  $j$ , is given as the product of the voltage conjugate and current:<sup>16</sup>

$$S_{ij} = V_i^* I_i = V_i^* Y_{ij} (V_i - V_j) \quad (19.26)$$

This can be expanded:

$$S_{ij} = V_i^* Y_{ij} (V_i - V_j) = V_i^* V_i Y_{ij} - Y_{ij} V_i^* V_j = Y_{ij} |V_i|^2 - Y_{ij} V_i^* V_j \quad (19.27)$$

This transformation can be explained by recalling that  $z^* z = (x + iy)(x - iy) = x^2 - i^2 y^2 = x^2 + y^2$ , as  $i^2 = -1$ . The term  $i^2$  is the square root of  $-1$  raised to power 2, or  $-1$ . The expression  $x^2 + y^2$  is the square of the vector norm  $|z|$ .

Equation 19.27 can be written using the polar representations of the complex numbers as follows:

$$S_{ij} = Y_{ij} |V_i|^2 - Y_{ij} V_i^* V_j = |Y_{ij}| |V_i|^2 \angle \theta_{ij} - |Y_{ij}| |V_i| |V_j| \angle (\theta_{ij} + \delta_j - \delta_i) \quad (19.28)$$

where  $\theta_{ij}$  is the angle of admittance and  $\delta_i$  ( $\delta_j$ ) represents the angle of voltage at node  $i$  ( $j$ ). The symbol  $\angle$  denotes an angle. This transformation becomes obvious, once we recognise that this is simply the change of one notation to another (one representation of complex numbers is changed into another). Also,  $|z| = |z^*|$ .

In case there are several power lines extending from a given node, the total power injected in a given node will be distributed over different lines, following the principles summarised in equation 19.28. Assuming that node 1 is connected directly to three nodes (2, 3, 4), the power injected at node 1 will be divided into flows  $S_{1,2}$ ,  $S_{1,3}$  and  $S_{1,4}$ , as follows (we shall drop the comma between node numbers to simplify the notation):

$$\begin{aligned}
S_1 = -(S_{12} + S_{13} + S_{14}) &= |Y_{12}| |V_1^2| \angle \theta_{12} - |Y_{12}| |V_1| |V_2| \angle (\theta_{12} + \delta_2 - \delta_1) + \\
&|Y_{13}| |V_1^2| \angle \theta_{13} - |Y_{13}| |V_1| |V_3| \angle (\theta_{13} + \delta_3 - \delta_1) + \\
&|Y_{14}| |V_1^2| \angle \theta_{14} - |Y_{14}| |V_1| |V_4| \angle (\theta_{14} + \delta_4 - \delta_1)
\end{aligned} \tag{19.29}$$

This can be rewritten as:

$$\begin{aligned}
S_1 = -(S_{12} + S_{13} + S_{14}) &= |Y_{11}| |V_1^2| \angle \theta_{11} - |Y_{12}| |V_1| |V_2| \angle (\theta_{12} + \delta_2 - \delta_1) + \\
&- |Y_{13}| |V_1| |V_3| \angle (\theta_{13} + \delta_3 - \delta_1) + |Y_{14}| |V_1| |V_4| \angle (\theta_{14} + \delta_4 - \delta_1)
\end{aligned} \tag{19.30}$$

$Y_{11}$  is the admittance at node 1 equal to the sum of admittances of the branches emanating from the node 1, ie  $Y_{12} + Y_{13} + Y_{14}$ . By convention, power flowing into the node ( $S_1$ ) is positive, power flowing out – negative. This is accomplished by aggregating the terms containing  $|V_1^2|$ , adding the admittance at node 1, recognising that the power injected at any node is equal to the power flowing out (this is a consequence of the Kirchoff law). The current injected at node 1 is given by:

$$\begin{aligned}
I_1 &= Y_{12}(V_1 - V_2) + Y_{13}(V_1 - V_3) + Y_{14}(V_1 - V_4) = \\
&(Y_{12} + Y_{13} + Y_{14})V_1 - Y_{12}V_2 - Y_{13}V_3 - Y_{14}V_4
\end{aligned} \tag{19.31}$$

This leads to the system of equations for  $n$  nodes, which is written in the matrix form as:

$$I = YV \tag{19.32}$$

where  $Y$  is the network admittance matrix and  $I$  and  $V$  are the vectors of injected power and voltage. The equation can be expanded as:

$$\begin{bmatrix} I_1 \\ I_2 \\ \dots \\ I_n \end{bmatrix} = \begin{bmatrix} \sum_{1,allbranches} Y_{1n} & -Y_{12} & -Y_{13} & \dots & -Y_{1n} \\ -Y_{21} & \sum_{1,allbranches} Y_{2n} & -Y_{23} & \dots & -Y_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ -Y_{n1} & -Y_{n2} & -Y_{n3} & \dots & \sum_{1,allbranches} Y_{nn} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ \dots \\ V_n \end{bmatrix} \tag{19.33}$$

The diagonal elements are the sum of admittances terminating at a given node, as seen in the four-node example of equation 19.31. Many elements of the matrix  $Y$  will be zero (as there are no connections between the two given nodes). Also, the sum of row elements will be zero, and this will make the matrix singular.<sup>17</sup> This problem can be addressed by choosing a reference node, which allows the removal of one row and one column from matrix  $Y$ . The resulting  $(n - 1) \times (n - 1)$  matrix can be inverted.

Equation (19.32) allows one to derive a system of equations for power. Power is given by the product of  $V^*I$ , which translates into system of equations:

$$S = \text{Diag}[V^*]I \quad (19.34)$$

where  $\text{Diag}[V^*]$  denotes a diagonal matrix with voltages at different nodes. We can replace  $I$  in equation 19.34 with the matrix multiplication from equation 19.33, which will produce the real ( $S$ ) and imaginary ( $Q$ ) parts of complex power. The imaginary component of complex power is known as reactive power. The equations are generalised for a node  $i$  connected to  $j$ ,  $k$  and  $l$ .

$$\begin{aligned} P_i = & |Y_{ii}| |V_i|^2 \cos(\theta_{ii}) - |Y_{ij}| |V_i| |V_j| \cos(\theta_{ij} + \delta_j - \delta_i) \\ & - |Y_{ik}| |V_i| |V_k| \cos(\theta_{ik} + \delta_k - \delta_i) + |Y_{il}| |V_i| |V_l| \cos(\theta_{il} + \delta_l - \delta_i) \end{aligned} \quad (19.35)$$

$$\begin{aligned} Q_i = & |Y_{ii}| |V_i|^2 \sin(\theta_{ii}) - |Y_{ij}| |V_i| |V_j| \sin(\theta_{ij} + \delta_j - \delta_i) + \\ & - |Y_{ik}| |V_i| |V_k| \sin(\theta_{ik} + \delta_k - \delta_i) + |Y_{il}| |V_i| |V_l| \sin(\theta_{il} + \delta_l - \delta_i) \end{aligned} \quad (19.36)$$

Equations 19.35 and 19.36 can be easily modified to describe a more complex system with multiple nodes and branches. The resulting system of equations will be non-linear and not easily solved using standard software. The practical solution is to linearise the system, effectively using a DC approximation of the alternating current system.

The DC approximation ignores the losses – ie, makes the assumption that the resistance of the transmission system is zero. One consequence will be the that the angle of admittances will be  $90^\circ$ , which means that<sup>18</sup>

$$\cos(\theta_{ij} + \delta_j - \delta_i) = \sin(\delta_i - \delta_j) \quad (19.37)$$

The equation for real power becomes:

$$\begin{aligned} P_i = & -|Y_{ij}| |V_i| |V_j| \sin(\delta_j - \delta_i) - |Y_{ik}| |V_i| \\ & |V_k| \sin(\delta_k - \delta_i) - |Y_{il}| |V_i| |V_l| \sin(\delta_l - \delta_i) \end{aligned} \quad (19.38)$$

For small values of the argument  $\theta$ ,  $\sin(\theta) \approx \theta$  (one can easily check that around 0, the sine function can be approximated through a straight line). This approximation simplifies equation 19.38 further:

$$\begin{aligned} P_i = & -|Y_{ij}| |V_i| |V_j| (\delta_j - \delta_i) - |Y_{ik}| |V_i| \\ & |V_k| (\delta_k - \delta_i) - |Y_{il}| |V_i| |V_l| (\delta_l - \delta_i) \end{aligned} \quad (19.39)$$

$$P_i = |Y_{ij}| |V_i| |V_j| \delta_i + |Y_{ik}| |V_i| |V_k| \delta_i + |Y_{il}| |V_i| |V_l| \delta_i - |Y_{ij}| |V_i| |V_j| \delta_j - |Y_{ik}| |V_i| |V_k| \delta_k - |Y_{il}| |V_i| |V_l| \delta_l \quad (19.40)$$

$$P_i = |Y_{ij}| + Y_{ik} + Y_{il} - |Y_{ij}| \delta_j - |Y_{ik}| \delta_k - |Y_{il}| \delta_l \quad (19.41)$$

The additional assumption that allows us to derive equation 19.41 is that, in a typical power system, the voltages will be practically equal at different nodes and can be normalised to have unit value. This means that  $|V_i| = 1$ , for all  $i$ .

Equation 19.41 implies a system of equation for all the nodes, which can be rewritten in matrix form as:

$$P = Y\delta \quad (19.42)$$

$$\delta = Y^{-1} P \quad (19.43)$$

In equation 19.42,  $Y$  denotes the admittance matrix, with nodal admittances on the diagonal,  $\delta$  is the vector of angles and  $P$  is the vector of power injected into the network at each node. Once the angles of the node voltages are calculated the line flows can be calculated based on the branch admittances

$$P_{ij} = |Y_{ij}| (\delta_i - \delta_j) \quad (19.44)$$

The final step is calculation of the flows across different branches of the network. This is accomplished by using the shift factor matrix that is used to translate the injections at specific nodes into the network flows.

$$P_{\text{flows}} = S_F \times P_{\text{node}} \quad (19.45)$$

where  $P_{\text{flows}}$  is the vector of flows across the branches of the network,  $S_F$  is the shift factor matrix and  $P_{\text{node}}$  is the vector of nodal injections of power. The matrix  $S_F$  is calculated as follows:

$$SF = Y \times A \times Z^T \quad (19.46)$$

where  $Y$  stands for the matrix of branch admittances,  $A$  is the incidence matrix of the network,<sup>19</sup>  $Z$  stands for the inverse of the reduced admittance matrix and  $T$  is the matrix transform operator. These concepts will be easier to grasp with a numerical example in Chapter 21.

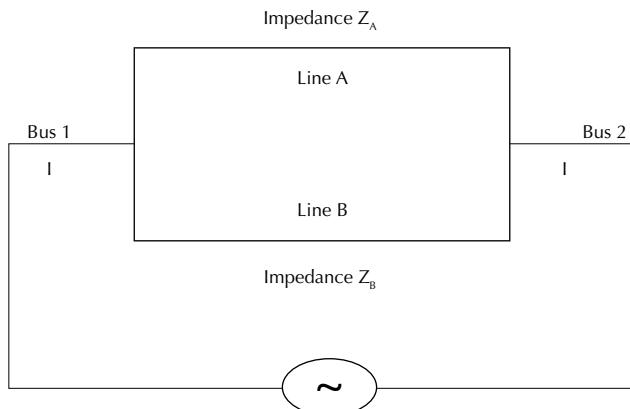
## NUMERICAL EXAMPLES

The examples below illustrate some of the concepts we have discussed.<sup>20</sup> We shall revisit them later when locational marginal prices are introduced. The concept of impedance is very important in modelling power flows in the networks and analysing physical generation/transmission systems. In practice, voltage, apparent power and impedance are normalised and expressed as dimensionless units. This is done by selecting a reference bus. All relevant quantities are expressed as multiples of the base values, and expressed in pu (per unit) levels. This approach is equivalent to measuring altitude by deciding on an arbitrary reference level (ie, the sea level in practice, although any arbitrary point would do). For example, in a balanced three-phase network, the base impedance will be given as:

$$Z_0 = \frac{V_0^2}{S_0} \quad (19.47)$$

with  $V_0$  and  $S_0$  denoting base voltage and base apparent power, respectively. The practical significance of impedance can be illustrated with an example that illustrates power flows on parallel lines and brings together all the concepts discussed above. In Figure 19.4, current flowing into Bus 2 from Bus 1 travels along two lines, A and

**Figure 19.4** Power flows on parallel lines



Source: Based on Daniel Kirschen, Goran Strbac, "Fundamentals of Power System Economics," John Wiley & Sons, Ltd, 2004

B, with impedances  $Z_A$  and  $Z_B$ . According to Ohm's Law and Kirchoff's Voltage Law, the current flows will be proportional to voltage drop and inversely proportional to impedances.

$$I_A = \frac{V_{12}}{Z_A} \text{ and } I_B = \frac{V_{12}}{Z_B} \quad (19.48)$$

It follows:<sup>21</sup>

$$Z_A I_A = Z_B (I - I_A) \quad (19.49)$$

$$Z_A I_A = Z_B I - Z_B I_A \quad (19.50)$$

$$(Z_A + Z_B) I_A = Z_B I \quad (19.51)$$

$$I_A = I \frac{Z_B}{Z_A + Z_B} \quad (19.52)$$

$$I_B = I \frac{Z_A}{Z_A + Z_B} \quad (19.53)$$

By analogy, if we ignore reactive power flows and line losses, we can derive power transfer distribution factors based on the reactances of the two lines ( $X_A$  and  $X_B$ ):

$$\frac{X_B}{X_A + X_B} \text{ and } \frac{X_A}{X_A + X_B} \quad (19.54)$$

To derive 19.54 we assume that resistance of every line is much smaller than reactance. This means that the impedances are reduced to reactances, as we can omit the resistance term (R) in (19.14).

The practical significance of these concepts can be illustrated with the ubiquitous example of a three-node network. Each node has both load and generation attached to it. Table 19.1 summarises the information about branches connecting the nodes. The reactances listed in the table are shown in hexagons in Figure 19.5. Suppose that Generator 1 sells 400 MW to Load 3. Power flows will distribute as follows:

$$\text{Path } 1-2-3 = \frac{0.2}{0.5} \times 400 = 160$$

$$\text{Path } 1-3 = \frac{0.3}{0.5} \times 400 = 240$$

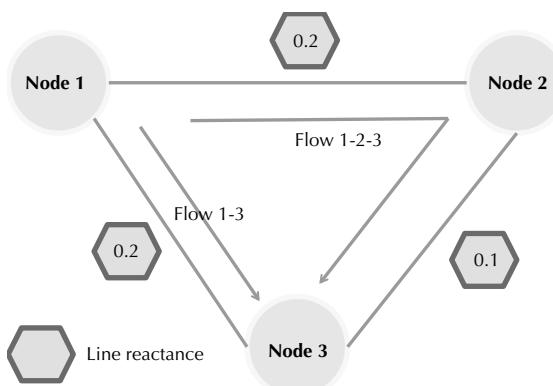
**Table 19.1** Three-node network description

Line	Reactance (pu)	Line capacity (MW)
1–2	0.2	126
1–3	0.2	250
2–3	0.1	130
Sum	0.5	

Source: Daniel Kirschen, Goran Strbac, "Fundamentals of Power System Economics," John Wiley & Sons, Ltd, 2004, page 145

These calculations can be explained as follows. The flow over Path 1–3 is calculated from the reactance of line 1–2–3 (.2+.1) and total reactance (0.5), and is equal to 240 MW. Reactances are additive, therefore the reactance for the network is 0.5 (.2+.2+.1). The flow over Path 1–2–3 is equal to 160 MW. The capacity of line 1–2 is only 126, and therefore the transaction is physically impossible. Power cannot be delivered to the load. What are the solutions? One obvious answer is to curtail the transaction to a viable level. For example, the sale can be reduced to 314 MW, and this will be roughly enough to satisfy the constraints of the network (as the reader can easily verify).

Another solution, with a much greater practical significance, can be illustrated by the assumption that Generator 3 sells 200 MW to Load 1. This will create a counterflow, which will offset the flows in

**Figure 19.5** Three-node network example

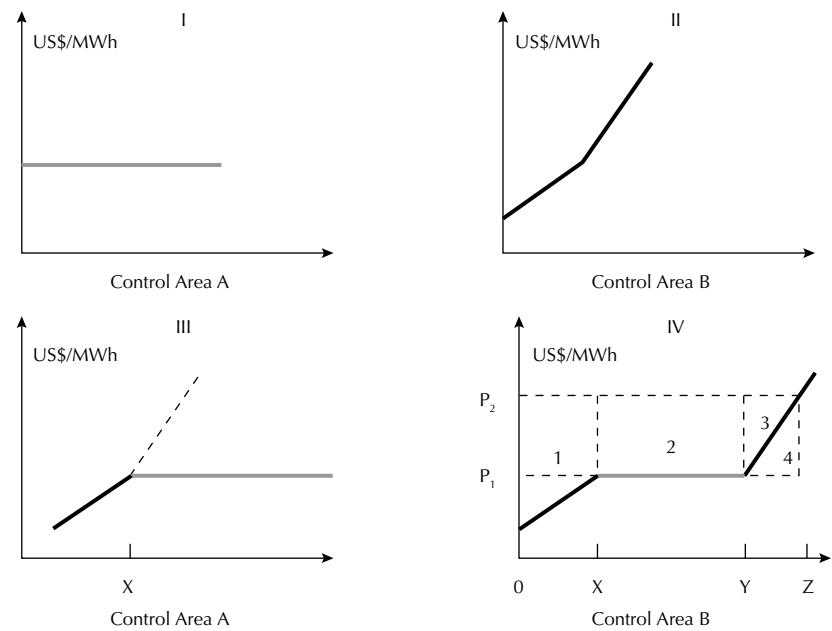
Based on Daniel Kirschen, Goran Strbac, "Fundamentals of Power System Economics," John Wiley & Sons, Ltd, 2004

the opposite direction. The flow of 80 MW on line 3–2–1 would partially offset the flow of 160 MW in the opposite direction, restoring the viability of the network.

### Uplift

A discussion of power flows would not be complete without exploring the concept of uplift, which is of significant practical importance to any power pool design.<sup>22</sup> This concept can be illustrated with a diagram showing two control areas, A and B: one with an upward sloping supply curve and the other with a horizontal supply curve if only locally installed generation is considered (see Figure 19.6). The supply curve for area B, with unlimited transmission capacity, is shown in Panel III of the figure. For loads below point X, local generation units are dispatched. Above point X, power is imported from control area A. The supply curve incorporating imports is initially upward sloping and then becomes horizontal as imports become competitive with local generation (the branch of the

**Figure 19.6** Uplift illustration



Source: U.S. Energy Information Administration

supply curve corresponding to imports is shaded). If the transmission capacity is limited to a level equal to the interval XY, and the load exceeds Y, additional units from B have to be dispatched. The supply curve consists of the upward sloping, then horizontal and then again upward-sloping segments. Assuming that market prices are formed in a competitive market and are equal to the marginal cost (represented by the supply curves), additional consumer expenditures (compared to the unconstrained case) are equal to the sum of the areas 1 + 2 + 3 + 4, or the product  $(P_2 - P_1) \times OZ$ . This area can be disaggregated into additional profits of the generators, area 1 + 3, payments received by the owners of transmission or the transmission rights (2) and the re-dispatch cost (4). The payments required to induce generators to re-dispatch are called uplift. The design of uplift charges is a very contentious issue in the design of any power pool.

### Units of measurement

Electricity is, for many reasons, the most challenging part of the energy system to a trader or analyst who is also not a physicist or engineer (as is usually the case). We cannot see electricity and cannot visualise how it propagates through the transmission/distribution system. At the same time, one cannot function in this industry without some basic understanding of how the physical layer of the power business works. This section will examine units of measurement used in electricity markets, with a discussion of some of the physical concepts used in the book. What is of concern here are the units one uses in pricing transactions and producing units, not the units used to describe and manage the physical flows.

Power, measured in watts, is the rate at which energy is produced, transmitted and used. Every electric appliance has a rating in watts that indicates the rate of energy consumption during a very short interval of time. As a biological organism, the author is roughly an equivalent of a 100-watt electric bulb (sometimes more, sometimes less). As a member of highly developed post-industrial society, we are an equivalent of an 11,000-watt bulb.<sup>23</sup> We live in a big house which we try to keep at 70° F in the summer and 74° F in the winter. We drive a big air-conditioned/heated car, and our diet is based on high protein input and exotic foods from around the world. A watt-hour is energy consumed at a steady rate of one watt for an hour. Most residential and commercial tariffs, and, by extension invoices,

are specified in kilowatt-hours (1 KWh = 1,000 watt-hours).<sup>24</sup> Energy trading relies on megawatt-hours (1 MWh = 1,000,000 watt-hours), and these are the units in which we transact. Installed generation capacity and transmission capacity is expressed in megawatts. The cost building a generation plant is often normalised and expressed in dollars per kilowatt. In the case of capacity markets, the prices are expressed in US\$/kW-year (kilowatt-year) or US\$/kW-month. In the power pools in which capacity markets are used and constitute a component of costs, one has to engage in some unit conversion to calculate the cost of providing electricity (capacity payments over time have to be distributed over the MWhs being delivered).

Many aspiring energy traders have been asked at their job interviews about what an MW is, with the expected (and presumably correct) answer being that it is an “amount” of electricity sufficient to supply 1,000 homes. The reality of the energy business is much more complicated. The common usage of megawatts and kilowatts to measure the generation and consumption of electricity should be exercised with caution given that both consumption and generation fluctuate over time. The fact that a given load may consume a MW of electricity at any given point in time does not mean that the rate of consumption is equal over time. The industry historically uses load factor to provide a concise description of a load generation profile. The load factor is described as the ratio of average load (generation) to maximum load (generation):

**Table 19.2** US Residential Electricity Consumption, 2001

Census division state	Number of consumers	Average monthly consumption (kWh)
New England	5,822,935	618
Middle Atlantic	15,045,495	641
East North Central	18,705,754	763
West North Central	8,287,837	903
South Atlantic	22,473,797	1,088
East South Central	7,356,975	1,193
West South Central	12,883,403	1,151
Mountain	7,368,280	847
West Coast	15,763,570	668
Hawaii & Alaska	609,661	642
<b>US total</b>	<b>114,317,707</b>	<b>877</b>

Source: Bob Bellemare, 2003, “What is a megawatt?” UtiliPoint’s *Daily Issue Alert*, June 24

$$\text{load factor} = \frac{\text{average load}}{\text{maximum load}} \quad (19.55)$$

This way of describing load is very imperfect as the same load factor (for example, 50%) may correspond to dramatically different load shapes. The industry developed this measure in the days when the load (generation) profile was quite stable and one summary statistic was sufficient to describe the load.

What can be said about the number of homes and one MW? The answer can be found in Table 19.2 which illustrates regional differences in average US residential electricity consumption.<sup>25</sup>

A power plant using coal may have a load factor of 75%. As Bob Bellemare explains:

Going through the math, a 1,000 megawatt rated coal generator with a 75 percent capacity factor generates about 6.6 billion kWhs<sup>26</sup> in a year, equivalent to the amount of power consumed by about 900,000 homes in the Northeast but only 460,000 homes in the South. In other words, each megawatt of rated capacity for a coal plant in the Northeast generates the equivalent amount of electricity consumed by 900 homes in the Northeast but only about 460 homes in the South. By comparison, a 30 percent capacity factor, 100 MW wind farm would generate the equivalent amount of power consumed by about 35,000 homes in the Northeast and 18,000 homes in the South. In other words, each megawatt of rated capacity for a wind farm in the Northeast generates the equivalent amount of electricity consumed by 350 homes in the Northeast and 180 homes in the South.

The sweeping statement about 1,000 homes to a MW turns out to be a simplification.

## CONCLUSIONS

This chapter is a good illustration of what every energy trader and analyst knows: it is a very complex industry and technological complexity translates into complexity of market design, market instruments and electricity-related contracts. Mastering this market requires a significant investment in learning the underlying technology, the laws and regulations that apply to this business, relationships with other parts of the energy industry, primarily natural gas and coal. The potential rewards are huge: the pool of talent in electricity trading is quite shallow and a competent trader can be very successful. It is a relatively immature market, with many barriers to entry reflecting the time and difficulty of acquiring the necessary skills.

Having learned what electricity is and how it propagates, we are ready to move on to the next topic: the many different ways of producing electricity.

- 1 It is imperfect as the electrons are not flowing through a conductor from one end to another, like molecules of gas through a pipeline. The energy is rather transmitted through the wire by electrons bumping into other electrons and displacing them.
- 2 The competition between the two competing solutions was very acrimonious and led to a personal clash between Westinghouse and Edison. The story of the early days of the US power industry is relayed in Jill Jonnes, 2003, *Empires of Light : Edison, Tesla, Westinghouse, and the Race to Electrify the World* (New York, NY: Random House); Tom McNichol, 2006, *AC/DC: The Savage Tale of the First Standards War* (San Francisco, CA: Jossey-Bass).
- 3 The choice between the trigonometric functions is a matter of convenience. One could use, for example, triangular wave functions.
- 4 Angular frequency is measured in radians per second. Angular frequency, multiplied by time, is equal to frequency,  $\phi$ , the number of complete oscillations of current per unit of time. The radian is a unit used to measure angles equal to  $180/\pi$ , or about 57.2958 degrees.
- 5 If the phase shift is equal to zero, the sine curve starts at point zero.
- 6 The interval  $T_2 - T_1$  is taken as the number of full cycles, the sine term in the integral is equal to zero (intuitively, the positive and negative values cancel out).
- 7 Inductive reactance is related to the ability of an inductor to delay or reshape alternating current.
- 8 In integrated transmission/generation systems, the most important examples of inductors are power lines.
- 9 Inductive reactance is associated with current that is lagging voltage ( $\phi$  positive). Capacitive reactance is associated with lagging current ( $\phi$  is negative).
- 10 The general formula for the division of complex numbers is given by  $(a + ib)/(c + id) = (ac + bd + i(bc - ad))/(c^2 + d^2)$ .
- 11 See also earlier in this chapter.
- 12 The electrons do not flow in the same sense a molecule of water passes through a pipe. A better analogy is electrons bumping into other electrons carrying energy this way.
- 13 Useful information on the issues related to reactive power can be found in "Principles for efficient and reliable reactive power supply and consumption," FERC, Staff Report AD05-1-000, February 4, 2005.
- 14 This section relies to a large extent on Narayan S. Rau, 2003, *Optimization Principles: Practical Applications to the Operation and Markets of the Electric Power Industry* (Piscataway, NJ: IEEE Press). The book combines a clear exposition of optimisation theory in combination with a concise introduction to power engineering concepts. See also Fred C. Scheppele, Michael C. Caramanis, Richard D. Tabors, Roger E. Bohn, "Spot Pricing of Electricity," Springer 1988.
- 15 The inverse of a complex number is easy to calculate using the polar form.
- 16 The conjugation of the complex number  $x + iy$  is  $x - iy$ .
- 17 A singular matrix is a square matrix with a determinant equal to zero. Such a matrix cannot be inverted.
- 18 This is based on the following identity:  $\sin(\theta) = \cos(90^\circ - \theta)$ . Note the change of sign.
- 19 The incidence matrix is explained in more detail in the section containing an example of the calculation of locational marginal prices.
- 20 The examples in this section are based on Daniel Kirschen, Goran Strbac, "Fundamentals of Power System Economics," John Wiley & Sons, Ltd, 2004.
- 21 This is because  $Z_A I_A = V_1 - V_2 = Z_B I_B$ . Power flowing from 1 to 2 splits into two flows:  $I_A + I_B = I$ .
- 22 Based on EIA, 2004, "Electricity transmission in a restructured industry: Data needs for public policy analysis," December, p 43.

- 23 We borrowed this example from Dr. Tad Patzek, UT Austin.
- 24 In the early days of the electricity industry billing, was based on the duration of connection. It was a solution feasible only in the systems in which electricity was flowing at the constant rate over time.
- 25 Bob Bellemare, 2003, "What is a megawatt?" UtiliPoint's *Daily Issue Alert*, June 24.
- 26 This is calculated as  $1,000 \times 0.75 \times 365 \times 24 \times 1,000$ .



# 20

## *Power Generation*

We will start our discussion of power generation with a review of the basic facts related to electricity generation, focusing on the composition of the generation fleet and electricity production in the US. This will be followed by a look at the different types of technology used in producing electricity. The key message of this section is the variety of different ways to generate electricity and the challenges of integrating different modes of generation.

Understanding generation technology is crucial for all traders and energy analysts. In competitive commodity markets, prices are set at the margin. If we ignore the potential for transmission constraints and line losses, the cost of generating electricity at the unit required to satisfy a marginal megawatt-hour of demand sets the market prices. This statement has to be further qualified once transmission issues and the potential for exercising market power are considered. However, one has to develop a good understanding of the technology behind electricity generation before more complex problems can be addressed.

Generation units vary with respect to size, unit cost of production, technology used to produce electricity and the time required to dispatch them (ie, the time elapsed from the moment the decision to generate electricity is taken to the moment of reaching full capacity output). From the point of view of a trader, units with more flexibility represent higher value. Controlling such units, either through direct ownership or through contractual arrangements, creates opportunities to take advantage of spikes in market prices. Most transactions in the electricity markets have a physical component and, in many cases, are unit-specific. Understanding the properties of a generation unit underlying a transaction is critical to proper contract valuation and risk management.

The early days of deregulated electricity markets saw many transactions reflecting a poor understanding of the generation side of the business. Many trading and marketing shops made significant investments in generation capacity (directly or through contractual arrangements) and in many cases suffered significant losses. The common denominator of such deals was a somewhat simplistic knowledge of the flexibilities embedded in many power plants – with a repeatedly used syllogism running as follows:

- ❑ certain power plants have a lot of embedded optionality;
- ❑ options are good; therefore
- ❑ buying or building power plants is good.

However, both of these premises could be often wrong. The optionality of some power plants was greatly overrated and good options acquired at excessive prices can wreck a career. Poor logic combined with a poor awareness of the facts resulted in some generation units being sold by companies on the edge of bankruptcy at 20 cents on the dollar. This illustrates the importance of understanding technology if one insists in investing in hard assets.

In this chapter, we will concentrate on thermal, nuclear, wind and hydropower units. Other renewable electricity sources (such as solar and geothermal) have limited significance from the point of view of electricity trading, except for being the source of many headaches. Transactions related to renewable energy sources that are not currently economic (ie, cannot operate without subsidies) have the huge embedded risk that the largesse of the state may be withheld at short notice.

### **US ELECTRICITY GENERATION: BASIC FACTS**

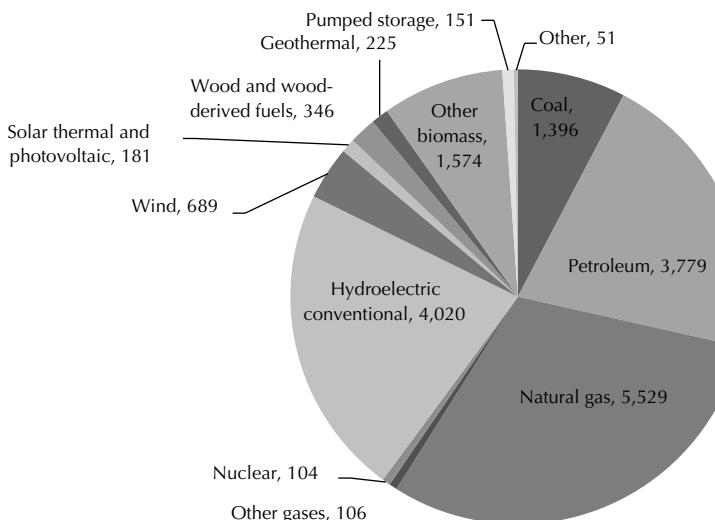
There are many different ways to classify and describe existing power plants. We can group them using criteria such as location, ownership and fuel type. To describe the size of generation fleet, we can look at the different types of plants, the nameplate capacity (the amount of electricity a plant can produce under optimal conditions) and the actual electricity output from different power plants. The answers will be different in each case: relatively small plants may increase the count but will not contribute significantly to production levels and total installed capacity. Many plants (for example, gas-

fired peakers and super peakers) are designed to operate for brief periods of time under conditions of peak demand, and the capacity they represent does not translate into a high output of electricity.

Figure 20.1 shows the number of power plants by different energy sources for 2010. The total number of generation plants in the US was 17,658, with natural gas plants accounting for 31% of the total, followed by hydroelectric plants (23%) and oil-fired plants (21%). Coal plants, surprisingly, accounted for only 8%, although they tended to be, on average, much bigger (as one can see from Figure 20.2). In terms of nameplate capacity, gas-fired power plants dominate (41%), followed by coal plants (30%) and nuclear power plants (10%). Hydropower plants represented just 7% of installed capacity, a clear indication that they tend to be numerous but, on average, quite small. The distribution of electricity actual output by plant type is dominated by the coal plants (48%), followed by natural gas (21%) and nuclear power plants (10%). The percentages reported here and in Figure 20.3 are based on the total generation of 4,119,388 million kilowatt hours in 2011.

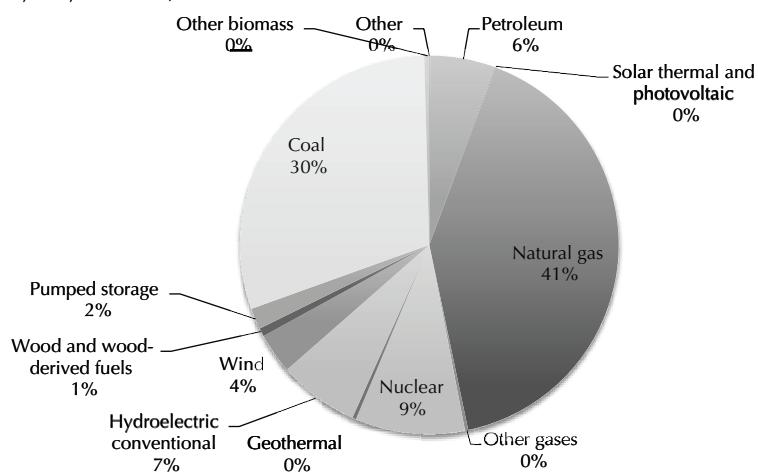
The discrepancies between the unit count, installed capacity and actual output will become clearer after the discussion of the

**Figure 20.1** US generation plants by type (nominal numbers), 2010



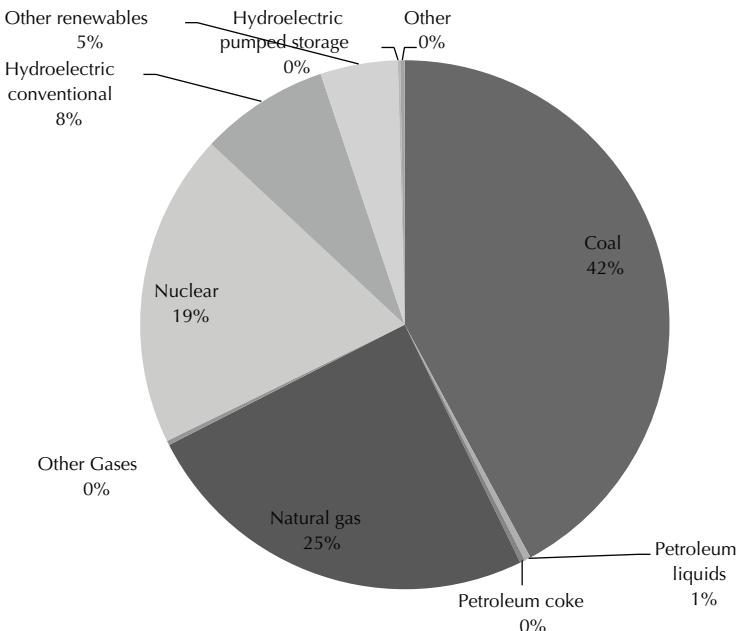
Source: U.S. Energy Information Administration (2010)

**Figure 20.2** US power plants, nameplate capacity, 2010 (total = 1,138,638 MW)



Source: U.S. Energy Information Administration (2010)

**Figure 20.3** US electricity generation by plant type (2011)



Source: U.S. Energy Information Administration

technology of different power plants later in this section. It suffices to say that nuclear and coal generation units operate continuously as base load plants over long time periods (as long as weeks and months), with some adjustments to the level of output during the off-peak hours. Gas-fired power plants may operate as base load units (combined cycle units, which will be explained later), but units such as peakers and super peakers operate infrequently and their capacity (ability to produce power) does not translate into the actual output.

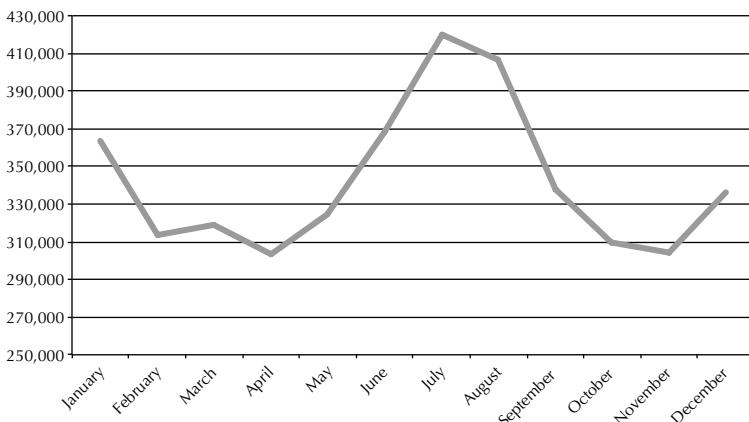
The operations of power plants are characterised by seasonality, as illustrated in Figure 20.4. Figure 20.4 shows total US output of electricity in 2011 by month. The output peaked in July. Output from coal, nuclear and gas-fired power plants increases in the summer months to meet air-conditioning loads. In other parts of the world, the seasonal pattern is different: in Europe, generation output peaks in winter.

## POWER GENERATION PLANTS

### Thermal power plants

Thermal power plants convert the chemical energy of fuel into electric energy through a number of processes, which involve the conversion of chemical energy into heat through its combustion and conversion into mechanical energy. Most thermal plants use fossil

**Figure 20.4** US electricity output, thousand megawatthours (2011)



Source: U.S. Energy Information Administration

fuels such as coal, natural gas or oil (processed into products such as #5 or #6 fuels), although there is a small but growing number of plants that use biomass (hay, straw, wood, etc). We shall start the review of thermal power plants with a look at coal plants.

### **Coal plants**

Coal arrives at power plants usually by train or barge, and is stored at the yard next to a plant. A train may contain about 100 cars (which translates into over a mile of length) and carry a load of 10,000 tonnes. The transportation cost represents most of the price paid by the power plants for coal. A large power plant may receive between three and five trains a day, with the unloading of each train taking a few hours.<sup>1</sup> The power traders should pay attention to any disruptions in the transportation system that would cause delays in deliveries of coal to the power plants (damaged railway lines, floods, frozen rivers and lakes) and result in shocks to electricity prices.<sup>2</sup>

Coal is transported to the boiler on a conveyor belt. Before it is burned, it has to be crushed to pieces of about two inches in size, or pulverised. One of the advantages of using pulverised coal is the increased flexibility of a plant. In the case of a stoker-fired plant, a significant amount of partly burned coal may accumulate on the grate, and it takes more time to bring the unit down during shutdown. Burning coal produces heat, which converts water into the steam that turns the turbine.

The steam passes through a turbine which has a large number of blades mounted on a common shaft, some moving and some stationary, organised in groups called stages. As the steam passes through successive stages, its pressure falls, explaining why the size of the blades and the diameter of a turbine are increasing in order to extract the energy of steam in the most efficient way. The rotating part of the turbine is very heavy (up to 200 tons) and has to continually rotate slowly, even if the plant is not operating, in order to avoid deformation. The temperature and pressure of steam decreases as it passes through the turbine. At the exit from the turbine, steam is converted into water in the unit called a condenser. Its purpose is not only to convert steam to water for re-circulation to the boiler, but also to accomplish this in a very short time to enhance the efficiency of a turbine by creating a near vacuum at its exit. Condensed steam accumulates at the hot well, and then can be pumped back to the boiler.

Along the way, water is reheated (exhaust heat is used for this process) before it enters the boiler. Of course, some water is lost through evaporation and inevitable leaks, and the losses have to be replaced by make-up water. This explains why thermal plants have to be located close to a source of water, such as lakes and rivers. Fresh water has to pass through an extensive process of treatment and removal of impurities. This removal is critical to the avoidance of accumulating scale-forming materials on the surfaces of the boilers, leading to reduced efficiency in the production of steam and potential overheating and equipment failures. Another important reason for the treatment of water is the removal of oxygen that is highly corrosive. The treatment of water may involve evaporation and condensation, the use of chemicals or ion-exchange demineralisers.

Air that is delivered to the boiler goes through a heating unit to increase its temperature in order to optimise burning of coal in the boiler. The heater uses turbine exhaust gases. Another important technological process is the handling of ash which accumulates in the boiler and has to be removed and disposed of.<sup>3</sup> The amount of ash depends on the type of coal burned by the plant. The ash content is one of the most important aspects of coal quality, and will be discussed in the section on coal. The removal of fly ash from the flue gases uses a number of different technological processes, such as electrostatic precipitators, cyclone separators and spraying with water. Ash recovered from the boilers and from the flue gases can be used in construction as a material in the production of cement.

Flue<sup>4</sup> gases contain certain acid gases (such as SO<sub>2</sub> and NO<sub>x</sub>) which pose serious health problems and create the problem known as acid rain and ozone formation. The details of environmental impact and process used to control this form of pollution are covered in the chapter on green markets.

The turbines are connected to the generation units which convert mechanical energy produced by a turbine into electricity. The physical principles underlying the generation of electricity are very simple. A prototype of a generator is a magnet spinning inside coils of wire. The spinning part is called a rotor and the stationary part is called a stator (or armature). An industrial generator does not use natural magnets in the rotor: their size and the strength of the magnetic field would be insufficient for the requirements of large-scale electricity production. A natural magnet is replaced with an

electromagnet which needs a separate source of power to energise, called an exciter. Exciters are very important components of a generation plant, as they are also used to control voltage.

The alternating current systems use three-phase generators. The stators have three separate sets of outgoing windings, each carrying current separated by 120° from the current flowing through the other two windings (the meaning of 120° has been explained in Chapter 19). There are two reasons for the use of such a solution. First, the rotor of a one-phase generator is exposed to a pulsating torque. This may be eliminated by using three sets of wires in the armature, called phases, which are spaced in such a way that the current and voltage in each phase are shifted in time by one third of a cycle. Three oscillating magnetic fields produced by the currents induced in the armature windings add up to a field constant in magnitude. The second reason is related to the transmission of alternating current. A three-phase system reduces the amount of wires needed to serve the load.

Pulverised coal (PC) generation plants operate using subcritical, supercritical or ultra-supercritical technology. The difference is primarily in the temperature at which boilers and turbine systems operate. Supercritical plants are more efficient and this translates into lower coal usage and reduced emissions of CO<sub>2</sub>. Table 20.1 summarises the technological parameters for these technologies. The statistics assume a 500 MW plant (net capacity) burning Illinois #6 coal (HHV<sup>5</sup> = 25,350 kJ/kg), 85% capacity factor. The last column contains performance data for a circulating fluidised bed (CFB) unit burning lignite with HHV = 17,400 kJ/kg and costing US\$1.00/MMBtu.

**Table 20.1** Performance characteristics, air-blown PC generation

	Subcritical PC	Supercritical PC	Ultra-supercritical	Subcritical CFB
Heat rate (Btu/KWh)	9,950	8,870	7,880	13,400
Generating efficiency	34.3	38.5	43.3	34.8
Coal feed (kg/hour)	208,000	185,000	164,000	297,000
CO <sub>2</sub> (kg/hour)	466,000	415,000	369,000	517,000
CO <sub>2</sub> (g/kWh)	931	830	738	1030
Plant cost (US\$/kW)	1,280	1,330	1,360	1,330
O&M (cents/kWh)	.75	.75	.75	1.00

Source: "The future of coal: Options for a carbon-constrained world," Massachusetts Institute of Technology, 2007, p 19.

## Gas-fired power plants

Gas turbines represent a very important development in the history of the power industry. Technological improvements in the design of natural gas turbines lowered the barriers to entry in the power generation business through combination of two factors: lower optimal scale, compared to coal and nuclear power plants, and lower capital cost and construction timelines. Easy access to the natural gas pipelines in the US and UK supported the “dash-to-gas” construction boom of natural gas power plants in the late 1990s, which was driven mostly by the merchant energy industry. Environmental considerations and measures designed to reduce SO<sub>2</sub> and NO<sub>x</sub> pollution also favoured natural gas generation, which not only benefitted from the deregulation of the electricity industry but was also one of the factors accelerating the trend towards liberalisation of the industry.

In most US power pools, gas-fired power plants are at the margin during the on-peak periods – ie, are run to satisfy the marginal unit of demand and, therefore, set market prices or at least have a major impact on how market prices are determined.<sup>6</sup> This explains why any power market participant should devote time and attention to understanding this technology.

Credit for the development of the design of gas turbines goes to a number of inventors, including the German engineer F. Stolze (1872) and an American engineer, George Brayton (c1870). The term “Brayton cycle” is still used in the US to describe the basic engineering principles underlying a gas turbine. The first turbine that could run in a sustained way was built in Paris in 1903, and in the same year, in Norway, by Aegidus Elling. The first modern natural gas turbine was demonstrated in 1939 at the Swiss National Exhibition in Zurich. Over the next several decades, gas turbines were used primarily for aviation as jet engines. The potential of gas turbines for power generation was limited due to the limited availability of materials able to withstand high pressure and temperatures. The thermal efficiency of the turbines was low, generally below 20%. The first gas turbine for power generation, produced by General Electric, was installed in 1949 in Oklahoma in a combined cycle configuration, with capacity of 3.5 MW.<sup>7</sup> In the 1970s and 1980s, significant progress was made in their design, primarily through advances in material technology.

A gas turbine has two basic components: a compressor and turbine proper, which use the same single shaft (although some designs use separate shafts for each). The compressor is driven by the power generated by the turbine, which means that the turbine has to be very efficient. Air drawn from the atmosphere is compressed at 15–19 times the normal level and sent to the combustion chamber, where it is mixed with fuel. The fuel is usually natural gas, although other gases such as butane may be used as well. Liquids used in power generation include gasoline, kerosene and light diesel oil up to heavy residual oil (Bunker C or No. 6 fuel oil). Ignition causes the combustion of about one third of the air, with the temperature of the remaining increasing to as much as 1400° C. The technological challenges here include the effective mixing of the air and fuel, and controlling the process of combustion to minimise the quantity of NO<sub>x</sub> that is produced. This may be accomplished by controlling the combustion in such a way that the supply of oxygen is used in burning the fuel and is not reacting with nitrogen. Combustion chambers are usually separate from the turbines, although in some designs they may be placed around the turbine itself.

The turbines, which operate effectively as windmills, contain between three and five stages of blades that have to withstand very high pressures and temperatures. The gas that passes through the turbine and cause rotary movement can be either released into the atmosphere (open cycle) or re-circulated through a heat exchanger to capture and utilise the heat. The heat may be used for increasing the temperature of air exiting the compressor (a process known as *recuperation*). Some turbines are divided into sections called spools. Air exiting the first section may be passed through the second combustion stage to increase its temperature. A similar process in reverse (inter-cooling) takes place in a compressor: air is cooled between its two sections in order to reduce its volume (the efficiency of the compressor is increased, because cooler air occupies less space).

Turbines that use a humid air turbine (HAT) cycle have equipment that allows for water vapour injection into the air entering the gas combustion chamber. The water vapour increases the mass of gases delivered to the turbine and reduces the amount of energy used to operate the compressor.

The processes described above (reheating, water vapour injections, inter-cooling) used in combination may eventually increase the

efficiency of the turbines to as much as 60%, although this level of efficiency seems to be a remote possibility. Gas exiting the turbine is still very hot and this energy is lost. One solution is to capture this energy and use it for other industrial processes. An alternative is to use the combined cycle process.

Combined cycle power technology was critical to the construction boom of power plants in the late 1990s. In addition to relatively low capital costs and short construction cycles, combined cycle units offer efficiency approaching 60%. This is accomplished by capturing exhaust heat from a combustion turbine in a heat recovery steam generator (HRSG) and using it for the production of steam used to produce electricity. A typical design includes two gas turbines with a heat recovery unit connected to a single steam turbine. Such plants are typically the size of between 300–400 MW, reaching an efficiency of 57%, and potentially 60%.

A diagram of a typical arrangement for a combined cycle plant can be found in a document by D.L. Chase, P.T. Kehoe, "GE Combined-Cycle Product Line and Performance."<sup>8</sup> Two combustion turbines are connected to two separate HRSG units, which support one steam turbine. This plant may operate in different modes, and this greatly increases the level of complexity related to the dispatch decisions. The following configurations are possible in practice:

- CT 1 alone;
- CT 2 alone;
- CT 1 and CT 2 together;
- CT 1 and steam turbine;
- CT 2 and steam turbine; and
- CT 1 and 2 and steam turbine.

This means that an optimal dispatch requires careful evaluation of the costs and benefits of running the plant in different configurations. Another trade-off the operator has to consider is the production of electricity versus heat, if a given unit is running as a cogeneration plant.

It is worth noting that thermal power plants operate with an efficiency that has seasonal characteristics. The units are more efficient in the winter with lower ambient temperatures, when air is cooler and, therefore, denser, and when water used for cooling has a lower

temperature at the intake pipes. This is why gas turbines have summer and winter ratings.

### Nuclear power plants<sup>9</sup>

Nuclear power plants are in many ways similar to coal power plants: electricity is produced in practically all the commercial units by turbines driven by steam. The steam is produced through a controlled fission reaction. The nuclear fuel used is uranium-235. Uranium<sup>10</sup> is an element (chemical symbol U) with a nucleus containing 92 protons, with a number of isotopes – ie, atoms with a different number of neutrons.

Uranium exists in nature as uranium-238 (99.284%), uranium-235 (0.711%) and uranium-234 (0.0058%). Uranium-235 is a naturally occurring fissile isotope, a material capable of sustaining a chain reaction. In the chain reaction, when the nucleus of an atom of uranium-235 is hit by a neutron, it is transformed momentarily into uranium-236 and then into two smaller nuclei, releasing energy and more neutrons. The released neutrons hit other atoms of uranium, producing even more energy and free neutrons. If there is no material that can absorb a certain number of neutrons, an explosion may occur: this technology is used in nuclear bombs.

The worldwide production of uranium in 2010 amounted to 54,200 tons, of which 25% was mined in Canada. Other important uranium mining countries are Australia, Russia, Niger, Namibia, Kazakhstan, Uzbekistan, South Africa, the US and Portugal.

Uranium is mined using a number of different techniques, including open pit mining and leaching with special liquids. The source rock containing uranium is then processed, typically on-site, and – after passing through several stages – uranium yellow ( $\text{U}_3\text{O}_8$ ) cake is obtained. The next stage is the enrichment process, which is designed to increase the concentration of naturally fissile isotope U-235 to a level which can sustain chain reaction. Power plants require relatively low levels of concentration (typically 3–4%, no more than 5%), as opposed to much higher levels of concentration required for military weapons.<sup>11</sup> Enrichment techniques include the gaseous diffusion process, the gas centrifuge method and enrichment by laser manipulation. A typical nuclear plant with 1,000 MW capacity requires about 75 tons of enriched uranium to operate, and about a third of this amount has to be replaced on an annual basis. The

**Table 20.2** Production of uranium by country, 2010 (tons, metal content)

Country	2010
Czech Republic	259
Germany	8
Romania	77
Russia	3,562
Ukraine	850
Malawi	671
Namibia	4,965
Niger	4,198
South Africa	583
Canada	9,684
USA	1,628
Brazil	148
China	827
India	400
Kazakhstan	17,803
Pakistan	45
Uzbekistan	2,400
Australia	6,085
World total	54,200
World total U <sub>3</sub> O <sub>8</sub> equivalent	63,900

Source: "World Mineral Production 2006–2010," British Geological Survey 2012

production of uranium, given the growing interest in nuclear power, remains below demand, with the balance being provided through uranium from decommissioned weapons. A treaty between the US and Russia, and under which uranium from processed warheads is imported to the US by American industry, is due to expire in 2013, and the new contract will be negotiated with a strong bargaining position for the exporter. The fear of possible shortages of uranium has resulted in rapidly rising prices. This statement has to be qualified given the uncertainty regarding the future of nuclear power in a number of developed countries (Sweden, Germany, Switzerland) following the Fukushima disaster (2011) and the potential market impact of these developments.

The increasing interest in nuclear power and the anticipation of uranium shortages has led to the introduction of a futures contract for uranium. The specification of the uranium Nymex contract is shown in Table 20.3.

Practically all nuclear reactors rely on a number of similar components; what is different is what materials are used and how they are combined in a given design.

**Table 20.3** Nymex uranium contract specification

Trading unit	250 pounds of U <sub>3</sub> O <sub>8</sub>
Price quotation	US\$ per pound
Minimum price fluctuation	US\$0.05
Trading hours	The contracts are available for trading on the CME Globex and Nymex ClearPort electronic trading systems from 18.00 Sundays through 17.15 Fridays, Eastern Time, with a 45-minute break each day between 17.15 and 18.00.
Trading months	60 consecutive months
Last day of trading	Trading terminates at the close of business on the last Monday of the contract month. If the last Monday in the contract month is not a business day, trading shall terminate on the last business day prior to the Monday that is not a business day.
Settlement	Financial, based on the spot month-end U <sub>3</sub> O <sub>8</sub> price published by Ux Consulting Company, LLC.
Margin requirement	Margins required, subject to change
Trading symbol	UX

Source: [www.nymex.com](http://www.nymex.com)

Fuel is mostly uranium oxide (UO<sub>2</sub>), compressed into pellets kept inside tubes made of zirconium alloy.<sup>12</sup> Some reactors can run on natural uranium – ie, uranium composed of isotopes in proportions that occur in nature. Moderator (in most reactors it is light water, heavy water or graphite) slows down the neutrons released during the fission process. Neutrons with excessive velocity cannot sustain nuclear chain reaction.

The rate of the fission reaction is controlled through the control rods, which are made of neutron-absorbing materials that can be dropped into the core. Coolant is a medium used to transfer heat from the core of a reactor to the steam generator. In some reactors, the coolant and moderator may be the same. Coolants perform two functions: they help to maintain the temperature of a reactor at a stable level and to extract heat from the core of the reactor to produce steam. The most commonly used coolants are water or certain industrial gases (CO<sub>2</sub> and helium). The reactors using water as coolant are classified as pressurised water reactors (PWRs) or boiling water reactors (BWRs). In the PWRs, water is maintained under a high pressure that increases its boiling point. In the BWRs, there is no need for a separate steam generation unit. The core of the reactor is contained

inside a pressure vessel. In addition, most reactors have containment structures (reinforced concrete or steel structures) designed to prevent the escape of radiation or contaminated materials, and also to protect the reactor from intentional or accidental damage.

Historically, the technology used in the design of nuclear reactors for power generation evolved towards solutions that make operations safer and reduce the cost of construction. The first and second generations of nuclear power plants were designed using gases (in France and the UK) or water (in the US) as coolants. The first power plant in the UK, Calder Hall, was commissioned in 1956 and used Magnox technology and relied on pressurised CO<sub>2</sub> as a coolant and graphite rods as moderators; magnesium was used for fuel cladding. Magnox technology was also used by North Korea and inspired the French solution known as Uranium Naturel Graphite Gaz (UNG). Magnox technology was eventually superseded by technology known as advanced gas-cooled reactors (AGR) – which use steel instead of magnesium for fuel cladding. The second-generation water-cooled reactors, Canada Deuterium Uranium (CANDU), which were developed in the US and Canada, were eventually applied by most countries developing nuclear energy. Table 20.4 provides summary statistics about nuclear reactors in operation. Given the growth of interest in nuclear technology, several countries are pushing ahead with a number of more advanced fourth-generation nuclear reactors, which are at the design stage. Table 20.5 summarises different solutions that are being developed.

Fourth-generation reactors rely more on the natural forces of gravity, convection and compression, and have fewer moving parts that can fail. The safety features of these new types of reactors are very important, given the very strong public opposition to nuclear power.

Fast neutron reactors represent a special technology that has not yet been deployed for commercial applications. These reactors rely on fast neutrons that cause fission in the fuel rods using U-238 and U-235 as well as plutonium, and in the process can produce more plutonium from U-238. If this is the case, such reactors are called breeders. These reactors represent the technological frontier of the nuclear power industry.

In the US, nuclear energy is regulated by the Nuclear Regulatory Commission (NRC), which was established by the Energy Reorganization Act of 1974. The mission of NRC is:

**Table 20.4** Nuclear power reactors in operation (2011)

Country	In operation		Under construction	
	Number	Electricity net output MW	Number	Electricity net output MW
Argentina	2	935	1	692
Armenia	1	375	—	—
Belgium	7	5,927	—	—
Brazil	2	1,884	1	1,245
Bulgaria	2	1,906	2	1,906
Canada	18	12,569	—	—
China				
• Mainland	15	11,078	27	27,230
• Taiwan	6	4,982	2	2,600
Czech Republic	6	3,678	—	—
Finland	4	2,716	1	1,600
France	58	63,130	1	1,600
Germany	9	12,068	—	—
Hungary	4	1,889	—	—
India	20	4,391	6	4,194
Iran	1	915	—	—
Japan	50	44,215	2	2,650
Korea, Republic	21	18,698	5	5,560
Mexico	2	1,300	—	—
Netherlands	1	482	—	—
Pakistan	3	725	1	315
Romania	2	1,300	—	—
Russian Federation	32	22,693	11	9,153
Slovakian Republic	4	1,816	2	782
Slovenia	1	688	—	—
South Africa	2	1,800	—	—
Spain	8	7,567	—	—
Sweden	10	9,298	—	—
Switzerland	5	3,263	—	—
Ukraine	15	13,107	2	1,900
United Kingdom	19	10,137	—	—
USA	104	101,240	1	1,165
<b>Total</b>	<b>433</b>	<b>366,555</b>	<b>65</b>	<b>62,592</b>

Source: <http://www.euronuclear.org/info/encyclopedia/n/nuclear-power-plant-world-wide.htm>

To regulate the nation's civilian use of byproduct, source, and special nuclear materials to ensure adequate protection of public health and safety, to promote the common defense and security, and to protect the environment.

The NRC's regulatory mission covers three main areas:

**Table 20.5** Generation IV nuclear reactors

Design name	ACR	AP1000	EPR
Manufacturer	AECL	Westinghouse – BNFL	Framatome ANP
Country of origin	Canada	USA	France/Germany
Type	Pressurised heavy water (HWR)	Pressurised light water (PWR)	Pressurised light water (PWR)
Output	700 MWe	1117–1154 MWe	1750 MWe
Capacity	93%	93%	92%
Lifetime	60 years	60 years	60 years
Burn-up efficiency	21 GWd/te	21 GWd/te	> 60 GWd/te
Construction period	36 months	36 months	57 months
Current construction	–	–	Finland
Potential market	Japan/USA/Canada	USA/UK	US/China/Europe

Source: "An introduction to nuclear power – science, technology and UK policy context," Sustainable Development Commission, March 2006, [www.sd-commission.org.uk](http://www.sd-commission.org.uk)

- reactors* – commercial reactors for generating electric power and research and test reactors used for research, testing, and training;
- materials* – uses of nuclear materials in medical, industrial, and academic settings and facilities that produce nuclear fuel; and
- waste* – transportation, storage, and disposal of nuclear materials and waste, and decommissioning of nuclear facilities from service.

The NRC is a very important source of fundamental data about the nuclear industry. One of the most important reports is the "Current Power Reactor Status Report," available every morning from its website. This report shows the level at which each of the US-based nuclear power plants operate. The outages may be due to a scheduled maintenance (including refuelling) or a forced outage. One important task performed by the fundamental analysis group (in addition to the acquisition and dissemination of the report) is assessing the impact of a nuclear power outage. Depending on a number of factors (such as the nuclear power plant location, duration of the outage, availability of transmission, load levels), an outage may translate into incremental demand for power from alternative sources, including coal and natural gas-fired units. In turn, this triggers additional demand for fuel, transportation and transmission services with a potential for a significant market impact. Information about the duration of an outage may be quite important and the competitive intelligence firms always try to obtain such information from various public sources.<sup>13</sup>

## Wind power

Wind power is one of the most important renewable sources of electricity and also one of the most difficult to manage in the context of the entire power system. The paradox of wind power is that it is seen as an alternative to other sources of electricity that are characterised by negative externalities,<sup>14</sup> such as pollution, noise and landscape destruction. At the same time, wind power has significant unintended impacts on an integrated transmission–generation system, which are often overlooked by supporters of this type of generation.

Wind is a source of renewable energy that is ubiquitous and available in large amounts. According to one estimate, the total wind power potential that could be realistically harnessed is equal to 53,000 TWh/year, roughly twice the estimated world electricity demand in 2020.<sup>15</sup> The available resources by region are shown in Table 20.6.

Wind resources are difficult to estimate, as they depend not only on technology (the capacity of wind turbines has been steadily increasing), but also on the social perception of what are acceptable locations for clusters of wind turbines, known as the wind farms or wind parks (in Europe).

For a given wind turbine design, the amount of power generated depends on the wind speed measured at the turbine height.<sup>16</sup> The two quantities are linked through the following equation:

$$P = \frac{\rho}{2} c_p(\lambda, \theta) A_r v^3 \quad (20.1)$$

where  $\rho$  stands for density of air ( $\text{kg}/\text{m}^3$ ) and  $c_p$  is the power coefficient of the wind turbine. The coefficient  $\lambda$  is the tip speed ratio, the

**Table 20.6** Available wind resources

Region	Wind resources (TWh/year)
Western Europe	4,800
Eastern Europe and FSU	10,600
North America	14,800
Australia	3,000
Africa	10,600
Latin America	5,400
Asia	4,600
Total	53,000

Source: Paul Breeze, 2005, *Power Generation Technologies* (Amsterdam, the Netherlands: Elsevier).

ratio between the turbine blade tip speed  $v_t$  (m/sec) and the wind speed upstream from the rotor, denoted by  $v$ ,  $\theta$  is the blade pitch angle, and  $A_r$  is the swept area of the turbine rotor blades. Turbine designers control output by adjusting the blade angle,  $\theta$ , and the tip speed ratio,  $\lambda$ . The maximum power coefficient is 16/27, a number known as the Lanchester–Betz–Joukovsky (also spelled as Zhukovsky) limit.<sup>17</sup>

The wind speed is measured by sensors located at a known level above ground (no more than 10 metres), but much lower than a typical height of a wind turbine (70–120 metres). Weather forecasts provide wind speed prediction close to the ground; in order to estimate the speed of wind at the turbine top one can use the following formula to carry out conversion:

$$\mu(z_h) = \mu(z_s) + \sigma(z_s) \left( \ln\left(\frac{z_h}{z_s}\right) + 5 \frac{z_h - z_s}{L} \right) \quad (20.2)$$

where  $\mu(z_s)$  denotes average wind speed at the sensor level  $z_s$  (measured at the chosen time intervals – for example, 10 minutes),  $\sigma(z_s)$  is the corresponding standard deviation and  $L$  is the Monin–Obukhov length.<sup>18</sup> It is assumed that  $\sigma(z_s)$  is the estimate of the standard deviation of the wind speed at the sensor height. The parameter  $z_h$  is the level of the turbine hub.

Over offshore locations, the standard deviation used in equation 20.2 may not be available. In this case, one can use the approximation given by:

$$\sigma(z_h) = 2.5\mu_* \quad (20.3)$$

where  $\mu_*$  is the friction velocity calculated solving the equation below.

$$\mu(z_s) - 2.5u_* \left( \ln\left(\frac{z_s g}{K u_*^2}\right) + 5 \frac{z_s}{L} \right) = 0 \quad (20.4)$$

where  $K$  is the Charnock's constant for surface roughness of sea (0.011) and  $g$  is the gravitational acceleration constant.

Another consideration related to wind characteristics at a given location is turbulence. The ground surface is seldom a perfectly flat plain and the relief of the surface and vegetation may cause a significant amount of turbulence. Turbulence complicates the operation of a wind turbine and may contribute to material fatigue and reduce the useful life of the equipment. This problem can be mitigated by more

robust design of the mechanical components of a wind turbine and by placing it above the turbulent layer of the atmosphere. More turbulence at a given location translates into higher required height of a tower and a higher development cost.

#### *Wind turbines*

The amount of power produced by a wind turbine depends on its rated capacity and wind speed. A 1.5 MW turbine will generate about 1,000 MWh/year with wind blowing at 5.5 m/sec. At a speed of 8.5 (10.5) m/sec, the output increases to 4,500 (8,000) MWh.<sup>19</sup> The relationship between wind speed and power output is non-linear.<sup>20</sup> As explained above, for a given location, the output will depend on the height above ground at which a turbine is located.

Wind turbines can be classified using a number of criteria, such as type of generator:

- fixed speed with induction generator (type A);
- variable speed with variable rotor resistance (type B);
- variable speed with induction generator; and
- direct drive turbine with permanent magnet generator.

By location:

- onshore; and
- offshore.

By design:

- horizontal; and
- vertical.

The choice of a location for a wind farm is based on multiple considerations. Onshore locations have an advantage due to a lower cost of construction and typically lower cost of connecting wind farms to the grid. The offshore locations are much more expensive to build and connections to the grid are more challenging, but have a number of advantages. Offshore clusters of wind towers are usually hardly visible from the coast or are hidden behind the curvature of the earth, reducing opposition to the development of wind farms that threaten the pristine beauty of unspoiled sites. The wind blowing over the sea is faster and also less turbulent, as the surface is not as rough as in some land locations.

The wind turbines can be alternatively classified as horizontal or vertical. The distinction between horizontal and vertical turbines is based on the location of the rotor shaft. The horizontal turbines have a shaft that is placed horizontally and is usually connected to the generator through a gearbox. The rotor is composed of one or more blades, with a typical design based on three blades, rotating in a vertical plane. The speed at which the rotor turns depends on the wind speed and is much lower than the speed of a typical generator. This explains the reason behind the use of a gearbox, although some designs rely on a rotor connected directly to a generator. The generator and the gearbox are housed within a nacelle.

One potential design issue is speed control to prevent turbine damage due to excessive wind speed.<sup>21</sup> This may be accomplished through stall control or by changing the pitch of a turbine. A related problem arises from different wind speeds at different elevations. The blade at the top of its orbit is subjected to higher stress than the blade at the bottom. This causes stress that is transmitted to the other components of a turbine, with a potential for inflicting some damage on the equipment.

Vertical wind turbines have a shaft placed at a right angle with respect to the land or a rooftop, with the generator and a gearbox placed at the base. If a turbine is located at the ground, it loses the benefits of large wind speed at higher elevations. This is why such turbines are placed on the roofs of high-rise buildings. The most popular design of a vertical turbine is based on the model developed by a French engineer, G. J. M. Darrieus.

#### *Wind power: Controversy*

Wind power has many enthusiastic, not to say fanatical, supporters, who face a large number of sceptics. We shall cover briefly a number of highly controversial issues or areas that require more research and are often ignored by the profession. These include:

- ❑ wind power potential;
- ❑ externalities of wind farms;
- ❑ the related issue of environmental benefits of wind farms; and
- ❑ integration of wind farms into the grid.

*Potential.* Wind is seen often as an inexhaustible, albeit unpredictable, source of energy that could be a solution to future energy challenges – provided the problem of storage can be successfully addressed. However, one can question the ultimate potential for wind power. There is a difference between the amount of energy contained in the atmosphere and the amount of energy that can be effectively harnessed. Fortunately, there are some estimates available in the literature that the reader is encouraged to review.<sup>22</sup>

The externalities of wind farms are well documented, known to the public and not much can be done about them: wind farms are eyesores, are associated with unpleasant noise pollution and kill birds flying into the rotors.<sup>23</sup> Another well-known fact is the wake effect: the impact of the first row of the turbines on the turbines located in the back of a farm.<sup>24</sup> Rotating blades cause local turbulence that affects the efficiency of other turbines. The same is true of wind farms located in close proximity (5–10 miles).<sup>25</sup> A technical discussion of this topic is beyond the scope of this book, but any practitioner in this area has to recognise the magnitude of modelling and legal challenges in siting wind farms due to this effect. Hopefully, understanding the effect will lead to a better spacing and positioning of wind mills, as well as their greater efficiency. There are many externalities of wind power that are often overlooked by its supporters and which can explain why the estimates of costs of wind power are so drastically different from one study to another.

A more serious problem arises from integrating wind farms into the power system. The intermittency of wind means that it can be predicted only with a margin of error; it may materialise and/or die off suddenly, and often blows mostly at night when demand is low. A sudden jump in wind power output means that some thermal plants have to be backed-off. Depending on the size and flexibility of the thermal generation fleet, this may be relatively expensive. This is especially true at night when the power plants usually operating are either combined cycle gas-fired power plants or coal plants which are characterised by lower flexibility compared to combustion turbines. Reducing output from such plants reduces their thermal efficiency and, if this is done repeatedly, increases maintenance costs. Sudden drops in wind power levels create similar problems related to bringing additional thermal units on line and/or lowering load by

disconnecting certain customers. The industry old-timers are aware of many cases of close calls related to the fluctuation of output from wind farms.<sup>26</sup> The assessment of the full social costs related to the intermittency and limited predictability of the wind power is a difficult task and requires running a complicated simulation model, combining stochastic load, weather conditions, transmission grid status, fuel prices and other variables. Such studies are quite difficult and expensive and very sensitive to the underlying assumptions.

The issue discussed above is closely related to the environmental benefits of wind farms, which are associated primarily with the potential for reducing emissions of CO<sub>2</sub>, SO<sub>x</sub> and NO<sub>x</sub>. This position has been challenged in a number of studies, and is currently one of the most controversial issues in the renewable energy debates. The central issue is the extent of emission reductions due to expansion of electricity output from wind farms, given the complex issues of integration of this source of electricity with the rest of the generation/transmission system. A recent study by BentekEnergy concluded:<sup>27</sup>

The results of this study suggest that wind energy constitutes a significant paradox: Generation of power from wind, per se, yields no emissions. However, integration of wind power into a number of complex utility systems has led to little or no emissions reductions on those systems, and has significantly increased costs to power producers, grid operators and electricity consumers.

This paradox is due to intermittency of wind, which requires thermal generation back-up, and the lower efficiency of steam and natural gas turbines which are required to cycle (ie, adjust output at very short notice) to react to changes in wind farm operations. Turbines operating at sub-optimal conditions (with respect to temperatures and speed) may produce more emissions than turbines operating at a steady state.<sup>28</sup> This is equivalent to mileage degradation in a car that is driven over a hilly terrain at variable speed.<sup>29</sup> Studies offering a different point of view are focused on the potential for reducing back-up generation through a better design of the overall system.<sup>30</sup>

*Reactive power* is another issue. Many wind turbines do not produce reactive power (a term previously explained in Chapter 19) and this means that the entity (an independent system operator, ISO, or a

local utility) responsible for managing the grid has to procure sufficient supply of reactive power. This can be done by installing devices producing reactive power or buying it from generators. This problem is magnified by the location of many wind farms at remote locations, usually at the end of long transmission lines, where the availability of reactive power sources is limited.

### **Hydropower**

Hydropower is one of the most important sources of electricity and the most widely used sources of renewable energy. The popularity of hydropower plants is matched by the objections raised by critics who point out the environmental and social consequences of massive projects undertaken in different parts of the world. One example of such a controversy is the construction of the Three Gorges Dam in China, which resulted in the flooding of large areas, relocation of the local population and serious environmental impact.

Hydropower is a broad term that captures many different technologies for producing power using energy of rivers, waves, ocean tides or energy stored in water accumulated behind a dam.

Water turbines were discovered in antiquity and have been used to power water mills for centuries. Modern turbines follow two basic engineering designs, known as impulse turbines and reaction turbines. The most frequently used impulse turbines are known as Pelton and Turgo. The term impulse is derived from the use of the principle of Newton's second law to extract energy from a jet of water projected from a nozzle on buckets that are mounted on the edge of the wheel causing its rotary motion. These turbines have to rotate in the air; their efficiency is close to 95%. The Turgot turbines look like Pelton wheels sliced in half. Both turbines are used for water heads above 450 metres, though Pelton turbines can be designed for smaller heads.

Reaction turbines are used for heads of water below 450 metres. Most such turbines are classified as either Francis or Kaplan designs. Francis turbines are known as mixed-flow turbines because water changes direction as it passes through the turbine. It enters the turbine in the radial direction and exits in the axial direction. The term reaction is justified by the fact that the water changes pressure as it passes through the turbine. Kaplan turbine uses a propeller design with a variable angle blade.

*Run-of-river* power plants are not only technologically the least complex and cheapest projects, but also they are relatively less controversial as they are associated with a more limited environmental impact. A run-of-river project often does not need a dam and a reservoir. Water is diverted into a pipeline called a penstock that feeds a turbine. In case a dam across a river is required, the flooding is limited. Also, the water, unless used for irrigation or consumption, is returned to the river downstream of a turbine. In some cases, the structure may be hardly visible. The disadvantage of this technology is that it requires a relatively steady water flow and the availability of potential sites may be limited.

*Reservoir* projects require the construction of dams across rivers and the creation of artificial lakes behind them. The creation of a lake may result in serious social and environmental disruptions, adverse visual impact, destruction of species and cultural sites and is often very controversial. Another reason why a potential location has to be carefully examined are the soil and rock characteristics, and the likelihood of a potential disaster (such as landslides or man-made earthquakes) should the structure fail. The construction of a dam and creation of an artificial lake gives the power plant operator a degree of control over the operations of the plant. On the other hand, the flexibility of a plant is constrained by a number of considerations related to other water uses, such as irrigation, fishing and tourism. Other constraints are related to fish migration and the preservation of species.

There are three types of dams: concrete straight dams, concrete arch dams and embankment dams.<sup>31</sup> The first two types are massive concrete structures, located often in a ravine whose sides provide a support. An arch dam points upstream in the same way an arch of a gothic church points upward. The bow of the dam distributes the weight of the water behind the dam across the structure, making it very strong and reducing its required thickness.

A typical reservoir hydropower plant uses water stored behind a dam built across a river, with water accumulating in what is called a forebay.<sup>32</sup> The water is released through the pipe known as penstock and spins a turbine, which is connected to a generator.

Hydropower plants using dams convert potential energy of water into kinetic energy. The amount of potential energy is equal to the

gravitational force multiplied by the height (distance from the surface of the earth). The potential energy of a volume of mass of water M kilograms that falls through a height of H metres can be calculated by using the SI system (International System) of units.<sup>33</sup>

$$\text{Potential energy (joules)} = M \times H \times G \quad (20.6)$$

using gravitational acceleration constant,  $G = 9.80 \text{ m/sec}^2$ .

Power is energy produced per unit of time (T).

$$\text{Power} = (M/T) \times H \times G$$

For one litre of water the mass is equal to one kilogram, and if  $H = 90$  metres with a stream flow of 882 litres/second, we get:

$$\text{Power} = 882 \text{ kg/sec} \times 9.80 \text{ m/sec}^2 \times 90 \text{ m} = 79,380 \text{ joules/sec} \quad (20.7)$$

or 79,380 watts. A watt, in the SI units, is one joule per second. As a reminder, Joule =  $\text{kg} \times \text{m}^2 / \text{sec}^2$ . The analysis does not account for the efficiency of a generation unit.

### Solar energy

Conversion of solar energy into electricity is based on two radically different techniques: thermal and photovoltaic (PV).

Thermal solar electricity is produced through concentrated solar power (CSP) technology used to collect sun radiation through arrays of mirrors (flat or parabolic) or lenses to reach temperatures at which electricity generation becomes feasible. Different media are used for heat transfer, ranging from steam and gas to melted fluoride salt.<sup>34</sup> Given that the position of the sun changes during the day and depends on the season, a tracking system adjusting the position of lenses and mirrors is practically mandatory, increasing the cost and complexity of thermal plants. The advantage of thermal solar electricity is that it can be produced at night as energy can be stored during the day in a medium such as molten salt.

PV cells convert sunlight directly into electricity.<sup>35</sup> Cells are made of semiconductor materials, such as silicon, compressed into a wafer. When sunlight hits the cell, electrons are knocked out of their orbits and, if conductors are attached to both sides of the cell, a circuit is formed and electric current will flow. The cells may be designed as a single junction or multiple junction devices. Multiple junction PV cells are made of layers of different materials that absorb photons of

different levels of energy. The cells are organised into modules, which are organised into arrays. Solar cells produce direct current that can be converted into AC current using a device called an inverter.

PV installations can be classified as behind-the-meter systems or utility-level systems. As has been explained in a report, “behind-the-meter PV refers to systems that are connected on the customer side of the meter, typically under a net metering arrangement. Conversely, utility-level PV consists of systems connected directly to the utility system, and may therefore include wholesale distributed generation projects.”<sup>36</sup> The costs of both systems are falling very fast as documented in the cited report from the Lawrence Berkley Lab.

- ❑ The capacity-weighted average installed cost of all behind-the-meter systems installed in 2010 – in terms of real 2010 dollars per installed watt (DC-STC) and prior to receipt of any direct financial incentives or tax credits – was US\$6.2/Watt, and was US\$1.3/W (17%) below the average for systems installed in 2009.<sup>37</sup>
- ❑ Among the 20 utility-sector projects in the data sample completed in 2010, installed costs ranged from US\$2.9/W to US\$7.4/W.

In spite of unquestionable technological progress and falling costs, solar electricity still remains much more expensive than traditional sources of electricity. Table 20.7 lists the estimated levelised cost of electricity (LCOE) for several different energy sources. The table is based on a paper by Ken Zweibel.<sup>38</sup>

Solar energy has unquestionable potential, as documented by the frequently replicated illustration showing the areas that could replace the entire world production of electricity at 8% efficiency factor. As explained by Mathias Loster:

**Table 20.7** Year-one annualised cost of electricity

Energy plant type	Lifetime cost ¢ per Kwh
Baseload natural gas	6.4
Wind turbine	7.5
Conventional coal	8.0
Advanced nuclear	10.0
Solar PV	15.0
CCS	12.0

*Source:* Ken Zweibel, 2010, “Should solar photovoltaics be deployed sooner because of long operating life at low, predictable cost?” *Energy Policy*, p 5.

Sunlight hitting the dark discs could power the whole world: If installed in areas marked by the six discs in the map, solar cells with a conversion efficiency of only 8% would produce, on average, 18 TW electrical power. That is more than the total power currently available from all our primary energy sources, including coal, oil, gas, nuclear, and hydro. The colors show a three-year average of solar irradiance, including nights and cloud coverage.

Of course, the obstacles to harnessing this potential are the cost (compared to traditional energy sources) and transmission. One cannot also expect that most nations would choose to satisfy their energy needs from a few sources located in countries which may be, or may not be, reliable, and which could exploit their bargaining power as a few suppliers of energy to the rest of the planet. Conventional sources of energy will be around for a long time.

## CONCLUSIONS

The system of generation plants within given interconnections is a technological marvel. Generation units are not only technologically very complex devices: they have to operate as one integrated, synchronised machine. Outputs of electricity and demand have to be exactly balanced second by second and this creates a number of unique challenges for system operators. An equally difficult task is planning expansion of the generation fleet. Understanding generation trends, the logic behind plant dispatch decisions and flexibilities embedded in different generation units is a precondition of coming up with reliable price forecasts and trading decisions.

- 1 The operations of the power plants and the unit dispatch decisions may look sometimes to be counterintuitive, unless the constraints related to coal deliveries are understood. When demand for power drops (due to mild weather or economic recession), power plants continue to burn coal that arrives on a daily basis according to long-term schedules. Electricity from the coal plants is occasionally at the margin, although this is not justified by relative fuel prices. In the late 1990s, during disruptions following the merger of two railways, the use of natural gas for power generation increased during the off-peak period, as a way of saving coal. This is further proof that familiarity with even minor details of physical asset operations is critical to energy trading. Logistics of coal deliveries is a factor limiting a trend triggered by low prices of natural gas: the substitution at the margin of natural gas for coal in some parts of the US.
- 2 One example was a traffic jam in China. "China is driving up world coal prices as clogged roads and railways from Beijing to Tibet restrict deliveries in the world's fastest-growing major economy while the country tries to build stockpiles ahead of winter. A jam held up traffic for as many as 10 days along the country's main east-west highway in August, underscoring a crisis that may buoy prices for the next two years, according to Daniel Brebner, an

analyst at Deutsche Bank AG in London. The China Coal Transport and Distribution Association say it may take up to four years to ease the gridlock." See <http://www.bloomberg.com/news/2010-10-27/coal-climbs-to-four-month-high-as-china-builds-stockpiles-energy-markets.html>.

- 3 Coal ash is one of the biggest – and at the same time unrecognised – environmental hazards related to the burning of coal, as demonstrated by an industrial accident in Tennessee in the early 2010s. See the following article from the *New York Times*: Shaila Dewan, 2009, "Huge coal ash spills contaminating US water," *New York Times*, January 7.
- 4 A flue is a conduit, such as a pipe or a chimney, for removing exhaust gases from a boiler or a furnace.
- 5 HHV stands for higher heating value. "The difference between LCV and HCV (or Lower and Higher Heating Value, or Net and Gross) [...] is the latent heat of condensation of the water vapour in the combustion exhaust gas, which in turn depends on the hydrogen content of the fuel being burned. The difference is minimal for coal, significant for natural gas and largest for pure hydrogen fuel." See <http://www.claverton-energy.com/>.
- 6 Low prices of natural gas in the US changed the dispatch order to some extent in some parts of the US (primarily SERC), putting coal plants at the margin during certain months.
- 7 Denise Lane "Brayton cycle: The ideal cycle for gas-turbine engines in relation to power plants" (<http://www.scribd.com/doc/37415474/Brayton-Cycle>).
- 8 D.L. Chase, P.T. Kehoe, "GE Combined-Cycle Product Line and Performance," GE Power Systems Schenectady, NY. <http://physics.oregonstate.edu/~hetheriw/energy/topics/doc/elec/natgas/cc/combined%20cycle%20product%20line%20and%20performance%20GER3574g.pdf>
- 9 This section is based primarily on the report "An introduction to nuclear power – science, technology and UK policy context," Sustainable Development Commission, March 2006.
- 10 Based on <http://en.wikipedia.org/wiki/Uranium>.
- 11 The nuclear non-proliferation treaty allows signatory countries to enrich uranium for peaceful purposes. The objections to the development of this potential by Iran are dictated by the fears that once the technology for low-level uranium enrichment is developed, it can be used to develop weapon grade uranium (given sufficient time and investment in equipment).
- 12 Zirconium is a chemical element with the symbol Zr and atomic number 40. It is highly corrosive-resistant and has low neutron-capture properties and, therefore, does not interfere with controlled fission reactions (see <http://en.wikipedia.org/wiki/Zirconium>).
- 13 The formal definition for "externality" was introduced by Tibor Scitovsky in 1954, although the concept was discussed in the economic literature for a long time prior to that (including by such economists as Arthur Cecil Pigou and Friedrich Hayek). Externality is an unintended impact of an economic activity on a party that is not involved in market transactions related to this activity. Thus impact may be positive or negative. Scitovsky's formal definition is based on inclusion in the production (utility) functions of producers (consumers) of the arguments corresponding to the levels of production (consumption) by other economic agents. Externalities result in the decoupling of social and private costs and benefits. Flowers planted by our neighbour benefit us, but we are not compensating our neighbour for his good efforts (except for an occasional compliment). Tibor de Scitovsky, 1954, "Two concepts of external economies," *Journal of Political Economy*, pp 43–51.
- 14 Paul Breeze, 2005, *Power Generation Technologies* (Amsterdam, the Netherlands: Elsevier).
- 15 The discussion here is based on Bart Christian Ummels, 2009, "Power system operation with large-scale wind power in liberalised environments," PhD Dissertation, Technische Universiteit Delft.
- 16 The maximum efficiency of an ideal wind turbine rotor is well known as the "Betz limit", named after the German scientist Albert Betz – who formulated it in 1920, the same year as

the Russian aerodynamic engineer Nikolay Zhukovsky. A British scientist, Frederick Lanchester, had already derived the same maximum in 1915.

- 17 The Monin–Obukhov length is defined as the height at which turbulence is generated more by buoyancy than by wind shear. Typical length is between 1 and 50 metres. An on-line calculator is available at [http://www.shodor.org/os411/courses/\\_master/tools/calculators/moninobukhov/index.html](http://www.shodor.org/os411/courses/_master/tools/calculators/moninobukhov/index.html).
- 18 Paul Breeze, op.cit.
- 19 B. C. Ummels op. cit., p 32.
- 20 One can learn about the consequences of an excessive wind speed by watching the following videos: <http://www.youtube.com/watch?v=CqEccgR0q-o> or <http://www.youtube.com/watch?v=MWjKUr5j5p8>.
- 21 Carlos de Castro, Margarita Mediavilla, Luis Javier Miguel and Fernando Frechoso, "Global wind power potential: Physical and technological limits", *Energy Policy*. See also the summary available at <http://www.theoildrum.com/node/8322#more>.
- 22 There is some evidence that the birds learn to co-exist with wind turbines and avoid collisions.
- 23 One can see an illustration of the wake effect (interactions between wind turbines) at <http://withoutthatair.blogspot.com/2010/01/wind-farm-wakes.html>.
- 24 An example of a study of such effects can be found in Sten Frandsen, Rebecca Barthelmie, Ole Rathmann, Hans E. Jørgensen, Jake Badger, Kurt Hansen, Søren Ott, Pierre-Elouan Rethore, Søren E. Larsen and Leo E. Jensen, 2007, "Summary report: The shadow effect of large wind farms: measurements, data analysis and modeling," Risø National Laboratory, Technical University of Denmark, Roskilde, Denmark, October.
- 25 One such event happened in Texas on February, 2008. "A drop in wind generation late on Tuesday, coupled with colder weather, triggered an electric emergency that caused the Texas grid operator to cut service to some large customers, the grid agency said on Wednesday. ERCOT said a decline in wind energy production in West Texas occurred at the same time evening electric demand was building as colder temperatures moved into the state [...] System operators curtailed power to interruptible customers to shave 1,100 megawatts of demand within 10 minutes, ERCOT said. Interruptible customers are generally large industrial customers who are paid to reduce power use when emergencies occur." See <http://www.reuters.com/article/2008/02/28/us-utilities-ercot-wind-idUSN27495229> 20080228.
- 26 "The Wind Power Paradox," BENTEK Energy, July 19, 2011. The author was a consultant to BENTEK Energy.
- 27 An example of the study of this effect and counter arguments can be found at: Warren Katzenstein, Jay Apt: "Air Emissions Due to Wind and Solar Power," Environmental Science Technology, 2009, vol. 43(2), page 253–258; "Air Emissions Due To Wind and Solar Power Supporting Information," [http://www.sustainable.gatech.edu/sustspeak/apt\\_papers/51%20Air%20Emissions%20Due%20To%20Wind%20And%20Solar%20Power.pdf](http://www.sustainable.gatech.edu/sustspeak/apt_papers/51%20Air%20Emissions%20Due%20To%20Wind%20And%20Solar%20Power.pdf); Andrew Mills, Ryan Wiser, Michael Milligan, Mark O'Malley, "<http://nrepubs.nrel.gov/Webtop/ws/nich/www/public/Record;jsessionid=0743E7B158BC7C16D7C4B145F4A34A92?rpp=25&upp=0&cm=85&w=NATIVE%28%27AUTHOR+ph+words+%27%27milligan%27%27%27%29&order=native%28%27pubyear%2FDescend%27%29>" "Comment on 'Air Emissions Due to Wind and Solar Power,'" Environmental Science and Technology, 2009, vol. 43(15), 2009, page 6106–6107.
- 28 A study based on Irish experience reached similar conclusions: it "was found that wind generation could be used as a tool for reducing CO<sub>2</sub> emissions but alone, it was not effective in curbing SO<sub>2</sub> and NO<sub>x</sub> emissions." See Eleanor Denny and Mark O'Malley, 2006, "Wind generation, power system operation, and emissions reduction," *IEEE Transactions On Power Systems*, (21)1, February.
- 29 See, for example, D. Jacobson and Virginia C. High, 2008, "Wind energy and air emission reduction benefits: A primer," *Subcontract Report NREL/SR-500-42616* February.

- 30 Paul Breeze, 2005, *Power Generation Technologies* (Amsterdam, the Netherlands: Elsevier).
- 31 Brian K. Edwards, 2003, *The Economics of Hydropower* (Cheltenham, England: Edward Elgar Publishing).
- 32 <http://esa21.kennesaw.edu/activities/hydroelectric/hydroactivity.pdf>.
- 33 Information about the only solar thermal energy plant operating in the US is available from [http://esolar.com/our\\_projects/](http://esolar.com/our_projects/).
- 34 <http://science.nasa.gov/science-news/science-at-nasa/2002/solarcells/>.
- 35 Galen Barbose, Nâim Darghouth, Ryan Wiser and Joachim Seel, 2011, "Tracking the Sun IV. An historical summary of the installed cost of photovoltaics in the United States from 1998 to 2010," Lawrence Berkeley National Laboratory, September.
- 36 General Electric reported that it hoped to bring this cost down to US\$3/watt (see <http://www.earthtechling.com/2011/11/pv-cost-down-to-3-per-watt-is-ge-goal/>).
- 37 Ken Zweibel, 2010, "Should solar photovoltaics be deployed sooner because of long operating life at low, predictable cost?" *Energy Policy*.
- 38 Source: [http://www.ez2c.de/ml/solar\\_land\\_area/](http://www.ez2c.de/ml/solar_land_area/).



## *Transmission, Loads and Power Pools*

This chapter will cover the transmission of electricity and review the characteristics of different loads.<sup>1</sup> The last section will look at the design principles of power pools, the entities responsible for coordinating the activities of generators, transporters and consumers of electricity. The production and consumption of electricity have to be balanced at each point in time (electricity is not storable from a system point of view), and understanding how this happens is critical not only for physical electricity traders but also for investors in physical assets and electricity derivative traders. This is not an easy task, which explains why electricity trading is an occupation for the best and the brightest, as well as for those with a lot of courage. Integrated generation-transmission systems, very complex and difficult to understand from the technological point of view, are managed by organisations operating within a system of rules which are sometimes inconsistent and often contradictory. The great economist Friedrich Hayek once said that the market form of economic organisation is the result of human action but not of human design. This is not true, unfortunately, of the modern electricity markets. Power pools are the product of efforts of an army of industry professionals with backgrounds in many technical disciplines beyond the reach of mere mortals. Most relevant documents are written in a language that is full of technical jargon, with references to other (equally obscure) documents. There is, however, a good reason to invest time in developing an understanding of the machinery of power markets. The trading results may be spectacular. The author still has recollections of electricity traders in the late 1990s poring over the documents of different power pools, looking for possible inconsistencies, creating potential for profitable trades.

Another good reason to study power markets, even if one never trades a single megawatt-hour, is the critical impact they have on the natural gas and coal markets. The demand for natural gas from peaking power units, and gas-on-coal competition as generation fuels, introduce the most uncertainty into these two markets. Any coal or natural gas trader has to develop a deep understanding of the electricity industry. The days of isolated markets for different energy commodities are over.

This chapter is the most important part of the section devoted to electricity markets. It explains the organisation of power pools and the concept of locational marginal prices (LMPs), the cornerstones of many modern electricity markets. It will also explore the alternative model of organisation of electricity markets, used in Europe, which has evolved around a system of linked power exchanges. The discussion of loads is crucial for a number of reasons. A forecast of electric load is a very important component of models used to forecast electricity prices. It is also a feed to models used in predicting natural gas consumption (and, by extension, storage levels) and demand for other fuels (such as coal) used in power generation.

## TRANSMISSION AND DISTRIBUTION

The system of underground and above ground lines, and associated transformers, monitoring and safety devices, used to deliver electricity to end users, is known as the transmission and distribution (T&D) system. Transmission is the term reserved for high-voltage lines used to wheel electricity over long distances. Lower voltage transmission lines (under 100 kV) are often referred to as the sub-transmission system. The distribution system, connecting end users to the grid, is divided sometimes into primary (high-volume industrial and commercial users connected directly to the transmission system) and secondary (households and small commercials). The demarcation lines between the transmission and distribution systems may be sometimes quite fuzzy, especially if both are operated by the same company. The point of interface between the two is a distribution substation.

The transmission system can be compared to the US interstate highway system or the trunk pipeline grid for natural gas. The transmission system is composed of high-voltage lines, with the cut-off voltage around 40–115 kV, depending on the source.<sup>2</sup> Typical volt-

ages used in the US power transmission grid are equal to 69, 115, 128, 230, 345, 500 and 765 kV. The starting point of a transmission line is a switchyard outside a power plant, with a step-up transformer increasing voltage to the level corresponding to the technical characteristics of a transmission line. At the end of a transmission line, there is a substation with a step-down transformer, which reduces voltage to the levels at which power can be injected into the local distribution systems. Some big energy users, such as large industrial plants or military bases, can be connected directly to the transmission system through their own dedicated substations.

The historical trend in transmission is towards the development of increasingly bigger interconnections. In the US, there are three interconnections, which effectively operate as three big machines. These interconnections are subdivided into reliability councils coordinated by the North American Electric Reliability Corporation (NERC). The NERC was established in response to the first serious blackout in the US in 1965.<sup>3</sup> In Europe, most member countries of the EU rely on the integrated, synchronised system.

All the power plants within an interconnection are synchronised and operate within very narrow frequency limits. Bigger interconnections produce significant economic benefits – they allow for the pooling of diverse generation and load resources, and make it possible to lower (in percentage terms) reserve margins while preserving the same level of reliability. For example, an area with a surplus of generation could provide a fallback in case of outages in a different region. It is also possible to dispatch the least expensive resources across the entire interconnection before moving up the supply stack to more expensive sources. In a deregulated, market-

**Table 21.1** High-voltage transmission lines in the US

Operating (kV)	FRCC	MRO	NPCC	RFC	SERC	SPP	TRE	WECC	Contiguous U.S.
100–199	–	–	–	–	–	–	–	–	–
200–299	5,922	7,241	1,521	6,949	21,100	2,776	–	36,810	82,319
300–399	–	11,468	5,064	13,610	3,538	4,934	9,500	10,301	58,415
400–599	1,201	473	–	2,551	8,617	47	–	12,729	25,618
600+	–	–	190	2,226	–	–	–	–	2,416
	<b>7,123</b>	<b>19,182</b>	<b>6,774</b>	<b>25,336</b>	<b>33,255</b>	<b>7,757</b>	<b>9,500</b>	<b>59,840</b>	<b>168,768</b>

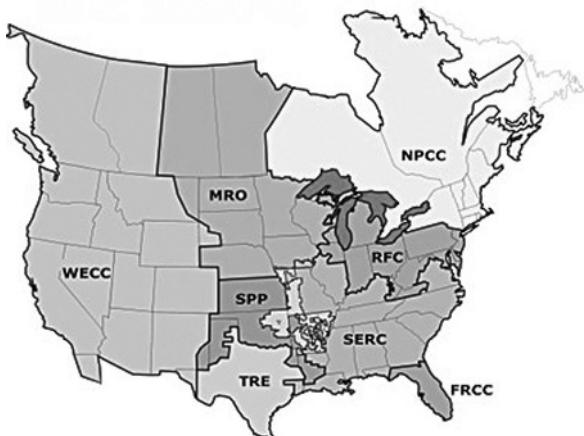
Source: U.S. Energy Information Administration

based system, access to transmission is a critical precondition of electricity trading.

Understanding transmission and the tracking of transmission-related developments is critical to any energy trading organisation. The availability of transmission is critical to any transaction involving the physical wheeling of energy or dependent on the level of prices at two different locations. This is also true for a number of reasons beyond any specific deals, including security, flexibility and the efficiency of electricity supplies. Transmission is also critical to the long-term viability of the industry model based on multiple types of generation ownership (utilities, independents, Federal agencies, municipalities).

Security of supply has two aspects: adequacy and security. Adequacy is the ability to mobilise resources to meet customer demand during peak times. Security is the ability of the system to survive a sudden shock, a loss of one or more important components – such as a big generation plant or an important transmission line. This is referred to sometimes as N-1 (N-2) contingency, withstanding the loss of one (two) critical component(s) of the system. Efficient, well-designed transmission allows for the pooling of diverse resources over a larger area and improves the chances of surviving a sudden shock. However, this statement will be seen by many industry analysts as controversial. Expanding the size of the interconnection may facilitate propagation of shocks, leading in extreme cases to a collapse of the system. An example of such a collapse was the blackout of 2003,<sup>4</sup> which left many millions of American and Canadians without electricity for several days, and has provided a counterargument against increasing ties between different regional power systems. Some transmission experts believe that a reliance on a smaller, isolated system, or systems, that could be quickly disconnected, would allow problems to be isolated and would localise potential system-wide outages.

Following that blackout, many industry observers pointed out the adverse impact of energy trading on energy security. Energy transactions may be executed at very short notice and produce significant shifts in electricity flows across the transmission system. Unexpected developments may overwhelm the system operator's ability to react to the developments.

**Figure 21.1** NERC Regional Entities

Source: <http://www.nerc.com/page.php?cid=1|9|119>

Legend:

- Florida Reliability Coordinating Council (FRCC)
- Midwest Reliability Organization (MRO)
- Northeast Power Coordinating Council (NPCC)
- Reliability First Corporation (RFC)
- SERC Reliability Corporation (SERC)
- Southwest Power Pool, RE (SPP)
- Texas Reliability Entity (TRE)
- Western Electricity Coordinating Council (WECC)

### The loop flow problem

One of the technical issues related to transmission is loop flow. Loop flow occurs when the actual flows of power diverge from the path on which the flow is scheduled. This happens because power flows (between two points connected by more than one physical path) are distributed according to the laws of physics and not contractual arrangements. Loop flows are an important problem from the point of view of market design and the bottom lines of many market participants, not just a curiosity for physics aficionados.

In the US markets, one of the best-documented cases of loop flow is that of transmission problems around Lake Erie. This phenomenon has been summarised as follows:<sup>5</sup>

For years, electricity flow around Lake Erie has been split, with much of the power going in a counterclockwise direction. But sometime in 2007 – no one is sure exactly when – that began to change. More electricity started flowing in a clockwise direction. Clockwise flow of

electricity around Lake Erie caused congestion on the lines of the New York Independent System Operator (NYISO), the state's grid operator. Because electricity flows along the path of least resistance, it doesn't always go where it has been scheduled. Although several NYISO market participants had scheduled power along transmission line paths running clockwise around the lake, about 80 percent of that power actually flowed directly from New York to the PJM Interconnection, which operates a transmission system covering parts of 13 states from Illinois to New Jersey. This unscheduled flow of electricity led to higher "uplift" costs – the costs incurred to relieve congestion within a zone and imposed on all grid customers – and, ultimately, higher electricity prices for consumers.

The case was the subject of extensive investigations by NYISO and FERC. The final conclusion was that this was due not to manipulation, but was caused by the actions of market participants reacting in a rational way to price signals. The modification of the market rules addressed the problem to some extent.

The local distribution systems are based on voltages equal to 110V (US) or 220V (most of the world). The design of the distribution systems used in the US has many implications for the evolution of the power industry and may create impediments to many highly innovative services. One of these is high-speed broadband delivered over power lines which connect practically every American home to the grid. The difficulty is that the point of interface between the broadband system and the distribution network has to be located behind the lowest-level transformers. Given the voltage used for residential customers in the US, the number of such transformers is relatively large and developing broadband connections is expensive.<sup>6</sup> European countries, using 220V distribution systems, have a relative advantage over the US, as the lowest-level transformers serve a larger number of end users.

## ELECTRIC LOADS

An electrical engineer looks at loads from the point of view of their physical characteristics and distinguishes between resistive and inductive loads, as well as capacitive loads. These terms will soon become clear. Power traders, and especially power marketers, look at loads from the point of view of their level, and statistical properties such as variability, seasonality and evolution over time. Above all, a power marketer is interested in the ability and willingness of loads to pay for reliable supply of electricity. The buyers of electricity pay

attention to the energy they consume over a period of time, measured in kWh or MWh. The electricity suppliers have to pay attention to the rate at which energy is consumed at a specific point in time (measured in watts or megawatts), the critical variable from the point of view of system reliability. A short-term spike in energy consumption may overload the system and cause serious disruptions.<sup>7</sup> One of the critical tasks performed by fundamental analysts supporting energy traders is the prediction of the load at the system level (for example, for ERCOT or the Pennsylvania–New Jersey–Maryland, PJM) or by a retail customer or a group of retail customers.

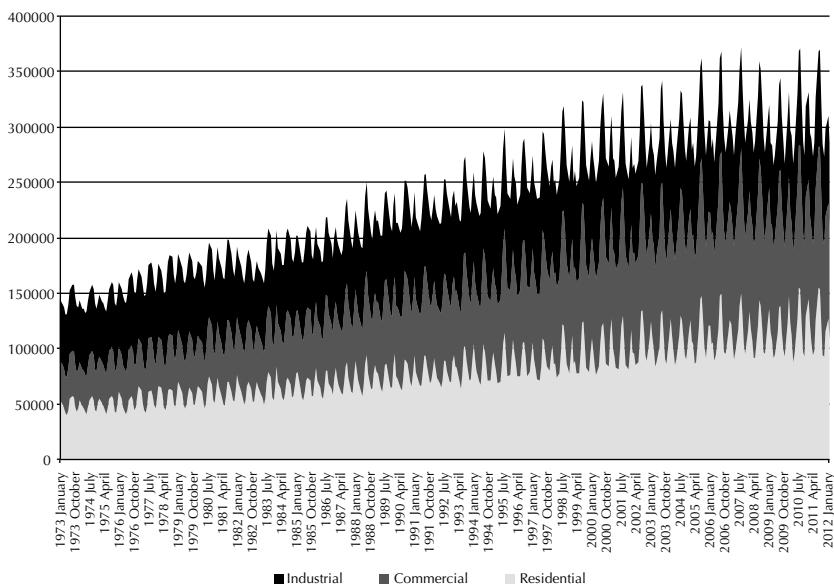
The term “loads” is used in the power industry in different contexts and may mean either specific physical devices consuming electricity or the aggregate power consumption by end users. The physical devices are classified as:<sup>8</sup>

- *resistive loads* (ie, loads containing a heated conductor), with the most obvious examples being light bulbs, ovens and toasters; resistive loads are not very sensitive to the quality of power (voltage levels or frequency);
- *inductive loads*, which include motors, fluorescent lights and transformers; and
- *capacitors* (devices that store electricity<sup>9</sup>), which are relatively unimportant from the point of the overall system behaviour.

The most important inductive loads are motors that consume around 60% of the electric energy produced in the US. This may be somewhat surprising, but one has to recognise that – in addition to industrial motors – American homes and commercial buildings are saturated with different types of devices that require motors (ceiling fans, air-conditioners, refrigerators, swimming pools, etc). Motors are classified as induction, synchronous or direct motors. In induction motors, the magnetic field in the rotor is induced by an electric current flowing through the stationary part (stator). One of the consequences of this is that starting a motor causes inrush current and a voltage drop. Synchronous motors contain either a permanent magnet or an electromagnet connected to an external power source (an exciter). This design allows a motor to operate at constant speed, a feature that is important to most industrial operations.

As shown in Figure 21.2, electricity consumption in the US is characterised by pronounced seasonality, both in the aggregate and by user class. Demand peaks in the summer, primarily due to the air-conditioning load. The shoulder months (spring and autumn) are used by the industry to maintain and repair the units, which are shut down according to pre-established schedules. Following scheduled outages is an important task for fundamental analysts. As one can see from Figure 21.2, the US consumption of electricity has been growing steadily over time, although this trend was halted during the 2008–2012 recession. Another important trend is the increasing electricity intensity of the US economy, as measured by consumption of electricity per unit of GDP. A typical trend in a mature economy undergoing a transformation to an economy dominated by the service sectors is the drop in general energy intensity. This is true of the US, but electricity bucks this general tendency. While electricity usage per unit of GDP is decreasing, electricity consumption per capita has been generally increasing since 1973, with a small drop in the years 2007–09 (due to recession). This is due to a number of

**Figure 21.2** US electricity end use (kWh millions)



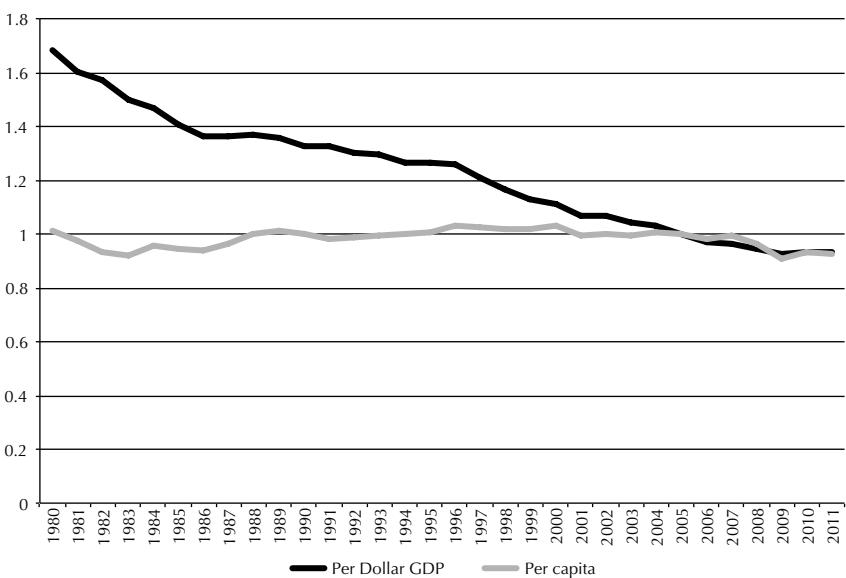
Source: U.S. Energy Information Administration (May 2012)

factors, including larger homes, proliferation of vampire electric appliances<sup>10</sup> and, primarily, the shift to an economy based on information processing and increased use of computers and other electronic devices. These trends discussed here are illustrated in Figures 21.3 and 21.4. Figure 21.3 demonstrates that energy input per one dollar of GDP has been steadily decreasing, while the overall energy per capita use has dropped slightly. This figure is based on statistics contained in the EIA "Annual Energy Outlook 2012." Figure 21.4 demonstrates trends in electricity usage per capita and per unit of GDP. Electricity usage is defined as total electricity retail sales.

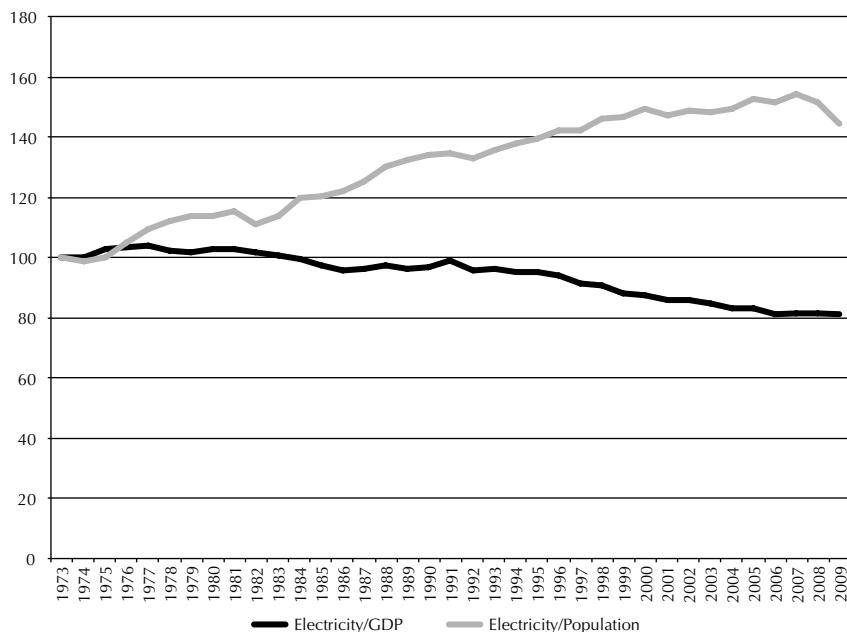
### Load profile

Electricity consumption measured over shorter time periods displays a similar cyclical pattern, during a day, a week and a month. The seasonal pattern varies from location to location, and depends on the time of the year. Load profiles – ie, patterns of electricity consumption (for example intraday or intraweek) – are very important in predicting prices and scheduling retail transactions.

**Figure 21.3** Energy use per capita and per one dollar of GDP (2005 = 1)



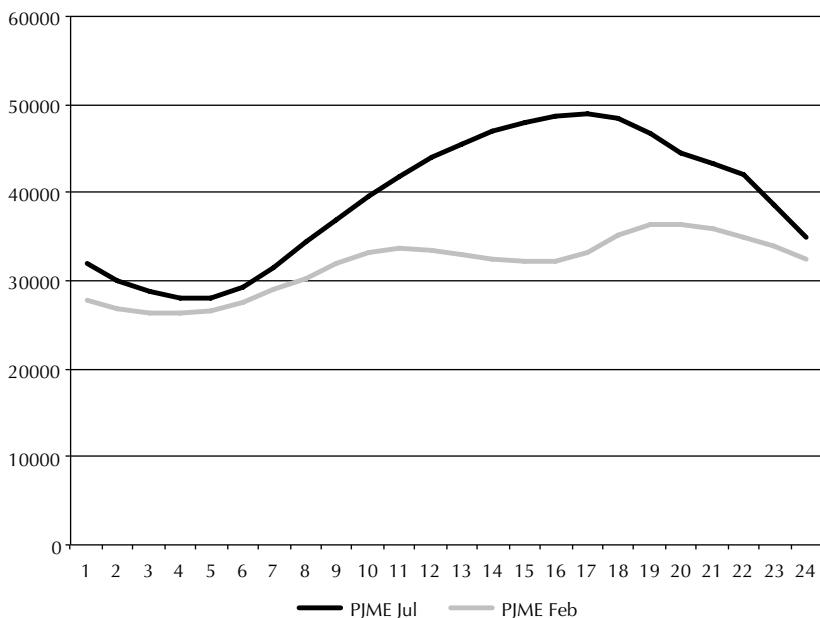
Source: U.S. Energy Information Administration (2012)

**Figure 21.4** Electricity intensity of the US economy (1973 = 100)

Source: U.S. Energy Information Administration (May 2012)

Two typical daily profiles for PJM East (11/07/2012 and 11/02/2012) are shown in Figure 21.5. The summer profile has one peak related to growing air-conditioning load as temperatures increase through the day. The winter profile has a morning peak and an afternoon peak, with a relative demand drop in the afternoon. It is important to recognise that the loads of different customers follow a different profile.

The demand for electricity can be classified as coincident or non-coincident demand. Coincident peak demand corresponds to the consumption of electricity by a group of customers coinciding with the system peak. For example, if the system peak load is registered at 18:45 pm on a specific day, the demand of different system customers at this time is called coincident. This may or not be their peak demand for the day. Non-coincident peak demand for a given customer is the highest level of their electricity usage in a given time period. It is measured without making any reference to the level of the system load. In practice, one can expect that individual

**Figure 21.5** Summer and winter load profiles, PJM East (2012)

customers will schedule their consumption differently over time, with individual demands peaking at different times. Under special circumstances, load patterns may change. For example, following an outage all the appliances may be turned on at the same time, either automatically, or by users resetting the clocks, cooking and showering, and catching up with the news.

Analysing load profile is a critical skill for power marketers offering full requirements deals (see Chapter 23). The holy grail of power marketers is a combination of diverse loads into a shape approximating a rectangle, which can be more effectively hedged with the instruments available in the marketplace. One growing area of research is the analysis of electricity consumption at the level of a single building or a cluster of buildings, such as a university campus or a hospital. There are many specialised companies looking at consumption patterns over the time of the day and the day of the year.

The importance of load profiling has increased with the deregulation of electricity markets and the introduction of competition at the

retail level. For many low-voltage customers (residential customers, small commercials), it is impractical to install interval meters to monitor and report actual electricity usage in real time.<sup>11</sup> Instead, customers are assigned to different load profiles and their aggregate electricity consumption for the period is distributed over time using load factors associated with their class.

Load profiles should meet a number of criteria, including:<sup>12</sup>

- homogeneity* – end users in a given class should have similar characteristics and close usage patterns over different timescales;
- assignability* – end users should share easily identifiable and measurable attributes facilitating assignment of a new user to a given class;
- differentiability* – the load profile should have distinct characteristics making any two profiles easily distinguishable;
- parsimonious character* – the number of distinct profiles should be low;<sup>13</sup> and
- accuracy*.

Several different methods are used in practice to determine the classes of load profiles.<sup>14</sup> Dynamic metering is based on metering with interval meters a sample of customers in each pre-established load class. The advantage of this approach is reliance on actual data reflecting existing market and environmental conditions, although the cost of installing interval meters and reading the data may be quite high. Dynamic modelling relies on a model linking the load to certain environmental variables, such as weather (temperature, dew point, time of day and the day of week). Proxy modelling is based on a selection of data from history for a day closely resembling a given day.

A good understanding of the load profiles of different customer classes and the ability to predict their evolution is a critical fundamental skill required in supporting power marketers and traders offering and hedging, respectively, retail electricity programmes and full requirements deals (covered in Chapter 23). An example will explain the importance of understanding loads. During the second half of the 1990s, the author worked for a company that offered one of the first full requirements deals (ie, a transaction under which a power marketer is obligated to satisfy full electricity needs of a client – of

course, for a price) under a contract executed with a collection of rural cooperatives in Georgia. The transaction was profitable for the first six months (although not excessively) and incurred big losses in the seventh month. What happened? The seventh month was the month of the Olympic Games in Atlanta (in 1996), and the load profile of the residential buyers of electricity (who were clients of the power cooperatives) changed. Instead of driving back to their homes in the suburbs of Atlanta after work, like good Christians, the residents of the service territory under this specific contract would stay in Downtown Atlanta in the sports bars or would attend the games. They would come back home after midnight, and would start cooking and showering, driving electricity consumption up in what was normally the low consumption off-peak time of the day. This pattern was reinforced by the guests in a large number of hotels built in the area just in time for the Olympics. Of course, the hedges for the normal peak hours of consumption were largely useless, and the temporary consumption peak in the early hours of the day was not covered.

The management and optimisation of load profiles could well be the next frontier in energy trading and marketing. Inefficiencies in electricity consumption are huge, and the elimination of waste is the cheapest and easiest way to increase power supply. Such progress is likely to happen through a combination of information technology and creative power supply contracts. Specialised electronic devices allow for the monitoring of power consumption in real time and help to identify waste. Contractual arrangements will combine services of monitoring electricity use, servicing of heating and cooling equipment with power procurement. The potential savings are huge. If one examines the consumption of electricity in a commercial building by day and hour, one can observe wasteful consumption on certain days, especially if the building is consuming a lot of power between midnight and early morning without any obvious reason.

## **POWER POOLS AND EXCHANGES**

### **Electricity market design**

Having covered generation, transmission and demand, it is now time to tie the different components of the electric industry together. The physical properties of electricity discussed in the previous chapters, especially the factor of no meaningful storability at the system level, require the balancing of generation and demand on a

second-by-second basis, while resolving operational issues related to transmission constraints, unit outages and sudden spikes in demand in real time. In the deregulated electricity systems, two prevailing models have emerged: a power pool and an electricity exchange. The power pool solution is associated with the US and has been replicated in a number of countries, including Canada, Australia, New Zealand and Russia.<sup>15</sup> The exchange model, on the other hand, is closely associated with many European countries. Both paradigms are constantly evolving, driven by the need to address shortcomings of different solutions as they are discovered and by the ability of financial markets to innovate and introduce new types of physical and financial transactions. To the best of our knowledge, there is no comprehensive up-to-date comparative analysis of pools and exchanges, and this is not surprising: trying to write such a study can be compared to jumping on the shadow of one's head. The target keeps moving all the time.

In principle, both solutions have a lot in common. They have to deal with the same laws of physics, limitations of physical infrastructure and operational challenges. They share the same objective: design of processes and business practices conducive to the maximisation of social welfare. The objective of maximising social welfare requires the quantification of this elusive concept. As John Maynard Keynes once observed "even the most practical man of affairs is usually in the thrall of the ideas of some long-dead economist." The problem at hand is not an exception. The answer is found in a tool known as consumer surplus, which was invented by the French engineer Jules Dupuit and reinvented independently by the British economist Alfred Marshall. Consumer surplus may be defined as the difference between the amount of money the consumers are willing to pay collectively for a given commodity minus the amount actually paid, and can be illustrated graphically as the area under the demand curve and above the horizontal line representing the market price. The logic behind the consumer surplus is that consumers are willing to pay more for the infra-marginal units of the commodity than the going market price – ie, the price determined through the intersection of demand and supply curves. Of course, this concept applies to a perfectly competitive market. It is possible that the consumer surplus may be extracted through price discrimination, ie, arrangements that allow the suppliers to charge varying prices to different

groups of consumers through clever marketing strategies resulting in market segmentation.

The concept of consumer surplus can be supplemented with the concept of a producer surplus that represents the difference between the market price and the costs incurred by the infra-marginal producers. Graphically, the producer surplus is represented by the area above the supply curve and below the horizontal price curve. The combined consumer and producer surpluses add up to the so-called social surplus. The social surplus is the area contained between the demand and supply curves. Under perfect competition, with no externalities in consumption and production, in the absence of the monopoly power, the social surplus is maximised through the decisions of firms and households, engaging, respectively, in profit and utility maximisation.

The objective of the power pool operator is to replicate the outcome of a market mechanism, ie, the maximisation of the social surplus through the optimal dispatch of available generation resources, given the demand for power, and the constraints of the existing transmission grid. In addition, the optimal combination of loads and generation should satisfy reliability requirements that are often described as N-1 or N-2 conditions. The N-1 (N-2) condition means that an integrated transmission/generation system will survive the failure of one significant component (two significant components). This task requires solving a large optimisation system.

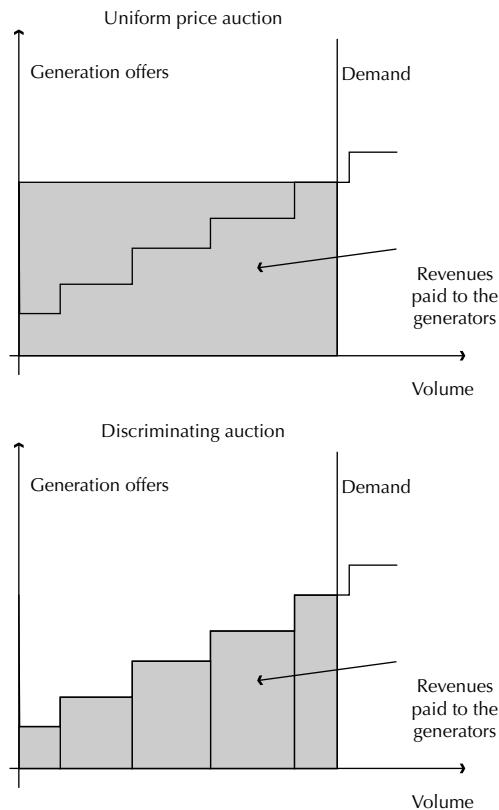
### **Electricity markets as auctions**

Decentralised electricity markets are organised as auctions. The auction is a mechanism for both price discovery and market clearing. The owners of generation and entities representing loads (power marketers, distribution companies) submit bids to the auction administrator (an exchange or a power pool operator). The bids represent offers to generate or bids to buy power. The administrator uses the bids to construct the demand and supply curves, and determines the price at which electricity output and electricity consumption are equalised (the market is cleared). In a complex and a very big electricity market, this process relies on powerful computers and sophisticated matching and optimisation algorithms. Hopefully, this process will replicate the outcome corresponding to what an efficient competitive market would produce (at least this is the idea). This is a

very stylised, bird's eye view description of actual market solutions. In practice, we encounter many very different implementations of this general idea, which are always a work in progress, subject to frequent modifications and experiments with new approaches.

In practice, two types of auctions are used: uniform price auction versus pay-as-bid auction. The uniform auction (or a single price auction) assumes that all successful (ie, accepted) bidders receive the same price, equal usually to the bid of the marginal participant. Under the pay-as-bid auction (discriminating auction), successful bidders receive the prices they bid. Figure 21.6 illustrates the difference between uniform and pay-as-bid (called also single price and discriminating) auctions.<sup>16</sup>

**Figure 21.6** Uniform and discriminating price auctions



The main objection to a uniform auction is the potential for the economic and physical withholding of generation capacity, with the objective of influencing the electricity prices upward. This topic and potential countermeasures are discussed in Chapter 24, in which gaming is discussed. Supporters of pay-as-bid auctions claim that the potential for withholding is eliminated because the payment accruing to one generation unit is not affected, by design, by the payments received by other units. The shortcoming of this auction is that it creates counter-incentives to bidding according to the marginal cost of generation. Low-cost generators are likely to anticipate the market clearing price and bid accordingly. A generator with marginal cost of US\$50/MWh who anticipates an outcome of US\$100/MWh at the auction is likely to bid near this price. It is conceivable that, with a sufficient number of incorrect guesses, higher cost generators may be dispatched ahead of more efficient producers. If the generators succumb to an inexplicable and devastating epidemic of honesty and bid just marginal costs, as hoped for, they will be unable to cover overheads (ie, fixed costs) and an alternative mechanism will have to be designed to finance incremental generation and the replacement of retiring units. The choice of different types of auctions is a very controversial issue, with the consensus shifting towards uniform auction design. A comprehensive comparative analysis of these two types of auctions produced a very strong recommendation in favour of uniform price auction:<sup>17</sup>

1. Do not switch from uniform-pricing to pay-as-bid pricing in the energy spot market. Hopes of saving money through price discrimination are naïve. Such a switch likely will increase consumer costs.
2. Do not attempt a regulatory taking of windfall profits and a regulatory allowing of windfall losses. Even if such a strategy achieved some short-run cost relief, it would destroy investment incentives, and thus, in the long run, destroy the market.
3. Do look for sensible ways to encourage long-term forward contracts that hedge fuel price shifts. Long-term contracts are the only market mechanism available to address the present concern.

Pay-as-bid auctions have another shortcoming, critical from the point of view of a trader. They fail to produce a well-defined reference price that can be used as a benchmark for settling financial derivative transactions and for marking-to-market long-term contracts. For readers of this book, this in itself should be an overriding argument.

---

Other types of auctions are mentioned sometimes in the context of electricity trading. Under the Vickrey auction, the highest bidder wins but the price paid is equal to the second highest bid. This is analogous to an antique auction – if each bidder believes that only they have the full appreciation of the true value of the object being auctioned (and other bidders are not informed), the likely outcome is that the bids will be artificially low. If, however, the Vickrey proposal is implemented, each bidder will have an incentive to bid truthfully and reveal their preferences, still hoping that others do not understand the true value and bid accordingly.

Another auction design challenge arises in connection to path dependence and the complementarities of dispatch decisions. Dispatch decisions require incurring costs during start-up (a different fuel has to be used, the technological heat rate is below the optimal level) and shutdown. The cost of dispatching a unit for an additional hour or a few consecutive hours may depend on whether the unit is up-and-running (hot start), has been recently shutdown (warm start) or has been down for a longer time (cold start). A generator may be willing to dispatch a unit under the condition that it may run for a few hours in order to solve the problem of complementarities and path dependency. An auction in which one bids on one hour at a time would not address this problem. It can be solved through the combinatorial auction – ie, an auction in which one bids on a combination (portfolio) of consecutive hours, subject to an “all-or-nothing” restriction. In practice, most auctions offer the option of block bids, as will be illustrated shortly in our review of selected European exchanges. The blocks can be pre-defined by an exchange or user-selected, subject to multiple conditions, including the size of a block, the span (hour slots included) and the number of blocks submitted by one participant. The latter restriction is related to the computational challenges of solving a combinatorial problem under time-constraint.<sup>18</sup> Spanish electricity exchange uses a different very ingenious solution (see below).

Complementarities are often referred to in technical papers as non-convexities. A short explanation must suffice. As one may remember from school mathematics, a convex function “holds water” (this is a statement with respect to its graph). A chord connecting *any* two points on the graph of the function is always on one side of the graph and above it. If the graph is wiggly, a chord *may*

cross the graph. Non-convexity is the curse of optimisation experts and micro-economists studying conditions for the optimal allocation of resources. Start-up and shutdown costs, lumpiness of dispatch decisions, variable ramp-up and ramp-down rates may result in cost curves that look like a saw with teeth of unequal size.

### **Electricity exchanges in Europe**

EU countries opted for a decentralised market-based electricity system based on an exchange model. This was not so much a result of government mandate but rather an evolution based on replication of successful solutions tested in some countries (especially in Scandinavia). One of the complications of this approach is the need to coordinate generation and transmission decisions. As long as the electricity markets were limited to the territory of one country, this was not a top priority: Europe has well-developed national transmission systems. The push towards the development of regional and eventually pan-European energy markets (which started in 1998), combined with limited cross-border transmission capacities, made the synchronisation of generation and transmission key to the success of electricity markets deregulation. The emerging solution to this generation/transmission synchronisation problem is known as market coupling: a congestion management method based either on the concept of an implicit or explicit auction.

In the case of an implicit auction, market participants submit bids on one exchange but their bids are incorporated into the outcomes of all coupled exchanges. The market prices are calculated jointly by all the coupled exchanges using an iterative algorithm, which considers exports and imports of electricity across national borders. The resulting prices reflect both the cost of generation and the cost of congestion. This process can be summarised as follows, with the following description from European Market Coupling Company (EMCC):<sup>19</sup>

At noon the exchanges receive bids from the players in their market area and transmit anonymous order books (OBK) to EMCC. On the basis of market coupling capacities and prices, EMCC calculates the optimal flow between the market areas. These flows are called market coupling flows (MCF). The algorithm uses the economic welfare criterion. After the calculation, EMCC submits additional price-independent bids/offers to the power exchanges. These bids and offers reflect the calculated market coupling flow. The exchanges

then calculate their own prices taking the bids from EMCC into account. [...] EMCC also nominates the market coupling flow on the interconnectors by sending it to the transmission system operators (TSOs). In hours with different prices on the power exchanges EMCC will collect a congestion rent equal to the MCF multiplied with the difference in prices. The congestion rent is subsequently paid to the owners of the interconnectors. It shall be used to enhance grid quality or extend the transmission network.

Under an explicit auction solution, transmission capacity is allocated independently of electricity trading, resulting often in suboptimal solutions. The mechanism used for the allocation of transmission is usually a sequence of annual, monthly and daily auctions. Under this solution, one can observe frequent price differences between two areas, with transmission capacity being available, or electricity flowing in a wrong direction, from a high-price area to a low-price area. Market splitting, used in the Nord Pool (and explained later), is a similar solution. The primary difference is that in the Nord Pool this mechanism is implemented under the auspices of one exchange, which controls multiple bidding areas.

Market coupling in Europe will happen in the context of several regional electricity markets defined by the European Regulators' Group for Electricity and Gas (ERGEG), an advisory group to the European Commission on internal energy market issues in Europe. In 2006, ERGEG launched the Electricity Regional Initiative (ERI) the aim of which is to speed up the integration of Europe's national electricity markets.

The explicit/implicit auction solutions are sometimes referred to as the available transmission capacity (ATC) versus flow-based trading. At the time of writing Europe was contemplating different solutions:<sup>20</sup>

Currently the central western European (CWE)-interim tight volume coupling market uses an ATC system, where transmission system operators (TSOs) are responsible for allocating the available transfer capacity. However, the target model proposed by the European Commission and Agency for the Cooperation of Energy Regulators sees flow-based trading replace ATC and the entirety of the north western European region market coupling will likely switch to flow-based in late 2013. With flow-based, the coupling algorithm takes into account the expanded and complex power grid and optimises flows so that it goes to where there is available capacity and lower pricing.

The European Council has targeted the end of 2014 as the deadline for achieving power market price coupling throughout Europe.

The blueprints for the emerging European electricity market will now be illustrated using examples of a few exchanges. The list of emerging exchanges to follow is long and the landscape of the power trading business is changing very fast. The section below also contains illustrations of the general principles reviewed so far in this chapter. A detailed review would be self-defeating: the rules and the entities evolve all the time.

#### *Nordpool*

Nordpool was created in 1993 as Power Exchange (PE), covering the territory of Norway.<sup>21</sup> In 1996, Norway and Sweden established a common electricity market and opened the first international power exchange, owned jointly by Svenska Kraftnät and Statnett, the Swedish and Norwegian transmission system operators. International expansion continued with Finland joining in 1998 and western and eastern Denmark in 1999 and 2000, respectively. In 2005, a Kontek bidding area in Germany was added and, in 2010, Nordpool expanded into Estonia.

In 2002, Nordpool was recognised as an organised exchange and clearinghouse. The clearinghouse was spun off as Nord Pool Clearing ASA and spot business was demerged into a company called Nord Pool Spot AS. In 2008, NASDAQ OMX acquired Nord Pool Clearing ASA and Nord Pool Consulting AS, as well as the international products from Nord Pool ASA. The acquired companies were merged into NASDAQ OMX Commodities AS. On March 17, 2010, NASDAQ OMX acquired Nord Pool ASA, as Statnett and Svenska Kraftnät decided to exercise their option to sell the shares in the company.<sup>22</sup> This acquisition did not extend to the physical market. As of April 2012, the products traded at NASDAQ OMX Commodities Europe's financial market include:

- ❑ Nordic, German, Dutch and UK power derivatives (base and peak load futures, forwards, options and contracts for difference);
- ❑ European Union allowances (EUA) and certified emission reductions (CER).

Nord Pool Spot AS, organised as an independent company in 2002, is a joint venture of the Nordic transmission system operators: Statnett

SF (30%), Svenska Kraftnät (30%), Fingrid Oyj (20%) and Energinet.dk (20%). Nord Pool has been actively involved in a number of emerging electricity and energy markets, either through direct participation or as a provider of technical expertise and trading solutions. On January 12, 2010, NASDAQ OMX Commodities and Nord Pool Spot launched N2EX, a market for UK energy contracts. In addition to electricity, Nord Pool is a trading platform for a number of other energy commodities.

The physical market in Nord Pool is organised into several markets:

- Elspot: day-ahead market;
- Elbas: hour-ahead market,<sup>23</sup> and
- balancing market, operated by respective system operators.

As mentioned, the financial market is operated by NASDAQ OMX Commodities. Elspot is a day-ahead market for power – with the system price set through an auction, covering Norway, Sweden, Finland, Denmark and Estonia. About 75% of power consumed in the Nordic countries is traded on this platform. The highlights of this market include:<sup>24</sup>

- Product: Hourly contracts with delivery of power;
- Delivery period: All 24 hours through the next day (12–36 hours ahead in time);
- Type of trading: Auction trading. Accumulated bids and offers that form equilibrium point prices that also reflects usage of available transmission capacities between bidding areas;
- Trading days: All days through the year;
- Bidding: Bids and offers in standard exchange format (Internet, EDIEL<sup>25</sup>);
- Trade currencies: Euro, NOK, SEK, DKK; and
- Price calculation: 12.00 am every day.

The platform supports both hourly and block bids (sale or purchase). Hourly bids cover all 24 hours and can be either price dependent or price independent. A price independent bid specifies volume. Block bids are aggregates of several consecutive hours (at least four), with an option to fill-or-kill the entire block. The starting and ending hours of a block can be freely chosen at the discretion of a bidder. Up to three blocks can be linked, making the following bid dependent on acceptance of the previous bid. Congestion (insufficient transmission

capacity between different bidding areas) is handled by splitting the market into different price areas and repeating price calculations for each area separately, allowing for trading between them. This process, called market splitting, is being repeated until available transmission capacity is exhausted and prices cannot fully converge. The system price is calculated ignoring the available transmission capacity between the bidding areas. The system price serves as a reference price for the trading and clearing of most financial contracts.

The Elbas market is a web-based, cross-border physical intraday market used for adjustments of generation and consumption schedules of electricity in Finland, Sweden, Denmark, Norway, Estonia and Germany. Elbas is licensed to APX-Endex for use as the intraday market in the Netherlands and Belgium.

The highlights of the Elbas market include:

- Product: Hourly contracts with delivery of power;
- Delivery period: All hours with Elspot price in present day and following day up to a few minutes before delivery (120 minutes in Norway, 30 minutes in Germany and 60 minutes in all other bidding areas);
- Type of trading: Continuous; and
- Trading days: Every day throughout the year, 24 hours a day.

Real-Time Market, operated by respective TSOs, is used for balancing output and consumption of electricity at any point in time. Market participants submit bids to the transmission operators after the Elspot market closes. A TSO uses the unit with the lowest price for upward regulation and a unit with the highest price for the downward regulation. If a supplier (ie, an entity marketing power) underestimated its needs, it effectively buys power from the system operator. In the opposite case, a supplier who under-consumes sells power to the system operator. The treatment of generators is different. As explained by Nord Pool:<sup>26</sup>

During an hour with up-regulation, producers producing too much will only get paid the market price (not the up-regulating price). During an hour with up-regulation, producers producing too little will be invoiced the up-regulating price (normally higher than the market price). During hours with down-regulation, producers producing too much will get paid the down-regulating price (normally lower than the market price). Producers producing too little will be invoiced the market price (not the down-regulating price).

*EEX*

Electricity restructuring in Germany started with the Law on Fuel and Electricity Industries (April 24, 1998), which represented the implementation of the EU Energy Directive 96/92/EC. The trading of electricity started with the creation of two power exchanges: LPX (Leipzig) and EEX (Frankfurt), which merged in 2002 to establish the new, Leipzig-based EEX AG. EEX has closely cooperated with French Powernext SA since 2008, and both companies are running an integrated electricity spot and derivative market. EPEX SPOT SE, based in Paris and jointly owned by EEX and Powernext, operates an electricity spot market for Germany, France, Austria and Switzerland. Electricity derivatives are traded in Leipzig through EEX Power Derivatives GmbH. EEX also owns 20% of EMCC GmbH (European Market Coupling Company).

Spot contracts are traded through daily auctions held at 12.00 pm for France, Germany/Austria and Switzerland. The contract specifications are very similar for all the markets. The underlying is hourly electricity for next day delivery, traded in increments of 1 MWh. Blocks trades are allowed, using either standard blocks, defined by the exchange, or customer-defined blocks. The mechanics of block trading is explained as follows:

Block orders are used to link several hours on an all-or-none basis, which means that either the bid is matched on all of the hours or it is entirely rejected. Block orders have a lower priority compared with single hourly orders. The quantity may be different for every hour of the block. A block order is executed for its full quantity only. A block order is executed or not by comparing its price with the volume-weighted average of the hourly market clearing prices related to the hours contained in the block.

Customised block orders link a minimum of two hours of the day together, with maximum of volume of 400 MW and a maximum of 40 block bids per customer's portfolio.

The intraday market allows for trading electricity delivered on the same or next day, as single hours or blocks of hours (Block covering hours 1–24, Block 9–20 or customised blocks). The market starts at 15.00, when all the hours of the following day can be traded. Trading stops 45 minutes before the delivery period begins. Customers can submit limit orders or market sweep orders, with the following restrictions:

- immediate-or-cancel (IOC);
- fill-or-kill (FOK);
- all-or-none (AON); and
- iceberg.<sup>27</sup>

#### *EEX futures*

The following futures contracts are offered by EEX:

- Phelix<sup>28</sup> base week futures (cash settlement);
- Phelix peak week futures (cash settlement);
- Phelix base year/quarter/month futures (cash settlement);
- Phelix peak year/quarter/month futures (cash settlement);
- Phelix off-peak year/quarter/month futures (cash settlement);
- German baseload futures (physical settlement);
- German peakload futures (physical settlement);
- French baseload futures (physical settlement);
- French peakload futures (physical settlement);
- French base week futures (cash settlement), and
- French peak week futures (cash settlement).

For physically settled contracts, the place of delivery are all admissible balancing zones of the EEX spot market (Phelix base futures, Phelix peak futures, Phelix off-peak futures), the balancing zone of the German RWE Transportnetz Strom GmbH (German baseload futures, German peakload futures) or the balancing zone of the French RTE<sup>29</sup> (French baseload futures, French peakload futures). Delivery periods are weeks (week futures), months (month futures), quarters (quarter futures) and years (year futures). The EEX futures provide an example of a cascading contract. According to the exchange, “year and quarter futures are cascaded into the respective quarter or month futures three exchange trading days before the beginning of delivery.”<sup>30</sup>

#### *MIBEL*

The Iberian Electricity Market (MIBEL – the Mercado Ibérico de Electricidad), is a joint venture by Spain and Portugal. The Spanish market started in 1998, and Portugal was added in 2007. The Lisbon Agreement, signed on January 20, 2004, after many years of preparatory work, outlined the basic principles of the common electricity markets:

- a combination of spot (day-ahead and intraday), bilateral contracts and derivatives markets;
- a Spanish-based platform for the spot market (OMEL);<sup>31</sup>
- a Portuguese-based market for derivatives (OMIP);<sup>32</sup> and
- a market for ancillary services.

The day-ahead OMEL market closes at 10.00 on the day prior to delivery day, with prices being published an hour later. The market clears positions including balances transferred from the futures markets going into delivery, and the results of capacity auctions for the transmission interconnections. Under the Royal Decree 485/2209 (April 3, 2009), market participants who signed bilateral contracts with physical delivery of electricity have to submit bids in the day-ahead market for the total contracted volumes at prices corresponding to the opportunity cost. The day-ahead market is followed by six intraday sessions, which allow for the elimination of imbalances developed after the day-ahead session is terminated.

OMIP is a trading platform for electricity futures, for both the Spanish and Portuguese markets, and a registration platform for forward and swap contracts. Futures traded on OMIP range in maturity from days, through weeks, months, quarters, to years. The contracts cover baseload (24 hours) and peak (12 hours), and allow either for cash settlement or settlement through physical delivery.

A special feature of the Spanish model is reliance on capacity payments to cover the fixed costs of generators who are available for dispatch 90% of the time in two peak-time bands. The payments are determined by multiplying a fixed rate (set equal to €5,150/MW for 2012) by a co-efficient ranging from 0.913 for combined cycle gas turbines through 0.912 for coal-fired plants, 0.877 for fuel-oil, to 0.237 for hydro and pumped storage.<sup>33</sup> Capacity payments explain high levels of participation in the spot market: a power producer has to bid into the wholesale market (which is in principle voluntary) in order to qualify.

OMEL addresses the problem of complementarity (non-convexity) using a solution that is unique in the European power markets, known as bids with “complex conditions,” including:

- minimum income;
- scheduled shutdown;

- ❑ indivisibility; and
- ❑ load gradients.

Complex bids are standard bids associated with additional conditions. Minimum income is a provision which results in cancellation of a producer's bid if a daily income from selling electricity falls short of a threshold defined as a fixed amount plus a variable component calculated by multiplying the volume (megawatt-hours times an hourly remuneration). This condition prevents a generator from being dispatched at a loss. Scheduled shutdown is a condition triggered if the minimum income provision results in the cancellation of a bid. If a producer was generating electricity during the last hour of the previous day, a shutdown becomes necessary. This is accomplished in an orderly way by bidding decreasing volumes for up to the first three hours of the next day. Two other conditions are explained by OMEL as follows:<sup>34</sup>

The indivisibility condition enables a minimum operating value to be fixed in the first block of each hour. This value may only be divided by the application of the load gradients declared by the same agent, or by applying distribution rules if the price is other than zero. The load gradient enables the maximum difference between the starting hourly power and final hourly power of the production unit to be established, limiting maximum matchable power by matching the previous hour and the following hour, in order to avoid sudden changes in the production units that the latter are unable to follow from a technical standpoint.

These conditions help to avoid discrete, excessive changes in dispatch levels or generation levels that are too small to be technologically viable. The intraday market uses a number of additional, more complicated, bid structures that any trader should fully understand.<sup>35</sup>

### **Power pool model**

One has to recognise that all power pool designs differ across the US and other countries, and are constantly evolving. This section will explore the principles underlying the organisation and responsibilities of a generic power pool. In the US, the term *power pool* is often used interchangeably with a regional transmission organisation (RTO) operating an ISO, or just an ISO. The responsibilities of an RTO are wider than the responsibilities of a generic power pool, as described below.<sup>36</sup>

A power pool is an entity responsible for the dispatch of generation units and the management of transmission resources in order to meet the end-user demand and ensure reliable service. The term power pool has historical ramifications. In the past, local utilities decided to pool resources (hence the term “pool”) in order to reduce the cost of providing electricity, procuring fuel and reducing outage risk.<sup>37</sup> The trend towards the creation of power pools has accelerated since the 1990s as many countries embarked upon a restructuring of the industry and provided direct political and institutional support to this form of industrial organisation. The pool typically does not own the generation and transmission assets, although it has a significant degree of control over their operations (however, some countries opted to create pools owning the transmission lines).

The difference in design of different power pools has profound implications for the electricity markets. One consequence is that the skills developed in trading one market are not directly transferable to a different region. A trader familiar with ERCOT may have to spend significant time learning the rules of another power pool before they can trade in a new market. In addition, one has to absorb additional information about the differences of physical infrastructure, load characteristics and the profiles of market participants. In 2002, FERC made an effort to develop a standard market design (SMD) for RTOs, but this initiative encountered strong resistance, particularly from stakeholders representing different regional interests.<sup>38</sup> The effort was effectively abandoned, although during the discussion of the FERC NOPR, significant progress was made in formulating the fundamental principles underlying any power pool and, by extension, an RTO. A review of the principles underlying the SMD is a useful exercise, as FERC addressed in this proposal most critical issues related to power pool design.

- ❑ *Objectives.* Power pool design should support several principal objectives, including the promotion of economic efficiency and reduction of the energy costs to the end users, enhancement of system reliability and mitigation of market power. The market framework should benefit the customers irrespectively of the state-specific solution related to retail market organisation (retail choice).
- ❑ *Transparency.* The rules should be fair and not excessively

complicated, beyond the ability of an average market participant to understand even the most intricate provisions.

- ❑ *Independence*. The governance system should guarantee the independence of a pool from the owners of the generation, transmission and distribution assets.
- ❑ *Freedom of choice*. The design should allow the market participants to exercise freedom of choice with respect to investment decisions related to physical assets, participation in the financial and bilateral markets, reliance on self-scheduling and management of the demand side.
- ❑ *Fuel and technology neutrality*. The pool mechanism should be neutral with respect to technology, fuel mix choices and the scale of load and generation resources. In other words, the pool design should not be skewed towards any specific outcome that would be different from decisions made in the context of a market mechanism.
- ❑ *Equal treatment of load and generation resources*. The pools should be based on the principle of equal treatment of load and generation. The demand response is an important factor contributing to improved economic efficiency and also to mitigation of market power.
- ❑ *Time and location-specific prices*. SMD-supported pool design based on locational marginal prices used not only as the tools of short-term dispatch decisions but also as important factors in investment decisions related to generation and transmission assets.
- ❑ *Demand response*. The English economist Alfred Marshall said that demand and supply determine prices in the same way that two blades of scissors cut paper. It does not make much sense to argue which blade is more important.<sup>39</sup> Power markets are often different in the sense that the end users of electricity are barred from receiving the price signals from the market in real time. In the past, most customers of regulated utilities were paying tariff prices based on average costs. Such solutions did not allow for adjustment of consumption to the instantaneous cost of producing electricity. FERC was supportive of the blueprints that would make at least part of the load price sensitive in real time. Interaction between supply and demand is important for a number of reasons, including the mitigation of market power

(end users may curtail consumption in response to price spikes) and expansion of the choices available to retail and wholesale customers.

- ❑ *Transmission cost recovery.* Transmission expansion should be based on merchant principles, with the costs shouldered by the investors, without cost-recovery assurances from the regulators. The pool design should offer the opportunities for recovery of historical and new expansion costs.
- ❑ *Access to transmission.* The FERC proposal honoured grandfathered transmission contracts. However, the unused transmission capacity should be made available to others.
- ❑ *Dynamic character of power pools.* The design of power pools should allow for the modifications of rules, as market realities evolve and to respond to regional requirements.

The specific issues that a power pool designer has to address include:

- ❑ *Price formation.* Most power pools currently in operation evolved towards a design based on LMPs, prices set for each node of the system (the mechanics of calculating LMPs and what is specifically meant by a node will be explained in the next section of this chapter). The difference between prices at two different nodes is known as congestion: transmission capacity constraints do not allow for the equalisation of prices across the grid. Some pools rely on the zonal blueprint with prices averaged over homogeneous clusters of nodes called zones. The implicit assumption is that the congestion within a zone is non-existent or minimal, and can be ignored. Some pools use a design based on a single system marginal price (SMP) across the entire pool (for example, the old England and Wales power pool).
- ❑ *Bids and offers.* The prices in a pool are based in most designs on the information about the supply and demand curves (bids and offers) communicated by the pool participants to the pool operator. The supply curve information may require submitting actual cost data, subject to verification by a pool. In most pool designs, the submitted bids are left to the discretion of the generation resources, with an implied assumption that the pool design architecture will induce them to reveal the true information in their best economic interest.

- ❑ *Energy-only pools versus energy-plus capacity pools.* Some pools rely on the pricing mechanism under which the generation units are receiving only a price for the energy they produce, and if they produce it. In addition, some pools have provisions for capacity payments (calculated usually per kW-year) the generation units may receive just for being around. The actual rules which determine who qualifies for such payments and how the payments are determined may be devilishly complicated and constantly evolving.
- ❑ *Closed or open pool.* A closed pool is restricted to a certain class of market participants. An open pool allows every generation and load resource to participate in the pool.
- ❑ *Bilateral contracts.* Most pools allow market participants to self-schedule generation (in the case of generation resources that are serving dedicated load) and rely on the bilateral contracts, subject to submitting relevant information to the pool.
- ❑ *Market mitigation.* Most power pools have design features put in place to address potential abuses of market power, which are often manifested through domination by some market participants of certain local markets, created by transmission constraints and limited supplies of power during some time periods. The range of different tools varies widely from market to market. A market participant should be well familiar with different market mitigation procedures used by different power pools, given the potential for retroactive adjustments of prices or the ability to override *ex post* prices determined by the pool.

#### *Locational marginal price*

As explained in the previous section, most power pools evolve in the direction of a design based on the nodal (or locational marginal prices). The specific algorithms and definitions of this concept vary from power pool to power pool, but one can identify a number of commonalities in the implementation of this concept. Here are the main highlights.

- ❑ Locational marginal prices (referred to as prices in this section to avoid repetitions) are calculated at the specific node of the integrated generation transmission system, although the definition of a node may vary in practice. The locational marginal prices are

used to charge the loads, compensate the generators and determine compensation to the holders of the firm (financial) transmission rights.

- ❑ Locational marginal prices are calculated as a solution to the linear programming problem and represent shadow prices, ie, the change in the value of the objective function due to relaxation of a specific constraint by a small number. For example, if the load at a specific node is increased by a unit, the locational marginal price will reflect the cost of serving this additional demand. This additional cost should reflect, in principle, system-wide, and not just local, conditions.
- ❑ The locational marginal prices are calculated through the solution of an optimisation programme that is designed to approximate the outcome of a competitive market process. The model is populated with the information supplied by the market participants in the form of offers (submitted by the generators) and bids (submitted by the loads), as well as additional information such as start-up and shutdown costs, technological ability to increase and decrease output (ramp rates) and minimum run times. The individual bid and offer curves are submitted as piecewise-linear functions. In other words, the blocks of power are associated with prices that the loads and generators are willing to accept. Any trader covering a given market should be familiar with the design of the optimisation system and technical details of the algorithm used for calculating the prices.
- ❑ The prices are calculated in most power pools for the day-ahead (DA prices) and real-time (RT prices). The real-time market is perceived as a market for last-moment adjustments, dictated by additional developments, such as weather forecast updates, changes to load estimates, transmission and generation outages.
- ❑ Prices in general are composed of three components:
  - energy prices that are flat across the power pool;
  - line loss components specific to each location; and
  - a congestion component (to be explained in detail in the next section).
- ❑ There are many differences in the technical details of calculating prices in different power pools. For example, for many years PJM ignored the line-loss component in its calculation of prices.

- ❑ An important difference between many power pools is the extent to which participants with no physical load and generation are allowed to participate in the day-ahead and real-time markets. In case virtual bids are allowed, a virtual sale in the day-ahead market becomes a purchase in the real-time market, and vice versa. In general, our experience of virtual bids is positive, as they represent a tool for financial players to arbitrage between the two markets, resulting in convergence of the DA and RT prices.
- ❑ The bids and offers in many cases may reflect strategic objectives of the market participants and the desire to influence market prices by taking advantage of local market power. For example, a generator holding a significant position in firm transmission rights which settle based on day-ahead prices may be tempted to avoid bidding into the day-ahead market in order to drive certain marginal prices up.<sup>40</sup>
- ❑ The system operator may choose to modify the prices derived from the solution of the optimisation problem. The interventions may consist in *ex post* interventions and adjustments to the prices, or *ex ante* designation of certain units as must-run for reasons related, for example, to voltage support.
- ❑ The paradox of most LMP systems is a gap between the voluminous data available in real time (often every five minutes) and a lack of transparency in the actual design of the system. The algorithms used by the power pools are usually developed by private consulting firms and may be proprietary. The technical details provided in official publicly available documents are contained in documents written in a language and style that make old Soviet documents on a planned economy an example of literary clarity. Some consulting firms provide commercial calculators of LMPs,<sup>41</sup> but a word of caution is required. A trader trying to guess the LMPs does not know the exact assumptions the system operator is using to initialise the calculation engine.

*LMPs: mathematical formulation and an example*

The locational marginal prices are derived as a solution of a constrained optimisation problem that represents the convergence of multiple theoretical threads and combines elements of economic theory and electrical engineering. We shall explain the model by starting with a discussion of the concept underlying the objective

function, before proceeding to examine constraints related to the available generation capacity, demand for power and the topology of the transmission grid.

A simplified version of a bid-based, security constrained, economic dispatch problem a system operator in a power solves in order to determine LMPs is shown below:<sup>42</sup>

$$\min_{G_i} \Sigma_i B_i G_i \quad (21.1)$$

subject to:

$$\Sigma_i G_i = \Sigma_i D_i \quad (21.2)$$

$$\Sigma_i A_{i,l} (G_i - D_i) \leq F_l^{\max} \quad (21.3)$$

$$G_i^{\min} \leq G_i \quad (21.4)$$

$$G_i \leq G_i^{\max} \quad (21.5)$$

We assume an integrated generation/transmission network with  $m$  nodes. Generation at node  $i$  is equal to  $G_i$ , demand to  $D_i$  ( $i = 1, 2, \dots, m$ ). Of course, there are nodes with no generation (nodes with just load attached to them, no generation units at this location) and nodes with no demand (just power generation plants). The cost of generation at node  $i$  is equal to  $G_i \times B_i$ , where  $B_i$  is a bid submitted by a generator located at this node. In principle, the bid may be based on marginal cost of the  $i$  th generator or may diverge from marginal cost and reflect the generator's bidding strategy. The objective function (equation 21.1) is a minimisation of the total cost of satisfying customers' demand. Constraint (equation 21.2) reflects the requirement that total generation (across all the nodes) is equal to total demand. Constraint (equation 21.3) captures transmission limits. The flow on line  $l$  ( $l = 1, \dots, k$ ) cannot exceed the maximum level given by the maximum line capacity of this line ( $F_l^{\max}$ ). The actual flow on line  $l$  is given by the left-hand side of equation 21.3, and is calculated by multiplying net injection of power at node  $i$  by the shift factor for node  $i$  and line  $l$ . This requires a few comments. Net injection captures how much power a given node exports to (or imports from) the rest of the integrated generation/transmission system. For example, if at given node we generate 300 MW and consume 200 MW, net export is equal to 100 MW. Net exports can be negative.

This is another way of saying that a given node imports electricity. Shift factor  $A_{i,l}$  is sensitivity of flows on line  $l$  to injections of power at node  $i$ . The left-hand side of equation 21.3 is a calculation of how much different nodes contribute in total to power flows on line  $l$ . Constraints (equation 21.4) and (equation 21.5) reflect the requirement that generation at node  $i$  cannot be below a minimum level (equation 21.4), or above a maximum level (equation 21.5) specified for this node.

An optimisation problem is solved by forming a so-called Lagrangian, which is equal to the objective function plus the constraints, with each constraint being weighted by a factor called a Lagrange multiplier. At the optimum point, the value of a multiplier measures the impact of the constraint on value of the objective function. If a constraint is not binding, the corresponding multiplier is zero (the constraint has no impact on the optimal solution). Such constraints do not get any respect. If a constraint is binding, the corresponding Lagrange multiplier measures how much relaxation of this constraint (by a small amount) would change the optimal value of the objective function. The Lagrangian corresponding to our problem is equal to:

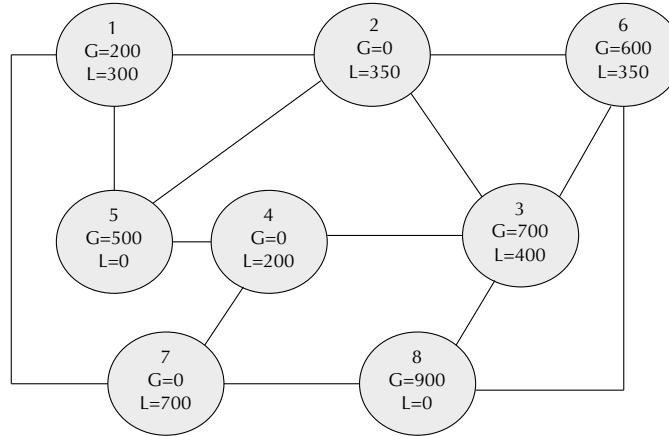
$$L = \sum_i B_i G_i + \lambda \left( \sum_i G_i - \sum_i D_i \right) + \sum_l \mu_l \left( \sum_i A_{i,l} (G_i - D_i) \leq F_l^{\max} \right) \\ + \sum_i \gamma_i^{\min} (G_i^{\min} - G_i) + \sum_i \gamma_i^{\max} (G_i^{\max} - G_i) \quad (21.6)$$

where the symbols  $\lambda, \mu, \gamma$  are used to represent the Lagrangian multipliers of the problem at hand.

This problem can be depicted using numerical examples, which contain a simplified illustration of a system with a few generation units and a few load sinks. The first example assumes a simple system with no transmission constraints. Generation capacity and load at each node are given in Table 21.2. Some nodes have only generation with no load attached to them; some nodes have load with no generation capacity. The system diagram is shown in Figure 21.7.<sup>43</sup>

The optimisation programme for this system is very simple. The decision variables include generation levels at each plant ( $G_i, i = 1, 2, \dots, 8$ ). The load at each node ( $D_i, i = 1, 2, \dots, 8$ ) is assumed to be fixed. The constraints can be divided into three sets:

**Figure 21.7** The eight-node system



Source: Narayan S. Rau, 2003, Optimization Principles: Practical Applications to the Operations and Markets of the Electric Power Industry (Piscataway, NJ: IEEE Press)

**Table 21.2** An eight-node system: load and generation capacity

<b>Node</b>	<b>Generation number</b>	<b>capacity (MW)</b>	<b>Load</b>	<b>Unit cost (4)</b>
	(2)		(3)	
1		200	300	50
2		0	350	0
3		700	400	22
4		0	200	0
5		500	0	40
6		600	350	35
7		0	700	0
8		900	0	32

$$G_i \leq capacity_i, i = 1, 2, \dots, 8 \quad (21.7)$$

$$D_i = load_i, i = 1, 2, \dots 8 \quad (21.8)$$

$$\sum_{i=1}^8 G_i - \sum_{i=1}^8 D_i = 0 \quad (21.9)$$

Constraints (equation 21.7) correspond to the requirements that generation at a given node cannot exceed existing capacity

(capacity<sub>i</sub>). Capacity at each node is listed in Table 21.2. Load at each node (load<sub>i</sub>) is fixed (see Table 21.2) and the system has to be dispatched to satisfy demand at each location (hence the equality constraints, equation 21.8). Constraint (equation 21.9) reflects the requirement that total load should be equal to total electricity output.

The objective function is minimisation of the total cost of satisfying demand. Unit costs are given in the last column of Table 21.3. Variable B<sub>i</sub> denotes unit generation cost at each node. It is assumed that the generators submit bids equal to their costs.

$$\sum_{i=1}^8 G_i \times B_i = \min \quad (21.10)$$

This very simple problem was solved using the Mathematica package (function LinearProgramming), but it can be handled with any decent program, including Solver in Excel. The solution is given in Table 21.3.

As one can see, total generation (2,300 MW) is equal to total demand and the cost of satisfying demand over the period of one hour is equal to US\$69,200. One can ask what LMP corresponds to this problem. There are no transmission constraints and we are ignoring line losses. Therefore, the costs of satisfying one additional unit of demand will be equal across the system, irrespective of where additional load is located. We can find this additional cost in a number of ways. The poor man's brute force solution is to perturb one of the load constraints (it does not matter which one, as the LMPs

**Table 21.3** Solution of Example I

Node number (1)	Optimal dispatch (MW) (2)	Unit cost (US\$/WWh) (3)	Cost (US\$) = (2) × (3)
1	0	50	0
2	0	0	0
3	700	22	15400
4	0	0	0
5	100	40	4000
6	600	35	21000
7	0	0	0
8	900	32	28800
Total	2300		US\$69,200

are equal across the system in the case of no congestion) and find a new optimal solution to the modified problem. Increasing the load at node 6 from 350 to 351 increases the total cost of satisfying demand to US\$69,240 (from US\$69,200). Therefore, the LMP is equal to US\$40/MWh. In this simple case, one could guess the result by observing that the marginal generation unit is located at node 5. This is the most expensive unit that is dispatched, and has additional capacity to satisfy additional units of load (it is initially dispatched at 100 MW level and has the capacity of 500 MW). Its cost is exactly US\$40/MWh.

The Mathematica package (Version 8) has a built-in function for solving the dual linear program (`DualLinearProgramming`). The solution to this problem yields locational marginal prices. What is a dual problem? If the primal problem is given (in vector notation) by:

$$\mathbf{c}^T \mathbf{x} = \max \quad (21.11)$$

subject to:

$$\mathbf{A} \mathbf{x} \leq \mathbf{b} \quad (21.12)$$

$$\mathbf{x} \geq \mathbf{0} \quad (21.13)$$

The dual is given by:

$$\mathbf{b}^T \mathbf{y} = \min \quad (21.14)$$

Subject to:

$$\mathbf{A}^T \mathbf{y} \geq \mathbf{c} \quad (21.15)$$

$$\mathbf{y} \geq \mathbf{0} \quad (21.16)$$

The dual of a dual is the original primal. If the primal has an optimal solution,  $\mathbf{x}^*$ , the dual has an optimal solution,  $\mathbf{y}^*$ , and

$$\mathbf{b}^T \mathbf{y}^* = \mathbf{c}^T \mathbf{x}^* \quad (21.17)$$

The new decision variables  $\mathbf{y}$ 's are the dual (shadow) prices of resources represented by the constraints of the primal problem. A shadow price is the change in the level of the objective function of the primal due to the change in the level of one constraint (with other constraints remaining the same) by one unit.

Finally, most packages (such as SAS) offer LP programs that

produce shadow prices as a solution of a linear optimisation program. For example, SAS legacy procedure LP (available in the SAS OR package) calculates shadow prices (if this is requested by the user).

The problem with transmission constraints is far more complicated. The next example is an expanded version of the first example, with additional constraints related to transmission limits. In order to solve the expanded problem, we have to go through the additional step of calculating the load shift factors. The shift factors measure the impact of additional demand/generation at one node on the rest of the system. The calculations start with the admittance matrix, which we assume to be equal to:

54	-20	0	0	-20	0	-14	0
-20	78	-15	0	-25	-18	0	0
0	-15	86	-40	0	-21	0	-10
0	0	-40	90	-20	0	-30	0
-20	-25	0	-20	65	0	0	0
0	-18	-21	0	0	74	0	-35
-14	0	0	-30	0	0	89	-45
0	0	-10	0	0	-35	-45	90

As explained in the chapter on electricity basics, in the equation the diagonal elements of this matrix are equal to the sum of the corresponding row elements, with the minus sign. The matrix is singular but the sub-matrix created by dropping  $m$ th row and column can be inverted. In our case, we can drop the 8th row and column. The resulting sub-matrix is inverted, and then expanded by adding to the right and the bottom a column and a row of zeros. This is equivalent to treating node 8 of the network as a reference bus, and calculating all the angles of voltage and current against the angle of this node (which may be set to zero for convenience). The inverted padded matrix is equal to:

0.039975424	0.023550632	0.013843307	0.015861677	0.026238582	0.0096570383	0.011634902	0
0.023550632	0.032707407	0.015993692	0.015561493	0.024614272	0.012494606	0.0089500411	0
0.013843307	0.015993692	0.02519674	0.017378945	0.015758267	0.011040784	0.0080356702	0
0.015861677	0.015561493	0.017378945	0.026952699	0.019158844	0.0087170909	0.011580274	0
0.026238582	0.024614272	0.015758267	0.019158844	0.038820082	0.010459196	0.010585454	0
0.0096570383	0.012494606	0.011040784	0.0087170909	0.010459196	0.019685937	0.0044574299	0
0.011634902	0.0089500411	0.0080356702	0.011580274	0.010585454	0.0044574299	0.016969627	0
0	0	0	0	0	0	0	0

The final step is pre-multiplication of this matrix by the branch admittance matrix multiplied by the incidence matrix,<sup>44</sup> as explained in Chapter 19.

The incidence matrix is:

1	-1	0	0	0	0	0	0
1	0	0	0	-1	0	0	0
1	0	0	0	0	0	-1	0
0	1	-1	0	0	0	0	0
0	1	0	0	-1	0	0	0
0	1	0	0	0	-1	0	0
0	0	1	-1	0	0	0	0
0	0	1	0	0	-1	0	0
0	0	1	0	0	0	0	-1
0	0	0	1	-1	0	0	0
0	0	0	1	0	0	-1	0
0	0	0	0	0	1	0	-1
0	0	0	0	0	0	1	-1

The branch admittance matrix is assumed to be equal to:

20	0	0	0	0	0	0	0	0	0	0	0	0	0
0	20	0	0	0	0	0	0	0	0	0	0	0	0
0	0	14	0	0	0	0	0	0	0	0	0	0	0
0	0	0	15	0	0	0	0	0	0	0	0	0	0
0	0	0	0	25	0	0	0	0	0	0	0	0	0
0	0	0	0	0	18	0	0	0	0	0	0	0	0
0	0	0	0	0	0	40	0	0	0	0	0	0	0
0	0	0	0	0	0	0	21	0	0	0	0	0	0
0	0	0	0	0	0	0	0	10	0	0	0	0	0
0	0	0	0	0	0	0	0	0	20	0	0	0	0
0	0	0	0	0	0	0	0	0	0	30	0	0	0
0	0	0	0	0	0	0	0	0	0	0	35	0	0
0	0	0	0	0	0	0	0	0	0	0	0	45	0

The final result (the load shift factors) is shown below:

0.32849584	-0.18313549	-0.043007714	0.0060036899	0.032486204	-0.056751363	0.053697218	0
0.27473684	-0.021272791	-0.038299201	-0.065943329	-0.25162999	-0.016043155	0.020988943	0
0.39676732	0.20440828	0.081306915	0.059939639	0.21914378	0.072794518	-0.074686162	0
0.14560988	0.25070572	-0.13804571	-0.027261778	0.13284007	0.021807337	0.013715562	0
-0.067198745	0.20232837	0.005885641	-0.089933774	-0.35514524	0.050885259	-0.040885344	0
0.2500847	0.36383041	0.089152358	0.12319924	0.25479137	-0.12944396	0.080867	0
-0.080734827	0.017287981	0.3127118	-0.38295016	-0.13602308	0.092947724	-0.14178418	0
0.087911642	0.073480811	0.29727508	0.18189893	0.11128049	-0.18154822	0.075143047	0
0.13843307	0.15993692	0.2519674	0.17378945	0.15758267	0.11040784	0.080356702	0
-0.20753809	-0.18105558	0.03241356	0.1558771	-0.39322476	-0.034842104	0.019896401	0
0.12680327	0.19834356	0.28029824	0.46117273	0.25720168	0.12778982	-0.16168058	0
0.33799634	0.43731122	0.38642744	0.30509818	0.36607186	0.68900781	0.15601004	0
0.52357059	0.40275184	0.36160516	0.52111237	0.47634547	0.20058434	0.76363325	0

In Table 21.4, Column (2) shows actual dispatch in the case of congestion. In the optimal solution some units may not be dispatched at all or dispatched at the levels below the maximum available capacity. Column 7 shows the total cost by unit and the total cost, which we

**Table 21.4** Optimal dispatch in the case of congestion

Node number	Optimal dispatch	Generation capacity (MW)	Demand	Unit cost (US\$/MWh)	Net node injection	Cost (US\$)
(1)	(2)	(3)	(4)	(5)	(6) = (2) - (4)	(7) = (2) × (5)
1	0	200	300	50	-300	0
2	0	0	350	0	-350	0
3	700	700	400	22	300	15400
4	0	0	200	0	-200	0
5	414.5331872	500	0	40	414.5331872	16581.32749
6	600	600	350	35	250	21000
7	0	0	700	0	-700	0
8	585.4668128	900	0	32	585.4668128	18734.93801
Total	2300		2300			US\$71,716

are seeking to minimise. The total cost of satisfying demand is equal to US\$71,716. Column 6 shows net injection of power by node: the difference between generation at a given node and demand at this node. For example, the generation unit at node 1 is not dispatched (this is the most expensive unit), but demand at this node that must be fully satisfied (by assumption) is equal to 300 MW. This results at net injection of -300 MW at this node.

Net injections at each node, in conjunction with the load shift factors, are used to calculate flows along each branch (starting with branch 1–2 through 6–8). These flows have to satisfy the constraints for each line (minimum and maximum levels). The flows along each line are calculated by post-multiplying the matrix of load shift factors by the vector of net injections. The resulting flows are shown in Table 21.5.

The congested case was created by restricting transmission capacity on line 38 (connecting nodes 3 and 8). The line flows and capacity limits for each line are shown in Table 21.5.

In the original case the line limits were (-500 MW, 1000 MW) and there was no congestion. By lowering the limits to (-20 MW, 20MW), we made line 38 full, and this required re-dispatch as one can see comparing the optimal generation levels in Tables 21.4 and 21.3.

The transmission constrained case was implemented using a number of packages (including SAS and Mathematica) and Excel (using Solver). We can calculate the LMP in a number of different ways. We can use again a poor man's solution – which is to run the optimisation problem for the baseline case, and then perturb

**Table 21.5** Line flows

<b>Limit lower</b>	<b>Branch flow</b>		<b>Limit upper</b>
-100	-86.8637	P12	100
-500	-196.289	P15	1000
-500	-16.8476	P17	1000
-500	-116.474	P23	1000
-500	-136.781	P25	1000
-500	-183.609	P26	1000
-500	254.673	P34	300
-500	51.14677	P36	1000
-20	20	P38	20
-500	81.46319	P45	1000
-500	-136.136	P47	1000
-500	-15.2446	P68	1000
-500	580.7114	P78	1000

constraints, one at a time, by a small amount and then re-run the optimisation problem. The difference between two values of the objective function, per unit change of constraint, is equal to the shadow price for a given constraint. For example, increasing required load at node 4 to 201 (from 200) increases the overall cost of satisfying total demand in the system to US\$71757.08827. The locational marginal price at this node is equal to (US\$71757.08827 – US\$71,716.2655 = US\$40.8277/MWh). One can obtain the same answer by requesting a sensitivity report for the baseline optimisation case,<sup>45</sup> which shows the reduced cost of cell B6 (corresponding to node 4) as US\$40.8227696590285/MWh. The small discrepancy between the two results is due to numerical problems (rounding and truncation errors).

## CONCLUSIONS

In this chapter, we covered a number of topics critical to the understanding of modern electricity systems, including a review of transmission grid and loads. These two components, combined with power plants, form an integrated and synchronised transmission/generation system. This is a marvel of modern technology that those of us fortunate enough to live in developed countries take for granted. Deregulated power markets are organised along two different alternative lines: power pools (primarily the US) and an exchange model (Europe). Both models keep evolving at a furious pace under the impact of regulatory pressures and technological

developments, and staying on top of these developments is almost an inhuman task. However, there is a silver lining. Complexity creates trading opportunities for those who can keep up with a market that changes practically every day.

- 1 The word *load* has multiple meanings in the power business. It refers either to aggregate electricity requirements of a given class of users or to specific devices consuming electricity.
- 2 Bob Shively and John Ferrare, 2004, *Understanding Today's Electricity Business* (San Francisco, CA: Enerdynamics).
- 3 NERC, which was created in 1968, started as a voluntary organization. "As of June 18, 2007, the U.S. Federal Energy Regulatory Commission (FERC) granted NERC the legal authority to enforce reliability standards with all users, owners, and operators of the bulk power system in the United States, and made compliance with those standards mandatory and enforceable. Reliability standards are also mandatory and enforceable in Ontario and New Brunswick, and NERC is seeking to achieve comparable results in the other Canadian provinces." See <http://www.nerc.com/page.php?cid=1>.
- 4 "On August 14, 2003, large portions of the Midwest and Northeast United States and Ontario, Canada, experienced an electric power blackout. The outage affected an area with an estimated 50 million people and 61,800 megawatts (MW) of electric load in the states of Ohio, Michigan, Pennsylvania, New York, Vermont, Massachusetts, Connecticut, and New Jersey and the Canadian province of Ontario. "Interim report: Causes of the August 14th blackout in the United States and Canada, November 2003" (<https://reports.energy.gov/814BlackoutReport.pdf>).
- 5 Alice Clamp, 2010, "The Lake Erie loop flow mystery," *Public Power*, January–February.
- 6 "But rural areas, which account for most of the 30 million US homes that don't have broadband access, provide an opportunity. Internet providers have avoided these locales because the population is too sparse for cable or phone companies to lay fiber or coaxial cable profitably, and hills and trees disrupt wireless networks." William M. Bulkeley, 2008, "IBM hired to develop power-line broadband," *Wall Street Journal*, November 12 (<http://Online.Wsj.Com/Article/Sb122644998675019183.Html>).
- 7 For example, the break during the Super Bowl causes a spike in electricity consumption as all the fans visit the toilet. This is followed by the water pumps starting in a synchronised way.
- 8 See Alexandra v. Meier, p 127, for a description of different physical devices, "Electric Power Systems: A Conceptual Introduction," John Wiley & Sons, July 11, 2006.
- 9 Capacitors are made of two metal plates separated by a non-conductive substance. Among other applications, they are used to improve the quality of power supply, in timing devices, or as storage devices.
- 10 "The typical American home has 20 electrical appliances that bleed consumers of money. That's because the appliances continue to suck electricity even when they're off, says a Cornell University energy expert. His studies estimate that these so-called "vampire" appliances cost consumers \$3 billion a year – or about \$200 per household. [...] Worldwide, standby power consumes an average of 7 percent of a home's total electricity bill, although that figure is as much as 25 percent in some homes. See *Science Daily*, September 27, 2002. Another source is David B. Floyd and Carrie Webber, 2012, "Leaking electricity: Individual field measurement of consumer electronics" (<http://enduse.lbl.gov/Info/ACEEE-Leaking.pdf>).
- 11 Technological progress lowering the price of interval meters and modern communications technology are likely to expand the scope of demand side management in the future. This will change the trend of electricity consumption with a corresponding impact on market prices. This is a development that electricity traders should follow closely.

- 12 Jeffrey Bailey, 2000, "Load profiling for retail choice: Examining a complex and crucial component of settlement," *The Electricity Journal*, December, p 69.
- 13 For example, in the power pool of England and Wales customers' electricity usage below 100 kW, maximum demand was settled using load profiles and readings from customers' existing electricity meters. Eight load profiles were used. See "Load profiles and their use in electricity settlement" ([http://www.elexon.co.uk/ELEXON%20Documents/load\\_profiles.pdf](http://www.elexon.co.uk/ELEXON%20Documents/load_profiles.pdf)).
- 14 Jeffrey Bailey, op.cit.
- 15 Axel Ockenfels, Veronika Grimm and Gregor Zoettl, 2008, "Electricity market design: The pricing mechanism of the day ahead electricity spot market auction on the EEX," March 11 (<http://www.eex.com/en/Download/Documentation/Expertises>).
- 16 A good survey of auction theory can be found in Paul Klemperer, 1999, "Auction theory: A guide to the literature," *Journal of Economic Surveys*, (13)3, July, pp 227–86.
- 17 Peter Cramton and Steven Stoft, 2006, "Uniform-price auctions in electricity markets," March 20 (<http://www.cramton.umd.edu/papers2005-2009/cramton-stoft-clearing-price-markets.pdf>).
- 18 A combinatorial problem is usually NP-complete (NPC) in the jargon of computational complexity theory. While a solution can be quickly validated, there is no feasible way of finding it in reasonable time.
- 19 <http://www.marketcoupling.com/market-coupling/concept-of-market-coupling>.
- 20 <http://www.risk.net/energy-risk/feature/2144493/european-power-market-coupling-marches>.
- 21 <http://www.nordpoolspot.com/About-us/History/>.
- 22 <http://ir.nasdaqomx.com/releasedetail.cfm?releaseid=452675>.
- 23 The hour-ahead market does not operate in Norway (a hydro-based system).
- 24 [http://www.nordpoolspot.com/Global/Download%20Center/Annual-report/Nord-Pool-Spot-Company-Presentation\\_2011.pdf](http://www.nordpoolspot.com/Global/Download%20Center/Annual-report/Nord-Pool-Spot-Company-Presentation_2011.pdf).
- 25 Nordic Ediel Group was formed in 2003 to address questions regarding data exchange in the energy market and to create electronic data interchange (EDI) within the Nordic power market.
- 26 [http://www.nordpoolspot.com/Global/Download%20Center/TSO/how-does-it-work\\_balancing-power.pdf](http://www.nordpoolspot.com/Global/Download%20Center/TSO/how-does-it-work_balancing-power.pdf).
- 27 An iceberg order is an order of an unknown size, ie, with partial volume disclosed to other market participants one at a time. Such an order may be entered if a trader does not want to reveal the full volume they intend to buy or sell, in order to avoid front running by other market participants or to avoid being taken advantage of if the full scale of their needs becomes known.
- 28 Physical Electricity Index. Calculated on a daily basis by EPEX SPOT, the Phelix is the average price for base load (Phelix day base) and peak load (Phelix day peak) electricity traded on the German/Austrian auction.
- 29 RTE is Réseau de Transport d'Électricité, the largest European electricity transmission system operator. RTE is a wholly owned subsidiary of Électricité de France (EdF).
- 30 [http://www.eex.com/en/EEX/Products%20%26%20Fees/Power/Phelix\\_Futures](http://www.eex.com/en/EEX/Products%20%26%20Fees/Power/Phelix_Futures).
- 31 Operador del Mercado Ibérico de Energia.
- 32 Operador do Mercado Ibérico, SGPS, SA.
- 33 Platts, 2011, "Spain's government passes decree on 2012 power capacity payments," November 21.
- 34 <http://www.omel.es/en/home/markets-and-products/electricity-market/daily-and-intradaily/daily-market/daily-market>. See also Néstor Sánchez María, 2010, "Day-ahead electricity market: Proposals to adapt complex conditions in OMEL," Universidad Pontificia Comillas. Escuela Técnica Superior De Ingeniería (Icai), Madrid.
- 35 A full review of the intraday market would take too much space. The bid structures include,

in addition to minimum income and load gradient, "Complete acceptance in the matching process of the first block of the sale bid, complete acceptance in each hour in the matching period of the first block of the sale bid, minimum number of consecutive hours of complete acceptance of the first block of the sale bid, maximum matched power." See <http://www.omeles/en/home/markets-and-products/electricity-market/daily-and-intradaily/intraday-market>.

- 36 The terms "power pool," RTO and ISO are seldom defined precisely in the literature and tend to be a bit nebulous.
- 37 The voluntary pool was based on the principle of voluntary surrender of dispatch decisions by individual utilities to the pool in return for participation in savings created through productivity enhancements. The split-savings approach was quite cumbersome and is no longer in use. See John Chandley, 2007, "How RTOs set spot market prices (and how it helps keep the lights On)," LECC, LLC, September.
- 38 Notice of Proposed Rulemaking, Docket No. RM01-12-000, issued July 31, 2002.
- 39 "We might as reasonably dispute whether it is the upper or the under blade of a pair of scissors that cuts a piece of paper as whether value is governed by utility or cost of production. It is true that when one blade is held still, and the cutting is effected by moving the other, we may say with careless brevity that the cutting is done by the second; but the statement is not strictly accurate, and is to be excused only so long as it claims to be merely a popular and not a strictly scientific account of what happens." Alfred Marshall, 1890, *Principles of Economics* (London, England: Macmillan) – "Book Five: General Relations of Demand, Supply and Value, Chapter 3, Equilibrium of Normal Demand and Supply."
- 40 In case a generator is obliged to bid, they may choose a bidding strategy designed to throw them into the real-time market.
- 41 See, for example, Ventyx Promod IV (<http://www.ventyx.com/analytics/promod.asp>), ABB Grid View (<http://www.abb.com/industries/db0003db004333/c12573e7003305cbc12570060069fe77.aspx>), Energy Exemplar PLEXOS (<http://www.plexosolutions.com/>), Power World (<http://www.powerworld.com/>).
- 42 Based on José Arce and Scott Wilson, 2004, "Managing congestion risk in electricity markets," Carnegie Mellon Conference on Electricity Transmission in Deregulated Markets, December 15–16.
- 43 The system topology is based on an example in Narayan S. Rau, 2003, *Optimization Principles: Practical Applications to the Operations and Markets of the Electric Power Industry* (Piscataway, NJ: IEEE Press). The numerical example is a modified version of a numerical model developed by Professor Rau and available from his book's website.
- 44 The node-branch incidence matrix is defined as follows: "Element node incidence matrix A shows the incidence of elements to nodes in the connected graph. The incidence or connectivity is indicated by the operator as follows:

$a_{pq} = 1$  if the p<sup>th</sup> element is incident to and directed away from the q<sup>th</sup> node.

$a_{pq} = -1$  if the p<sup>th</sup> element is incident to and directed towards the q<sup>th</sup> the node.

$a_{pq} = 0$  if the p<sup>th</sup> element is not incident to the q<sup>th</sup> node."

P. S. R. Murthy, 2007, *Power System Analysis* (Hyderabad, India: BS Publications), p 19. The word "incident" in this context means an element appertaining or attaching to another element.

- 45 This can be done from the dialogue box after the Solver produces information that the solution was found. The spreadsheet is available from [vkaminski@aol.com](mailto:vkaminski@aol.com).



## 22

### *Analytical Tools*

This chapter will cover a number of the analytical tools and data sources a trader or analyst can use to monitor developments in the electricity markets and make trading decisions. Some of the tools, such as the complex software packages used for the analysis of integrated generation transmission systems are beyond the scope of the book. Using such systems requires dedicated staff with a strong background in electrical engineering and a good understanding of available data sources and their potential shortcomings. Many of these packages go back to the models used by regulated utilities in their resource planning, and were overhauled to incorporate features of the deregulated electricity markets, including the potential for strategic bidding. These modifications have sometimes been less than fully successful – but this should not be taken as a criticism. It is rather as recognition of what is state of the art in this field and the difficulty of modelling the power markets. We are dealing with very complex and evolving systems, and it will take time before we can accumulate enough experience to translate it into stylised representations of reality using mathematical tools.

Some of the available packages include:<sup>1</sup>

- ❑ Cambridge Energy Solutions's (CES) Day Ahead Analyzer (Dayzer) and Transmission Analyzer (Tranzer);
- ❑ GE-MAPS,<sup>2</sup>
- ❑ PROMOD IV;<sup>3</sup>
- ❑ Power World;<sup>4</sup> and
- ❑ AURORAXmp.<sup>5</sup>

These packages are fairly expensive and are used in-house by large companies that can afford the cost (licence and necessary staffing).

However, many consulting firms lease the software and use these and other packages in their client engagements.

In this chapter we will examine simpler (but still very powerful) tools that can be used to analyse electricity markets and support trading decisions. More specifically, we will explain such concepts as load factors, technological and market heat rates, spark spread and generation stack. In the second part of the chapter, we will cover some of the sources of the data for calibration of analytical models and for measuring the pulse of the power markets.

## TECHNICAL CHARACTERISTICS OF GENERATION AND LOAD

In this section we will discuss different analytical tools and quantitative concepts used to describe and analyse individual generation units, as well as the entire portfolio of generation units. The levels of power plant generation are often summarised using indicators known as the capacity factor, heat rate, and availability factor.

### Capacity factors

The capacity factor is defined as the actual output of energy over a certain period of time divided by the potential output if the unit is operated at full nameplate capacity 24 hours a day with no interruptions during the entire period. Two formulas for capacity factors are shown below:

$$\text{Quarterly capacity factor} = \frac{\text{Energy produced during the quarter}}{\text{Quarterly capacity} \times 2190} \quad (22.1)$$

$$\text{Annual capacity factor} = \frac{\text{Energy produced during the year}}{\text{Quarter 4 capacity} \times 4 \times 2190} \quad (22.2)$$

The number 2,190 used in the definition of the capacity factor corresponds to the average quarter length of 91.25 days. Quarter 4 capacity means capacity at the end of the last quarter of the year. An average annual capacity is an alternative representation of maximum potential production. An example of such a calculation is shown below, assuming a coal-fired power plant with a capacity of 1,500 MW might produce 972,000 megawatt-hours in a 30-day month. The full capacity output is given by:

$$1,500 \text{ MW} \times 30 \text{ days} \times 24 \text{ hours/day} = 1,080,000 \text{ MWh.}$$

The capacity factor is given by:

$$\text{Monthly capacity factor} = \frac{972,000\text{MWh}}{30 \text{ days} \times \left( 24 \frac{\text{hours}}{\text{day}} \times 1,500\text{MW} \right)} = 90\% \quad (22.3)$$

Average capacity factors for the US power plants are reported by the EIA, with the data for 2009 shown in Table 22.1. One important trend illustrated by this table is the increasing utilisation of combined cycle natural gas power plants (capacity factors increasing from 33.5 to 42.2 between 2003 and 2009). This development reflects low prices for natural gas in the US and the substitution of natural gas for coal.

Capacity factors for individual plants and across regions and countries vary considerably for a number of reasons, including the technological characteristics of the generation fleet, its age, load profile and composition, prices and the availability of other fuels. The differences are especially big for the renewable sources of electricity, such as wind and solar, given the intermittency of wind and cloud cover in different locations.

Capacity factor is a very crude statistic ignoring the time profile of generation. The same capacity factor may correspond to continuous operations for 10 days in a month, with a down period for 20 days, and operations characterised by frequent starts and shutdowns.

The availability factor of a power plant is the fraction of time it is available to produce electricity. No power plant can be called upon

**Table 22.1** US power plants annual capacity factors (%)

Year	Coal	Petroleum	Natural Gas CC	Natural Gas Other	Nuclear	Hydroelectric Conventional	Other Renewables	All Energy Sources
1998	67.7	22.2	–	34.2	79.2	46.6	57.0	54.6
1999	68.1	22.4	–	33.2	85.3	45.9	56.9	54.9
2000	71.0	20.5	–	37.1	87.7	39.5	59.1	54.6
2001	69.2	21.5	–	35.7	89.4	31.4	50.2	51.4
2002	70.0	18.1	–	38.2	90.3	38.0	54.0	49.7
2003	72.0	22.4	33.5	12.1	87.9	40.0	50.0	47.7
2004	71.9	23.3	35.5	10.7	90.1	39.4	50.5	47.9
2005	73.3	23.8	36.8	10.6	89.3	39.8	47.0	48.3
2006	72.6	12.6	38.8	10.7	89.6	42.4	45.7	48.0
2007	73.6	13.4	42.0	11.4	91.8	36.3	40.0	48.7
2008	72.2	9.2	40.6	10.6	91.1	37.2	37.3	47.4
2009	63.8	7.8	42.2	10.1	90.3	39.8	33.9	44.9

Source: U.S. Energy Information Administration (April 2011)

to operate 365 days a year, 24 hours a day. Scheduled maintenance, unplanned (forced) outages and starting failures reduce plant availability. Additionally, plant availability may be affected by curtailment of fuel or water supply. Availability factor is calculated by dividing operation hours + stand by hours (when the plant can operate) by the number of hours in the period. The availability factors should not be confused with the capacity factors. Thermal power plants have an availability factor ranging from 80–90%+, but some peaking units may operate only for a small fraction of hours during a year. Solar panels and wind turbines may have, by frequently used convention, availability factors approach 100%, but the capacity factors may be of an order of magnitude of 15–20%.

Similar measures can be used to describe the characteristics of loads. Load factor measures how steady electricity generation (consumption) is. It is defined as the ratio of the average load supplied during designated time to the peak load. Power producers and marketers have a strong preference for customers with high load factors, as it translates into a higher predictability of load. Load factor depends on the time period for which it is defined. Streetlights during the night operate roughly at the same level, so the average load is very close to the peak load, resulting in a load factor close to 100%. If one looks at the 24-hour period, streetlights consume electricity during the daytime at a very low level (primarily for the traffic control), and the overall load factor is likely to be between 15 and 30%, depending on the geographical location and season.

Demand factor is the ratio of maximum demand divided by the total load connected to the system. For example, in the case of a residence, total connected load may be 10,000 W (all the TV sets, kitchen appliances, lights, swimming pool pumps, etc) but all these devices do not run at the same time. Maximum demand may be only 4,000 W, and this translates into the demand factor of 40%.

Load characteristics are very important in highly structured, complex transactions such as full requirements deals. Many power marketers have suffered big losses because they either did not understand the properties of load they sought to serve or did not predict correctly the impact of the changing economic and demographic conditions on the load profile.

## HEAT RATE AND THERMAL EFFICIENCY

The energy efficiency of a thermal generation unit is expressed in terms of percentage of theoretical maximum level of energy that can be produced, or in terms of the heat rate. Both measures are closely related, and one can be converted into another through a simple transformation and adjustment of units. In the US, the industry has a preference for the use of the heat rate; in Europe, they use thermal efficiency expressed in percentage terms.

Thermal efficiency in the first sense is defined as:<sup>6</sup>

$$\eta = \frac{\text{Electricity produced}}{\text{Energy input}} \quad (22.4)$$

For example, if a unit produces electricity equivalent to 3,500Btus from burning coal containing 10,000Btus, the thermal efficiency of the plant is equal to 35%.

Given the second law of thermodynamics, this ratio cannot exceed one and is typically much lower than the maximum value. For thermal plants, this ratio varies typically between 30 and 50%, with combined cycle plants reaching 60%. In other words, a power plant can convert into electricity only a fraction of energy contained in fuel. The rest is wasted through frictions, radiated heat, etc. However, one should be very careful comparing reported efficiency of different power plants. As explained by Eike Roth:<sup>7</sup>

Upon closer examination, one can see that the numerical value for efficiency is based on a combination of the laws of physics with definitions made by man. [...] Based on these different definitions, it is impossible to determine by that single value which is "better": a hydro power station with 85% efficiency, a coal-fired thermal power plant with 45% efficiency, or a solar power plant with 15% efficiency. Which power plant is 'better' depends on how it fulfills its role to produce reliable, cheap and environmentally sustainable power in the best way. This could very well be a plant with a relatively low efficiency.

As always, the devil is in the detail. For example, one can ask the obvious question of how to measure the energy content of fuel. In practice, engineers distinguish between lower and higher heating values. As explained by the EPA, the two concepts are related through the following equation:<sup>8</sup>

$$\text{LHV} = \text{HHV} - 10.55(\text{W} + 9\text{H})$$

where LHV is the lower heating value of fuel in Btu/lb, HHV is the higher heating value of fuel in Btu/lb, W is the weight % of moisture in fuel and H is the weight % of hydrogen in fuel.

Technological heat rate is energy input (ie, energy contained in fuel) required to produce one kWh (or one MWh) of electricity in a given generation unit. The energy markets rely almost exclusively on the heat rate approach, whereas the engineers rely on the first approach (ie, thermal efficiency).

The heat rate is defined as the Btu input per 1 kWh of electricity produced (or MMBtu per one MWh). A highly efficient combined cycle power plant may have a heat rate as low as 6,300 (according to the first definition using kWh units), with older combustion turbines operating at the heat rates between 14,000 and 16,000. A heat of 6,300 Btus means that a plant burns 6,300 Btu of natural gas to produce 1 kWh of electricity. If we use different units, this is equivalent to 6.3 MMBtus per 1 MWh of electricity. At prices of natural gas of around US\$2.5/MWh, this translates into a marginal cost of US\$15.75/MWh.<sup>9</sup> One should not be surprised that in the US natural gas competes very effectively against coal as a fuel of choice in electricity generation. Historically, in most US power pools, gas-fired units were at the margin – ie, were used to satisfy marginal demand. This has been reversed over time, and in some power pools and some months, coal has become the marginal fuel.<sup>10</sup>

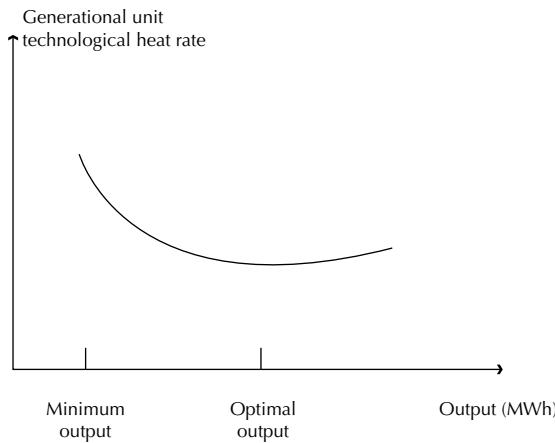
It is important to recognise that the heat rate is quoted for optimal level of operation of a plant. One can diverge from the optimal level of output by sacrificing efficiency, ie, accepting higher heat rate (as shown in Figure 22.1). This means that a generation unit can operate below the optimal level (ie, generate less) or above the optimal level (ie, generate more), but this translates into a higher fuel usage per unit of electricity produced. Most coal plants cycle down during the night, producing at 70–80% of their optimal level. This results in the marginal cost increasing when market prices drop to off-peak levels. This is still cheaper than shutting down the plant and restarting it in the morning. This is particularly important for coal power plants that operate continuously for long time periods but have to reduce output during the periods of lower load (at night or due to weather conditions). Also, the heat rate is higher during the ramp-up period or when a power plant shuts down. This is an important consideration for decisions related to unit dispatch. Every decision to start up a

plant or shut it down is associated with costs that have to be considered to achieve an optimal mode of plant operations.

The term market heat rate denotes the ratio of electricity price to the fuel price. For example, if the forward price of electricity is equal to US\$36/MWh, and the price of natural gas is equal to US\$4/MMBtu, the implied market heat rate is equal to nine. The market heat rate is related to the technological heat rate of a marginal generator (ie, the generator required to satisfy marginal load) but can diverge from the latter for a number of reasons.

The technological heat rate of a marginal generation unit determines marginal cost – ie, the cost of satisfying additional units of demand. In a perfectly competitive market, the market price is determined by the marginal cost. In this case, a market heat rate would be roughly equal to the technological heat rate. The difference between the two would arise because marginal cost includes a small component in addition to fuel: variable operation and maintenance (O&M) cost. In the US, this cost is typically assumed to be of the order of magnitude of US\$2–3/MWh, and is often ignored in the economic and financial analysis of generation decisions. In practice, however, electricity markets do not function as textbook visions of a perfectly efficient market. For example, if the marginal generator has market power (defined as the ability to set prices above the marginal costs), the market and technological heat rates will diverge. Heat

**Figure 22.1** Heat rate versus output level



**Table 22.2** Average operating heat rate for selected energy sources (2001–10, Btu per kWh)

Period	Coal	Petroleum	Natural gas	Nuclear
2001	10,378	10,742	10,051	10,443
2002	10,314	10,641	9,533	10,442
2003	10,297	10,610	9,207	10,421
2004	10,331	10,571	8,647	10,427
2005	10,373	10,631	8,551	10,436
2006	10,351	10,809	8,471	10,436
2007	10,375	10,794	8,403	10,485
2008	10,378	11,015	8,305	10,453
2009	10,414	10,923	8,160	10,460
2010	10,415	10,984	8,185	10,452

Source: U.S. Energy Information Administration

rate-related transactions are very popular in the energy markets and will be covered in the next book considered by the author.

Both thermal efficiency and heat rate may be very elusive concepts. The comments below explain some details in the concept of a heat rate. First, some electricity produced by a plant is used by the plant itself (for example, for lighting or security and control systems) and one has to specify if the heat rate applies to the net or gross output. Net power output is gross output adjusted for internal electricity consumption. Also, the concept of energy input may be imprecise because some energy may be used for the treatment of fuel (for example, removing excess water from coal). Two identical plants may have different thermal efficiency if the fuel arrives to the site in different conditions. There is a difference between receiving dry coal and the same quality coal saturated with water.

Thermal efficiency (heat rates) are affected by a number of different factors.<sup>11</sup>

- ❑ *Design choices.* For example, higher turbine pressure and temperature increase efficiency but require more expensive materials.
- ❑ *Operations procedures.*
- ❑ *Fuel quality.* There are many trade-offs between coal quality (for example, ash and sulphur content) and energy content.
- ❑ *Pollution abatement.* Emission controls and CO<sub>2</sub> sequestration (ie, capture in underground rock formations) is a good social policy but it has to be paid for.

- *Ambient conditions.* Heat rates of thermal power plants are lower (ie, efficiency is higher) in winter. Air used for combustion is denser (burning processes are more efficient) and steam (gases) exiting a turbine decompress more effectively if the ambient temperature is lower. Also, plants located at lower altitudes are more efficient, everything else being equal.

### **Spark spread and its uses in valuation models**

The merchant power industry relies extensively on a valuation model known as the spark spread to determine the market value of thermal power plants, especially units that have significant embedded optionality. For a number of reasons (listed below), the way this technique is used may result in a significant overvaluation of the physical generation assets and excessive claims regarding their financial performance. The spark spread model is also used occasionally as a planning and forecasting tool.

A power plant may be represented as a portfolio of spark spreads, defined as the difference of power prices and fuel prices. Fuel prices are multiplied by the heat rate to convert them to the same units in which electricity is priced. Such models have been widely used since the 1990s to manage and value power plants. Given that gas-fired power plants or dual-burner plants have most optionality, this model applies typically to valuation of such units.

To be more specific, the underlying assumption is that a generation unit can be looked at as a portfolio of options to generate electricity or standby, depending on the market conditions. The payout of this option is defined using a spark spread, which is effectively a profit margin from producing power, if one ignores the fixed costs (depreciation, taxes, interest on debt, etc). The strike of this option is equal to the variable non-fuel unit cost,<sup>12</sup> which is sometimes ignored in the analysis. In the latter case, the strike price is assumed to be zero. A plant is dispatched into the grid when the option is in the money – ie, the spark spread exceeds the strike price. In other words, a plant operates if the revenues from selling power can cover the fuel cost and variable non-fuel cost. This does not mean that the fixed cost is necessarily recovered.

### Alternative spark spread definitions

The industry developed a number of different spark spread definitions evolving around the same basic equation:

$$\text{Spark spread} = \text{Power price} - \text{Heat rate} \times \text{Fuel price} \quad (22.5)$$

where the heat rate corresponds to the technological heat rate of a given power plant.

Initially, the term “spark spread” was used exclusively for natural gas-fired power plants and most people in the industry still define it in this way. Spark spreads for certain standardised levels of heat rates are reported in many industry newsletters and news services. The example below was borrowed from *Megawatt Daily*.

The spark spread at Mass Hub (a trading hub in Massachusetts) for a 7,000 heat rate hypothetical natural gas turbine is calculated as follows:

$$45.75 \text{ US\$/MWh} - 7\text{MMBtu/MWh} \times 4.187 \text{ US\$/MMBtu} \approx \$16.44$$

where US\$45.75/MWh is the price index for Mass Hub reported in the issue of *Megawatt Daily* for Monday, November 8, 2010, the price

**Table 22.2** Spark spread calculation

Marginal heat rate	Spark spreads				
	@7k	@8k	@10k	@12k	@15k
<b>East</b>					
Mass Hub	10925	16.44	12.25	3.88	-4.50
N.Y. Zone-G	12861	22.33	18.52	10.90	3.28
N.Y. Zone-J	12710	22.13	18.25	10.50	2.75
N.Y. Zone-A	8994	8.43	4.20	-4.25	-12.70
Ontario*	7867	3.67	-0.56	-9.01	-17.47
PJM West	11355	16.88	13.00	5.25	-2.50
TVA, into	9648	9.40	5.85	-1.25	-8.35
<b>Central</b>					
Cinergy Hub	9189	8.10	4.40	-3.00	-10.40
NI Hub	8707	6.28	2.60	-4.75	-12.10
Entergy, into	8863	6.31	2.92	-3.85	-10.62
ERCOT, Houston	8515	5.21	1.77	-5.10	-11.97
<b>West</b>					
Mid-C	8839	6.36	2.90	-4.01	-10.92
Palo Verde	9393	8.04	4.68	-2.04	-8.76
NP15	9352	8.80	5.06	-2.43	-9.91
SP15	10247	11.17	7.73	0.85	-6.03

Source: *Megawatt Daily*, Monday, November 8, 2010

of natural gas is equal to US\$4.187/MMBtu (Gas Daily price for the closest natural gas market hub).<sup>13</sup>

Other definitions of spark spreads available in the literature include dark spread and clean spark spread, which is popular in the European markets. Dark spread is the spread between electricity prices and the prices of thermal coal burnt at the plant. Quark spread is the spread calculated for nuclear power plants.

The clean spark spread (otherwise known as a green spark spread) represents the profit margin of a generator who sells power produced in a gas-fired power plant, after buying carbon credits corresponding to the volume of CO<sub>2</sub> emissions. The clean spark spread is calculated as follows (we use the formula from the publication *Tendance Carbone*, from Caisse des Dépôts):<sup>14</sup>

Clean spark spread = Electricity price

$$-\left( \frac{\text{Price of natural gas}}{\text{Net thermal efficiency of power plan}} + \text{Price of CO}_2 \times \text{CO}_2 \text{ Emission factor} \right) \quad (22.6)$$

The clean dark spread (dark green spread) is calculated in a similar way:

clean dark spark spread = Electricity price

$$-\left( \frac{\text{Price of coal} \times \frac{1}{\text{Net thermal efficiency of power plan}}}{+ \text{Price of CO}_2 \times \text{CO}_2 \text{ Emission factor}} \right) \quad (22.7)$$

$$\text{Clean Dark Spread} = P_{elec} - \left( P_{Coal} \times \frac{1}{\rho_{Coal}} + P_{CO_2} \times FE_{Coal} \right)$$

With:

$P_{elec}$	Price of electricity Powernext Futures Month Ahead Peak, in €/MWh;
$P_{Coal}$	Price of coal CIF ARA Month Ahead, in €/MWh;
$\rho_{Coal}$	Net thermal Efficiency of a conventional coal-fired plant (40%);
$P_{CO_2}$	Price of CO <sub>2</sub> of BlueNext, in €;
$FE_{Coal}$	CO <sub>2</sub> emission factor of a conventional coal-fired plant in tCO <sub>2</sub> /MWh.

In the specific examples available from Caisse des Dépôts, the price of electricity is taken as Powernext Futures month ahead peak

(€/MWh), price of natural gas is taken at Zeebrugge market (month ahead) ((€/MWh)),<sup>15</sup> net thermal efficiency of conventional gas-fired power plant is taken as 55% (40% for the coal plant), the price of carbon (one ton of CO<sub>2</sub>) is taken from Bluenext. CO<sub>2</sub> emission factor of a conventional power plant is measured in tons of CO<sub>2</sub> per MWh of electricity output. The CO<sub>2</sub> intensity factor is assumed by convention to be equal to 0.411 tons of CO<sub>2</sub> per MWh for gas plants, and 0.86 for coal plants. The efficiency levels of a typical gas-fired power plant are assumed in most publications to be between 0.5 and 0.55, and between 0.38 and 0.4 for coal plants. In the UK and German spark spread tables, the gas-fired plant efficiency is assumed often to be equal to 0.4913. The UK and German calculations of dark spreads assume a coal plant efficiency of 35%. The use of this number is dictated by strictly pragmatic reasons and preference for round numbers: 25,000 therms of natural gas translate into 15 MWh of power at this efficiency level. The price of coal is taken as the price of coal at ARA<sup>16</sup> one month ahead and is expressed in €/MWh (€/MWh = (€/GJ)/0.27777 = ((€/t)/29.31)/0.2777).

Climate spread is defined as the difference between clean dark spread and clean spark spread. This spread captures the climate benefits of using natural gas as a fuel instead of coal. The climate spread is a major fundamental driver behind the prices of carbon emission credits in the European Emission Trading Scheme. We shall revisit this topic in the chapter on environmental markets.

### **Spark spread models: Shortcomings**

The concept of a spark spread is used to value thermal power plants (especially gas-fired power plants), which are looked at as portfolios of spread options. The basic approach is very sound, but this tool was at some point in the history of the US merchant power industry drastically misused in order to justify massive investments in the gas-fired generation, with the final losses measured in billions of dollars. The US “dash to gas” in the late 1990s was supported by slogans about the “convergence” of natural gas and electricity industries, a term that makes little sense when looked at in hindsight. The valuation models, based on a spark spread concept, were greatly simplified – resulting in excessively optimistic valuations. Additional mispricing was caused by using unrealistic correlation assumptions, sometimes as low as 30%. Suffice to say that high corre-

lations between electricity and fuel prices (technically, we are talking about correlation of price returns) translate into low values of spread options. If both prices move exactly in sync, there is little potential for the spread to increase. Lower correlations mean more potential variability of the spread, hence a higher option value. When more realistic correlation levels (between 70–90%, depending on the season) were used, billions of dollars evaporated instantly from balance sheets.

Excessively simplistic valuation models based on the spark spread concept had a number of different flaws, which unfortunately survive in many practical applications. One can identify three major types of shortcomings:

- ❑ unrealistic assumptions used for modelling price dynamics;
- ❑ a primitive dispatch model that ignores many rigidities in the operations of a power plant; and
- ❑ modelling a power plant's operations as if it existed in a spatial and institutional vacuum.

The mathematics and calibration of spread options and the spark spread model will be covered in the next book. At this point we shall discuss technological and institutional considerations affecting performance and, by extension, the valuation of natural gas-fired power plants (points 2 and 3 above).

Simplistic spark spread models rely on an unrealistic dispatch model and treat power plants similarly to desk lamps that can be turned off and on at will. In practice, a power plant's dispatch decisions have to recognise a number of important constraints that limit the ability to exploit all short-term market opportunities. Understanding these constraints is important to any electricity and natural gas trader. The generation-related component of demand for natural gas is a primary factor behind the volatility of gas prices, and also is notorious for being difficult to model and forecast.

In the case of natural gas power plants, one has to recognise limitations related to the following.

- ❑ *Number of starts.* A number of starts per year exceeding a certain level (typically, around 50) results in the excessive wear and tear of the turbine and increases maintenance and servicing costs.

- ❑ *Ramp-up<sup>17</sup> and ramp-down costs.* Depending on the initial conditions (warm or cold start), it takes time to bring the unit to full capacity or to shut it down. During the ramp-up and ramp-down time, emission rates and heat rates increase compared to the optimal levels of operation.
- ❑ *Minimum up- and downtime.* After it is shut down, a unit may have to remain idle for a period of time dependent on the turbine design. This is related, among other factors, to the requirement to cool down (warm up) different components at an optimal rate to avoid thermal stress.
- ❑ *Start-up and shutdown costs.* Depending on the generation unit characteristics and design, some additional costs may be incurred during start-up and shutdown, and these costs have to be recognised in economic analysis.
- ❑ *Limitations on the total number of hours of operation.* National and local environmental laws translate into either limitations on the total numbers the units can run in a given period of time and/or the need to purchase emission credits (CO<sub>2</sub>, SO<sub>2</sub>, etc).

Most models we have examined in our career as a consultant, or found in the literature, tend to ignore the obvious fact that power plants (as with other energy assets) do not exist in a vacuum and cannot be modelled as if their operations were independent of the rest of the surrounding infrastructure and market organisation. A few examples illustrate the fallacy of this approach. Complications that are usually ignored include:

- ❑ access to fuel supply at short notice due to technological factors;
- ❑ an ability to coordinate dispatch decisions with the fuel suppliers; and
- ❑ the potential for fuel suppliers to extract value from a power plant.

The ability of a peaking unit to start at short notice depends on the ability to obtain fuel, and this in turn depends on a number of factors such as access to a pipeline (or multiple pipelines) or to the LDC distribution network, availability of storage, direct access to a production field and contracting practices with gas suppliers. Technological constraint is related to the ability of a pipeline to

deliver sufficient volume of natural gas to a unit that, by design, was put in place to be dispatched at very short notice. This is due to technological factors as well as differences in operational procedures followed by the power and gas industries.

### **Interactions between natural gas and electricity industries**

A NERC document released in 2011<sup>18</sup> identifies the following technological challenges natural gas pipelines face when servicing power plants.

- *High-point loads.* The off-takes of power plants per unit of time are large relative to those of other customers. For example, a 500 MW unit with an 8 MMBtu/MWh technological heat rate will require 4,000 MMBtus of natural gas per hour, sometimes at a moment's notice. On a daily basis, some bigger power plants rank closely behind LDCs of reasonable size and may exceed the daily loads of small LDCs.<sup>19</sup>
- *High-pressure loads.* Modern gas turbines require delivery pressures of 450 to 475 psi at the fuel skid.<sup>20</sup> Higher gas pressure is one of the factors behind improvement of technological heat rates of the recent vintages of turbines.
- *Large variation loads.* As explained above, many gas-fired generation units are used for servicing peaking loads and this means that they fluctuate between zero and a maximum level of fuel requirements. This does not mesh well with pipelines that are designed to provide uniform service and rely on operational procedures developed to optimise performance based on expectations of ratable takes.

The biggest potential problem is that a big gas-fired unit can drain line pack very quickly, once started, and this may reduce its ability to operate for an extended period of time. In many cases, pipeline pressure may drop quickly below the levels at which the plant can continue operating. For example, a modern combustion turbine unit of 500 MW can exhaust the line pack of a 36" line operating at 800 psi pressure in about four hours.<sup>21</sup> The lesson to be learned from this is that, in any valuation exercise, one has to evaluate carefully the technological specification of the pipeline supplying a plant (the diameter, operating pressure, age, operational procedures, etc).

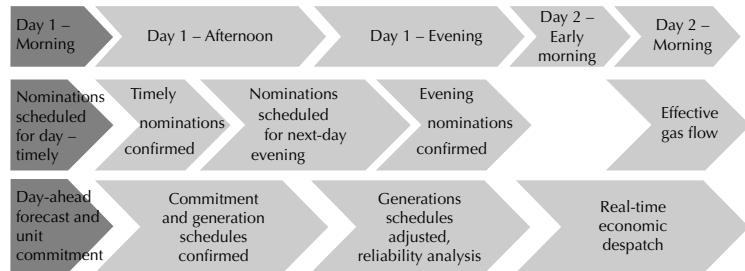
As explained in the chapters on natural gas, pipeline operations evolve around the gas day, which is defined as a 24-hour period starting at 9:00 am CPT. The electric day is from midnight until midnight. This means that not only communications between the operators of the two systems are difficult, but it is also problematic to synchronise the decision-making process. Pipeline schedulers have to make decisions without having full details of the actions the power pool operators and electric utilities are likely to take for the following day. As explained in the NERC document, “[W]hile the completed electric utility plan identifies which electric units will run the next day, which in turn provides the basic information to project the next day’s fuel consumption, the pipeline deadlines for nominations historically have been at 11:30 am CPT of the current day (the day before the gas flows). Thus, there is a time gap, possibly up to eight or more hours, of incompatibility between the two traditional approaches to planning and scheduling.”<sup>22</sup> As shown in Figure 22.2, a gas supply based on nominations completed by 11:30 CPT is the most reliable. Subsequent revisions to nominations – required by, for example, weather conditions or plant outages – are less certain.

Contractual arrangements with the pipelines are equally important, and any analyst engaging in a valuation exercise should ask if a plant has connection to an interstate or intrastate pipeline. The interstate pipelines have to offer services at tariff rates and it is important to review the tariff. Tariff design changes over time as the pipeline responds to the needs of their customers. Examples include additions of swing optionality, the ability to increase volumes and more frequent intraday nominations. A very important development has been the growing popularity of no-notice service (explained in the chapter on pipeline transportation), which tends to be expensive. The intrastate pipelines have much wider contracting flexibility that may translate into the ability to extract more value from a power plant. Negotiating a contract with a gas supplier often requires help from an experienced law and/or consulting firm. Ignoring this obvious constraint in financial analysis of a proposed power plant is a major flaw.

### **Supply stack**

The supply stack (or generation stack) is a tool used by electricity trading desks to develop projections of electricity prices over short-

**Figure 22.2** Electric planning and operations days for gas and electric planning



Source: [http://www.nerc.com/files/Gas\\_Electric\\_Interdependencies\\_Phase\\_I.pdf](http://www.nerc.com/files/Gas_Electric_Interdependencies_Phase_I.pdf)

and medium-term horizons, given the information about the technological characteristics of the generation units, fuel prices and the level of anticipated load. The starting point in construction of the stack is a database containing information about the generation units operating in a given region and the heat rates. It is assumed that the power plants are dispatched based on the short-term marginal cost, calculated as the cost of fuel and marginal non-fuel costs (although this item is often ignored), typically abbreviated as VOM (variable operations and maintenance). In other words, the marginal cost for a given unit is equal to the fuel price multiplied by the technological heat rate plus VOM. The generation units are ordered with respect to the marginal cost in ascending order, resulting typically in an upward sloping supply curve, with segments corresponding to different types of power generation technologies. A typical stack begins with a segment corresponding to hydro units, followed by nuclear units, coal units, combined cycle plants and combustion turbines. Oil-fired units are at this point in time the most expensive. In practice, this clear separation of different types of generation may be less than perfect, with different classes of units overlapping to some extent.

In practice, the generation stack is used in conjunction with a load forecast. Market prices are set by the least efficient unit required to satisfy anticipated load. The stack is a very popular tool among energy traders, and the construction and daily updates of a stack are two of the main responsibilities of any fundamental analysis group. This does not mean that this tool is without flaws. As a matter of fact, given the complexity of the electricity industry, this is a very blunt

tool that has to be used with extreme care. The limitations of the supply stack include the following.

- ❑ The supply stack ignores completely the spatial aspect of the power industry. A given region is treated as if intra-zone congestion and line losses did not exist and all units had unrestricted access to the transmission grid. In practice, some units may be located behind congestion points, barring them from injecting electricity into the system. Other units may be dispatched “out-of-merit” for voltage support or to replace units with no access to the grid.
- ❑ The dispatch model underlying the supply stack is excessively primitive. In practice, the decisions to dispatch a unit are not binary decisions (the unit runs or not) but often require deciding at what percentage of the rated capacity the unit should run and whether a unit should operate at some hours of the day at a loss to avoid costly shutdowns and frequent starts. For example, a coal unit typically runs for weeks and months unless there is forced outage. At night, the output of the plant may be reduced, even if the plant is generating at a loss, given that shutting down and restarting the plant incurs significant costs. The efficiency of the plant suffers and the plant equipment is exposed to thermal stress, which increases the maintenance costs. The same is true of combined cycle gas units, which are usually dispatched as base load units. Combustion turbines have much more flexibility.
- ❑ Some units may be offline due to forced or planned outages. Information about the status of different generation units is often unavailable or becomes available with a delay. The sources of information about unit availability are discussed in the section on the sources of fundamental information for the power markets. At this point we will just point out that some units may be offline not due to outages but because they are reaching the limits of allowed maximum hours of generation due to environmental considerations, such as emission limits. Incorporating this information on an on-going basis in a rapidly evolving market is not only technically difficult but also costly in terms of information requirements.
- ❑ The treatment of hydropower is the one of the most questionable aspects of the generation stack tool. Hydro units are treated as

units with zero marginal cost and represent the least-expensive segment of the supply portfolio. In practice, water behind the dam cannot be treated as if it were free. The value of water should be derived from an inter-temporal model balancing the current proceeds from generating electricity against the opportunity cost – ie, the reduced ability to produce electricity in the future. The discussion of a quantitative model supporting hydro-dispatch decisions is deferred until the next book.

- ❑ No US region (except for Hawaii) is a perfect island when it comes to electricity generation. Even ERCOT is connected by a few (low-capacity) DC lines to other parts of the US. Exports of electricity are treated in the stack model by creating a “dummy” generation unit with a certain implied marginal cost.
- ❑ The stack model ignores the potential for strategic behaviour of power plant operators, bidding into the pool at the levels diverging from their marginal costs. The chapter on market gaming covers this topic in more detail.

## DATA SOURCES

Electricity trading is one of the most information-intensive activities in the commodity markets due to the sheer volumes, variety of types and data sources, and complexity of the models required for the interpretation of the information. Traders and fundamental analysts rely on multiple data sources, some of which will be reviewed in this section. Complete coverage is next to impossible, given the variety of available publications, databases and news organisations. In principle, every website of an energy company is a potential location from which useful information can be extracted by a clever analyst.

A word of caution has to be offered at this point. Data collection is a sterile and useless exercise unless a formal network exists to draw conclusions from the hundreds of thousands of new data records becoming available daily. A pile of bricks is not a home, and megabytes of disconnected information are not knowledge. The ultimate objective of fundamental analysis is to make statements about future energy prices, with the forecasting horizon extending from a few hours to many years. This effort, as sophisticated as it is, still follows the century-old template proposed by Alfred Marshall – prices are determined at the point of intersection of supply (generation stack in the industry jargon) and demand curves and

fundamental models are designed to approximate such curves at the locations where power is traded.

The fundamental information for the US markets is very uneven quality, with many blind spots, which exist for a number of reasons. A significant part of the data is collected by different government agencies, using machinery and procedures designed in the past and applicable to a system that no longer exists. In many cases, these procedures have become obsolete but have not been updated. We have documented many such problems in other parts of this book. They are primarily due to budgeting problems and the reluctance of the US Congress to appropriate necessary funds to overhaul the statistical systems of the government agencies. The budget-cutting zeal demonstrated by many members of Congress does not help in funding efforts to improve the reporting of economic data, which is quite important to the private sector. This has, in our view, a long-term impact on market efficiency and causes a significant, although difficult to quantify, impairment of market mechanism. The agencies involved in data collection include FERC (covering investor-owned utilities, IOUs, and other jurisdictional entities), complemented by EIA in the case of entities outside FERC's jurisdiction (independent power producers, IPPs, cooperatives). Other important sources of data, some of them covered in this section, are DoE (for international power flows across the US borders), EPA (for emission data), NERC (for reliability related issues) and the Department of Agriculture, for cooperatives that receive financing from Rural Utilities Services (RUS). Different state agencies (primarily public utilities commissions or equivalent bodies) are important for local distribution and retail business. Financial filings with the SEC for IOUs, IPPs and other electricity related businesses that are publicly traded, contain real gems of information.

ISOs/RTOs are the second important category of data sources, with very high quality and timeliness of data. The information available in real time includes prices (day ahead and real time), load and system-related data. The challenge arises from the differences in power pool designs that make it difficult to transport skills acquired in working with one power pool to another location. Finally, one cannot ignore a large number of highly innovative companies which specialise in the collection and delivery of data used by power traders. Some companies organise and distribute information

collected by the government, others generate information through their network of industry contacts and other means.

As always, it is impossible to cover everything (it would make the book a few thousand pages long) and we have to exercise judgement in the selection of material. Different types of information can be classified as follows:

- generation units and their technological characteristics;
- operational status of generation units;
- transmission lines status;
- load information;
- spot and forward prices of electricity;
- spot and forward prices of fuels;
- hydro conditions; and
- regulatory developments.

### **Generation units**

The information related to capacity and technological characteristics is collected by FERC and EIA, and is usually identified through the names of the forms used by respondents to report the data. Different forms have usually a long and convoluted history of evolving scope and of different institutions responsible for the administration of the programmes.

Form EIA-923, "Power Plant Operations Report," is used to collect information on:<sup>23</sup>

- receipts and costs of fossil fuel;
- fuels stocks;
- generation;
- consumption of fuel for generation; and
- environmental data.

This information is collected monthly and covers about 1,600 plants. Additional information is provided on annual basis by non-utilities (source and disposition of electric power and environmental data). About 3,300 generators who are not included in the monthly sample report data on the entire form (the form consists of schedules 1–8F).

The fossil fuel data (including type, cost and quality) are available at the plant level from independent power producers, electric

utilities, and commercial and industrial combined heat and power plants (CHPs) with nameplate capacity equal to, or exceeding, 50 MW.

The preliminary calculation of estimated emissions of NO<sub>x</sub> and SO<sub>2</sub> is accomplished through the application of an emission factor (average quantity of a pollutant assuming no use of emission controlling equipment) to fuel consumption data:

$$\text{Emissions} = \text{Quantity of fuel consumed} \times \text{Emission factor}$$

For the fuels containing sulphur, the formula is modified as follows:

$$\begin{aligned}\text{Emissions} &= \text{Quantity of fuel consumed} \times \text{Emission factor} \\ &\quad \times \text{Sulphur content}\end{aligned}$$

The emission factor depends not only on the type of fuel, but also on the combustion system and firing configuration (for example, cyclone boiler versus fluidised bed boiler).<sup>24</sup> Preliminary estimates are adjusted to account for the plant's pollution abatement equipment, for which information is available from the EIA-860 survey. These modified estimates are replaced with actual emission for the plants for which detailed emission information is available from the Continuous Emission Monitoring System (CEMS) database (to be discussed shortly). The estimates are used to allocate emissions by different source fuel, as this breakdown is not available from the CEMS database. CO<sub>2</sub> emission information is based on the data regarding fuel consumption and heat content of fuel, extracted from other sections of form EIA-923.

The task of estimating emissions is a complicated one, due to inconsistencies of data collection across different reporting systems. Many plants use different flue gas desulphurisation (FGD) equipment and low-NO<sub>x</sub> burners, and selective catalytic reduction (SCR) systems for reducing SO<sub>2</sub> and NO<sub>x</sub> output. Control equipment information is available for the units that reported using the Form EIA-923 (the history is available from EIA-767).

CEMS is a real treasure trove of information that the industry does not use to its full potential. The data is collected under Title IV of Clean Air Act Amendments of 1990 (CAAA), which requires electric utilities to monitor and report hourly NO<sub>x</sub>, SO<sub>2</sub> and CO<sub>2</sub> emissions and heat input to the EPA. Data collection requirements vary depending on the generation unit characteristics. Coal units have to

monitor physical emissions of CO<sub>2</sub>; oil- and gas-fired units can use estimates (the use of physical equipment is optional). Formula-based estimates rely on the measurements of hourly fuel flows. CEMS data can be downloaded from: <ftp://ftp.epa.gov/dmdownload/emissions/hourly/monthly/> (this character string should be entered into Windows Explorer).

CEMS data files contain following information:

- STATE
- FACILITY\_NAME
- ORISPL\_CODE
- UNITID
- OP\_DATE
- OP\_HOUR
- OP\_TIME
- GLOAD
- SLOAD
- S02\_MASS
- S02\_MASS\_MEASURE\_FLG
- S02\_RATE
- S02\_RATE\_MEASURE\_FLG
- NOX\_RATE
- NOX\_RATE\_MEASURE\_FLG
- NOX\_MASS
- NOX\_MASS\_MEASURE\_FLG
- C02\_MASS
- C02\_MASS\_MEASURE\_FLG
- C02\_RATE
- C02\_RATE\_MEASURE\_FLG
- HEAT\_INPUT

Most of these acronyms are self-explanatory. ORISPL stands for ORIS (Office of Regulatory Information Systems) Plant Location number – ie, EIA Plant Code for power plants (facility); OPTIME is unit operating time, GLOAD is gross unit load during unit operation (MW), SLOAD is steam load during unit operation and HEATINP is the hourly heat input rate during unit operation for all fuels (MMBtu/hour).

CEMS data and its importance cannot be estimated. In addition to

providing information about emissions, as envisaged in the design of this system, can be obtained invaluable insights into the operations of power plants, including scheduled and forced (involuntary) outages as well as dispatch patterns. An example is provided by a study showing the dispatch, as implied by a quantitative model, and by actual operations as inferred from the historical CEMS data. It compares the hypothetical and actual dispatch of specific generation units located in the Pacific Northwest.<sup>25</sup> The hypothetical dispatch diverged significantly from the empirical data. This study is a very good illustration of the deficiencies of many models used for decision support in energy trading: a tendency to ignore rigidities embedded in the physical assets and the costs of excessive cycling, restricting the ability to operate a plant through frequent starts and shutdowns.

Form EIA-860<sup>26</sup> is used to collect data on the status of existing electric generating plants and associated equipment (including generators, boilers, cooling systems and FGD systems) in the US, and those scheduled for initial commercial operation within five years of the filing of this form.<sup>27</sup> The respondents include all existing plants (or plants expected to be built within the next five years) with a nameplate capacity of 1 MW or more, which are connected to the grid, enabling them to draw/deliver power from/into the local transmission or distribution system. The reporting plants included facilities operated both by the utilities and non-utilities, such as independent power producers, combined heat and power and industrial installations. The data items reported in this form include:<sup>28</sup>

- initial date of commercial operation;
- prime movers;
- generating capacity;
- energy sources;
- status of existing and proposed generators;
- proposed changes to existing generators;
- county and state location (including power plant address);
- ownership;
- FERC qualifying facility status; and
- ability to use multiple fuels (data on co-firing and fuel switching are included).

Form EIA-861 is used to collect information related to peak load, generation, electric purchases, sales, revenues, customer counts and demand-side management programmes, green pricing and net metering programmes, and distributed generation capacity.<sup>29</sup> The information available is organised into six files, containing the following information:<sup>30</sup>

- ❑ “File 1. Aggregate operational data such as energy balance and revenue information from each electric utility in the country, including power marketers, and Federal power marketing administrations. File1\_CAO contains information on control area operators.
- ❑ File 2. Retail revenue, sales, and customer counts, by State and class of service, for each electric distribution utility, for all consumer provided fully bundled electric service (both energy and delivery service) by a single electric utility or energy service provider in all 50 states, the District of Columbia, the Dominion of Puerto Rico, and the Territories of American Samoa, Guam and the Virgin Islands.
- ❑ File 3. Electric utility demand-side management programmes, including energy efficiency and load management effects and expenditures.
- ❑ File 4. The names of the counties, by state, in which the respondent has equipment for the distribution of electricity to ultimate consumers.
- ❑ File 5. Aggregate data of the number of customers by state and customer class for green pricing and net metering programmes.
- ❑ File 6. Utility or customer-owned distributed generation capacity such as the number, capacity and types of generators, as well as the types of prime movers used.”

Information collected by FERC is distributed by a number of consulting firm and marketed as searchable databases. A word of caution is required here. The databases may contain sometimes information that is not correct and may result in serious errors when used uncritically in complex models, and aggregated in a way that makes it difficult to identify outliers. The errors result mostly from reporting mistakes and mistakes during the transfer of information to the databases.

---

### Forward prices

The availability and quality of the price information varies from markets to markets. In most power pools, price information for the day-ahead and real-time markets, as well as the prices of firm (financial) transmission rights, are available for download from the websites of RTOs at no charge.

Forward price information is available from internal sources (electricity trading desks) with external assessments of forward prices available from different sources. One of these is the Platts–ICE Forward Curve–Electricity (North America).<sup>31</sup> The forward prices are disseminated under an agreement reached in October 2007 between Platts and the IntercontinentalExchange. Under this arrangement, forward prices for electricity and natural gas are provided by ICE to Platts, which uses its market expertise to generate forward price curves. The information from ICE is combined with information collected from brokers, traders and back offices of companies active in the electricity markets. These three data sources are combined by Platts into an assessment of forward price curves. The forward prices are time-stamped as of 2:30 Eastern Prevailing Time, when the Nymex natural gas futures contract closes. This timing allows Platts to align forward prices of electricity with the prices of natural gas, in recognition of the importance of this fuel to price formation in the US power markets. These prices represent market-on-close assessment and may diverge from actual reported transaction prices as market conditions may evolve throughout the day. It is critical to recognise the role of judgement in arriving at the final posted prices.

The definitions of forward prices follow market conventions used in different regions of the US. On-peak forward prices in Eastern and Central markets represent  $5 \times 16$  packages ( $6 \times 16$  packages in Western markets), with NERC holidays excluded. The off-peak forward prices in the New England, New York, Ontario, PJM, MISO, ERCOT South, Into Entergy, Into Southern and Into TVA markets represent a  $5 \times 8$  plus a  $2 \times 24$  package, called a wrap.<sup>32</sup> Platts makes an effort to publish price assessments even in the absence of trades. This is done through the cross-calibration of prices at different locations and the use of calendar and locational spreads to determine prices for inactive future time periods and/or locations.

The forward prices are available for the prompt month, second

and third months, balance of the year by month, two full years (by season or month) and four calendar years.<sup>33</sup> Western markets forward electricity prices, both on-peak and off-peak, are available for the prompt month, second month, balance-of-the-year in quarters, two full years in quarters and four calendar years. The prices quoted for individual months and for the periods including these months are synchronised. As of January 2010, the forward prices were available for 48 hubs.<sup>34</sup>

The M2M system for generation of forward price curves supplements the system described above and is based on statistical analysis. Regression equations generate forward prices for 42 hubs from the forward curves derived from market observations, relying on relationships between spot prices and lagged forward prices.<sup>35</sup>

$$F_{tod}(T) = a + b \times S_{tod} + c \times F_{yest}(T) + \dots$$

where  $F_{tod}(T)$  is forward price as of time  $tod$  (today) with maturity equal to  $T$ ;  $F_{yest}(T)$  is lagged forward price,  $S_{tod}$  is a spot price as of today. Forward prices extending for 36 months out are produced daily. Forward curves 20 years out are produced on a monthly basis. M2M curves are classified as editorial M2M Hubs (same as those from Platt-ICE Forward Curves), market M2M Hubs (liquid forward trading but not covered as editorial), and proxy M2M Hubs (no or limited forward trading, liquid spot markets).

Argus services include assessments of forward curves at 28 different locations and extend a minimum of seven years into the future. The forward curves are constructed using monthly granularity, with additional calendar, seasonal and quarterly block pricing. Forward curves represent value as of 2:30 pm Eastern Time, the close of the CME natural gas futures contract. The rationale for using this convention is that the industry traditionally relies on more liquid natural gas markets to construct forward curves. Historically, gas-fired power plants were at the margin (ie, were setting market prices) in most US power pools. As in the case of Platts, Argus relies on multiple data sources to derive prices. In the case of illiquid markets or markets with no transaction information available on a given day, Argus uses cross-triangulation against other hubs or a proprietary quantitative model. In addition to pricing forward price curve assessments, Argus offers daily electricity/natural gas correlation curves and heat rates for 26 electricity trading hubs.

### Bilateral transactions: Cash markets

Platts collects each-day information about fixed price and financial transactions for delivery of electricity in North America. The information required includes:<sup>36</sup>

- location (for example, tie point or congestion management zone);
- trade date;
- start flow date;
- end flow date;
- shape (peak or off-peak);
- deal type (physical or financial);
- firm or non-firm flag;
- price (US\$/MWh);
- volume (MW);
- buy or sell information flag;
- counterparty name; and
- intermediary name (broker or trading platform).

The price should not include estimated transmission costs to adjust for delivery at one location against the price at one of the standard Platts' locations (ie, actual locations should be reported, even if Platts does not publish a specific price index for a given location). Given sensitivity regarding the counterparty information, Platts accepts the submission without this specific detail, although it communicates its strong preference for having this information. The deadline for inclusion of the submitted price data, including forward packages, is 2:30 Eastern Prevailing Time (4:30 EPT is the deadline for sending the data). Deals executed after option expirations are included, as long as the prices do not diverge from those in other transactions.<sup>37</sup>

Daily index prices are calculated as volume-weighted average of reported prices (after eliminations of outliers). The reported items include the index price, the change from the previous day, the low, the high, the volume, the number of transactions the index is based on and the running average for the index price for the month. Index prices, lows and highs are expressed in US\$/MWh.<sup>38</sup>

For low-liquidity locations (fewer than five transactions or three counterparties), Platts publishes assessments that are based on differentials to other locations, physical bid-ask spreads, derivatives trading and other information. Near-term assessments (balance-of-the-week, balance-of-the-month and next-week) are reported as

a high-low range based on available transactional data (or as US\$0.50 range around the price if only one transaction is available for pricing).

Argus US electricity service provides daily spot prices (peak and off-peak prices, as well as gas and coal spark spreads). Markets covered by Argus include:

- East: NY Zone G; PJM West; NE Pool;
- ERCOT: Houston; North; South; West;
- Midwest: Cinergy; Northern Illinois; PJM AEP-Dayton;
- Southeast: Entergy; Southern; and
- West: California-Oregon Border (COB); Four Corners; Mead; Mid-C (Mid-Columbia); Mona; NP 15; Palo Verde; SP 15.

The indexes published by Argus are calculated as volume-weighted averages of physical trade data<sup>39</sup> for firm pre-scheduled day-ahead transactions, reported voluntarily by market participants. The indexes for low-liquidity markets are assessments based on consummated transactions, bids and offers, historical price relationships and other relevant information.

### **Power/fuel price spreads**

Spark spread is defined as the difference between prices of electricity produced in thermal power plants and the prices of fuel used for power generation. Given that electricity and fuel prices are expressed in different units, the prices of fuel are adjusted by a coefficient called a heat rate that represents the efficiency at which energy contained in the fuel is converted into electricity. The heat rate, as explained in other sections of this chapter, is measured in Btu/kWh or MMBtu/MWh.

$$\text{Spark spread} = \text{Electricity price} - \text{Heat rate} \times \text{Fuel price}$$

In calculating and reporting the spark spread, the critical issues are:

- the selection of a meaningful power price representing market conditions in a given region;
- the selection of a meaningful fuel price underlying the dispatch decisions made by the power plant operators; and
- the determination of a relevant heat rate.

These issues will be illustrated using the example of the spark spread, calculated using natural gas prices. Natural gas-fired power plants are at the margin (ie, represent the marginal cost of pricing electricity required to satisfy given demand) and, therefore, regional spark spreads are very closely watched by traders.

The electricity prices used by Platts are prices for next-day delivery (Monday prices in the case of Friday-dated transactions), published for each relevant market hub in Platts' *Energy Trader*. Natural gas prices are the next-day spot prices for the most active trading point critical to a given electricity market and published in *Energy Trader*.

Oil prices are represented by spot cargoes for 1% sulphur residual HP (High Performance) at New York port for the East region and 0.7% sulphur residual fuel oil at Gulf Coast ports for the Central region. Coal prices are used for calculations of the dark spread are next-month delivery bilateral prices. Prices used for New England, Mid-Atlantic, Southern and Midwestern hubs are those of Nymex-spec Appalachian coal.<sup>40</sup> For the Entergy, Chicago, Houston and Northwestern hubs prices of Powder River basin coal are used. The heat rates used in calculations are either based on the median heat rates calculated over three-year rolling windows or the lowest efficiency quartile of the power plants. Where warranted, an assessment of transportation cost is added to the fuel price.

### **Power plant outage information**

Information about scheduled and forced outages of power plants is available from a number of sources. Industrial Info Resources<sup>41</sup> provide information about past, current and expected outages (for scheduled outages) based on the information released by the power producers. Additionally, the company distributes to subscribers non-public outage information obtained by contacting the plants directly. An alternative is to subscribe to the service offered by Genscape,<sup>42</sup> which relies on remote sensors distributed around the country and monitoring electromagnetic fields around the power transmission lines. The primary objective is to infer from these data the condition of the generating units and, especially, unscheduled outages resulting in sudden changes of power flows, as the integrated transmission/generation system adjusts to the shock. As explained in one of the patents available from the company:<sup>43</sup>

A method for remotely monitoring the magnitude and direction of net electrical power and current flow to or from a facility over a prolonged period of time. The method includes detecting and measuring the magnetic field emanating from the monitored line(s), and detecting a signal synchronized to the power system frequency, typically the electric field, emanating from the power lines. The method further includes evaluating, storing, and transmitting the data on the electromagnetic (EM) fields that are recorded.

The electronic sensors are located under the transmission lines within approximately 200 feet, outside the transmission line right-of-way. The information collected by the sensors is processed by Genscape and delivered to customers over their website. The information about the operational status of the power plants is also leveraged by the company into additional reports about emission levels<sup>44</sup> and coal usage.<sup>45</sup>

Other methods for monitoring power plants include techniques such as:

- infrared or conventional optical cameras connected to the Internet and located in the line of sight of power plant's smoke stacks;
- seismic devices monitoring vibrations from the generation turbines; and
- a proverbial grandmother on her porch armed with binoculars.

The obvious question is why power traders are willing to invest in this and similar services. In the 1990s (a few years before Genscape), Enron developed a system of sensors that monitored transmission lines and deployed optical and infrared cameras – at a considerable cost to the company. One can probably still find units abandoned by the company post-bankruptcy, rusting happily in the fields. The answer is that unexpected outages can shock not only the physical power markets but also natural gas markets. Transmission constraints tend to create areas with precarious balance between load and generation resources, and limited ability to import power. Sudden removal of a section of the generation stack may send power prices to very high levels as power marketers and distribution companies scramble to replace lost supply. In many cases, additional purchases may be quite small from the perspective of a large integrated utility and this translates into a readiness to pay a very high price for additional megawatt-hours. To a trader or a poorly hedged

power marketer, this price spike may spell disaster. A power plant outage may also propagate to other markets. A nuclear or coal unit going down may, for example, require replacement power from a natural gas plant – and this translates into an increase in short-term demand for this fuel. Entities who have direct access to plant-status information (for example, the local utility operating the plant) may quietly buy additional volumes of natural gas before other market participants can react. Access to the Genscape information eliminates this advantage and creates equality of conditions for all market participants.

Historical information about plant operations may be also obtained from the Generating Availability Data System (GADS) database. GADS Services, part of NERC, supports a cluster of databases containing historical information regarding the reliability of different types of electrical equipment. The reports produced by GADS include the Generating Availability Report (GAR), a related summary statistical brochure and a Historical Availability Statistics (HAS) report. HAS contains annual, five-year, 10-year and multi-year interval reports for 63 generator unit groups.<sup>46</sup>

### **Hydro conditions**

Hydropower plants are a critical component of the generation fleet in some regions of the US. Analysis of the hydro systems evolves around two critical issues:

- how much water will be available in the reservoirs for power generation or what will be the river flows; and
- how hydro power plants will be operated, given the water availability.

Hydropower plants represent a unique challenge for an analyst supporting a power trading desk. A difficulty arises from the level of complexity of most hydro systems and the multiple, and often conflicting, objectives that guide decisions with respect to system operations. Typical considerations include flood control, fish and wildlife management, power generation, navigation, irrigation, tourism and water quality. The decision process requires the involvement of many different state and federal agencies, and in many cases foreign governments, as water use is subject to interna-

tional treaties negotiated with Mexico and Canada.<sup>47</sup> The political, social and economic aspects of hydro system management are not codified and are learned through immersion and the accumulation of experience: by working for one of many different agencies and utilities in a region heavily dependent on electricity generated from water. An additional level of difficulty arises from the need to operate the hydro river system as an integrated entity and reconcile many conflicting objectives. A poor analyst with no such background has to start somewhere, and most data collection efforts evolve around weather forecasting.

The starting point is one of the National Weather Service Rivers Forecast Centers, with the most important being that covering US Northwest and parts of British Columbia. The website for the Northwest centre contains a wealth of information on the Columbia River system, including current weather conditions, short-term weather forecast and long-term prospects driven by weather anomalies, such as El Niño, La Niña and the MJO. The website contains also information about current operations of the system. An example of the rivers condition summary can be found at the referred website.<sup>48</sup>

Even a perfect weather forecast is only a starting point for the analysis. For a short-term forecast, one has to recognise multiple objectives that decision-makers try to reconcile. The long-term forecast has to recognise the complexity of the physical system. Having a perfect forecast of precipitation is not sufficient. The reservoirs are filled during the winter months and early spring. The Columbia River system is divided by the Cascades Mountain Range, which separates maritime areas with their moist and mild climate from the desert conditions in the interior. East of the Cascades, precipitation happens as snow, and the density of snow, the rate at which it melts, the run-off speed dependent on temperatures and soil conditions may result in very different outcomes for the entire system. Many traders rely on consulting firms offering proprietary hydrological models for the region.

The comments made for the Columbia River system apply in general to other regions with a heavy dependence on hydropower plants. Analysts face challenges of a similar level of complexity, with specific differences reflecting geography, climate and political and social pressures. In many cases, data collection efforts are very

ingenious, ranging from airplanes overflying reservoirs before critical reports are released through satellite telemetry and customised optical devices zeroed in on dams.

## CONCLUSIONS

Electricity is the most data-intensive segment of energy trading. This can be explained by the organisation of the electricity markets and the underlying technology. Prices are available at multiple locations (PJM has about 8,000 nodes) at an hourly frequency (and sometimes even at shorter time intervals). Analysis of the supply side requires the collection of data about thousands of generation plants, an understanding of transmission grid topology and the physics of electricity flows. The specific features of multiple power pools, as well as regional differences, make fundamental analysis both labour- and time-intensive. I can hardly think of any other market where a trader is so dependent on their support staff. The obvious advice is to hire highly qualified analysts and keep at any cost those tested by fire and able to prove their worth.

- 1 This is by no means a complete list. The author has no business relationships with the companies mentioned in this section.
- 2 "The GE Multi-Area Production Simulation Software (GE MAPS) is a detailed, chronological simulation model that calculates hour-by-hour production costs while recognising the constraints on generation dispatch imposed by the transmission system. GE MAPS performs a transmission-constrained production simulation, which uses a detailed electrical model of the entire transmission network, along with generation shift factors determined from a solved AC load flow, to calculate the real power flows for each generation dispatch." See <http://www.crai.com/consultingexpertise/Content.aspx?tID=828&subtID=842&tertID=896>.
- 3 "Ventyx PROMOD IV is a Fundamental Electric Market Simulation solution which incorporates extensive details in generating unit operating characteristics, transmission grid topology and constraints, and market system operations to support economic transmission planning." See <http://www.ventyx.com/en/enterprise/business-operations/business-products/promod-iv> for more information.
- 4 <http://www.powerworld.com/Default.asp>.
- 5 The AURORAxmp Electric Market Model was developed by EPIS, Inc (<http://epis.com/company/>).
- 6 Eike Roth, 2005, "Why thermal power plants have a relatively low efficiency" (<http://www.leonardo-energy.org/repository/Library/Papers/EfficiencyThermal.pdf>).
- 7 Eike Roth, op.cit. This is a good example of why one should never accept any reported number at its face value. The conventions behind any number matter.
- 8 [http://www.epa.gov/ttn/atw/utility/fnl\\_biomass\\_cogen\\_TSD\\_04\\_19\\_07.pdf](http://www.epa.gov/ttn/atw/utility/fnl_biomass_cogen_TSD_04_19_07.pdf).
- 9 We ignore in this example a variable O&M cost.
- 10 See, for example, Susan F. Tierney, 2012, "Why coal plants retire: Power market fundamentals as of 2012," Analysis Group, Inc, February 24.
- 11 See National Petroleum Council, 2007, "Topic paper #4. Electric generation efficiency," July 18. The reader is encouraged to consult this document.

- 12 As explained earlier, this cost in most cases tends to be rather insignificant (by a rule of thumb, it is assumed typically to be about US\$2–3/MWh).
- 13 A spark spread calculator is available at [http://www.cmegroup.com/tools-information/calc\\_spark.html](http://www.cmegroup.com/tools-information/calc_spark.html).
- 14 [http://www.caissedesdepots.fr/fileadmin/PDF/finance\\_carbone/document\\_methodologie\\_tendances\\_carbone\\_en\\_v4.pdf](http://www.caissedesdepots.fr/fileadmin/PDF/finance_carbone/document_methodologie_tendances_carbone_en_v4.pdf). One has always to consult the technical notes for every spark spread table. The conventions vary from publications to publication and market to market.
- 15 As explained in the chapters on natural gas, the energy unit used for natural gas in continental Europe is MWh or kWh.
- 16 Amsterdam–Rotterdam–Antwerp, an energy trading hub in Europe.
- 17 Increasing load on a generating unit at a rate called the ramp rate (see [http://www.teainc.org/glossary\\_qr.html](http://www.teainc.org/glossary_qr.html) for more definitions).
- 18 “A primer of the natural gas and electric power interdependency in the United States,” December 2011 ([http://www.nerc.com/files/Gas\\_Electric\\_Interdependencies\\_Phase\\_I.pdf](http://www.nerc.com/files/Gas_Electric_Interdependencies_Phase_I.pdf)).
- 19 The details can be found in the cited NERC document, p 85. For example, a 500 MW power plant is equivalent to 116,000 residences (on a zero HDD basis).
- 20 Fuel gas skid is a conditioning system for the treatment of turbine or combustion turbine fuel gas. The treatment may include decontamination, heating and pressurisation.
- 21 Ibid, p 89.
- 22 Ibid, p 98.
- 23 EIA, “Electric Power Annual 2008.” Form EIA-923 replaced in 2008 EIA-423 Form and FERC Form 423. FERC Form 423 covered about 600 regulated plants. Form EIA-423 was implemented to fill the gap in data collection created by transfer of many regulated plants to the non-utility sector.
- 24 See Appendix A2 and A3 to “Electric Power Annual 2008.”
- 25 Maury Galbraith, 2008, “Update: Modeled operation of combined-cycle combustion turbines in the Pacific Northwest, Northwest Power and Conservation Council.” September 25 (available from [www.nwcouncil.org](http://www.nwcouncil.org)).
- 26 New Form EIA-860 includes Schedule 6 for the collection of data previously complied using form EIA-767.
- 27 <http://www.eia.doe.gov/cneaf/electricity/forms/eia860.pdf>.
- 28 <http://www.eia.doe.gov/cneaf/electricity/page/eia860.html>.
- 29 <http://www.eia.doe.gov/cneaf/electricity/page/eia861.html>.
- 30 F861 File Layout-2007\_rev.doc downloaded from the website listed in the previous footnote.
- 31 [http://www.platts.com/IM.Platts.Content/MethodologyReferences/MethodologySpecs/platts\\_ice\\_forwardcurve\\_power.pdf](http://www.platts.com/IM.Platts.Content/MethodologyReferences/MethodologySpecs/platts_ice_forwardcurve_power.pdf).
- 32 Eight-hour off-peak blocks on weekdays plus two 24-hour blocks on weekends.
- 33 The usual breakdown for two years following the current is January–February winter package, the March–April spring package, May, June, the July–August summer package, September and the fourth quarter.
- 34 The list is available at <http://www.platts.com/Products/m2mpower/CoverageDetails>.
- 35 Anguel Grigorov, 2011, “Introduction to Platts forward curves,” presentation on August 31.
- 36 “Methodology and specifications guide: North American electricity” ([http://www.platts.com/IM.Platts.Content/MethodologyReferences/MethodologySpecs/na\\_power\\_method.pdf](http://www.platts.com/IM.Platts.Content/MethodologyReferences/MethodologySpecs/na_power_method.pdf)).
- 37 The issue here is that the prices in deals executed after options expire may represent abnormal market conditions (distressed situations of either buyer or seller). The difficulty in identifying such deals is that many counterparties do not report time stamps.
- 38 Ibid, p 4.
- 39 Swaps, contracts for difference and derivative-linked deals are used for the NY Zone-G and California markets.

- 40** See the section on coal for the contract specification. Appalachian coal has higher heat content (around 12,000 MMBtu/lb) and load sulphur content <1%. Western coal has lower calorific value (8,500 Btu/lb), lower sulphur content (<0.5%).
- 41** <http://www.industrialinfo.com/index.jsp?qiSessionId=CC0C69FA5722C007597E4E00D8B9F304.wolf>.
- 42** As explained on the company's website, "Genscape Inc. is the originator of real-time power supply information to support decision making for power marketers, regulators, utilities, distributors, and other energy market participants. The company was founded in 2000 in Louisville, Kentucky and expanded to the European market in 2004.[...] In December 2007 Genscape acquired Enva, the leading provider of (real-time) power market intelligence. [...] Genscape's parent company is DMG Information, the business information division of the Daily Mail and General Trust." See <http://www.genscape.com/pages.php?uid=2>.
- 43** Patent No. US 6,714,000 B2, March 30, 2004.
- 44** "Weekly Emissions Report" tracks SO<sub>2</sub> and NOx emissions. The report contains additional information about the YTD balance of emissions versus allowances held for SO<sub>2</sub> and NOx, emissions held by the parent company, a weekly ranking of the top 25 companies by volume transactions and YTD trades by vintage for SO<sub>2</sub> and NOx.
- 45** "Coal Burn Report" summarises coal usage by the US power plants on a weekly basis.
- 46** <http://www.nerc.com/page.php?cid=4%7C43>.
- 47** For example, the 1944 treaty with Mexico to share the waters of the Rio Grande, and the Columbia River Treaty (1961) and Protocol (1964) with Canada.
- 48** [http://www.nwrfc.noaa.gov/river/river\\_summary.php?sort=s1&ss=name&ss=COLUMBIA](http://www.nwrfc.noaa.gov/river/river_summary.php?sort=s1&ss=name&ss=COLUMBIA).

## *Electricity Markets Transactions*

Electricity markets cover a number of specialised products and services that are largely invisible to the public. An average consumer of electricity is naturally concerned with the price of kWh, but does not realise how complex is the machinery that delivers energy to their house with an exceptional level of reliability (at least in the developed countries). We find it sometimes surprising that most people in Western Europe and the US seem to assume that electricity has been available forever. In reality, electricity has become available relatively recently in many places. The author's mother remembered the year (1926) when power lines were extended to the country house in central Poland where she grew up. An executive of one of the largest utilities in the US told me about his experience of growing up as a boy without electricity in rural Texas. The Hill Country in central Texas was eventually electrified in the late 1930s and early 1940s through the efforts of Lyndon Johnson. Prior to electrification, the lives of farmers in this part of the US were quite difficult:<sup>1</sup>

Running water was impossible without electricity. Hill Country residents, therefore, depended upon water wells for their needs. By their thirties, women had stooped shoulders from hoisting four gallon buckets of well water weighing thirty-two pounds apiece up from average depths of fifty to one hundred feet. And everything, whether for house or field, required water. The average rural family in Central Texas went through two hundred gallons of water a day – 73,000 gallons a year (weighing 300 tons).

Sadly, this is still the plight of hundreds of millions of people in developing countries.

This chapter will cover the US markets for a number of different electricity-related services and products. We start with a discussion of ancillary services, followed by capacity markets, before reviewing the most common types of transactions in the energy markets (ie, the

markets for megawatt-hours). Another segment of the power industry covered in this chapter will be the market for transmission rights.

### ANCILLARY SERVICES

Ancillary (auxiliary) services are a highly specialised and poorly understood niche of the power markets. They represent the specialised operations of power plants, transmission networks and power pool organisations that are necessary to support the smooth, uninterrupted operations of the system. The FERC defines ancillary services as:<sup>2</sup>

those services necessary to support the transmission of electric power from seller to purchaser given the obligations of control areas and transmitting utilities within those control areas to maintain reliable operations of the interconnected transmission system.

Provision of these services may require investment in specialised equipment (for, example, power plants with special characteristics), operating existing equipment in a special way and developing decision-making capabilities and procedures by the power pools in order to support the operations of the entire system.

A broad overview of ancillary services that is not specific to power pools is difficult because the industry has not developed a consistent set of terms for these services. The same (or very similar) services may be provided under different names by different pools, or the same names will have different meanings across geographical locations. An example is provided by Yann Rebours and Daniel Kirschen:<sup>3</sup>

In general, reserve can thus be defined as the amount of generation capacity that can be used to produce active power over a given period of time and which has not yet been committed to the production of energy during this period. In practice, different types of reserve services are required to respond to different types of events over different time frames. In particular, while the term "spinning reserve" is widely used in literature, this service can be defined in different ways. This may lead to some confusion.

In some pools certain ancillary services will be considered to be a part of fundamental market design, while in other pools they may be offered as specialised functions from outside of the basic system. Within the same power pool, certain services may be bundled differently when offered to end users or when purchased from providers.

Ancillary services may be provided through a market based auction mechanism, may be cost-based (ie, paid for using prices based on the cost-plus principle), may be self-provided or acquired through bilateral transactions under obligatory targets issued by the pool.

The FERC identified the following types of ancillary services:

- reactive power and voltage control;
- loss compensation;
- scheduling and dispatch;
- load following;
- system protection; and
- energy imbalance.

Other documents and articles identify a larger number of ancillary services that amount to as many as 19 different categories. One possible way to avoid theorising through classification is to identify the fundamental reasons for ancillary services and use them as a map through the maze of competing lists.<sup>4</sup> These reasons can be categorised as follows.

- Load following. The supply and consumption of electricity has to be balanced at any instant (given that electricity is not storable), but load (defined as the sum of electricity requirements of all customers connected to the system) is subject to random, short-term variations that are often unpredictable. If we ignore the possibility of shutting down load, the solution is the development of a generation system that has a sufficient degree of flexibility and ability to react quickly to load variations. Some generation units with the required technical characteristics will be designated, by fiat or through contractual arrangements, to adjust output quickly to the fluctuation of load. Over a very brief period of time, the load fluctuation may be handled through the inertia of the integrated generation/transmission system.<sup>5</sup> The generation units equipped with a governor will adjust automatically to load fluctuations. Over longer time periods (one minute to 30 minutes), certain units must increase output or come on line at the request of the system operator. The reaction length requirements vary from power pool to power pool and result in different specifications for ancillary services. The reserves for

load that follow are often referred to as spinning reserve, although the terms used vary from pool to pool and author to author.

- ❑ Maintaining frequency within required limits (ie, within a narrow bound of around 60 or 50 Hz, depending on the country). The frequency is maintained if the load is equal to generation (ie, if the generation and imports of power by a given control area are equal to the load, exports and losses). Should the frequency exceed 60 Hz, this would be defined as an excess generation, and as a consequence the control area would become an inadvertent exporter of electricity. Also, the generators would momentarily speed up. The opposite happens if load exceeds generation. These two aspects of imbalances are captured in the so called Area Control Error (ACE), given by:<sup>6</sup>

$$\text{ACE} = (\text{NI}_A - \text{NI}_S) - 10 b (F_A - F_s) - I_{ME}$$

where  $\text{NI}_A$  stands for the net interchange actual,  $\text{NI}_S$  stands for the net interchange scheduled,  $b$  is frequency bias ( $<0$ ),  $F_A$  is frequency actual,  $F_s$  is frequency scheduled and  $I_{ME}$  stands for meter error. NERC formulated two criteria for ACE: A1 and A2. Under A1, ACE must return to zero within 10 minutes of the previous zero. Under A2, averaged ACE for each 10-minute period must be within limits.

- ❑ Operating reserves. Operating reserves may be looked at as the equivalent of load-following services for imbalances (as described above) caused by unexpected changes in generation (such as outages of generation units or transmission lines).<sup>7</sup>
- ❑ Energy imbalance. This service corresponds to an inability to match generation and load perfectly in a given service area, with such condition leading to more or less inadvertent exchanges (ie, unexpected electricity flows between two areas, causing potential disruptions in power system operations). The market rules should define the acceptable band (ie, levels that can be routinely tolerated by system operators) for such exchanges and a compensation mechanism to provide market settlements for such exchanges.<sup>8</sup>
- ❑ Reactive power and voltage support. Reactive power and voltage support is the most complex service and the least understood in the industry. This service received a lot of attention after the

August 2004 outage which affected the US North East and a part of Canada, which was partly attributed by the joint US Canadian task force investigating the event to a shortage of reactive power, as its supply failed to increase with the system load.<sup>9</sup> FERC issued a report on this topic<sup>10</sup> in 2005 in anticipation of a technical conference on the subject. The discussion following the report cast light on the fundamental difficulty of handling the supply of reactive power under a market-based system. This shortcoming of a market-based systems favours a reliance on cost-based solutions to assure a sufficient level of supply of reactive power. "Cost-based" in this context means compensating the market participants for the costs (plus approved profit margin) of providing required amounts of reactive power. One can ask, of course, the question why the supply of both types of power could not be controlled through the price system. The answer to this requires repeating a few basic facts about reactive power.

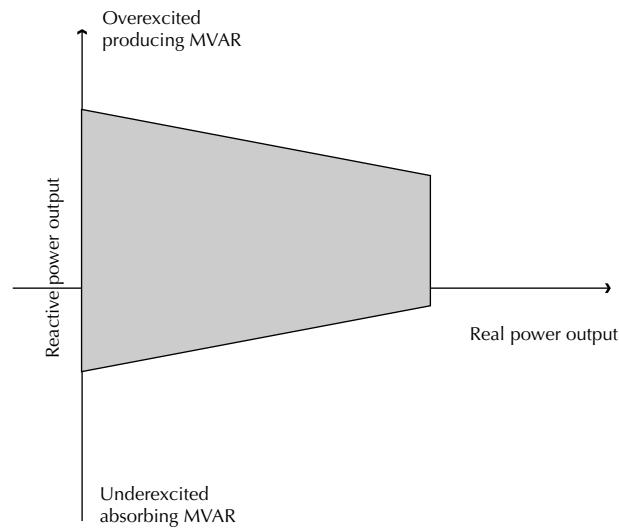
Reactive power provides the support to the electromagnetic fields that transport alternating current through the transmission and distribution grids. It may be both supplied and consumed by three components of the electric system: generation, transmission and load, which may switch from positive to negative output, depending on the conditions of the network. Its supply is critical to providing voltage support at different locations; however, at the same time, it cannot be effectively transported over longer distances due to high losses, exceeding those of real power. This means that the local supply of reactive power near load centres is critical to alleviate transmission constraints.

Reactive power can be produced as either static or dynamic. Dynamic reactive power is produced by generators, whereas static reactive power comes from capacitors. In the case of the generators, one has to recognise a trade-off between reactive and real power, summarised by the generator capability curve. A generator consists of a stationary part, called an armature, and a moving part called a field or a rotor. Both parts have slots through which coils of wire are pulled. The rotor is an electromagnet (a natural magnet would not create a magnetic field strong enough for any practical applications), with the power source called an exciter (this creates another reason why most power plants need an external supply of power to

start). The physical limits of a generator's output are related, including to the thermal limits of the stators and the rotors (overheating to either can not only degrade performance but also cause an outage) and the capacity of the mover (turbine). The generators can both supply and consume reactive power, depending on their mode of operation. The generators supplying reactive power have a lagging power factor and operate in the overexcited mode. In the under-excited mode, a generator has a leading power factor.<sup>11</sup>

Figure 23.1 shows a typical generator capability curve. The generator capability curve describes the trade-offs in producing real and reactive power by a given generation unit. Typically, a generator will operate somewhere close to the boundary of the shaded area to avoid inadvertent equipment damage. One has always to weigh a small increment in output against the potential cost of repairing a generator (measured in millions of dollars). As one can see, at the branch of the curve determined by the armature heating limit, one can obtain a significant increase in the output of reactive power, at no or at a small cost to the output of real power. If a generator operates inside the capability frontier, an increase in output of reactive power can be obtained at practically no cost. For generators operating at the heating limit, an increase in reactive power output will require a

**Figure 23.1** Generator capability curve (stylised representation)



reduction in the active power production. This explains why most schemes for pricing reactive power incorporate an opportunity cost in the compensation of reactive power units.

The difficulty in designing a market-based system for reactive power can be explained by the dependence of the capability boundary on the system condition. This means that the boundary may vary with the system load and available generation (and the configuration of transformers and other compensation devices). Given these externalities, it is difficult to design a market-based incentive system that would be non-discriminatory (ie, would not reward or penalise generators for the circumstances outside their control).

The FERC document listed in endnote 10 of this chapter reviews the history and potential approaches to pricing reactive power. A detailed review of this approach, country by country, pool by pool, is beyond the scope of this book.

## CAPACITY PAYMENTS

Under the traditional vertically integrated utility model, capacity expansion was handled through the regulatory process and adjustments to tariff rates. Utilities were obligated to carry out periodic assessments of expected demand and required investments in generation and transmission capacity. Investments approved by regulators would be included in the rate base, which, in conjunction with approved rate of return, would determine the utility's revenue stream and after-tax profit. This approach, which evolved gradually over the first few decades of the 20th century and was finally completed in the US in 1935,<sup>12</sup> worked reasonably well in the US and Western Europe, and end users took relatively inexpensive and reliable power supply for granted. If anything, the utilities were being often accused of excessive investments in the quest to increase the rate base. Deregulation created the need to address the issue of funding investments in new capacity, and we can now see a number of different solutions applied across the world under the system of capacity payments.

Capacity payments represent a relatively obscure but important part of the power markets. They are also quite controversial, as many industry theoreticians and practitioners believe that the market should be based exclusively on payments for energy, but not for

capacity. Under this solution, prices would be equal to marginal cost, which, with some simplification, can be assumed to be determined by the fuel cost plus variable O&M cost incurred at the last unit required to satisfy the demand. Of course, under such market architecture, many generation units would not generate (most of the time) sufficient revenues to cover fixed costs (the cost of capital, taxes, administrative overhead, etc). The resulting revenue shortfall would then be recovered during the periods of price spikes corresponding to the periods of high demand and unexpected shortages of generation capacity.<sup>13</sup> This market design often clashes with the public expectations of reliable supply of electricity at stable prices. The public preference for a market design that shields consumers from short-term fluctuations of electricity prices (ie, spikes) destroys the logic of a market-based system for resource allocation. The smooth functioning of the system requires that price signals about relative scarcity of different commodities be transmitted through the price system to both consumers and producers. The practical difficulty of assuring public acceptance of the high price volatility of a product treated as a public good, and an entitlement, has created the need for alternative market solutions.

A two-tier price system, based on (i) closely monitored and controlled energy prices, and (ii) capacity payments, reconciles the reliability concerns of the public and sensitivity to political pressures. This solution has a long tradition in the electricity industry, and has its roots in the traditional approach to rate design used under most regulatory regimes. This solution has also a theoretical justification. A branch of normative economics known as welfare economics formulated a number of prescriptions regarding the optimal allocation of resources, with equality of prices and marginal costs being the cornerstone of its recommendations. In industries that are capital-intensive, such as the electricity business, the marginal cost is likely to be below the average cost and implementation of this rule would result in a deficit. A deficit could be covered through taxation and the revenues allocated to money-losing industries through lump-sum subsidy payments.

Capacity payments are an attempt to address the fundamental conflict between the market-based provision of what is perceived by most citizens as the public good and a birthright – electricity, the critical form of energy in a post-industrial society – based on the

processing of information. Consumers want reliable service but do not want price volatility. These two objectives may be mutually inconsistent: price spikes may be necessary to create incentives for investments in generation and guarantee reliability. Capacity payments represent a compromise. Practical implementations of this solution through regulatory decisions are imperfect by definition: they represent an effort to reconcile the sometimes-conflicting objectives of economic efficiency and social justice. Once the imperfection of adopted solutions become obvious, patches and revisions are applied. In most cases, they represent a design by a committee<sup>14</sup> and lead inevitably to additional patches.

### **Market design issues**

The implementation of a capacity payments system requires making decisions about a number of market design issues. The most important questions to be asked are detailed here.

- ❑ In the case of an energy-only market design, what are the regulatory backstops if the market fails to produce a desired level of investment, and what are the tools, if any, the system operator can use to suppress price spikes seen as excessive?
- ❑ Who should receive the capacity payments? The most frequently debated demarcation line is between offering the capacity payments to all the generators, or only to new generation plants or to power plants about to be retired.
- ❑ What is the optimal solution with respect to the locational differentiation of capacity payments? In other words, should the capacity payments be uniform across the power pool or should they be differentiated across different sub-areas with limited transmission capacity across their borders? If the answer to the latter question is that the capacity payments should have a locational flavour, how should different regions be defined in practice?
- ❑ Should capacity payments be based on current generation-load balance or should they be forward-looking (ie, should they be based on anticipated imbalances predicted by forecasters)?

Detailed answers to these would require a separate book, but there are several excellent reviews of this subject a reader can consult.<sup>15</sup> We

shall make only a few comments on the most important aspects of these problems.

With respect to a willingness to tolerate price spikes, regulators seek to walk a line between the two extremes: unlimited price spike versus stable prices over long time periods. Excessive fluctuations in energy prices are politically unpopular, and it takes a regulator who is very committed to the principles of free markets to suffer the heat from end users of electricity exposed to wide price swings. Most energy-only power pools include provisions for price caps and ad hoc interventions. This, in turn, produces a problem known in the industry as missing revenues – ie, insufficient anticipated revenues to support expansion of the existing generation capacity or prevent retirements of the existing units.<sup>16</sup> Available regulatory backstops to counter generation capacity shortages include the emergency purchase of power from adjacent power pools, out-of-market capacity purchases, dispatch of units that were shut down for environmental reasons and, ultimately, rolling brownouts<sup>17</sup> and blackouts. Of course, such temporary measures do not address the basic issue of shortage of generating capacity. In the interest of fairness, one should not overlook the positive experience of some power pools based on the energy-only model, one being the power pool in Alberta, Canada.

Capacity payments to existing generators are very controversial and are often opposed by end-user coalitions. The critics point out that the owners of already-built generation plants may be interested in the perpetuation of conditions of scarcity (ie, insufficient size of installed generation capacity) in order to maximise the economic rents they receive in the form of capacity payments. In other words, if the generation fleet is expanded, capacity payments are likely to fall and this will hurt the owners of existing power plants. Making capacity payments only to new and retiring units (to keep them online) creates a different problem: existing generators may threaten to retire the units in order to qualify for the payments.

Spatially uniform capacity payments (ie, the same payments across a power pool, irrespective of location) are often criticised for failing to promote generation additions in the load pockets – areas of insufficient generation, with limited ability to wheel power from outside sources. This is a valid criticism. To use a purely hypothetical example, it is difficult to blame a developer who does not pursue

projects in Southwest Connecticut, the area with high concentration of wealth and political influence and with potentially significant local opposition, if capacity payments are the same across the New England power pool.

Table 23.1 is adapted from a source offering a useful classification of different solutions with respect to capacity payments.

*Example: PJM Reliability pricing model*

The PJM Reliability Pricing Model (RMP) represents a synthesis of a number of patches implemented across the industry in order to

**Table 23.1** Capacity payments: Alternative market designs

Description	Highlights	Examples
Energy-only markets	No capacity payments. No pure energy-only pools exist; the regulators use different tools to intervene in cases of capacity shortfalls.	Alberta, Australia NEM, <sup>18</sup> ERCOT (initial design), UK, Nordpool, Ontario
Energy markets with capacity payments determined through regulators/pool administrators	Capacity payments supplement energy-related revenues to cover the costs capacity expansion.	Argentina, Chile, Colombia, UK and Wales Power Pool, <sup>19</sup> Peru, Spain, South Korea
Energy markets with reserve requirements	Administratively determined reserve margins are met by load serving entities (LSEs) through bilateral contracting or investments in generation capacity	SPP, discontinued designs (NYPP, PJM, NEEPPOOL)
Energy markets with reserve requirements and centralised capacity markets	Centralised capacity market administered by power pool allows LSEs to meet capacity shortfalls.	PJM (prior to RPM), <sup>20</sup> MISO, NYISO, SWIS <sup>21</sup> market in Australia
Energy markets with forward reserve requirements	Reserve requirements defined and enforced on a forward-looking basis.	California
Energy markets + forward reserve requirements + centralised capacity markets	Centralised markets supports capacity requirements based on expected load.	ISO-NE FCM, <sup>22</sup> PJM RPM, Brazil, NYISO (proposed)

Source: Johannes Pfeifenberger, Kathleen Spees and Adam Schumacher, 2009, "A comparison of PJM's RPM with alternative energy and capacity market designs," Brattle Group, September.

address the defects of previous designs of capacity markets. It has a number of features discussed above, including:

- ❑ A *forward-looking definition of reliability*. Annual bid-based auctions allow the market participants to acquire capacity three years in advance. Base residual auctions<sup>23</sup> started in 2007 and covered successive 12-months periods beginning from June 2007. For example, the May 2010 auction covered capacity requirements from June 2013 through to the end of May 2014.
- ❑ The *capacity products have a locational flavour*. Each auction is preceded with determination of locational deliverability areas (LDAs) that are treated separately in the auction from the rest of PJM. Three commonly defined LDAs include:<sup>24</sup>
  - Eastern Mid-Atlantic Area Council (EMAAC): Atlantic City Electric, Delmarva Power, Jersey Central Power and Light Company, PECO Energy Co, Public Service Electric and Gas Company, and Rockland Electric Company;
  - Southwestern MAAC (SWMAAC): Baltimore Gas and Electric Company and Potomac Electric Power Co (PEPCO); and
  - MAAC: All of SWMAAC and EMAAC, plus Metropolitan Edison Company, Pennsylvania Electric Company and Pennsylvania Power & Light (PPL).
- ❑ The PJM RPM uses an administratively defined, downward-sloping demand curve. The curve is designed using the cost-of-service of the least expensive generation to build. The guiding principle behind it is capping total energy and capacity payments at the cost-of-service level.<sup>25</sup> This effectively amounts to reconstituting, at least by intention, the regulated electricity markets, with its concept of allowed return.
- ❑ Demand response is treated as a capacity resource. This means that measures undertaken to achieve permanent efficiency improvements and reduce electricity consumption count as increased generation capacity.

As of 2012, PJM's RPM received mixed reviews. In general, PJM emphasises the growing contribution of demand response to the auctions. Critics<sup>26</sup> point out the huge disparities between capacity prices at the LDAs and the rest of the RTOs. For example, in the 2013/2014 auction, the net capacity price<sup>27</sup> for RTOs was US\$27.73-

/MW-day, whereas the prices for the cleared LDAs defined for this auction were equal to 223.85 (MAAC), 240.32 (EMAAC), 236.93 (PEPCO). At the same time, the percentage of payments going to new generation was equal, respectively, to 5.4%, 9.5%, 5% and 0%. This seems to confirm the fears of the capacity markets critics who argue that this solution fails to trigger supply response and favours the incumbents – ie, the owners of existing generation fleets.

## FTR MARKET

Firm transmission rights (FTRs) are financial instruments with payouts based on the difference between two locational prices: the price at the sink and at the source (at the injection node). The FTRs have been designed to perform two critical related functions:

- redistribution of congestion fees; and
- hedging congestion risk.

A simple example illustrates the concept of FTRs. Suppose that an independent power marketer has a generation unit injecting power at node A, and a load obligation at node B. The locational marginal price at node A is US\$60/MWh, the locational marginal price at node B is US\$70/MWh. The generator receives the locational marginal price at node A, but has to pay the price at node B to cover their load obligations. In this example, the generator loses US\$10/MWh. The risk of decoupling of prices between these two nodes can be hedged with FTRs, which are available in different US and non-US power pools, but vary with respect to their design details and also the names under which they are known and which are often changed when a power pool is redesigned. Some of the terms used by these different pools are:

- financial (of firm) transmission rights (PJM);
- transmission congestion contracts (NY ISO);
- congestion revenue rights (CA ISO);
- financial transmission rights (NEPOOL, New Zealand); and
- transmission congestion rights (ERCOT).

FTRs can be defined as obligations that may produce either revenue or liability for the holder, depending on the definition of an FTR and

the direction of congestion – ie, the sign of the difference between the price at the sink and at the injection point. If an FTR is defined from point A to B, and the price at point B is higher than the price at point A, the holder will receive the congestion revenue. If the price at point A is higher than at point B, the holder pays a congestion cost.

The payout for an FTR defined from *injection* node to *sink* node is equal to:

$$\begin{aligned} \text{FTR MW} \times & (\text{Locational marginal price (sink)} \\ & - \text{Locational marginal price (injection)}) \end{aligned}$$

In our example, the payout per MW would be equal to US\$10 per hour. Should the direction of congestion be reversed, the revenue would be negative. For example, if the price at node A were equal to US\$55/MWh, and the price at node B to US\$50/MWh, a 1 MW FTR would result in a negative cashflow of US\$5 per hour to the holder.

If a prevailing flow FTR is acquired at a positive price: the buyer expects positive revenue from holding this right, after accounting for the cost of acquiring the instrument. An FTR with a negative purchase price (the buyer is paid for accepting the congestion risk)<sup>28</sup> is called a counterflow FTR. The buyer hopes that the actual congestion (ie, price differential) will be lower than expected. Many power pools offer FTRs defined as options, which will not result in a negative obligation to the holder.

FTRs can be defined as point-to-point contracts or as zonal contracts. Point-to-point contracts can be, in principle, defined for any two nodes of the system, even if there is no *direct* transmission line between the two specific nodes. In a power pool with n nodes, there are potentially  $n^2-n$  point-to-point transmission lines. One has to remember, however, that only  $n-1$  FTR contracts are independent<sup>29</sup> – ie, cannot be reconstructed as combinations of other contracts.

Zonal FTRs are defined in terms of zonal prices, representing averages of nodal prices,<sup>30</sup> appropriately defined. In zonal pools, the FTRs are defined and auctioned on a zone-to-zone basis. Sometimes the definition of the FTRs may be quite complex, as in the case of ERCOT prior to the transition from the zonal to the nodal system. The ERCOT market under the zonal system (in existence until November 30, 2010) was divided annually into congestion zones. These were defined in such a way that every load or generation

resource had a similar impact on transmission interfaces between the zones. A statistical technique called cluster analysis was used to identify homogeneous groups of nodes with similar characteristics. The main interfaces at which transmission congestion rights were defined were called commercially significant constraints (CSCs). Zonal generation weighted-average shift factors (explained earlier in this section) were used to manage congestion on the CSC, with every node within a zone being treated as having the same impact on the CSC. Transmission congestion rights were distributed through the annual and monthly combinatorial auctions.<sup>31</sup> Pre-assigned congestion rights (PCRs) were allocated to municipal utilities and cooperatives.

FTRs are very popular with hedge funds. This can be explained by the purely financial nature of FTR (no headaches related to physical delivery) and the ability to use quantitative trading strategies. The directional trades are designed using fast computers running complicated statistical and power flow models. The level of participation of hedge funds in the FTR markets is not without controversies. PJM executives complained about the high volume of bids submitted by hedge funds, overwhelming the computer systems operated by the pool. Some market participants complain that hedge funds are siphoning out the funds that should be distributed among physical market participants and used to upgrade the transmission and generation assets. One aspect of the FTR market is that it is perfectly transparent from the point of view of reported trades (and financial outcomes of the trades). This allows the markets participants to watch each other and study the strategies. PJM operates an FTR market and a related auction revenue rights (ARRs) market, enabling the distribution of revenues from the annual FTR auction to eligible parties. Both FTRs and ARRs can be used as a hedge against congestion and also as instruments to make speculative bets.

PJM FTRs are defined as obligations as well as options. The highlights of PJM FTRs include:

- ❑ FTRs are based on day-ahead prices;
- ❑ adjusted for marginal line losses (since June 1, 2007);
- ❑ available for any pricing node in the system, including hubs, control zones, aggregates, generator buses, load buses and interface pricing points;

- may be bought or sold or self-scheduled,<sup>32</sup>
- available as 24-hour, on-peak, off-peak financial products;<sup>33</sup>
- options are available for all the three product classes listed above;
- calculated hourly; and
- available at denominations to the nearest 0.1 MW.

FTRs were originally allocated by PJM to the market participants. Since June 1, 2003, the allocation mechanism has been replaced with allocation of ARRIs (to be explained below) and auctions for the rights with different tenors (ie, maturities). The long-term auction is carried out for three consecutive planning years (June–May) immediately following the planning year in which the auction is conducted. In addition, PJM conducts annual FTR auctions and monthly balance of planning period auctions. PJM facilitates also the secondary market that allows market participants to buy or sell existing FTRs outside the regular auction mechanism.

FTR auctions are carried out through an optimisation programme that maximises net revenues, using the offer-based values of FTRs. Nodal prices are determined as a function of FTR bids and binding transmission constraints. Each market participant is subject to credit risk requirements determined as a function of FTR cost and an estimate of potential congestion (a measure of historical congestion discounted by 30%, except for counterflow FTRs). PJM calculates the credit requirement for each participant.

The credit risk arises from the inability of the market participants to meet their obligations. This happens if the prevailing direction of the congestion on a given path reverses and the holder of an obligation has to make payments (receiving a negative cashflow is the same as paying cash out). Typically, the cases of default can be attributed to counterflow FTRs where the owner pays the congestion of a given path. The owner collects the payment for accepting this obligation, hoping that the price received will exceed the FTR collections.

ARRIs are not financial instruments but entitlements that can be converted to FTRs through the mechanism described above. In other words, ARRIs provide revenue to the firm transmission customers to offset purchase price of FTRs. The term “entitlement” means that ARRIs are not acquired through market transactions but are allocated

annually to network transmission service customers and firm point-to-point transmission customers, who receive an allocation of the revenues from the annual FTR auction. ARR revenues reflect the expectation of auction participants regarding locational differences in the day-ahead PJM energy markets. ARRs also are available only as 24-hour, seven-day a week obligation-type products. Long-term ARRs that cover 10 consecutive planning periods are available.

ARR entitlements are allocated to network transmission services customers (NIS or FPPS customers) through a three-round process. The ARRs are based on the historical load and service levels.

One of the challenges of the FTR market in PJM and elsewhere is revenue adequacy. The congestion fees collected by the power pool may become insufficient to cover revenues due to the holders of FTRs. This may happen for a number of reasons. One example is the purchase of FTRs on a transmission line of 500 MW capacity. If the line is subsequently de-rated to 450 MW, the congestion fees will be insufficient to meet the pool obligations with respect to this specific FTR contract. Revenue adequacy can become a serious problem in PJM. FTR underfunding in March and April 2010 was equal to 26% and 31%, respectively. This was caused by planned outages that were related to upgrades of the transmission lines. The paradox is that the actions taken by PJM to improve the reliability of the transmission system undermined the hedging efficiency of the FTR contracts in the short run.

## **POWER MARKET TRANSACTIONS**

### **US power pools market transactions**

Electricity-related transactions can be classified in a number of different ways. One frequent distinction is between straightforward, standardised transactions with a single underlying, and highly structured transactions. As always, the demarcation line between simple and complex transactions is in the eye of the beholder. Everybody can agree, however, that transactions requiring long gestation periods, both in terms of modelling effort and negotiations, extend over a long time period, are supported by long, laboriously prepared and highly customised legal documents, and result in a significant financial exposure, can be classified without any ambiguity as highly structured. In this book, we will cover the following types of complex deals:

- full requirements deals; and
- tolling transactions.

Other, more complicated transactions will be covered in the next book.

Another classification of these transactions is based on the underlying. On one hand, it is possible to distinguish between energy-related transactions, and, on the other, between ancillary services, transmission (congestion) and capacity deals. We have explained the design of the specialised non-energy markets (such as capacity markets and FTRs) in previous sections. Energy transactions can be classified in a number of ways, starting with time period for which the underlying price is defined:

- real-time (market settled hours or minutes ahead of delivery);
- day-ahead (market settled ahead of the delivery day); and
- term (from prompt month to several years).

The locational aspect of the underlying leads to a distinction between nodal and zonal deals, with the underlying design of the power pool determining the types of prevailing transactions. Nodal transactions are defined with respect to a specific point (a node, bus or location), and transactions settle on the price calculated at this point at the time agreed in the contract. Zonal transactions settle on prices defined over larger territories (the old ERCOT power pool is a good example). Of course, there are no pure nodal and zonal markets. In nodal markets, the transactions are often based on hub prices, defined as averages calculated over predefined clusters of locations. The most popular hubs are:

- PJM West;
- Mass Hub in ISO NE; and
- Zones A to L in the NYISO.

In a zonal pool, prices are sometimes available for a specific location, such as, for example, the STP (nuclear South Texas Plant) price in the old ERCOT pool.

A unique feature of the electricity markets is the availability of prices for hourly and sub-hourly periods combined with intraday

fluctuations of electricity loads (see Figure 23.2), which leads to transactions for power supplied over hourly blocks, with the most popular deals in the US markets being the following.<sup>34</sup>

- ❑ On-peak:
  - West HE7-HE22, Monday – Saturday, except NERC holidays;
  - East HE8-HE23, Monday – Friday, except NERC holidays; and
  - ERCOT HE7-HE22, Monday – Friday, except NERC holidays.
- ❑ Off-peak:
  - Hours which are not on-peak.
- ❑ Other time buckets:
  - $5 \times 16$  same as on-peak in the East and ERCOT;
  - $6 \times 16$  same as on-peak in the West, which includes Saturday;
  - $2 \times 16$  HE8-HE23 on Saturday, Sunday and NERC holidays in the East;  
HE7-HE22 on Saturday, Sunday and NERC holidays in ERCOT;
  - $1 \times 16$  HE7-HE22 on Sunday and NERC holidays in the West;
  - $7 \times 8$  HE1-HE6 and HE23-HE24 (HE1-HE7 and HE24 in the East), every day;
  - Wrap same as off-peak (combination of  $2 \times 16$  and  $7 \times 8$ , or  $1 \times 16$  and  $7 \times 8$ );
  - $24 \times 7$  same as ATC (around-the-clock), which equals all hours.

HE stands for hour ending. For example, HE7 is the hour from 6 to 7 am, ending at 7:00 am. NERC holidays change slightly from year to year (although in a highly predictable way) and are available from its website.

In the established power pools such as PJM, MISO, NY ISO, NE ISO include:

- ❑ wholesale size – 50 MW:
  - block types:
    - $7 \times 24$ ,  $5 \times 16$ , off-peak wrap ( $7 \times 8$  and  $2 \times 16$ )

- products:
  - fixed price power;
  - heat rates (including options);
  - options (vanilla put/call); and
  - fixed-for-floating swaps.

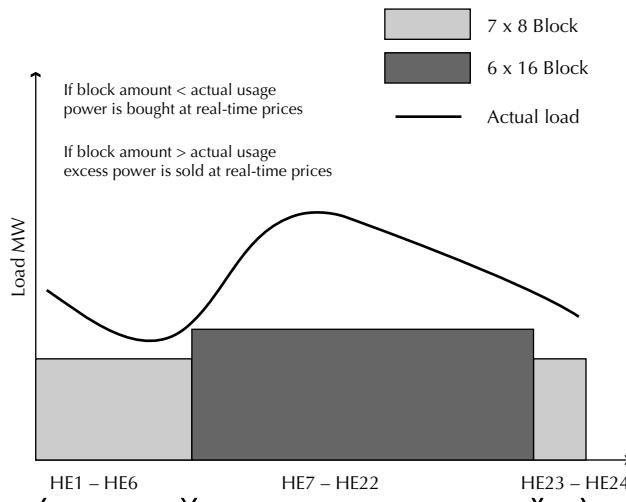
PJM has a number of active trading hubs:

- PJM Western Hub;
- NI (Northern Illinois) Hub (connection to MISO); and
- AD (AEP Dayton) Hub (connection to MISO).

In the New York ISO, trading evolves around zonal transactions, with most activity being at:

- Zone A (West);
- Zone G (Hudson Valley); and
- Zone J (NYC).

**Figure 23.2** Average weekday usage versus block hedge amounts



Source: Martin Lin, Class Slides, Rice University MGMT610

In the Midwest ISO (MISO), trading evolves around zonal transactions, with most activity being at:

- Cinergy Hub;
- Michigan Hub; and
- Illinois Hub.

In the New England ISO (NE ISO), trading evolves around zonal transactions, with most activity being at the Central Massachusetts Hub.

## STRUCTURED TRANSACTIONS

### Full requirements deals

A full requirements transaction is a contract under which a power marketer commits to satisfy the full electric energy and capacity requirements<sup>35</sup> of the counterparty. In some cases, the power marketer takes over the control of the physical generation resources and of the portfolio of supply contracts of the client, and provides additional management services related to their operations and dispatch. Such transactions have several defining features, which are detailed below.

#### Services provided

Practically every full requirements contract guarantees supply of electricity. In addition, the power marketer may be obligated to provide additional commodities, including ancillary services and fuel for the generation units.

Early full requirements deals negotiated in the late 1990s typically included additional provisions under which a power marketer would manage the generation resources of a client (including maintenance and dispatch of generation units), as well as the supply portfolio of power purchase contracts. Sometimes such services would be extended to the management of fuel supply contracts and related physical and contractual assets (transportation contracts, natural gas and fuel oil storage facilities). Since the late 1990s, many local distribution companies divested their generation fleets. In such cases, these companies depend completely on the full requirements transactions. Some distribution companies rely on the full requirements deals to support the provider of last resort (POLR) service.<sup>36</sup>

The providers are selected through the process known as a request for proposal or through auctions.

### Volumetric commitments

A full requirements transaction puts a power marketer under an obligation to satisfy the full electricity needs of a client (subject to negotiated limits) (or a portion of client's needs) at a fixed price or a floating price based on a formula. The pricing rule typically references market prices of electricity and fuel, and may be quite complicated. Demand is satisfied by using the client's generation resources (if such resources are available) and through market transactions (buying electricity from third parties under contracts negotiated by the client and managed by the power marketer or under contracts negotiated by the power marketer). A power marketer has an option to supply electricity produced from the client's generating plants or to shut them down selectively if power can be procured from the market at a lower cost. In case the client's demand for power drops below the level of available resources (the generating plants, as well as the rights to electric energy under contractual arrangements), the power marketer has an option to sell surplus power into the market (ie, to third parties), assuming that this can be accomplished at a profit. The ability to optimise the existing portfolio of the physical and contractual assets in the context of a deregulated market was seen as one of the main drivers behind full requirements deals. Many power marketers believed or maintained that they had unique asset optimisation and trading skills that could produce savings big enough to cover their required profit margins and produce significant benefits to their clients.

### Contract risks

Full requirements transactions have many potential risks with primary exposure related to price and volumetric uncertainty.

#### *Price risk (aka market risk)*

Market risk in the case of full requirements deals has many dimensions. The biggest risk is related to the potential of electricity prices to spike and reach levels measured in hundreds and, sometimes, thousands of dollars per MWh. Spiking prices may become very costly to a power marketer obligated under a full requirements contract to

satisfy the entire load of the client. A power marketer has no discretion to delay purchases when prices spike. As a consequence, if electricity prices increase significantly above the price the marketer receives under the contract, the marketer can sustain substantial losses.

#### *Volumetric risk*

Volumetric risk is related to the fact that demand for electricity may be very difficult to predict and is highly variable. Variability of load (ie, the demand that must be served) has many dimensions.

- ❑ Short-term variations in load are related to weather fluctuations and other developments that may be either non-recurring or seasonal. Some load variations are related to the load profile (a curve that represents fluctuations in demand, typically during daily, weekly, monthly or annual periods). Such changes may be due to special circumstances<sup>37</sup> or evolution of the customer base (discussed further below). The weather may have a long-term impact if a power marketer has the bad luck to enter into a contract covering a spell of a few very hot summers and cold winters.
- ❑ Long-term variations in load are related to organic demand growth due to economic development, demographic changes and evolving consumption patterns. Some parts of the US have witnessed very rapid growth of demand for electricity since the start of the 2000s related to population migration and an increased preference for large dwellings.
- ❑ Load fluctuations can be amplified by customer migration due to economic reasons. Sometimes, a power marketer competes for load with retail suppliers of electricity who have the right to operate in their contract area, and this competition may lead to a vicious circle of adverse selection by customers. Suppose that a power marketer commits to a contract with a fixed price of power and then the prices of electricity fall. This may trigger outbound migration of customers who will select alternative providers of electricity offering lower prices. Power marketers may find themselves in the situation of having an over-hedged contract (more power was bought forward than necessary and, in addition, probably at excessive prices). If market prices go up,

the outbound migration of customers slows down and, sometimes, additional customers may decide to migrate in (if the contract gives them this option).<sup>38</sup> A power marketer who made certain assumptions regarding customer attrition may find out it has to serve a load greater than expected and that it is under-hedged – ie, does not have protection against the cost consequences of having to serve a larger customer base.

#### *Compounded price and volume risk*

A power marketer with a full requirements transaction is likely to be exposed to combined price and volumetric risks. A marketer is likely to be short power when the demand spikes or when there are supply shocks. This means that the power marketer has to buy power when everybody else in a given area is short power and prices spike. When a power marketer has a surplus of electricity that it can dump on the market, other market participants are likely to be selling as well.

#### *Market structure risk*

Many full requirements transactions in the 1990s were negotiated with the expectation that the electricity markets would quickly mature in terms of trading volumes, liquidity and transparency. In many cases, this proved to be wishful thinking, as efficient markets did not materialise in time and marketers had to rely on very shallow and illiquid markets. Additional complications arose from the fact that most full requirements transactions have to be disclosed (or are known) to the market and this means that other trading operations will be able to extract a premium because they realise that a power marketer with a full requirements deal is often in a desperate position and has to acquire power at very high prices, given very high penalties in the contract for a failure to deliver.

#### **Tolling transactions**

Tolling transactions are a typical deal structure used in the commodity markets. In the most general sense, it is a contract for the conversion, processing or transportation of a raw material through a production or transportation facility for a fee. The owner of production facility does not have responsibility for the procurement of inputs and marketing the final product.<sup>39</sup> These transactions are

particularly important in the energy industry, and also for many reasons in the electricity business.

Tolling transactions may be either physical (ie, involve a physical assets, which is managed with respect to its dispatch by the holder of a tolling contract) or financial (virtual).<sup>40</sup> For the latter, a transaction settles in cash, based on the market prices of natural gas and electricity and a formula agreed in the contract. A buyer of a tolling contract goes long the price spread between electricity and fuel prices. In the case of a financial tolling transaction, their exposure is limited to two agreed price indexes (for power and fuel) and the heat rate specified in the contract. This is effectively an option on a spread, with the fee paid to the counterparty (the “owner” of the virtual power plant) representing the option premium. An example will help to explain this contract. Suppose the agreed heat rate is equal to 8 MMBtu/MWh, with the price indexes being the Gas Daily Houston Ship Channel next day natural gas price and the Megawatt Daily next day on-peak electricity price in the Houston zone. If the price of natural gas is equal to US\$4.00/MMBtu, the implied price of electricity is equal US\$32/MWh (given a heat rate of 8). If the Megawatt Daily electricity price prints at US\$35/MWh, the buyer of this tolling contract receives US\$3/MWh multiplied by the number of megawatts specified in the contract. If the contract is for 100 megawatts, the daily revenue is equal to  $100 \times 16 \times 3 = \text{US\$4,800}$ . Of course, this is a gross amount, which ignores the premium paid for entering into this contract. If the electricity price prints below US\$32/MWh, the buyer receives nothing. The underlying price indexes are left to the discretion of the two parties, and they can use, for example, the Nymex settlement price of natural gas prompt contracts or the price of electricity in the ERCOT market.

A physical tolling contract is unit-specific. Again, we can use here an example involving natural gas. The buyer of the contract is responsible for delivering natural gas to the plant and marketing electricity. The buyer will typically absorb the risk of the infrastructure conditions upstream and downstream of the plant. What varies from contract to contract is the extent to which the buyer accepts responsibility for the physical conditions of the plant and the *force majeure* events related, for example, to forced outages or labour disputes, or to unexpected costs such as environmental penalties. The fees received by the plant owner may be designed in many

different ways, and every tolling contract is effectively a snowflake. From the point of view of the plant owner, it is important to account in the contract for the pattern of option exercise decisions by the contract buyer. Frequent starts and shutdowns result in increased maintenance costs and may shorten the useful life of the plant. The fees should in principle vary with the number of starts per month and should account for the start-up and shutdown costs.

## CONCLUSIONS

Power market transactions are the most complicated and risky part of the energy markets. This is a result of the technological complexity of electricity business, overlapping and constantly evolving regulatory regimes and, most of all, the volatility of electricity prices. Successful power trading and marketing requires a cohesive team combining multiple skills and experience in the power industry. Many times in my career we have witnessed teams representing sophisticated financial institutions losing to presumably unpolished provincials. What the latter had in common was an unmatched familiarity with local systems, economic conditions, prevailing laws and, especially, a lot of commonsense.

1 <http://www.austincc.edu/lpatrick/his1693/lcra.html>.

2 US Federal Energy Regulatory Commission, 1995, "Promoting wholesale competition through open access non-discriminatory transmission services by public utilities," Docket RM95-8-000, Washington, DC, March 29.

3 Yann Rebours and Daniel Kirschen, 2005, "What is spinning reserve?" working paper, University of Manchester, September 19.

4 Eric Hirst and Brendan Kirby, undated, "Ancillary services," Oak Ridge National Laboratory (available at [http://consultkirby.com/files/Ancillary\\_Services\\_APPC\\_1996.PDF](http://consultkirby.com/files/Ancillary_Services_APPC_1996.PDF)).

5 "The system inertia is the ability of power system to oppose changes in frequency. Physically, it is loosely defined by the mass of all the synchronous rotating generators and motors connected to the system. If system inertia is high, then frequency will fall slowly during a system casualty such as a generator tripping off line. If system inertia is low, then frequency will fall faster during this casualty." See B. J. Kirby, J. Dyer, C. Martinez, A. Rahmat Shoureshi, R. Guttromson and J. Dagle, 2002, "Control concerns In the North American electric power system," Oak Ridge National Laboratory, December, for more information.

6 Mike Terbrueggen, "Control performance standards," presentation, North American Electric Reliability Council.

7 This distinction between operating and load following reserves has been proposed by Eric Hirst and Brendan Kirby (see footnote 4).

8 For example, if one system operator incurs losses because of unexpected power flows into this area, there should be a mechanism for the resolution of resulting claims.

9 Lori A. Burkhart, 2005, "FERC takes on reactive power," *Utilities Fortnightly Letter* # 15

- (SPARK), March. See also G. Chicco and G. Gross, 2008, "Current issues in reactive power management: A critical overview," Proceedings of the IEEE PES General Meeting, Pittsburgh, PA, July 20–24.
- 10** FERC, 2005, "Principles for efficient and reliable reactive power supply and consumption," February 4, Docket No. AD05-1-000.
- 11** *Phase angle* is a quantity that indicates the difference in time of peaks of sinusoid waveforms. *Power factor* is a measure of real power in relation to reactive power; mathematically, it is defined as the cosine of the phase angle between voltage and current. When the power factor is leading, the current phase angle is greater than the voltage phase angle; when the power factor is lagging, the current phase angle is smaller than the voltage phase angle. *Capacitors* supply reactive power and have leading power factors, while *inductors* consume reactive power and have lagging power factors.
- 12** The passage of The Public Utility Holding Company Act of 1935 (PUHCA), also known as the Wheeler-Rayburn Act. This act was effectively repealed on August 8, 2005, when the Energy Policy Act of 2005 was signed into law after clearing both houses of Congress.
- 13** The price spikes during in the summer of 1998 and 1999 (due to immaturity of the power markets in the US at the time) and during the power crisis in California (2000–01), due to market manipulation, convinced the industry that one cannot go wrong in building a peaking power plant. The perception was that inevitable price spikes would allow recovering fixed costs of a unit within a few days (if not hours), leaving the rest of the annual net revenues to pure profits. The resulting construction boom resulted in many poorly designed and located plants, and excessive reserve margins in many parts of the US. When overbuild and the poor quality of many plants became obvious, liquidation prices reached 20 cents on a dollar in some cases. In the markets, the day of reckoning will arrive sooner or later.
- 14** Sir Alec Issigonis is remembered as the author of the maxim "a camel is a horse designed by a committee." To expand on the maxim by Georg Christoph Lichtenberg ("A *donkey* looks to me like a horse translated into Dutch"), most official documents on capacity payments remind me of treaties by Keynes or Ricardo translated into modern bureaucratic language.
- 15** Graham Shuttleworth, Jonathan Falk, Eugene Meehan, Michael Rosenzweig and Hamish Fraser, 2002, "Electricity markets and capacity obligations. A report for the Department of Trade and Industry," National Economic Research Associates, December 13; Johannes Pfeifenberger, Kathleen Spees and Adam Schumacher, 2009, "A comparison of PJM's RPM with alternative energy and capacity market designs," The Brattle Group, September.
- 16** Concerns about potential shortages include the situation in the UK, where the supply of electricity may fall short of demand by 2015. Many critics attribute this problem to the lack of capacity payments in the British Electricity Trading and Transmission Agreements (and its predecessor prior to 2005, NETA). On the day when this section was updated, Texas experienced rolling blackouts due to exceptionally cold weather. ERCOT is an energy-only pool, without capacity payments, and many industry observers worried for some time about threats to electricity supply reliability in Texas.
- 17** Brownout is an intentional drop in voltage that can be noticed by consumers as the lights start to dim and flicker (although it still better than losing power completely).
- 18** Australian National Electricity Market.
- 19** England and Wales Power Pool was replaced by NETA (New Electricity Trading Arrangement) in 2001.
- 20** RPM was implemented in 2007.
- 21** South West Interconnected System.
- 22** Forward Capacity Market.
- 23** Transitional BRAs were initial auctions for less than the three-year horizon.
- 24** Additional smaller LDAs were defined and cleared in some auctions. In the 2011/2012 auction, just one RTO cleared; in the 2012/2013 auction, five LDAs did so. A discussion of the results, year by year, is beyond the scope of the book. See Electric Market Reform

Initiative and American Public Power Association, 2010, "A Review of PJM's Reliability Pricing Model," September.

- 25 As explained by Andre Ott, "prices realized under RPM are limited by cost-of-service principles, so that over time, total market revenues from PJM's energy, operating reserve and capacity markets (RPM) cannot exceed the total revenue requirements that would be determined under well-functioning cost-of-service regulation." See Andrew L. Ott, 2008, "The PJM's reliability pricing mechanism (why it's needed and how it works)," March 2008 (<http://www.pjm.com/~/media/documents/reports/pjms-rpm-j-chandley.ashx>).
- 26 See Electric Market Reform Initiative and American Public Power Association, 2010, "A Review of PJM's Reliability Pricing Model," September; "Direct Testimony of James F. Wilson in Support of First Brief of the Joint Filing Supporters," July 1, 2010, FERC Docket No. ER10-787, <http://www.wilsonenec.com/page1.php>.
- 27 Net capacity price after deduction of the CTR (Capacity Transit Right) credit (reflecting the ability of an LDA to access lower cost resources from the rest of the RTO).
- 28 Paying a negative price is the same as receiving a revenue.
- 29 This can be proved by induction, demonstrating this property starting with three nodes, assuming that this is true for n nodes, and showing that this implies that the hypothesis holds for n+1 nodes.
- 30 As a reminder, FTRs are defined in terms of differences of nodal prices (ie, prices established for specific components of generation/transmission system). FTRs may be defined also for clusters of nodes (zones).
- 31 In combinatorial auctions, the participants may bid on portfolios (packages) in addition to single items.
- 32 Self-scheduled FTRs are created as obligations by bidding in the auction the rights obtained through ARRs (24-hour class only). Self-scheduled FTRs have to be consistent, with respect to the path definitions, with associated ARRs.
- 33 The standard definition of on-peak and off-peak applies. On peak is HE 0800–2300 EPT, Monday–Friday, excluding NERC holidays. Off-peak are the remaining hours of a week.
- 34 Martin Lin, class slides for the Rice University MGMT610 Course.
- 35 Many power pools are based on a market design that includes prices for energy and the capacity payments for the right to use generation resources.
- 36 Different names are used from state to state for this service (for example, a standard offer). POLR is intended for consumers whose contracts with retail electricity providers expired for whatever reason, and need an alternative provider before they can find a new supplier.
- 37 Enron incurred significant losses in full requirements deal with Oglethorpe (a rural cooperative in Georgia) when the load profile changed dramatically during the Olympic Games in the summer of 1996.
- 38 For example, a power marketer may be serving the standard offer, a default service for residential customers that did not choose an alternative electricity supplier and decided to stay with an incumbent utility. In some cases, the customers may drop out and then opt for the standard offer again. A utility may choose to subcontract this load (or a part of it) to a power marketer.
- 39 <http://www.businessdictionary.com/definition/tolling-contract.html>.
- 40 Michael L. McWilliams and John M. T. Wilkinson, 2003, "Tolling valuation and strategies for Powerex Corp" Simon Fraser University, September.

## *Manipulation and Gaming of Energy Markets*

Price discovery – the process through which prices are determined and communicated to market participants – may be affected by occasional manipulation. Given the very vigorous measures taken by regulators in the US and the EU in the last few years, in our view the occurrences of manipulation have been infrequent (or the manipulators have learned new tricks). Still, the potential for manipulation exists, and traders and risk managers should be aware of this issue. A trader who engages in manipulation puts themselves and their firm in serious danger. A trader who is a victim of manipulation can incur significant losses, which are even more painful if their trade was based on solid market fundamentals and good research.

In this chapter, we will discuss some historical instances of market manipulation in the gas and electricity markets. What is interesting is that, in many cases, the manipulation in the gas markets was driven by a desire to influence electricity prices. For these, we have the benefit of ample documentation due to actions taken by the regulators. We shall also cover what seems to be an emerging pattern of manipulation in the energy markets: active trading in the markets that set benchmarks for the settlement of derivative contracts. The leverage inherent in derivative contracts is a key component of this strategy.

### **PRICE MANIPULATION: THE MECHANICS**

The reader should be warned that market manipulation is a very controversial and emotional topic, and this short section does not even attempt to offer a final word on the topic.<sup>1</sup> The subject has many

ramifications, ranging from complex legal and regulatory issues to a fundamental debate about the efficiency and equity of the commodity markets. Some economists argue that manipulation is practically impossible and self-defeating in the contexts of highly efficient markets, and is a financial equivalent of shooting yourself in the foot. Even if market manipulation happens occasionally, the argument goes, its scope and social cost do not justify the allocation of resources to prevent it: the market performs the function of enforcement more effectively than any district attorney or any regulator.<sup>2</sup> Other economists (typically those of a left-of-centre political persuasion, or those with some practical experience who have decided to enter academia away from the tough realities of the trading floors) sometimes see the commodity markets as inherently corrupt and depraved. The truth is somewhere in between: market manipulation happens often enough to be a factor a trader should be aware of and consider in their decision-making process. Also, irrespective of a theoretical and ideological position on market manipulation, one has to recognise that people (in some cases very junior and inexperienced employees) have been sent to prison and some companies fined hundreds of millions of dollars for relatively minor offences.<sup>3</sup> This is why the topic deserves close attention from any practitioner. All supervisors have the fiduciary duty to protect the resources and reputation of their company and the future of the employees reporting to them. We have always treated this responsibility seriously when working in the merchant energy business and for financial institutions.

A discussion of market manipulation represents a number of challenges, as the topic straddles two complex fields: economics and law. An energy practitioner should be aware of two major difficulties related to dealing with this issue. First, the law does not provide a clear definition of manipulation, leaving the implementation of some very vague anti-manipulation statutes to the courts. Historically, the US courts would rely on four tests:

- ❑ the ability of the alleged manipulator to affect market prices;
- ❑ occurrence of intentional actions designed to affect the price;
- ❑ the occurrence of an artificial price;<sup>4</sup> and
- ❑ causality (the alleged manipulator made the artificial price happen).

The implementation of these tests in practice is extremely difficult, and criminal and administrative trials often evolve into duels of competing economic modellers and lawyers debating opposing points of view. It is quite obvious, even to a layman, that proving intention in the absence of wiretaps or co-operating witnesses is almost impossible, and an alleged manipulator can always use the defence of absent-mindedness, incompetence and forgetfulness.<sup>5</sup> Equally difficult is proving what the market price would be in the absence of manipulation. It is not surprising that a 2007 manipulation case was dismissed by the court.<sup>6</sup>

Second, the enforcement of the anti-manipulative provisions of US laws is fragmented between the criminal justice system and several federal agencies. The CFTC has enforcement powers with respect to the futures markets. The FERC has oversight responsibility with respect to cash markets for natural gas and electricity. The oil and refined product markets are covered by the Federal Trade Commission (FTC).

For a long time, the primary form of market manipulation was a corner, a strategy consisting of acquiring control over the physical supply of a commodity traded in the futures markets, denying the shorts the ability to make deliveries under the contract. Corners were a common occurrence in the late 19th and early 20th century in the Chicago markets. Benjamin Hutchinson, known as "Old Hutch," provided the inspiration for what is probably the best book on trading and speculation in American literature: *The Pit*, by Frank Norris,<sup>7</sup> the story of a speculator who makes and loses a fortune, leaving behind a trail of enriched friends and bankrupted competitors, before leaving Chicago without much money but much more experience.<sup>8</sup>

The corners do still occur in the commodity markets, but new techniques of market manipulation have become more pervasive. The mechanism of price discovery that evolved in the natural gas markets (and with some differences in other energy markets), as well as the proliferation of multiple electronic exchanges and derivative contracts, has created unique opportunities for price manipulation by dishonest market participants. Price manipulation techniques evolve all the time and assume many forms as market microstructure changes. The section below will examine briefly the potential for manipulation in the natural gas markets:

- manipulation of price indexes; and
- manipulation on electronic exchanges or across different types of exchanges/markets.

It is important to recognise at this stage that price manipulation is not equivalent to speculation, although these terms are often used interchangeably in the media. Speculation is not only legal but is a necessary component of any market. Speculators render an important economic service as providers of liquidity and a critical link in the price discovery process. Manipulation not only undermines market efficiency, but is also morally wrong: it is equivalent to stealing from hard-working and honest traders.

### **Manipulation of price indexes**

The manipulation of price indexes may be classified as unilateral and bilateral. Unilateral price manipulation consists of reporting false, non-existent transactions to industry newsletters in order to influence the calculated index. This practice had become endemic<sup>9</sup> in the US natural gas markets by the late 1990s, and was eventually investigated and documented extensively by the FERC and the CFTC.<sup>10</sup> What is puzzling is that, with the wealth of evidence freely available online, there are still academics willing to deny the existence of price manipulation.<sup>11</sup> A more sophisticated practice is to report transactions selectively, submitting only those that are likely to push the final index price in the desired direction. Sometimes, transaction volumes are reported incorrectly in order to achieve such objectives through the weights used in calculating the index price. Another scheme relied on comingling different types of transactions where only one type was supposed to be reported (for example, combining physical and financial transactions in reports related to physical transactions).

Bilateral price manipulation, on the other hand, involves a conspiracy with another trading operation, in order to create supporting information for data being submitted to the index publishers. Such practices became more frequent when the price publishers became more suspicious of improper practices in transaction reporting and engaged in efforts to validate the information they were receiving, asking for additional information about the transactions' counterparties. Under a bilateral scheme, two traders in different organisations create false schedules of transactions they

supposedly consummated, validating their respective representations to the index publishers.

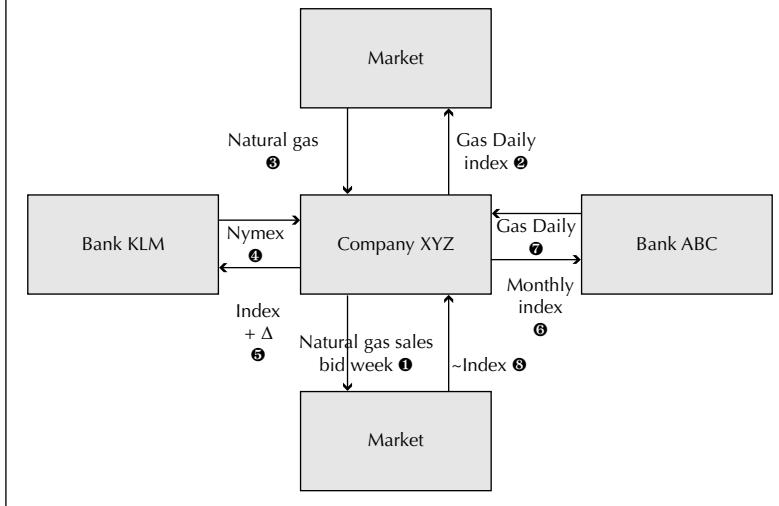
A more sophisticated version of a bilateral scheme was a transaction known as a round trip. Such transactions were more popular in the electricity markets where traders were under an additional obligation to report transaction volumes to the FERC. A round trip (a wash sale) is a simultaneous buy/sell transaction of the same volumes of a commodity at the same price and with identical terms with respect to other aspects of the contractual arrangements. The transaction has no bottom-line impact (as the two parallel transactions cancel each other), but achieves other important objectives:

- ❑ leaves an audit trail of a transaction that can be reported to the index publishers and to the government authorities; and
- ❑ can create the illusion of high trading volumes (the motivation is discussed below).

In some cases, traders in two different companies may engage in multiple transactions that do not result in net changes in their mutual obligations but increase the volume reported to the index publishers. This scheme exploits the procedure for calculating monthly indexes based on volume-weighted prices. In some cases, the pretended transactions were executed on the electronic platform known as Enron Online. This platform allowed some market participants to transact the same packages of natural gas, back and forth, moving the price higher each time.<sup>12</sup>

A hypothetical manipulative transaction (and this should not be construed as advice) consists of transacting in the fixed price markets for monthly gas and entering into simultaneous transactions through Gas Daily swaps that reverse the volumetric consequences of the first set of transactions (see Figure 24.1). Such arrangements open the traders to price risk, which can be hedged with financial swaps exchanging cashflows based on the Gas Daily and monthly index prices.

This hypothetical scheme may be used by a company dominating a given market hub. Local market power may be exploited to dump significant volume of natural gas in compressed timescales on the market during bid week under fixed price transactions (❶). These transactions, reported to the index publishers, can effectively set the

**Figure 24.1** Manipulation of monthly index prices

monthly index (if only a few other companies report). The price company XYZ receives under these transactions is, on average, equal or close to the index (❸). The objective of large volume sales is to reduce the level of the monthly index calculated by the index publishers. XYZ benefits from the lower level of the index through the basis swaps under which it receives the Nymex settlement price (❹) and pays the index + differential  $\Delta$  (❺). XYZ benefits by lowering the index it pays out. This scheme may be further enhanced by selling more gas than the company effectively controls. XYZ may buy natural gas from the market under Gas Daily swaps. Under these arrangements, XYZ pays Gas Daily price (❻) and receives physical natural gas (❽), which is flown to the customers to satisfy the obligations acquired under bid week transactions. XYZ sells gas and buys gas at the same time under different arrangements, enhancing its ability to impact the monthly index and balancing its volumetric positions at the same time. The risk is that the daily prices spike during the month. This risk is hedged by entering into a financial floating for floating swap under which XYZ receives the Gas Daily price (❾) and pays monthly index (❻).<sup>13</sup>

The Gas Daily price received under (❾) is flown to the sellers of natural gas (❽). The flows (❽) and (❻) will be roughly offsetting, assuming that XYZ will be able to influence to a significant

degree the monthly index. The profits will be realised on the basis swap.

### Motivation

The reporting of false transaction prices and/or false transactions was carried out in order to improve the profitability of the trading operation. Traders who participated in this scheme were motivated by a desire to maximise their trading profits and increase their bonuses, directly related to the profits generated by their units. The specific situation of each trading desk varied from time to time and from location to location, and traders were sometimes motivated to push the prices down and sometimes to push the prices up. Occasionally, the motivation of different groups of traders in the same organisation pointed in different directions and conflict resolution required the involvement of senior management. Some traders were motivated to participate in the illegal schemes out of a sense of camaraderie and misplaced loyalty to their organisation.

There are many specific situations creating economic incentives to engage in the reporting of false prices, and it is impossible to identify and describe all of them. However, some typical examples represent an overwhelming majority of them.

In many cases, a company has a significant volumetric commitment buying or selling huge volumes of natural gas at monthly index (ie floating) prices. In many locations, the trading volume during the bid week is relatively small, but the outcome affects transactions that often have a much bigger value. The design of the US natural gas markets has created a situation in which the tail is often wagging the dog. This characteristic extends to other commodity markets: small transaction volumes set the prices for a large number of long-term contracts. This may give companies reporting to index publishers enormous leverage over the market.

In many instances, the producers of natural gas are selling under contracts that give them prices pegged to a natural gas index at a specific location. For example, producers may receive a percentage of the index, ranging from the high 90s to as low as 75%. A marketer who aggregates gas flows in the producing region for transportation to other parts of the country has an incentive to push the index down. Producers receive the short end of the stick. It is very difficult to argue in this case that price manipulation is a victimless crime.<sup>14</sup>

Efforts to influence a price index through false reporting may be influenced by certain derivative transactions (such as basis swaps), which are settled based on monthly index prices. A natural gas trader may accumulate over time a position in basis swaps under which they receive a fixed price and pay a floating price. The floating price in this case is the basis, defined as the difference between the monthly index price and the Nymex natural gas contract final settlement price for the same month. By reporting unrealistically low prices transacted during the bid week, a trader can push down the index and lower the reported basis, the price they have to pay. This is a low-risk manipulation game, because the trader does not even have to take the risk of trying to bring down prices through aggressive selling. A trader can accomplish their objectives by sending a fax or email. Of course, an alternative is to engage in very aggressive selling in the fixed price market during the bid week.

Finally, the false reporting of transactions may have an additional collateral benefit of artificially inflating the volume of transactions the firms pretended to engage in. During the merchant energy boom of the late 1990s, many newly created trading entities were just clusters of a few traders and fax machines, and had little to report in terms of actual profits and cashflows. Equity analysts had no choice but to use non-conventional measures of the economic potential of the energy trading companies, in the same way new approaches had to be developed to value Internet-based companies. Daily trading volume was one of the valuation metrics used for the merchant energy business. Reporting a large daily transaction volume was seen as a strategy to influence directly the stock price.

### **Two parallel markets of different levels of transparency**

Many market observers who question the viability of market manipulation still use the old frame of reference of the 19th century markets, where the most sophisticated form of manipulation was a commodity corner. This was a world with no derivatives and no electronic trading platforms, instant communications or social networks. The new trading venues and trading instruments created the preconditions for new types of manipulation. These manipulative strategies are well understood by practitioners and took many years before they caught the attention of economists.

What is often of critical importance is that the prices are formed in

one market (which also happens to be more transparent) and are used to settle transactions executed in another market. A trading desk can accumulate quietly positions in the less-transparent market and engage in manipulative trading activities using the platform that offers higher visibility of their actions and, therefore, the ability to influence the prices. Operations carried out in the full light of the day result in losses that are more than offset by gains in the other market. This strategy may additionally exploit the benefits of leverage inherent in the derivative instruments and market microstructure (ie, price formation mechanisms). Shrewd traders can use this form of market organisation to manipulate the market. There are many indications that this type of manipulation is currently most widely used in the markets. Aggressive trading in the physical markets is used to influence the level of price indexes on which derivative contracts settle. Losses in physical trades can be easily recovered in the derivative markets, given the leverage inherent in these transactions.

A paper by two economists has looked at the potential for price manipulation in the Italian bond markets.<sup>15</sup> The authors focus on “modelling the strategic behaviour of informed dealers in the presence of parallel markets of the same asset, but with very different degree of transparency. In particular, [they] focus on dealers behaviour in the secondary market of T-bonds [Italian Treasury bonds] at the times when the primary market (ie, auction of T-bonds) is open.”<sup>16</sup> Their approach is important because they investigate manipulation across markets, not just manipulation inside a given market, the traditional object of interest of economists. They conclude that:<sup>17</sup>

the informed dealer may use the more transparent market in order to manipulate prices. That is, he can take advantage of the fact that the other dealers react to his trading in the more transparent market in order to shape their expectations about the value of the asset about which he has superior information. If he knows that the value is higher than the one currently impounded in the price and it is due to rise, he may want to short sell the asset in the more transparent market. In this way he would further reduce competitors’ expectations. The dealer would, at the same time, aggressively stock himself up in the less transparent market where his transaction would go largely unnoticed.

Losses accrued in one market can be recovered through a much bigger (or much more leveraged) position established in a parallel market.

It is obvious that the cited paper does not address the case of manipulation in a commodity market. However, it does point to a potential abuse of fragmentation of a given market, different regulatory treatment and the various levels of transparency of competing trading platforms. One historical example is the potential use of two critical US natural gas trading platforms: ICE and Nymex. Lower levels of transparency on ICE (for example, the reporting of large trading positions discussed in Chapter 6 did not apply in the past to ICE-traded contracts) could allow a trader to accumulate a position on ICE in instruments which settled based on the final Nymex prices of the expiring contract.<sup>18</sup> In the case of the electricity markets, such strategies could involve strategic bidding in order to influence day-ahead prices underlying firm transmission rights.<sup>19</sup>

The potential for this type of manipulation is very important to any practitioner. It is often true that market manipulation is self-defeating, as was pointed out by Daniel Fischel and David Ross in the paper quoted earlier in this chapter.<sup>20</sup> For example, a market manipulator trying to drive a price to a lower level will sell a large volume of a given financial instrument or commodity. Once the market fundamentals re-establish themselves and prices return to the level supported by supply and demand, they may be left with significant losses.<sup>21</sup> Proliferation of parallel trading platforms and derivative contracts drives a spike through the heart of this argument. Small losses incurred in trading on one platform (or in the physical market) can be easily recovered in a parallel market due to leverage offered by derivative contracts. Manipulative behaviour consists typically in the following trading patterns:

- selling or buying unusually large quantities of a given commodity in extremely short periods of time, overwhelming the ability of the market to absorb the increased volume;
- entering the market very early in the trading day, posting multiple offers to sell and bids to buy; this may create the perception of the price level for the current day and condition the market to trade around certain levels; and
- establishing a leveraged derivative position benefitting from a price move in the manipulated market.

The second technique exploits certain behavioural patterns known as framing. Framing is defined as influencing the context of decisions made by individuals.<sup>22</sup> Conversations with many traders have led me to believe that this type of behaviour is quite frequent. One can sometimes see transactions executed for small volumes in the pre-dawn hours on electronic trading platforms. These transactions may be very small but they may define the climate for a given trading day. Other traders check the market when they sip their morning coffee or on the way to work and assume that there is information in the market they are not aware of, and act accordingly. It takes many years of experience to identify such decoys.

Spoofing and layering are related strategies. Spoofing is placing buy or sell orders with no intention of executing them, with the objective to confuse and mislead the market-maker and other market participants. Layering is submitting multiple orders on one side of the trading book to create a false impression of mounting buying or selling pressure. When the market moves in the desired direction, the orders are quickly cancelled and new layers are established.

This section relied to a large extent on examples from the US natural gas market. The strategies described above can be easily illustrated with case studies from other energy markets. The central message is that it is sometimes difficult to understand certain trades, unless one can see the entire portfolio across many units of the firm. For example, it is sometimes impossible to rationalise generators' bidding strategies into the power pools, without considering the established positions in transmission contracts (FTRs) or derivative instruments. Given the large size of trading books and the volume of daily transactions, a regulator is often outgunned and made impotent by the sheer volume of information that has to be processed. It is likely that the tracks may be covered by intentionally inflating the size of the trading book to make it less transparent.<sup>23</sup> The next section will discuss some forms of market manipulation unique to the electricity markets.

## **MANIPULATION AND GAMING OF ELECTRICITY MARKETS**

Electricity markets share a number of common characteristics that make them exceptionally vulnerable to market manipulation and gaming. The concept of gaming includes actions that may or may not

meet the legal criteria of manipulation but are nevertheless pernicious in their impact on market efficiency and integrity, leading to a suppression of correct price signals. The features of electricity markets that combine to facilitate manipulation and gaming include the following.

- ❑ Non-storability of electricity as a product (apart from water behind the dam and the units which can be dispatched at short notice, functioning indirectly as a very expensive inventory). This means that there is no buffer that can be used to remedy short-term shortages resulting from outages of generation units or a surge in demand.
- ❑ Low short-term price and income elasticity of demand for electricity, resulting from the unique role electricity plays in our lives, both as a consumption good and an input to production processes, critical to our standard of living, safety and national security. Any power pool operator is willing to go an extra mile to avoid an outage leading to blackouts, inconvenience to customers and production losses, and this translates into a willingness to pay a very high price for the marginal unit of output, and a readiness to engage in transactions that set the market price but represent a very small percentage of overall volume delivered to customers. According to the old proverb, “The true taste of water reveals itself only in the desert.” As in other energy markets, a small volume of transactions may drive the prices of record, used for settlements of physical and financial transactions.
- ❑ The complexity of the rules can overwhelm most practitioners of the electricity markets. Technical documents are written in jargon inaccessible to people who do not have technical training or relevant experience, and, as with all products of committees, represent compromises and last minute efforts to meet deadlines. Rules are often created in organisational silos responsible for different segments of the system, and the gaps and contradictions at the seams are not immediately obvious. However, smart and motivated market participants can identify the weak points in the system design, sometimes through pure brain power, sometimes through experimentation.<sup>24</sup> When the problems are identified, the remedy is usually modification of the rules, which

often create new opportunities for gaming. Unfortunately, we are condemned to live in the world in which detailed rules are being constantly replaced with even more detailed rules.

Manipulative and gaming schemes can be classified as follows:

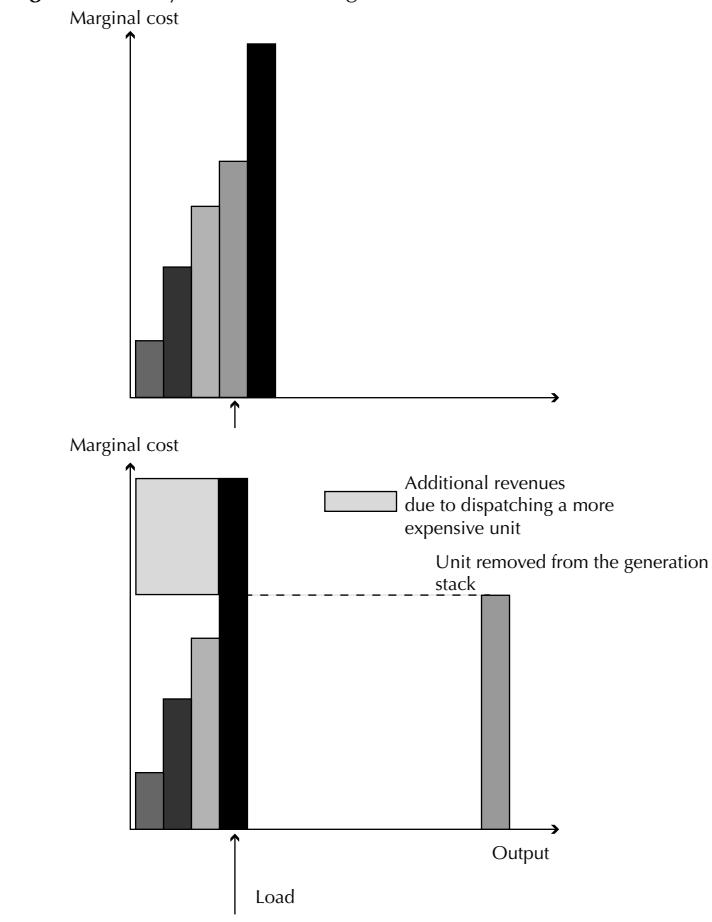
- ❑ physical and economic withholding;
- ❑ manipulation of market A in order to reap benefits in market B; and
- ❑ taking advantage of loopholes and the complexity of market design blueprints.

### **Physical and economic withholding**

Physical withholding can be illustrated with a highly stylised example (see Figure 24.2). A power producer who controls several units removes one or more of them from the generation stack. It is accomplished by claiming that the units require maintenance or often by offering no excuse at all. Given the price formation process, the market price will be set at the margin by a more expensive unit which would not have been dispatched if a more efficient unit(s) were available. The producer loses potential profit from the unit that was shut down and incurs a higher cost dispatching a unit with higher cost, but, hopefully, gains through increased revenues from higher prices resulting from their actions. This scheme works if the additional revenues from sub-marginal units (shown in Figure 24.2) exceed the higher cost of production in the most expensive unit.

Economic withholding takes place not through a direct act of shutting down a generation unit but by engaging in bidding strategy that makes the dispatch of a unit unlikely. There may be different reasons for using this strategy. One reason would be to shift generation from day-ahead market to real-time market, characterised by lower price elasticity of demand. This, in turn, allows a power producer to design other strategies, exploiting these characteristics of demand.

Physical and economic withholding happens typically when the balance between available resources and load is very tenuous. This makes the strategy not only more likely to succeed but also provides a convenient smokescreen to cover up market manipulation. The Western Electricity Markets Crisis of 2000–01 was explained at the

**Figure 24.2** Physical withholding

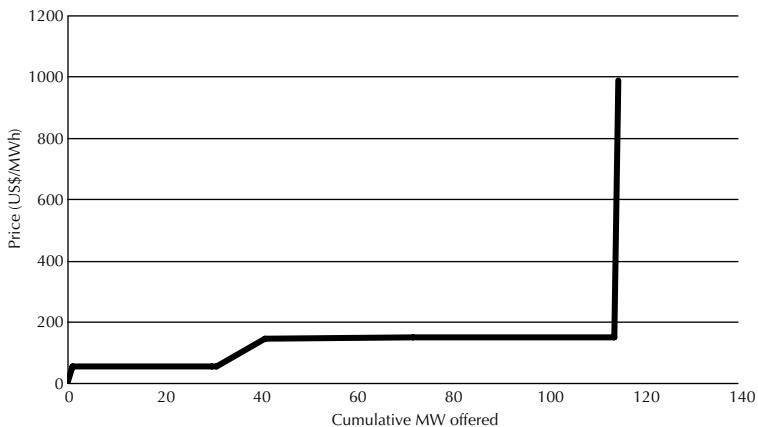
time as a “perfect storm” of hot weather, high demand and exceptionally high rates of generation outages, due to a history of poor maintenance of units which had recently changed hands (were sold by regulated utilities to independent power producers). In reality, those claims turned out to be bogus or greatly exaggerated.

### Hockey stick bidding strategy

A hockey stick bidding strategy may be compared to buying a lottery ticket. It consists of bidding a very high price (sometimes at, or close to, the cap established by a power pool) for the last few units of generation capacity (see Figure 24.3 for an example).

In a pool with ample generation capacity and high reserve margins, this leads on most days to foregone profits as the high bid resources are not dispatched. Occasionally, unusual circumstances can result in a demand spike or generation capacity shortage, and a very high bid sets the market price, not only for the company but for other owners of generation fleet. An article<sup>25</sup> about price spikes in Texas in February 2003 contains some very interesting statistics, showing standard 16-hour daily products in the ERCOT North Zone jumping from US\$49/MWh to US\$340/MWh during the episode (a 594% increase). The article discusses the remedies undertaken by ERCOT and the PUC of Texas at the time. The moral suasion approach includes disclosing the identities of companies behind unusually high price bids.<sup>26</sup> The Competitive Solution Method (CSM) was effectively a procedure for the disciplined overwriting of the outcome of ERCOT market processes in case they produce an outlier. Another solution, we may add, is to have faith in the invisible hand of the market. In an ideal world of a perfectly competitive market, new companies would materialise out of nowhere and negate the potential profits from those gaming the market. However, the practicalities of power markets are such that building a competing power plant takes years and may be impossible given the

**Figure 24.3** An example of hockey stick bidding



Source: Based on David Hurlbut, Keith Rogas and Shmuel Oren, 2004, "Protecting the market from 'hockey stick' pricing: How the Public Utility Commission of Texas is dealing with potential price gouging," *The Electricity Journal*, April, pp 26–33.

limited availability of convenient sites with access to transmission, water, pipelines, labour, etc. One can also ask the more fundamental question: why would any rational business man make an investment to deny other market participant the opportunity to buy a lottery ticket from time to time (especially if they were playing the lottery themselves)?

### **Manipulation of settlement prices of outstanding contracts**

The second type of a strategy, designed to influence price A in order to benefit from a position in a different financial instrument or in a different physical product (price B), can be illustrated with an example of a power pool combining two increasingly popular design features, towards which almost all existing markets are converging: virtual bidding and FTRs.

Virtual bidding describes a solution under which one can bid into the pool without the back-up of actual physical generation or load resources. Selling in the day-ahead market is covered by buying at the prevailing real-time market price, and vice versa. Virtual bids are called, “decremental bids” (DECs, or “virtual load”); virtual offers are called “incremental bids,” “INCs” or “virtual supply”. Traders refer to such transactions as “incing” and “decing.”<sup>27</sup> The pool algorithms do not discriminate between real and virtual bids and offers, and virtuals impact the outcome of the price-setting process. The benefit of this pool design feature is an improvement in convergence between day-ahead and real-time prices. However, as with any good thing in life, this strategy can be abused.

FTRs are financial instruments designed to offer congestion hedges or a platform for speculation (as has been explained in previous chapters). An FTR designed as a swap produces an economic outcome equivalent to the difference between locational marginal prices at the source and the sink (the nodes at which power is injected and extracted from the grid). A highly recommended paper by Shaun D. Ledgerwood and Johannes P. Pfeifenberger<sup>28</sup> provides a theoretical framework to explain how virtual bidding can be used to enhance the value of an FTR portfolio and how to detect this form of market abuse. In addition to FTRs, a company may have a number of other reasons to attempt the manipulation of day-ahead versus real-

time prices. A company may have entered into a variety of contracts (full requirements deals, retail transactions, hedging transactions with unsuspecting financial firms) benefitting, for example, from lower real-time prices. For example, a company may be a party to a retail transaction under which it sells electricity at fixed prices and sources power in the real-time markets.

The main challenge regulators assessing the propriety of such strategies face is the complexity of the portfolios of physical and financial positions held by many market participants. Each specific transaction may be justified as a legitimate action when assessed in isolation, and contesting a company's motives may not overcome legal defences. The capacity of regulators to engage in multiple investigations is limited, especially when one considers the most binding resource: experienced market oversight and enforcement experts.

### Taking advantage of the market rules

An example can be provided by an obscure provision in the PJM tariff related to transmission loss credit. Trades challenged by the regulators happened in an opaque market called up-to congestion (UTC), which is unknown to most energy professionals.<sup>29</sup> The highlights of these transactions include the following.

- ❑ Transmission service has to be reserved in order to submit these transactions to the ISO.
- ❑ PJM allocates part of the surplus arising from collections of charges for marginal line losses to transmission users. Why this surplus exists at all is explained below. The logic behind this approach was that UTC transactions required transmission and, therefore, contributed to the cost of the transmission grid (by paying a transmission access charge).<sup>30</sup>
- ❑ The economic rationale for such transactions can be explained best with a quote from PJM documents:

“Up-to’ congestion bids permit transmission customers to specify how much they are willing to pay for congestion by bidding a certain maximum amount for congestion between the transaction source and sink. If the congestion charges are less than the amount specified in the bid, then the transaction will be scheduled in the day-ahead schedule. These ‘up-to’ bids protect transmission customers from paying uncertain congestion charges by guaranteeing that they will pay no more than the amount reflected in their

bids. Transmission customers also may use an increment and decrement bid pair to accomplish the same type of hedging strategy, which further enhances their price certainty options.”<sup>31</sup>

Marginal line losses are twice the average losses.<sup>32</sup> This is due to the physics of power flows: marginal line losses are proportional to the square of the power flows:

$$\begin{aligned} \text{Losses} &= a \times I^2 \\ \text{Average loss} &= \frac{a \times I^2}{I} = a \times I \\ \text{Marginal loss} &= \frac{\partial}{\partial I} a \times I^2 = 2 \times a \times I = 2 \times \text{Average loss} \end{aligned}$$

In the equations above,  $I$  stands for the power flows, coefficient  $a$  for a constant. Contested transactions in day-ahead markets happened in two different forms. One type of transaction consisted in trading at two points that were virtually identical, given their physical or network proximity, with flows going in both directions, effectively cancelling each other. The second type was based on trading at the same point – ie, submitting bids and offers at the same node.

To put things in perspective, in July 2010 the average marginal loss surplus allocation was US\$1.32/MWh, compared to the cost of non-firm transmission of US\$0.67/MWh; the average on-peak marginal loss surplus allocation was ~US\$1.85/MWh; the average off-peak marginal loss surplus allocation was ~US\$0.67/MWh.<sup>33</sup>

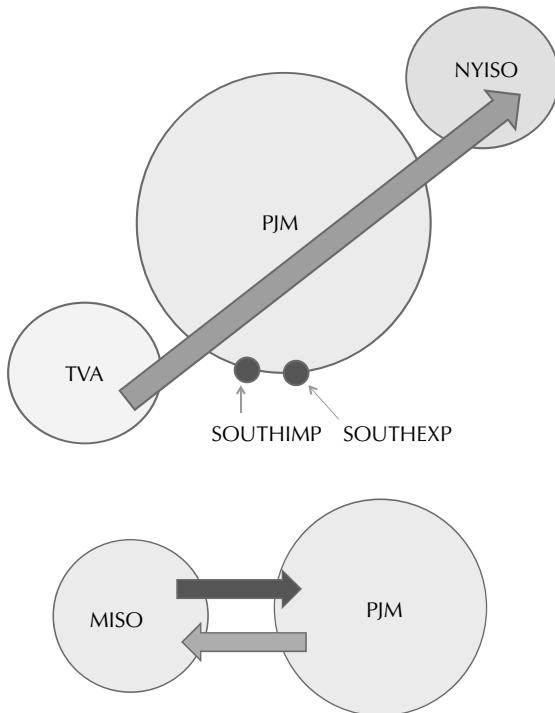
In this example, transmission is acquired for wheeling power from TVA into NYISO. Transactions are scheduled through interface points between TVA and PJM, SOUTHIMP and SOUTHEXP. The day-ahead locational marginal prices are equal at these points (the distinction between them is really an accounting artifact), transmission costs US\$0.67 (which is left on the table as it is not used) and marginal loss surplus allocation is equal to US\$1.85/MWh – leaving a trader with a net profit US\$1.18/MWh. It is not surprising that such transactions became as popular as hot rolls at the breakfast table. The counterflow example (see Figure 24.4) involves buying and selling power at the same bus (say X) in PJM against MISO. Congestion from X to MISO and from MISO to X are equal in absolute values and have opposite signs. This is a perfect wash. The cost of transmission and the loss surplus allocations are the same as in the first example, producing the same net profit for little trouble.

Whether UTC trades qualify as manipulation has to be left to regulators and the lawyers representing the traders involved. What is quite obvious is that these transactions have no economic significance and would not be structured this way in the absence of the loss surplus allocations. It is troubling that there are indications that a modification of the tariff was sought to make UTC transactions qualify for the payments out of the funds raised from the line loss charges. Time and money was invested in creating this market opportunity.

### Enron trading strategies in California

Famous Enron strategies<sup>34</sup> used during the California energy crisis of 2000–01 fall under the umbrella of exploitation of the rules of a power pool. The circumstances of the Enron bankruptcy and the colourful names used to identify the schemes made them household names.<sup>35</sup>

**Figure 24.4** Up-to congestion strategy



The strategy that captured most public attention was known as Death Star, not because of its importance but due primarily to its name, and was a good example of PR-myopic thinking (whether you are a customer or not, the memories of Alderaan are difficult to erase from the collective memory of people who grew up watching Star Wars). Death Star consisted in structuring circular transactions, partly outside the jurisdiction of the California ISO. For example, power in California would flow over a congested interface in the direction opposite to congestion, providing a relief. Then it would come back over a different line that the ISO in California could not see. For example, power would enter at Lake Mead, NV, and leave at Four Corners. The connection between Lake Mead and Four Corners was controlled by Arizona Public Service (APS).

What happens is that the compensation for the congestion relief is collected without effectively transmitting any power. The strategy is profitable if the cost of transmission through bordering areas is lower than the congestion payment.

The strategy has a number of negative consequences, including an adverse impact on the reliability of the system. Whether this constitutes market manipulation can also be left to the lawyers, but it is quite obvious that if we started collecting money through misrepresentation for the services we did not render, we would not be enjoying freedom for long.

The Cut Schedule strategy exploited weaknesses in the accounting system. A power marketer was paid for creating a counterflow and relieving congestion in the day-ahead or hour-ahead markets. After the payment is awarded, the counterflow was intentionally eliminated by the scheduling coordinator. The accounting system could not rescind the payment due to its flaws.<sup>36</sup> Similar congestion-related strategies are known as Forney Perpetual Loop, Cong Catcher and Red Congo. Another strategy evolving around congestion management included Load Shift, a strategy involving misrepresentation of load. Scheduling unrealistically high load levels would create the illusion of congestion that would, in turn, benefit transmission rights position.<sup>37</sup> Another way to benefit from this strategy would be to engage in actions designed to relieve the illusory congestion by restating the load for the real-time market, effectively putting it back where it belongs.

The Fat Boy strategy exploited one of the cornerstones of market

design in California, which required scheduling coordinators (SCs) to submit balanced schedules (load equal to generation). This rule was exploited to shift generation from day-ahead markets to real-time markets with lower price elasticity of demand (a technical term to describe system operators' desperation to avoid black-outs and brown-outs).<sup>38</sup> The load expectations of the ISO were often unrealistically low and the Fat Boy strategy would take advantage of this by scheduling excess generation against fictitious next-day load. In the real-time market, generation with no offsetting load would be treated as "uninstructed energy" and would receive real-time prices. It is clear again that this strategy depended heavily on flaws in market design. Utilities in California were displaying a tendency to underschedule load, in a futile effort to bring down the Power Exchange (PX) prices and creating inadvertently the conditions for intra-day shortages.<sup>39</sup> The gap in available generation would be filled from generation that was scheduled based on misrepresentation of the load levels. Thin Boy was the opposite strategy, designed to take advantage of expectations of lower prices in the real-time markets (lower than the day-ahead prices). This strategy required underscheduling next-day demand, potentially using a web of trading partners.

Ricochet (megawatt laundering) was a strategy of parking power outside the California market. At some point, price caps of US\$250/MWh were instituted by the ISO but the limits did not apply to power imported from other markets. The solution was to engage in fictitious exports of power in the day-ahead market and to re-import it in the real-time markets. In reality, power would never leave physically the California ISO territory. The export/import transactions were entirely paper entries and critically depended on collaboration of other entities.<sup>40</sup>

These examples are by no means an exhaustive list of the Enron western energy markets trading strategy. Many strategies have not been fully documented and are probably lost in the sands of time. Some were likely still on the drawing board and were never quite implemented.<sup>41</sup> These examples are intended to demonstrate the importance of understanding the rules of a power pool and the ability to think through the complicated chess games played in the energy markets.

## MARKET BEHAVIOUR RULES

The manipulation of price indexes and electricity markets became endemic in the late 1990s. These practices were extensively investigated by the FERC and CFTC in subsequent years, leading to a number of new rules being promulgated by these two agencies. FERC issued two orders related to market behaviour in the natural gas and electricity markets, respectively. The rules are similar, although those for the power markets contain a number of provisions related to the operations of physical infrastructure of the industry. On November 17, 2003, FERC issued "Order Amending Market-Based Rate Tariffs and Authorizations,"<sup>42</sup> which contained the following rule:

Actions or transactions that are without a legitimate business purpose and that are intended to or foreseeably could manipulate market prices, market conditions, or market rules for electric energy or electricity products are prohibited.

The Energy Policy Act enacted on August 8, 2005 added section 4A of the NGA, which prohibited market manipulation as follows:

SEC. 4A. It shall be unlawful for any entity, directly or indirectly, to use or employ, in connection with the purchase or sale of natural gas or the purchase or sale of transportation services subject to the jurisdiction of the Commission, any manipulative or deceptive device or contrivance (as those terms are used in section 10(b) of the Securities Exchange Act of 1934 (15 U.S.C. 78j(b))) in contravention of such rules and regulations as the Commission may prescribe as necessary in the public interest or for the protection of natural gas ratepayers. Nothing in this section shall be construed to create a private right of action.

FERC, under the authority of the Act, implemented additional rules in January 2006 (Order 670<sup>43</sup>), making it unlawful for:

any entity, directly or indirectly, in connection with the purchase or sale of natural gas or the purchase or sale of transportation services subject to the jurisdiction of the Commission, or in connection with the purchase or sale of electric energy or the purchase or sale of transmission services subject to the jurisdiction of the Commission, (1) to use or employ any device, scheme, or artifice to defraud, (2) to make any untrue statement of a material fact or to omit to state a material fact necessary in order to make the statements made, in the light of the circumstances under which they were made, not misleading, or (3) to engage in any act, practice, or course of business that operates or would operate as a fraud or deceit upon any person.

The anti-manipulation provisions were further reinforced in the Dodd–Frank Act. The CFTC in the past has developed through case law a test for manipulation composed of four parts (to be met jointly):

- ❑ the ability to influence market prices;
- ❑ an intention to create an artificial market price, diverging from the levels determined through the normal interaction of market forces;
- ❑ artificial prices were established; and
- ❑ artificial prices were caused by the alleged manipulator.

Providing an analytical proof that all the four criteria were jointly satisfied is quite difficult. Any quantitative model or empirical proof can be easily challenged, given the obvious limitations of econometric and statistical tools, noisiness or outright unavailability of empirical data. This may explain why the CFTC has always demonstrated a strong inclination to charge traders with attempted manipulation, a charge with a lower burden of proof.

Rules related to implementation of the Dodd–Frank Act are contained in CFTC rules 180.1 and 180.2.<sup>44</sup> Rule 180.1 addresses fraud-based manipulation, while Rule 180.2 is similar in its main outline to manipulation prohibition preceding the Dodd–Frank Act.

The highlights of Rule 180.1 (which every trader should be familiar with) include:

- ❑ it is similar to SEC Rule 10b-5 and follows the logic used by the FERC and the FTC.
- ❑ under the rule, it is unlawful to engage (intentionally or recklessly) in the following actions “in connection”<sup>45</sup> with any swap, futures or cash transaction.
  - (1) Use or employ, or attempt to use or employ, any manipulative device, scheme, or artifice to defraud.
  - (2) Make, or attempt to make, any untrue or misleading statement of a material fact or to omit to state a material fact necessary in order to make the statements made not untrue or misleading.
  - (3) Engage, or attempt to engage, in any act, practice, or course of business, which operates or would operate as a fraud or deceit upon any person.

- (4) Deliver or cause to be delivered, or attempt to deliver or cause to be delivered, for transmission through the mails or interstate commerce, by any means of communication whatsoever, a false or misleading or inaccurate report concerning crop or market information or conditions that affect or tend to affect the price of any commodity in interstate commerce, knowing, or acting in reckless disregard of the fact that such report is false, misleading or inaccurate. Notwithstanding the foregoing, no violation of this subsection shall exist where the person mistakenly transmits, in good faith, false or misleading or inaccurate information to a price reporting service.

In the interest of full disclosure, we have to advise any trader to consult a lawyer to gain a better understanding of the scope of this rule. We have no legal background, but as anybody who has had to function in our exceptionally complex society and navigate the often-uncharted waters of an industry being created in real time through the collective efforts of countless professionals, we can offer our interpretation of this rule.

The rule allows the CFTC to prosecute alleged manipulators even if their actions were not successful (recklessness is a sufficient charge).<sup>46</sup> A trader is not required to reveal non-public private information to a counterparty. This means that a trader who acquired proprietary information through legal devices is not required to disclose it. However, injecting false information into the marketplace by, for example, providing false transaction data to the PRAs or falsifying commodity inventory reports would count as manipulation.

Section 753 of the Dodd–Frank Act amended Section 6(c) of the CEA, and explicitly prohibits direct or indirect price manipulation or attempted price manipulation. This prohibition is implemented through Rule 180.2 as follows:

It shall be unlawful for any person, directly or indirectly, to manipulate or attempt to manipulate the price of any swap, or of any commodity in interstate commerce, or for future delivery on or subject to the rules of any registered entity.

The Commission reiterated that, in applying final Rule 180.2, it will continue to be guided by the traditional four-prong test described above. Attempted manipulation requires only proving the last two points of the four-part test:<sup>47</sup>

- (1) An intent to affect the market price; and
- (2) Some overt act in furtherance of that intent.

It remains to be seen how the two rules will be used in practice and how they will interact. Our expectation is that the CFTC will face an easier task in proving manipulation cases.

Current regulatory initiatives in the EU reach far beyond these US rules, inserting more stress into the lives of energy traders. In December 2010, the European Commission proposed the Regulation on Energy Market Integrity and Transparency (REMIT),<sup>48</sup> which expanded the scope of the Market Abuse Directive (MAD) to cover wholesale energy trading.<sup>49</sup> The key provisions of REMIT include the prohibition of market manipulation and attempted manipulation through the following activities.<sup>50</sup>

- (a) entering into any transaction or issuing any order to trade in wholesale energy products which:
  - (i) gives, or is likely to give, false or misleading signals as to the supply of, demand for, or price of wholesale energy products;
  - (ii) secures or attempts to secure, by a person, or persons acting in collaboration, the price of one or several wholesale energy products at an artificial level, unless the person who entered into the transaction or issued the order to trade establishes that his reasons for doing so are legitimate and that that transaction or order to trade conforms to accepted market practices on the wholesale energy market concerned; or
  - (iii) employs or attempts to employ a fictitious device or any other form of deception or contrivance which gives, or is likely to give, false or misleading signals regarding the supply of, demand for, or price of wholesale energy products;
- or:
- (b) disseminating information through the media, including the Internet, or by any other means, which gives, or is likely to give, false or misleading signals as to the supply of, demand for, or price of wholesale energy products, including the dissemination of rumours and false or misleading news, where the disseminating person knew, or ought to have known, that the information was false or misleading.

When information is disseminated for the purposes of journalism or artistic expression, such dissemination of information shall be assessed taking into account the rules governing the freedom of the press and freedom of expression in other media, unless:

- (i) “those persons derive, directly or indirectly, an advantage or profits from the dissemination of the information in question; or
- (ii) the disclosure or dissemination is made with the intention of misleading the market as to the supply of, demand for, or price of wholesale energy products.”

Other provisions of REMIT include:

- the prohibition of insider trading in the wholesale energy markets;
- an obligation to disclose information regarding facilities owned or controlled by a business entity engaged in energy trading; and
- a system of market monitoring and data collection by a newly established Agency for the Cooperation of Energy Regulators (ACER).

Limitations of space prevent us from a more in-depth discussion of the provisions of this very important rule, a task that is better left to the lawyers. Any trader or manager working in the wholesale energy business will undoubtedly eventually touch something that is related to the EU energy market, so familiarity with REMIT should be regarded as a top priority.

## CONCLUSIONS

Manipulation is a practice as old as the markets. It is ironic that the first reported derivative transaction can be interpreted as the first documented market corner.<sup>51</sup> Our advice to any trader is to avoid any activity that may qualify as market manipulation. Any manager in a trading operation should ruthlessly eradicate any attempts to boost profits through questionable strategies. Experience has taught us that behaviour tolerated for a long time may be criminalised retroactively, with those at the bottom of the food chain being punished out of proportion with their importance in a scheme. Other events (such as the Libor rates manipulation) will reinvigorate the regulators, who are now not only armed with more potent rules, but are also subject to pressures imposed by public opinion.

- 1 The US authority on market manipulation is Professor Craig Pirrong from the University of Houston. An interested reader should consult some of his papers on the topic (available at <http://www.cba.uh.edu/spirrong/>).
- 2 Daniel Fischel and David Ross are the best-known and vocal proponents of this point of view. They argue that “the concept of manipulation should be abandoned altogether ... Actual trades should not be prohibited as manipulative regardless of the intent of the trader.” See Daniel R. Fischel and David J. Ross, 1991, “Should the law prohibit ‘manipulation’ in financial markets?” *Harvard Law Review*. See also Steve Thelt, 1994, “US\$850,000 in six minutes – the mechanics of securities manipulation,” *Cornell Law Review*, pp 219–98, for a rebuttal.
- 3 “BP has agreed to pay US\$303 million to settle civil charges that it cornered the propane market three years ago and inflated heating and cooking costs for about 7 million mostly rural American households, a source familiar with the accord said.” See Steven Mufson, 2007, “BP settles propane price-fixing suit firm agrees to pay US\$303 million to end civil charges, submit to oversight,” *Washington Post*, October 24. See also Hanah Cho, 2012, “Constellation US\$245M settlement for trading is largest ever,” *Baltimore Sun*, March 12.
- 4 The artificial price is defined as the price diverging from the levels determined by unobstructed interaction of supply and demand forces.
- 5 A standard defence is to claim a level of incompetence and confusion that would make inspector Jacques Clouseau a paragon of efficiency.
- 6 See the ruling in the BP propane case, United States District Court, Southern District of Texas Houston Division Criminal Action H-08-411, and September 17, 2007.
- 7 Frank Norris, *The Pit: A Story of Chicago* (New York, NY: Doubleday & Co), available at <http://etext.virginia.edu/toc/modeng/public/NorPit.html>; Volume 2 of “The epic of the wheat.”
- 8 As the “Old Hutch” frequently mentioned in farmers’ prayers, the main character of *The Pit*, Curtis Jadwin, was very popular in the Farm Belt. “[A]ll through the Middle West, all through the wheat belts, a great wave of prosperity was rolling because of Jadwin’s corner. Mortgages were being paid off, new and improved farming implements were being bought, and new areas seeded new live stock acquired. The men were buying buggies again, the women parlor melodeons, houses and homes were going up; in short, the entire farming population of the Middle West was being daily enriched.”
- 9 A common excuse was: “Everybody was doing this and we had no choice. We had to defend our competitive position.” It is curious that the same defence was used in the case of Libor manipulation, which resulted in the imposition of significant penalties on one bank in June 2012, and investigations are still under way in a number of countries.
- 10 “Report on natural gas and electricity price indices,” FERC, May 5, 2004; Price Discovery in Natural Gas and Electric Markets, Docket No. PL03-3-004; Natural Gas Price Formation, Docket No. AD03-7-004.
- 11 One company effectively carried two sets of books with the false transaction data kept in a spreadsheet called “IFERC Bogus.” Even Homer Simpson would have known better.
- 12 “On January 31, 2001, for next-day gas at the Topock delivery point. On that day, there were 227 trades made on EOL for next-day gas at Topock. The price rose from US\$11.30/MMBtu to US\$15.00/MMBtu. Of the 227 trades, 174 were made with a single counterparty. The total volume on EOL for next-day Topock gas for the day was 2,240,000 MMBtu, of which 1,740,000 MMBtu was with that single counterparty.” See “2000 initial report on company-specific separate proceedings and generic reevaluations; Published natural gas price data; and Enron trading strategies. Fact-finding investigation of potential manipulation of electric and natural gas prices,” Docket no. PA02-2-000, FERC, August 13, 2002, p 53.
- 13 Once the monthly index is set, this swap mutates into a fixed-for-floating swap.
- 14 The very notion of a victimless crime is an oxymoron in our view.
- 15 F. Drudi and M. Massa, 2002, “Price manipulation in parallel markets with different trans-

- parency," working paper, January. Reprinted in the *Journal of Business*.
- 16 Ibid, p 1.
- 17 Ibid, p 2.
- 18 This type of manipulation is discussed in a paper by Shaun Ledgerwood, Gary Taylor, Romkaew Broehm and Dan Arthur, 2011, "Losing money to increase profits: A proposed framework for defining market manipulation," The Brattle Group.
- 19 Shaun D. Ledgerwood and Johannes P. Pfeifenberger, 2012, "Using virtual bids to manipulate the value of financial transmission rights," The Brattle Group, May 3.
- 20 "Profitable (successful) manipulations [...] require two conditions: first, trading must cause the price of the relevant security to rise; and second, the manipulator must be able to sell at a price higher than the price at which the manipulator purchased." See footnote 2 for the source.
- 21 There are many examples of market manipulation when losses incurred in manipulative actions are more than offset through gains obtained under different arrangements. For example, the market may be manipulated in order to benefit from a pending merger agreement in which the purchase price depends on the price of the stock on a given day.
- 22 In one frequently quoted experiment, participants were asked to choose randomly a number between, say, 0 and 100. Subsequently, they were asked to come up with an estimate of, say, the national income of Turkey. The random numbers and estimates showed positive correlation. A larger random number creates an attitude, or a frame of mind, leading to subconscious production of a larger guess for the GDP level.
- 23 "Where does a wise man hide a pebble? On the beach. Where does a wise man hide a leaf? In the forest." G. K. Chesterton, 1911, *The Innocence of Father Brown* (New York, NY: John Lane Co): "The Broken Sword."
- 24 "On May 25, 1999, Enron scheduled for the transmission of nearly 2,900 megawatts of electricity over a line known as the Silver Peak, which was able to carry just 15 megawatts. The power was never sent, but the bogus scheduling caused energy disruptions and higher prices." According to subsequent investigations, Enron "submitted the bogus bids and schedules on Silver Peak as an experiment." See Kurt Eichenwald, 2002, "A powerful, flawed witness against Enron," *New York Times*, October 21.
- 25 David Hurlbut, Keith Rogas and Shmuel Oren, 2004, "Protecting the market from 'hockey stick' pricing: How the Public Utility Commission of Texas is dealing with potential price gouging," *The Electricity Journal*, April, pp 26–33.
- 26 I happen to know some of the usual suspects, and doubt if this would have any impact on some of them. Even writing to their grandmothers would not work. Their girlfriends would be proud.
- 27 Pronounced as "ink" and "deck" (attachment to a house).
- 28 Shaun D. Ledgerwood, Johannes P. Pfeifenberger, "Using Virtual Bids to Manipulate the Value of Financial Transmission Rights," The Brattle Group, May 3, 2012
- 29 This demonstrates how really obscure rules can be used to benefit those who know how to navigate the system.
- 30 As reported by a law firm, it "[s]uccessfully obtained orders from FERC directing PJM to include Up To Congestion transactions in the allocation of the over \$1 billion/year in transmission line loss over collections. This case remains in litigation in order to potentially expand this relief to all types of virtual and export transactions." See [http://www.perkinscokie.com/energy\\_power\\_marketing/](http://www.perkinscokie.com/energy_power_marketing/). The FERC order referenced here can be found in Docket EL08-14.
- 31 PJM Interconnection, LLC, Compliance Filing, Docket No. ER00-1849-000, at 7 (March 10, 2000).
- 32 Leslie Liu and Assef Zobian, 2002, "The importance of marginal loss pricing in an RTO environment," *The Electricity Journal*, (15)8, pp 40–5 (<http://www.ces-us.com/download/Losses%20paper%20-%20web.pdf>).

- 33 Joseph Bowring, 2010, "Virtual transactions and marginal loss surplus allocations," *PJM*, August 5.
- 34 There are many public sources describing the details of Enron strategies. A good place to start is: "Analysis of trading and scheduling strategies described in Enron memos," California ISO, Department of Market Analysis October 4, 2002 ([http://www.hks.harvard.edu/hepg/Papers/CAISO\\_enron.trading.analysis\\_10-4-02.pdf](http://www.hks.harvard.edu/hepg/Papers/CAISO_enron.trading.analysis_10-4-02.pdf)).
- 35 Enron insisted during the crisis that it controlled no generation in California and, therefore, could not be blamed for what was described as a "perfect storm" of natural conditions and the ineptitude of power pool designers, operators and regulators. When the memos describing the schemes were discovered after the bankruptcy of Enron, the documents were passed to the FERC by the board of the bankrupt company (see <http://news.findlaw.com/legalnews/lit/enron/#documents>).
- 36 One of the justifications for the exploitation of weaknesses in power pool design in California was that the state brought it on itself by allowing a flawed system to be put in place. The latter is probably true, but that does not justify taking advantage of it. This reminds me of a hypothetical defence: "Yes, Your Honor. We steal and cheat. But we only cheat, and steal from, the stupid and the weak." A burglar taking advantage of a broken alarm system could equally argue that breaking and entering was a valuable social service and great help in identifying the flaws in the home protection equipment.
- 37 The gaming of day-ahead prices in order to benefit FTR positions seems to be an endemic practice. Day-ahead prices can be affected through virtual transactions.
- 38 Low price elasticity of demand means that buyers are less sensitive to prices and changes of demand are small (in relative terms), given price changes.
- 39 It is a good habit to avoid playing poker with professional players.
- 40 FERC, "Price manipulation in western markets," Docket No. PA02-2-000 V-1.
- 41 Some of the more colourful code names included *Big Tuna*, *Donkey Punch*, *Ping Pong*, *Sidewinder*, *Russian Roulette* and *Black Widow*.
- 42 Docket Nos. EL01-118-000 and EL01-118-001.
- 43 Docket No. RM06-3-000.
- 44 <http://www.cftc.gov/ucm/groups/public/@lrfederalregister/documents/file/2011-17549a.pdf>. Other quotes in this section are derived from this document.
- 45 Several legal reviews zeroed in on the use of the term "in connection" in the CFTC rule. This language can expand the scope of the rule and enhance the CFTC's existing authority. It makes sense to discuss this with your friendly lawyer.
- 46 In legal language, the scienter was expanded to include recklessness. Scienter refers to the knowledge of wrongdoing. An alleged manipulator should realise that the contested behaviour happened with an understanding of its unlawful character.
- 47 See footnote 129 in the referenced CFTC document.
- 48 See Dan Harris and Shaun Ledgerwood, 2012, "Completing the puzzle: Defining manipulation under REMIT," The Brattle Group. See also Shaun Ledgerwood and Dan Harris, 2012, "A comparison of anti-manipulation rules in US and EU electricity and natural gas markets: A proposal for a common standard," *Energy Law Journal*, (33)1.
- 49 The final rule was published in the *Official Journal of the European Union* on December 8, 2011, and became law 20 days following its publication (on December 28, 2011).
- 50 Article 2 of REMIT
- 51 According to Aristotle, Thales of Miletus reserved olive presses in anticipation of good harvest (acquired a call option). When the harvest materialised and the prices of presses went up, Thales rented them at a high profit.



## *Emission Markets*

This chapter will review the different market-based solutions developed to address the challenge posed by externalities generated in the process of producing and consuming energy commodities. We will limit its scope to the emissions of certain gases and particulate matter pollution by power plants. An externality, as has been mentioned several times in this book, is the impact of economic activities of some producers/consumers on other producers and consumers, which are not captured in market prices without government intervention. Externalities are often seen as one of the primary reasons of market failure – defined as a market equilibrium that does not represent an optimal allocation of resources. The price system fails sometime to internalise certain private costs that are imposed on others without a corresponding compensation. Economic theory suggests there are a number of public policy measures that can be used to correct market imperfections, and many of the solutions discussed here have previously been suggested by long-forgotten economists. This is another proof regarding Keynes' views on the ability of economists to influence public debate from beyond the grave.<sup>1</sup>

Policymakers in legislatures and executive governments have many arrows in their quiver when it comes to addressing the challenge of externalities, and not all of these solutions are market-based. A list of alternative policy measures includes:<sup>2</sup>

- ❑ *command and control*-type solutions;
- ❑ market-based solutions:
  - taxes and subsidies;
  - tradable permits; and
  - market friction reductions:

- market creation;
- liability rules;
- information programmes.

Every society relies on many command and control tools to address environmental issues – for instance, in banning the injection of toxic fumes into the atmosphere. Such direct rules affect prices indirectly by restricting the range of technologies that can be used. In this chapter we will cover widely defined market-based solutions, including taxes, the establishment of property rights and the creation of new tradable financial instruments and commodities (such as tradable pollution permits, to be defined in the next section), which are most important from the point of view of energy traders. A detailed discussion of all the measures would take at least an additional volume.

## ENVIRONMENTAL POLICY MEASURES

### Tradable permits

The policy measures discussed in this chapter are designed to reduce emissions of gases related to the production or consumption of energy commodities, and which have negative environmental consequences and affect adversely social welfare. As mentioned in the introduction, this could be accomplished through purely administrative measures. However, an alternative is to meet these objectives through a political process that relies on the market mechanism. This can be accomplished by establishing acceptable emission levels (reduced over time) and requiring that actual emissions are offset with permits (often called allowances), which can be created in a variety of ways. In most cases, emission permits are tradable and can be acquired either through a centralised auction or in the secondary markets. Often, the permits are initially allocated to the emission sources free of charge. Over time, the initial free allocations are phased out and a transition to a solution based purely on market forces can take place. Government involvement is then limited, in principle, to a determination of emission targets and occasional interventions when glitches are discovered in the adopted solutions.

Tradable permits are created when companies or individuals: (i) commit to reduce harmful activity below certain level determined through a political process; and (ii) engage in certain types of activity with positive externalities.

In principle, both solutions can be used at the same time. The first case mentioned above can be illustrated with a commitment to reduce emissions below historically attained levels in order to meet the ambient air quality standards imposed by regulatory authorities. The reductions in the level of emissions with respect to a historical benchmark translate into credits, which may be marketed. The first programme of this type was instituted by the US Environmental Energy Agency in 1974 under the Clean Air Act umbrella, with the objective of reducing emissions of volatile compounds, SO<sub>2</sub> and NOx.<sup>3</sup> The second way of gaining credits can be illustrated with an example of a company planting trees absorbing CO<sub>2</sub> and earning credits representing the volume of captured greenhouse gases.

The solution based on tradable emission permits is usually referred to as cap-and-trade (CaT), and is used widely in most markets and jurisdictions.<sup>4</sup> What follows is a description of a generic CaT programme, without going into any details on specific implementations (examples will follow later; here we shall discuss the features of a typical programme, although specific solutions vary from state to state and country to country). Under a CaT solution, the overall volume of emissions of certain types is determined through a political process, supported by scientific and economic studies seeking to quantify the environmental impact and social cost of different types of pollutants. The overall quota is allocated to different emission sources<sup>5</sup> in the form of allowances (called sometimes credits or rights), which effectively are permits to pollute. The total allocations are calibrated to be lower than emissions from targeted sources under "business as usual" solutions – ie, without measures designed to reduce emissions. The total amount of allowances starts at or below some historical total emissions, and is adjusted downward over time. Gradual turning of the screw helps the targeted sources (ie, the installations covered under the programme) to adjust to more rigid emission standards over time and to reduce the burden of compliance. The allowances are tradable. This means that the sources that do not use their allocation of allowances for different reasons (adjustments to the level of activity producing emissions or implementation of emission reduction measures) can sell them to sources that have emissions exceeding the cap.

This is, of course, a very general description of a generic cap-and-trade programme. The details of such a programme include, among other things:

- a list of specific emissions covered under the programme (for example, CO<sub>2</sub> or SO<sub>2</sub>);
- the institutions that determine the acceptable level of emission and its adjustments over time (ie, a parliament or an executive government acting within the limits of the law);<sup>6</sup>
- the historical emission benchmark against which the emissions are measured;
- the list of targeted emission sources<sup>7</sup> and its modifications over time;<sup>8</sup>
- penalties for non-compliance; and
- market design (a centralised auction, an exchange).

This is just a partial list, and we shall provide a few illustrations later. Our objective is to describe here not a specific solution but to identify characteristics common to many different programmes. What is important in practice is that these issues are addressed one way or another. It is also important to remember that the design of any cap-and-trade system is not static and evolves over time as conditions change.

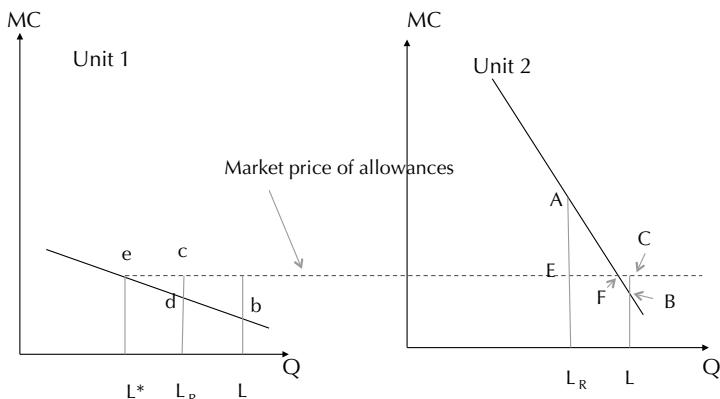
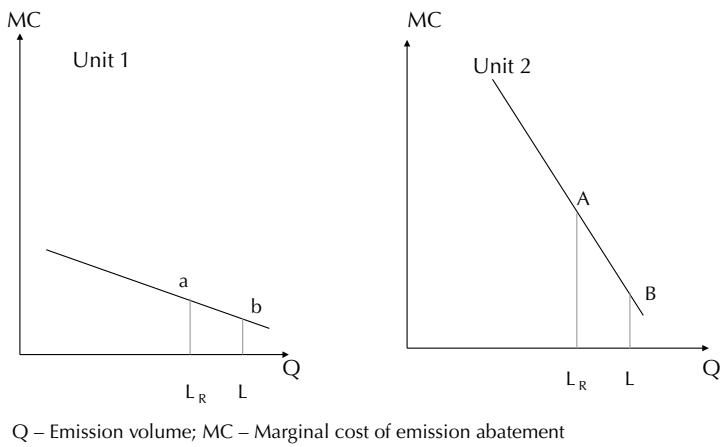
#### *Efficiency*

Economic efficiency requires that emissions be reduced to a level where the marginal cost inflicted on society is equal to the marginal cost of reducing the level of emissions. There is a general consensus among economists that cap-and-trade comes potentially closest to meeting this requirement, despite some of the shortcomings discussed below and subject to difficulties related to quantification of the impact of specific emissions in complex physical and economic systems.

A frequently used diagram (Figure 25.1) illustrates the mechanics of cap- and- trade. Suppose that two sources (identified as units 1 and 2 in Figure 25.1, Panel 1) are required to reduce emissions by the same amount, equal to  $\Delta L$  ( $= L - L_R$ ). This amount is equal to the difference between the initial level of emissions ( $L$ ) and the emission targeted under the programme ( $L_R$ ). We assume for simplicity that the initial and targeted emissions are the same in both installations. The marginal cost of reducing emissions is different in these two installations. The marginal cost of the abatement curve (ie, the curve illustrating the cost of reducing emissions by an additional unit)

faced by source 2 is much steeper than the corresponding cost curve in the case of 1. The total cost of reducing emissions by a specific source is equal to the area under its marginal cost over the interval corresponding to the emission reduction amount. The total abatement cost is represented by the area  $a b L L_R$  for 1, and  $A B L L_R$  for 2 (Figure 25.1, Panel 1). Under a cap-and-trade solution (Figure 25.1, Panel 2), source 1 reduces emissions to the level  $L^*$  by the additional amount  $L^* - L_R$ , equal to the reduction of emissions required of source 2. By reducing the level of emissions, source 1 effectively creates allowances that are marketable and can be sold to sources failing to reduce their emissions by the required amounts. Source 2 continues

**Figure 25.1** Cap-and-trade approach



to emit the same amount of emissions as before, buying the allowance from source 1. The overall cost of reducing emissions (from  $L$  to  $L^*$ ) by source 1 is represented by the area  $ebLL^*$ . If the market price of emission allowances settles at the level represented by the dashed line in Figure 25.1, source 1 sells the allowances for the amount represented by the area  $ecL_R L^*$  for the net gain of  $ecd$  (the incremental cost of reducing emissions from  $L_R$  to  $L^*$  is equal to  $edL_R L^*$ ). Source 2 pays for the permits the amount  $ECLL_R$  (equal to  $ecL_R L^*$ ) for the net savings of  $AEC$  less  $FBC$ .<sup>1</sup> Both sources benefit from the cap- and- trade solution.<sup>9</sup> Source 2 would have to incur the cost  $BLL_R A$  reducing emissions through a technological solution, which is greater than the cost of buying emission credits.

#### *Price volatility*

The price volatility of emission credits is the most frequent criticism against the CaT solution. As demonstrated below, this is a real, not an imaginary, problem. Huge swings in the prices of emission credits may be caused by a number of factors. The first is the controversial nature of CaT programmes and the potential for challenges by politicians and in the courts. Changes to a programme's design through legislation, suspension or even the repudiation of a programme due to a judicial decision, may cause a significant spike in price, up or down, which undermines the credibility of emission countermeasures and causes erosion of public support. It is not unusual to see a prudent utility, accumulating emission credits to meet future needs, taking a big hit to profits following a major shock to the market.<sup>10</sup> The prices of emission credits depend on a number of factors (as exemplified in the reviews of specific programmes below), such as the overall level of economic activity (affecting output of electricity, refinery operations, etc), the mix of different emission mitigation strategies used by emission sources and the reaction of consumers to changes in the price levels of different energy commodities. Activities of highly leveraged market participants making speculative bets on prices of emission permits may further contribute to volatility, further undermining public acceptance of a CaT programme.

A CaT programme may contain certain features designed to mitigate the consequences of price volatility, including price caps and floors, and the ability to bank credits or borrow them from future allocations. Banking credits is equivalent to postponing their use

until a future period, or hoarding them if prices are expected to increase. The benefits of banking credits are reduced if they are expected to expire eventually or if there is a time decay factor built into the balances held by the sources. Borrowing emissions from future allocations, ie, using in a given year allocations for future time periods, is a solution that can be used to avoid market purchases of credits at times of elevated prices.

### *Offsets*

Offsets are a common design feature of most CaT programmes. This feature allows the use of credits generated outside the system (in other jurisdictions or in industries not covered by the programme), often through special projects promoting the use of certain technologies. The objective of offsets is to increase the supply of credits to dampen price volatility, and also to support emission curtailment in other industries (not included in the CaT programme) or in other jurisdictions (especially in developing countries). In some cases, the offsets may be traded on a voluntary basis.<sup>11</sup> A major challenge to their wider use is protecting the integrity of the system behind their creation. Frequently cited abuses (in addition to outright fraud) are activities that would have been undertaken anyway being used to create offsets. An oil company injecting CO<sub>2</sub> into a reservoir under an enhanced oil recovery (EOR) programme and double dipping by claiming offsets for this activity is one obvious example.

### *Carbon leakage*

The term carbon leakage is used for CO<sub>2</sub> CaT programmes, although the challenge it describes can apply to any emission reduction programme. It defines the shifting of certain activities to jurisdictions with less restrictive emission reduction programmes (or none at all), with obvious negative consequences. To give a hypothetical example, an industrial company could shift steel production to a plant located outside Europe and not covered by the CO<sub>2</sub> emission measures in place in the European Union (EU). Global emissions are not reduced as expected, and the country with emission restrictions suffers through the loss of output and jobs. Some CaT programmes seek to address this complication by offering a more generous allocation of credits to industries that are more vulnerable to the threat of relocations and offshoring. An alternative is an import duty related

---

to the emission content of foreign goods imposed in order to offset unfair advantages accruing to countries with lax environmental regulations.

#### *Hot spots*

A frequent criticism of CaT programmes is that, in many instances, emissions tend to reach high levels of concentration at a specific location, exceeding legal norms, even if the overall cap (for a country or a region) is very restrictive. This argument is often made in the case of mercury pollution that tends to form deposits in the immediate vicinity of power plants burning coal. Mercury has a very serious impact on human health (especially on infants), and it is only a minor consolation to those affected that average mercury concentrations across the entire jurisdiction are low.

#### *Forecast of emission levels*

Practically all CaT programmes have a recurring weakness arising from the difficulty of aligning expected levels of emissions with the emission cap. Different factors, such as weather conditions and levels of economic activity, conspire to decouple the emission quota from actual emission levels. A recession, for example, reduces electricity production and industrial output, causing a drop in CO<sub>2</sub> and sulphur dioxide emissions. The total volumes of allowances are usually determined for multi-year compliance periods and cannot be easily adjusted without going through a protracted political process. Suffice to say that, in the EU, 27 nations have to consent to any revisions of CO<sub>2</sub> allowances. The outcome is a familiar scenario of high emission allowance prices followed by prices falling to very low levels (flirting sometimes with close to zero values) as the relative oversupply of allowances becomes obvious.<sup>12</sup> The high volatility of allowance prices is seen by many market participants as a highly negative feature of the system. Industrial companies and utilities seek price certainty before they can make commitment to specific abatement technologies and before they decide to locate production facilities in jurisdictions with different rules with respect to greenhouse gases or other types of emissions. The potential for speculative gains in a highly volatile market is seen as a transfer of resources to trading companies and away from technological solutions, leading to a permanent reduction in emissions.

## Emission taxes

Emission taxes<sup>13</sup> are often seen as an alternative to CaT programmes. An emission tax may be imposed upstream (ie, on the producers of energy commodities) or downstream (ie, on the emission sources). For example, a tax varying with sulphur content may be imposed on coal producers at the mine mouth. The main argument supporting imposing taxes upstream is the cost of collection: the number of producers is relatively limited. An alternative is taxation of the emitting sources, with the tax imposed per emission unit. The argument supporting this solution is that one can link the tax rates to the local levels of emissions. The tax rates can be adjusted over time to achieve the targeted level of emissions over a defined area.

Other arguments supporting the taxation solution are the lower overall implementation cost (with the ability to use some existing taxes to piggy-back emission charges on them), transparency and the predictability of the compliance cost. This last factor is very appealing to many executives of affected companies who complain often about huge swings in emission prices under CaT programmes, and are reluctant to make commitments to new technologies and make investments in pollution control equipment, sometimes with a very long life, given the uncertainty regarding the price of emission permits in the future. Tax revenues may be used in a number of different ways:

- ❑ distribution of tax revenues to citizens to offset the impact of emission tax on the cost of living – this would help to reduce the opposition to environmental tax measures;
- ❑ investments in clean technology;
- ❑ subsidies to firms using recommended technologies; and
- ❑ financial transfers to developing countries to support the use of environmentally friendly technologies and offset the costs of environmental compliance.

Arguments supporting the CaT solution emphasise its potential for minimising the overall social cost of compliance, subject to a given emission cap, by shifting mitigation efforts (ie, measures taken to reduce emissions) to the sources which are best suited for this activity. One can mention also that CaT creates at least one constituency supporting this measure (in addition to environmentally responsible citizens), for what is effectively a glorified tax:

companies engaged in trading or supporting the trading of emission permits (consulting firms, software developers, law firms). After all, there is no better market than a market mandated by government.

### **Command and control measures**

Another solution used to reduce the level of emissions and other forms of pollution, which has a long history<sup>14</sup> and is still widely used (although it has relatively limited support in the economic profession), is the outright ban on the use of certain technologies (or a directive to use certain technological processes) or certain inputs or mandates related to product specifications. An example of such a rule is the maximum sulphur content of certain fuels –for example, in diesel. Another example of command and control measures is the requirement that water-intake pipes are located downstream from pipes discharging waste into a river. This very simple action forces industrial plants to introduce measures to control their pollution volumes in their own best interest.

## **SO<sub>2</sub> AND NO<sub>x</sub> EMISSION REGULATION**

SO<sub>2</sub> stands for sulphur dioxide, the compound produced by burning sulphur. SO<sub>2</sub> occurs naturally in the atmosphere (primarily through volcano eruptions) and is produced by the chemical industry, primarily for further transformation into sulphuric acid, an important material in a number of industrial processes. It is also a byproduct of burning coal and refined products which contain sulphur compounds. The release of sulphur dioxide into the atmosphere is associated with harmful side effects: acid rain and a contribution to the formation of fine particles.<sup>15</sup> Fine particles (PM<sub>2.5</sub>, less than 2.5 micrometres in size) have been linked to premature death and serious illnesses such as chronic bronchitis and heart attacks, as well as respiratory problems. Acid rain, a serious threat to human, animal and plant life, is caused when SO<sub>2</sub> is transformed into sulphuric acid – which returns to earth with rain. The outcome is acidification of surface waters and soil, destruction of ecosystems and reduced visibility.

NO<sub>x</sub> is a generic terms for the nitrogen compounds NO and NO<sub>2</sub> (nitric oxide and nitrogen dioxide), which are produced during combustion, especially under conditions of high temperature. There

are many sources of NO<sub>x</sub>, both stationary and non-stationary, including cars, industrial boilers and power plants. It is estimated that, in the US, power plants contribute about 20% of NO<sub>x</sub> emissions.<sup>16</sup> NO<sub>x</sub> are associated with many environmentally adverse effects, including:

- ❑ the formation of ozone through mixing of NO<sub>x</sub> with volatile organic compounds (VOC); ozone can inflict damage on lung tissue and is particularly harmful to children and people with lung and heart conditions;
- ❑ NO<sub>x</sub>, through reactions with water and water vapour and ammonia, forms nitric acid (HNO<sub>3</sub>), which may cause or aggravate respiratory diseases, and is another cause of acid rain;
- ❑ NO<sub>x</sub> destroys the ozone layer in the stratosphere, which provides a protective layer against ultraviolet light; and
- ❑ NO<sub>x</sub> can react with different organic chemicals to form a number of different toxic substances.

Both SO<sub>2</sub> and NO<sub>x</sub> cause acidification of surface waters (rivers, lakes and seas), posing a serious risk to certain critical biological processes supporting life on earth.

Policy measures (such as the Clean Air Act (CAA) of 1970, discussed below) designed to mitigate SO<sub>2</sub> and NO<sub>x</sub> solutions that apply to power plants include the following.

- ❑ *The installation and optimised utilisation of emission control devices.* Such devices range from scrubbers<sup>17</sup> to selective catalytic reduction (SCR)<sup>18</sup> equipment.
- ❑ *Modification of the dispatch decisions.* This includes a number of measures related to the way the generation units are run and bid into the power pools, and modifications of maintenance schedules to meet seasonal emission levels, limiting the number of hours a unit runs annually.
- ❑ *Purchase of emission allowances.* Older, less-efficient plants for which the installation of emission control equipment is not a cost-effective solution could purchase allowances from more efficient sources. The overall volume of emission allowances would be maintained at the level of national cap, evolving over time.

- ❑ *Fuel switching.* Some generation plants with access to many different coal sources could switch to fuel with lower sulphur content.
- ❑ *Retiring inefficient older units.* The most drastic solution is retirement of older plants and construction of new units using the most advanced emission controls to replace lost generation capacity.

Regulation of SO<sub>2</sub> and NO<sub>x</sub> emissions started in the US with the Clean Air Act.<sup>19</sup> Title I of the CAA provided timetables and methods for states to reach compliance with the National Ambient Air Quality Standards (NAAQS). The Act required states to promulgate emission rate limits, both for NO<sub>x</sub>, SO<sub>2</sub> and VOCs, for the non-attainment regions.<sup>20</sup> Under Section 110 of the Act, each state is required to develop a state implementation plan (SIP) summarising different strategies and policies designed to reach and maintain the prescribed air quality levels. A SIP is subject to EPA review and approval. Should a state fail to produce a SIP, a federal implementation plan (FIP) would be promulgated by the EPA. Certain areas within each state for which the air quality standards are not being met are subject to additional state and federal oversight and regulation. In practice, this means that different areas within each state may be subject to different rules.

The Act and its subsequent revisions and legal challenges are a good example of the difficulties involved in environmental legislation and the complexities of the underlying issues. It also exemplifies the problems that analysts supporting energy traders face every day. The history of the Act is really a chronicle of constant updates and revisions designed to address the imperfections of its original formulation, and which reflect our improved understanding of environmental issues based on scientific progress. One of main difficulties haunting all efforts to implement the Act since the very beginning is the transport of pollutants across jurisdictional frontiers. The local standards for SO<sub>2</sub> concentrations were countered by utilities through the construction of taller smokestacks, with SO<sub>2</sub> emissions from the Midwestern plants affecting a number of Eastern states. In 1977, the Act was amended through the addition of a requirement that all new power plants install scrubbers.

*The 1990 CAA Amendment.* The Acid Rain Program (ARP) was initiated in 1995 under the provisions of the 1990 Clean Air Act Amendment. The scheme was designed to reduce electric plant emissions of SO<sub>2</sub> and NO<sub>x</sub>. The ARP SO<sub>2</sub> programme is based on the cap-and-trade approach. A limit of SO<sub>2</sub> emissions, set for the contiguous 48 states,<sup>21</sup> was translated into allowances (adding up to a cap), which are distributed to the regulated sources based on a formula. A small portion of the allowances is auctioned. Each allowance represents one ton of SO<sub>2</sub> and is year-specific (meaning it can be used in a given year or may be banked for future use). Actual emission of a ton of SO<sub>2</sub> means that one allowance has to be surrendered and retired. The regulated sources are required to have a sufficient number of allowances to offset actual emissions. Under the cap-and trade system, allowances may be traded. Sources that succeed in reducing their emissions through different measures, such as the installation of scrubbers, adjustments to the mix of fuels and changes in the level of operations, are sellers of allowances. There is a punishment for non-compliance equal to US\$3,152 per ton of SO<sub>2</sub> (as of 2006).

Phase I of ARP began in 1995, covering 263 generation units in 1,110 plants, located in 21 states – mostly in the Midwest and the Northeast. Additional units joined on a voluntary basis as substitution or compensating units.<sup>22</sup> Phase II started in 2000, with tighter emission limits and expansion of the list of regulated sources to coal-, oil- and natural gas-fired units with capacity exceeding 25 MW (and all new utility units). The emission cap for the US started at 9.97 million allowances and was gradually reduced to 8.95 million by 2010.

The NO<sub>x</sub> chapter of the ARP targeted NO<sub>x</sub> emission reductions through a programme based on establishing emission rates for different types of boilers. The reductions could be achieved by setting maximum rates for specific boilers, average emission limits for two or more units (this is equivalent to over reducing emissions in some units where it is easier or cheaper) and alternative emission limits (AELs).<sup>23</sup> Non-compliance would result in a fine equal to US\$3,152 per ton. Most fossil fuel-burning sources were not affected by Title I, as they were located outside the non-attainment areas.

The 1990 Amendment of the CAA established the Ozone Transport Commission (OTC), which was given the responsibility

for developing a multi-state response to ozone non-attainment challenges.<sup>24</sup> The OTC developed the NO<sub>x</sub> Budget Program to assess the magnitude of the problem. Based on an assessment done carried out by this programme, the OTC concluded that the cap-and-trade solution would be the best policy. The OTC jurisdictions followed up on these recommendations and signed a memorandum of understanding (MOU) to establish a CaT programme for NO<sub>x</sub>. EPA also developed a rule called the Nitrogen Oxides State Implementation Plan Call (NO<sub>x</sub> SIP Call), which required a state that significantly affected other states' air quality to achieve a reduction in emissions through "highly cost-effective controls."<sup>25</sup> States were given an option to join a trading programme, the NO<sub>x</sub> Budget Trading Program (NBP), administered by the EPA.

These initiatives led to a comprehensive programme, the Clean Air Interstate Rule (CAIR), proposed in January 2004 and promulgated on March 1, 2005, and expected at the time to be fully implemented by 2010. The programme covered a number of Eastern states (28 states + the District of Columbia), and was designed to address the pollution caused by power plant emissions drifting across state lines (hence the name). The CAIR programme was introduced under the "good neighbour" provision of the CAA, which recognised the difficulty of maintaining local ambient air quality standards given the transport of emissions across state boundaries from the upwind to the downwind states, and required the EPA to address this problem. The objective of the CAIR programme was to reduce emissions of SO<sub>2</sub> and NO<sub>x</sub> by 70% through a combination of different measures (described briefly below). The EPA expected the CAIR programme to reduce SO<sub>2</sub> emissions in the covered states from 9.4 million tons in 2003 to 2.5 million tons by 2025, and NO<sub>x</sub> emissions from 3.2 million tons to 1.3 million tons. Full implementation of the law was expected to produce tangible economic and public health benefits estimated at US\$85–100 billion annually, including prevention of 17,000 premature deaths, 22,000 non-fatal heart attacks, 12,300 hospital admissions, 1.7 million lost work days and 500,000 lost school days (all numbers on an annual basis). In total, the benefits were expected to exceed the cost of the programme by 25 times by 2015.<sup>26</sup>

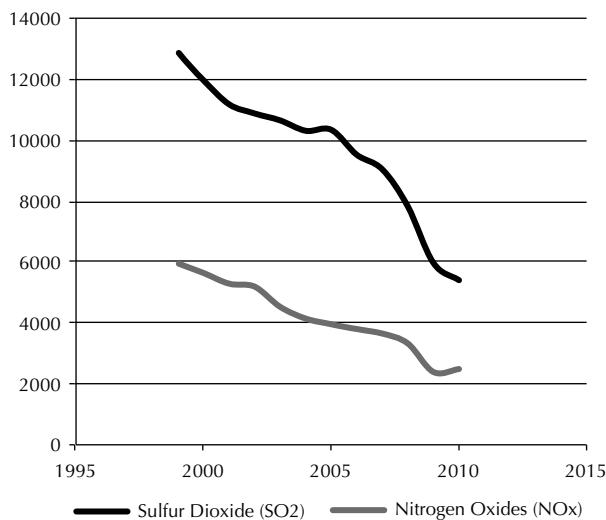
CAIR did not limit directly the emission sources. This decision was left to the states that were given the task of identifying the best mix of cost-effective policy measures. It is important to emphasise

that CAIR was not a substitute for other regulations related to ambient air quality: it was rather a programme designed to address the issues of cross-boundary emission transport and reconcile the conflicts between a regional problem and sub-regional mitigation measures. Other highlights of CAIR included:<sup>27</sup>

- ❑ an emission reduction requirement for each state, based on capping power plant emissions collectively at levels that the EPA believes are highly cost-effective to achieve;
- ❑ an optional cap-and-trade programme based on successful acid rain and NO<sub>x</sub> budget trading programmes as a method to implement the necessary reductions;
- ❑ a two-phase programme with declining power plant emission caps:
  - SO<sub>2</sub> annual caps: 3.6 million tons in 2010 and 2.5 million in 2015;
  - NO<sub>x</sub> annual caps: 1.5 million tons in 2009 and 1.3 million in 2015;
  - NO<sub>x</sub> ozone season caps: 580,000 tons in 2009 and 480,000 tons in 2015; and
  - emission caps are divided into state SO<sub>2</sub> and NO<sub>x</sub> budgets.
- ❑ flexibility on how to achieve the required reductions, including which sources to control and whether to join the trading programme.

The CAIR programme was successfully challenged in the courts by a coalition of utilities on the grounds that: "EPA's approach – region wide caps with no state-specific quantitative contribution determinations or emissions requirements – is fundamentally flawed."<sup>28</sup> One very important aspect of this decision was elimination of the provision requiring the Eastern states to surrender two SO<sub>2</sub> allowances for every ton of SO<sub>2</sub> emissions in 2010 and 2.86 allowances for every ton of SO<sub>2</sub> emissions starting in 2018 (as opposed to a one-to-one ratio). Given that annual SO<sub>2</sub> emissions have been below the cap for a number of years and that the allowances are bankable, restoration of the one-to-one surrender ratio caused a sharp drop in the prices of the allowances (which has persisted up until the time of writing).<sup>29</sup> The court remanded the CAIR programme back to the EPA with the direction to issue a replacement rule.<sup>30</sup>

**Figure 25.2** US emissions from energy consumption at conventional power plants and combined heat-and-power plants (1999–2010, thousand metric tons)



Source: U.S. Energy Information Administration

The overall assessment of the programmes reviewed here is generally positive. In spite of many challenges through the judicial and political system and frequent revisions, the level of emissions in the targeted states has been successfully reduced (see Figure 25.2), producing significant social benefits. Experience accumulated in the early stages of these programmes benefited not only future US environmental programmes, but also has become an important catalyst for the development of similar programmes abroad.

### US environmental regulations 2012–20

Since 2010, the EPA has issued or announced a number of pending new environmental regulations, which have contributed greatly to the uncertainty regarding future trends in the electricity markets, the profitability of regulated utilities and independent power producers, and the technological choices available to the industry. The EPA is in a very difficult position because it has to act under the law, given the mandates of multiple legislative acts and many court decisions. The proposed rules are highly controversial, given the large number of

different regulations that will be promulgated and become binding in a relatively narrow window of time during an era of great economic and political uncertainty. It suffices to mention just a few factors (as of 2012).

- ❑ The US is still recovering from a severe recession that affected levels of electricity consumption.
- ❑ Political and social polarisation is likely to result in legislative and court challenges to the new rules, with highly unpredictable outcomes and the potential for endless skirmishes delaying final implementation of the rules and the specific provisions they contain.
- ❑ Energy markets in the US are going through a period of great technological revolution (shale gas and oil) affecting relative prices of different energy commodities and changing the optimal technological choices.

These environmental rules promulgated by the EPA, or in their final stages of formulation, include:

- ❑ the Cross-State Air Pollution Rule (a revision of the CAIR programme, mandated by the court);
- ❑ cap-and-trade programmes for utility emissions of sulphur dioxide and nitrogen oxides;
- ❑ maximum achievable control technology standards for mercury and other hazardous air pollutants (“Utility MACT” or “HgHAPS”);
- ❑ review of NAAQS for ozone, sulphur dioxide, nitrogen dioxide and particulate matter;
- ❑ regulation of greenhouse gas emissions;
- ❑ cooling water intake regulations;
- ❑ clean water effluent guidelines; and
- ❑ coal combustion waste management rules (“Ash” or “CCBs Management”).

The combined impact of these rules is difficult to assess (explaining what it takes to develop a reasonable estimate is the number one job interview question for people in the industry). It takes a lot of analytical effort to predict the outcome of even one new rule in isolation from other policy measures. The combined effect, given a very high

---

level of political and regulatory uncertainty, can be guessed at only with a high margin of error.

There are several reasons why it is very difficult to come up with a reliable forecast of the consequences of the regulations listed above.

- ❑ As experience teaches us, all proposed rules are subject to court challenges.
- ❑ A decision to retire a specific plant will be based on the consideration of multiple rules and interactions between them. A change in one rule will start a chain reaction of changes to decisions that have been already taken. Another factor to consider will be competition from other coal-fired and natural gas-fired power in the vicinity of a plant under consideration.
- ❑ The same mitigation technologies can contribute to compliance with multiple rules.<sup>31</sup> Modification of one rule or a court challenge to a specific rule changes the outcome of the entire package of mitigation measures.
- ❑ Prices of coal and natural gas are likely to remain highly volatile, complicating the process of making optimal decisions.

The obvious consequences go well beyond the confines of electricity industry:

- ❑ The retirement of older and less-efficient coal-fired power plants, for which installation of emission control equipment is prohibitively expensive;
- ❑ expansion of gas-fired generation;
- ❑ impact on relative and absolute prices of electricity, as well as locational price differences of natural gas, electricity and coal;
- ❑ changes in the profitability of electric utility companies and independent power producers, with benefits accruing to the owners of nuclear power plants, companies with a modern generation fleet equipped with scrubbers, SCR technology and bag houses;

### **CSAPR**

The EPA acted on July 6, 2011, promulgating the final rule known as the Cross-State Air Pollution Rule (CSAPR) also referred to as the Transport Rule. In August 2012, the D.C. Circuit Court vacated the CSAPR and related EPA's Federal Implementation Plan. This means that the power plants in the states under this program would

continue to operate under the CAIR. EPA would have to go through the third attempt to incorporate the “good neighbor” policies, as mandated by the CAA, in specific rules. These rules have had to address the common thread through all the court decisions mentioned here: recognition of the rights of the states to have a voice in implementation of rules under federal laws. The highlights of this rule (at the time of writing) are detailed below.

- ❑ Coverage. Power plants in 27 eastern US states with respect to SO<sub>2</sub> and NO<sub>x</sub> emissions contributing to pollution in other states are covered by this rule. The number of affected coal-, natural gas- and oil-powered power plants is equal to 1,074 (3,632 generation units). State by state coverage is shown in Table 25.1.
- ❑ Timeline. The rule applies to SO<sub>2</sub> emissions, starting on January 1, 2012 and to NO<sub>x</sub> emissions starting May 1, 2012.
- ❑ Benefits. The EPA estimates that the rule will result in US\$120–280 billion in annual benefits, arising from, among other things, the prevention of:
  - 13,000 to 34,000 premature deaths;
  - 19,000 cases of acute bronchitis;
  - 15,000 non-fatal heart attacks;
  - 19,000 hospital and emergency room visits;
  - 1.8 million days when people miss work or school;
  - 400,000 cases of aggravated asthma; and
  - 420,000 cases of upper and lower respiratory symptoms.

### Utility MATS

The second rule currently under consideration by the EPA is related to mercury and toxic air pollution: Mercury and Air Toxics Standards (MATS).<sup>32</sup> The rule, issued on December 21, 2011, pursuant to Section 112<sup>33</sup> of the CAA, covers emissions of mercury and other toxic substances, such as arsenic, acid gas, nickel, selenium and cyanide. The affected sources have to comply by 2014, with the possibility of a one-year extension (primarily in the case of reliability concerns caused by plant deactivations and delays in the installation of controls).<sup>34</sup> Mercury is a metal present in varying concentrations in different types of coal. The biggest potential threat of mercury is the potential for entering into a food chain through transformation into

methyl mercury. This compound is formed when mercury is deposited in water, and then absorbed by microorganisms that attach to it a methyl group ( $\text{CH}_3$ ). This substance accumulates through the food chain in crustaceans and fish, as there are no effective processes for purging it from their bodies. Long-living fish, like tuna and swordfish, can reach a high level of concentration of mercury in their bodies by consuming other fish, and eventually become the main source of human exposure to this metal.<sup>35</sup> Mercury can cause irreversible damage to the central nervous system (especially in foetuses and young children) and has also been linked to cardiovascular problems.<sup>36</sup>

The MATS is expected to reduce mercury emissions from power plants by 90%. The rule applies to about 1,000 coal-fired plants and 300 oil-fired plants. According to EPA estimates, compliance with the new rules will prevent 11,000 premature deaths and 4,700 heart attacks each year. An additional benefit is the prevention of 130,000 cases of childhood asthma and 6,300 cases of acute bronchitis per year. The combined health benefits of MATS and CSAPR are estimated by the EPA<sup>37</sup> to equal US\$380 billion by 2016.

The CAA contains the provisions that require all sources of hazardous air pollutants to install maximum achievable control technology (MACT).<sup>38</sup> The MACT standard establishes a benchmark defined by the best-performing 12% of existing sources. This explains why the MATS rule is often referred to as "Utility MACT." Compliance with the rule will happen through either deactivation of existing plants (primarily older and smaller plants) or through the installation of emission control equipment, using a number of available technologies, such as flue gas desulphurisation (FGD), dry sorbent injection (DSI), activated carbon injection (ACI) and fabric filters.

### **Market impact**

The environmental regulations discussed above have created a cottage industry of consulting efforts and thinktank studies offering analysis of expected market impact, including the consequences for grid reliability.<sup>39</sup> The results of these studies should be treated with caution, not because we have any reservations regarding the quality of analytical effort, but most studies were published prior to promulgation of the final rules, at different times and under different initial

**Table 25.1** CSAPR coverage

State	Reducing emissions of NO <sub>x</sub> during the ozone Season (1997 Ozone NAAQS)	Reducing annual emissions of SO <sub>2</sub> and NO <sub>x</sub> (1997 Annual PM2.5 NAAQS)	Reducing annual of SO <sub>2</sub> and NO <sub>x</sub> (2006 24-hour PM2.5 NAAQS)
Alabama	X	X	X
Arkansas	X		
Florida	X		
Georgia	X	X	X
Illinois	X	X	X
Indiana	X	X	X
Iowa	X (proposed)	X	X
Kansas	X (proposed)		X
Kentucky	X	X	X
Louisiana	X		
Maryland	X	X	X
Michigan	X (proposed)	X	X
Minnesota	X		
Mississippi	X		
Missouri	X (proposed)	X	X
Nebraska	X		
New Jersey	X	X	
New York	X	X	X
North Carolina	X	X	X
Ohio	X	X	X
Oklahoma	X (Proposed)		
Pennsylvania	X	X	X
South Carolina	X	X	X
Tennessee	X	X	X
Texas	X	X	
Virginia	X	X	
West Virginia	X	X	X
Wisconsin	X (Proposed)	X	X

Source: <http://epa.gov/airtransport/pdfs/CSAPRFactsheet.pdf>

market conditions. The fundamental difficulty is estimation of the combined impact of different rules, with some of the mitigation measures addressing more than one of the new standards. We have given above several reasons why any forecasting effort is very difficult to exercise and has a wide margin of error.

The primary impact of the new rules is related to retirement of the older and smaller generation plants for which the cost of installation of control equipment may be prohibitively high. This is illustrated by Figures 25.3 and 25.4, which show the cost of installing SO<sub>x</sub> and NO<sub>x</sub> controls in coal-fired plants of different size.

The cost of FGD in a coal-fired plant of 50 MW capacity is equal to US\$1,137/kW, an amount exceeding the cost of a new combined cycle gas-fired unit. It is obvious that a rational decision would be to shut down such a coal unit.

The estimates of retired coal-fired capacity (in GW) by 2015 vary from study to study. Some estimates, provided here just to illustrate the order of magnitude of the problem, include:

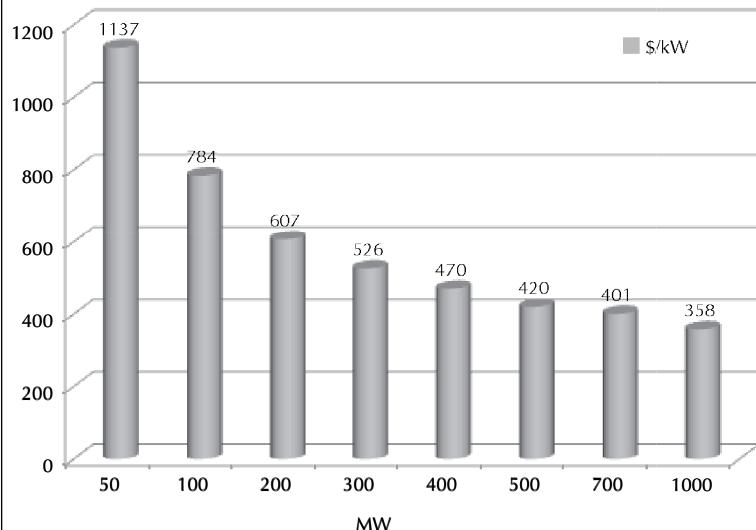
- 30–40 GW (PIRA Energy Group, 4/2010);
- 50 GW (The Interstate Natural Gas Association of America (INGAA), 5/2010);
- 25 GW (Edison Electric Institute (EEI), 5/2010);
- 50 GW (Credit Suisse, 7/2010);
- 65 GW Bernstein Research (10/2010);
- 6–25 GW (NERC, 10/2010); and
- 35 GW (Charles River Associates (CRA) 12/2010).

As a frame of reference, the total US-installed, coal-fired capacity is equal to 315 GW.

### CARBON MARKETS

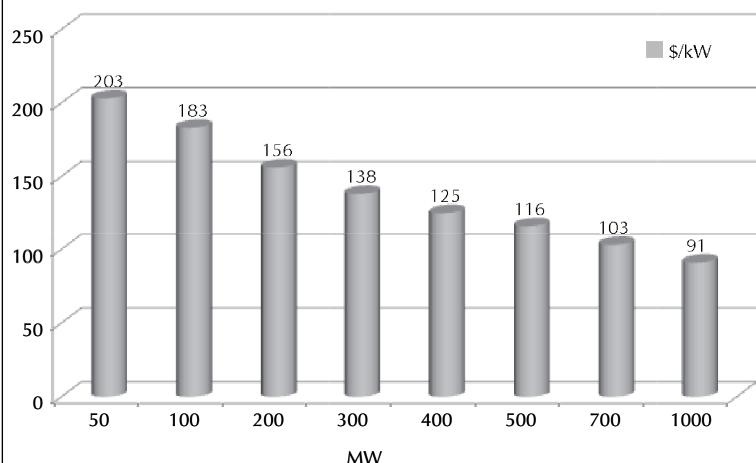
The greenhouse effect is the process of the warming of the earth's surface due to entrapment of heat by the atmosphere. We can relate to this natural phenomenon by recalling the experience of sitting at a desk next to a window on a sunny day. Light from the sun (short, visible waves) penetrates through the glass and is absorbed by the objects inside the room, but the return radiation, longer infrared waves, cannot pass through the window and is trapped inside. This increases the temperature inside the room and makes the experience of working at one's desk quite uncomfortable. Global warming happens exactly in the same way. Energy emitted by the sun penetrates the earth's atmosphere and warms up the surface of the planet. About 50% of energy is absorbed and the rest is reflected back into the space. Certain gases naturally present in the atmosphere absorb partially these infrared emissions, to our great relief. If this did not happen, the average temperature of the earth would be around minus 18°C. Increasing concentrations of certain gases in the atmosphere, due to human economic activity, reduce the portion of energy escaping into the space and cause what is called global warming.

**Figure 25.3** Capital cost of FGD retrofits by generation unit capacity (US\$/kW)



Source: Bernstein Research

**Figure 25.4** Capital costs of NO<sub>x</sub> controls by generation unit capacity (US\$/kW)



Source: Bernstein Research

Main greenhouse gases include steam, carbon dioxide ( $\text{CO}_2$ ), methane ( $\text{CH}_4$ ), chlorofluorocarbons (CFC), nitrogen dioxide ( $\text{N}_2\text{O}$ ) and hydro fluorocarbons (CFH).<sup>40</sup> Most greenhouse gases are much more potent than  $\text{CO}_2$  and may last in the atmosphere for a much longer time. In 1990, the International Panel on Climate Change (IPCC) produced a report<sup>41</sup> summarising the information about the potency of different greenhouse gases as a multiple of the global warming potential of the same amount by weight of  $\text{CO}_2$  (see Table 25.2).

**Table 25.2** Global warming potential, 100-year horizon

Gas	Global warming potential
$\text{CO}_2$	1
Methane ( $\text{CH}_4$ )	21
Nitrous Oxide ( $\text{N}_2\text{O}$ )	290
CFC-11	3,500
CFC-12	7,300
HCFC-22	1,500

Source: IPCC, 1990 Report, Page XXI.

It goes without saying that the global warming hypothesis is very controversial, in spite of the overwhelming majority of scientists (often-mentioned as a 97 to 3 ratio, give or take)<sup>42</sup> subscribing to this theory. The debate often degenerates into invectives, accusations of intellectual dishonesty (the supporters allegedly seeking research grants and publicity, the opponents selling out to powerful interests) and sometimes outright criminal activity (hacking of Internet accounts and stealing of emails). In the interest of full disclosure, for the record we subscribe to the global warming theory; having said that, we have a number of friends who are climate scientists and highly ethical individuals, and who have very strong doubts regarding global warming.<sup>43</sup> This demonstrates that it is imperative for every responsible citizen to educate himself about global warming and form an informed opinion. If this is a real threat to humanity, this advice is obvious. If it is an example of collective delusion by the scientific community, the resources likely to be allocated to address this problem on a planetary scale will be massive, and this waste should be prevented.

The critics of the global warming theory take a number of different positions.

- ❑ Global warming does not exist and is an invention (not to say hoax). This position is difficult to defend, however, in view of the incontrovertible information to the contrary, as illustrated by the history of global temperature data (temperature anomaly, or deviation from the mean) over the period 1850–2011 (see Figure 25.5).
- ❑ A less extreme position is to recognise global warming as an empirical fact but treat it as a natural phenomenon (not an anthropogenic consequence of human economic activity), most likely a manifestation of the natural cyclical oscillations of temperatures.
- ❑ Even if global warming is for real, an optimal strategy is to adjust rather than counteract.
- ❑ Global warming will produce net benefits to mankind (Scottish wine when sunbathing on a Siberian beach anybody?).
- ❑ Other counter-arguments against global warming include:<sup>44</sup>
  - CO<sub>2</sub> concentrations are not sufficient to cause the increase in the global temperatures;
  - CO<sub>2</sub> concentrations were high in the past (10 times as much as today), without causing catastrophic consequences);
  - surface warming exceeds atmospheric warming, contrary to the findings of the models behind the global warming hypothesis; and
  - global warming is likely to be caused by changes in solar activity.

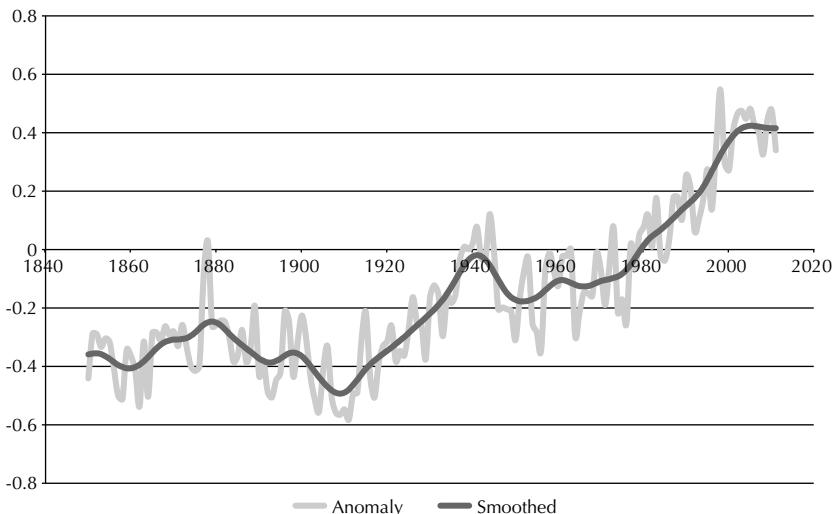
As we have said, in our view, global warming is a serious threat to our planet but the likelihood of an effective collective action remains quite small. Humans are not very rational creatures.<sup>45</sup> Our advice to anybody in the carbon trading business is to engage in this activity only if there is a true belief in the ethical and moral case for global warming mitigation measures. This will be always a nerve-wrecking business, one failed COP (to be explained shortly) meeting away from obliteration.

### **The legal framework**

The greenhouse effect has been known to scientists for a long time. President Lyndon Johnson was briefed about the phenomenon back in 1965 and we recall several books and articles on this topic in the

early 1970s. Initially, there was no sense of urgency in the scientific community as there was a general consensus that this process would unfold at a relatively low rate over a long time. The accumulation of more climate-related data and progress in the computer modelling and simulation of natural phenomena led to growing concern about the potential for adverse and irreversible change in the earth's climate. In 1988, the World Meteorological Organization and the United Nations Environmental Program established the Intergovernmental Panel on Climate Change (IPCC), with a mandate to review and assess "the most recent scientific, technical and socio-economic information produced worldwide relevant to the understanding of climate change."<sup>46</sup> The IPCC "does not conduct any research nor does it monitor climate related data or parameters."<sup>47</sup> Scientists from many countries contribute information on a voluntary basis, with 195 governments participating. Part of the mission of the IPCC is management of the process of collecting information and review of the submitted data to ensure quality control. At the Rio de Janeiro Earth Summit in 1992, the United Nations Convention on Climate Change (UNCCC) was adopted, with an objective "to cooperatively consider what they could do to limit average global temperature increases and the resulting climate change, and to cope with whatever impacts were, by then, inevitable."<sup>48</sup> It had three separate categories of signatories.

- ❑ *Annex I:* countries willing to limit the greenhouse gas (GHG) emissions and reduce the total volumes to the levels reached in earlier periods. Annex I countries include the industrialised countries which were the members of the OECD in 1992, and economies in transition (EIT), including the Russian Federation, the Baltic states and certain East European countries.
- ❑ *Annex II:* Annex I countries that agreed to help the developing countries to reduce emissions and adjust to climate change through different measures, including financial aid, development and the transfer of new technologies. The countries included in this category are OECD members, but not the EIT countries.
- ❑ *Non-Annex I:* no specific commitment with respect to climate change measures. This category includes primarily the developing countries. Non-Annex I countries are considered to be

**Figure 25.5** Global air temperature (1850–2011)

Source: <http://www.cru.uea.ac.uk/cru/info/warming>

especially vulnerable to climate change for two reasons. They are located in regions characterised by extreme climatic conditions that will be greatly aggravated by global warming. Also, many countries depend on mineral industries and may be on the receiving end of measures taken in the developed world to reduce GHG emissions.

The decision-making body of the convention is the Conference of Parties (COP), which convenes annually, starting with a Berlin meeting in 1995. Two other important auxiliary organs of the COP are the Subsidiary Body for Scientific and Technological Advice (SBSTA) and the Subsidiary Body for Implementation (SBI). Meetings of the COP (referred to as COP1, COP2, etc)<sup>49</sup> are the milestones of international cooperation regarding climate change, with the most important conferences being:

- ❑ COP3 (1997): Kyoto Protocol;
- ❑ COP7 (2001): The Marrakesh Accords, detailed rules for implementation of the Kyoto Protocol;
- ❑ COP13 (2007): the Bali Road Map;

- COP15 (2009): The Copenhagen Accord; and
- COP17 (2011): the Durban Platform for Enhanced Action.

The framework for international cooperation on climate change, due to last until the end of 2012, was established at the COP3 in Kyoto in 1997, when a Protocol to the United Nations Framework Convention on Climate Change was adopted through the consensus of member countries. The provisions of the Protocol included:

- legally binding targets for Annex I countries to reduce their emissions by at least 5% compared to the 1990 levels, between 2008 and 2012 (EU, Switzerland, most of Central and Eastern Europe: 8%, US: 7%, Canada, Hungary, Japan, Poland: 6%);
- a cap-and-trade framework with reduction targets translated into specific numerical goals, known as assigned amounts (AAs);
- the levels of assigned amount units (AAUs) were determined for the 2008–12 period; and
- the reduction of emissions through production cuts, new technologies or purchases of AAUs from surplus countries.

Emission reductions of three principal greenhouse gases ( $\text{CO}_2$ ,  $\text{CH}_4$  and  $\text{N}_2\text{O}$ ) are measured with respect to emissions during the reference year of 1990 (with the exception of certain EITs). Reductions of emissions of three industrial gases (HFC, PFC and  $\text{SF}_6$ ) are measured using either 1990 or 1995 as a year of reference.<sup>50</sup>

Target emission reductions could be achieved through three complimentary mechanisms:

- clean development mechanism (CDM);
- joint implementation (JI); and
- international emission trading (IET).

Under CDM, Annex I investors develop carbon-reducing projects in countries outside this group and use Certified Emission Reductions (CERs), credits generated through these projects to meet Annex I countries' emission reduction targets. Under JI, investment projects supported by Annex I countries are located in other Annex I countries or in countries in transition. JI projects create credits known as emission reduction units (ERUs).

## EU emission trading system

The most important cap-and-trade system was established by the EU under the European Emissions Directive (EED), which became law on October 25, 2003.<sup>51</sup> Member states were obliged to:

- translate Kyoto objectives into specific targets (caps);
- develop a system for the allocation of caps to designated sectors and specific installations;
- establish emissions inventories and develop procedures for emissions measuring and monitoring; and
- set up national registries to capture information about emissions, compliance and trades.

National quotas for greenhouse emissions were determined through a mechanism known as the national allocation plan (NAP). NAPs for the first (2005–07) and the second (2008–12) phase of the European Trading System, ETS, to be covered shortly) contained the targets (established by the EU member countries' governments) regarding the total level of allowances for each country for each phase of the ETS, and allocations to each covered installation (ie, sources of emissions subject to a CO<sub>2</sub> cap).

The deadline for the first NAP was March 31, 2004 (May 1, 2004, for the 10 countries joining the EU in 2004). The deadline for the second phase was June 30, 2006. There will be no NAP for the period beginning in 2013, as this decision (overall and installation specific emission targets) becomes the prerogative of the EU.

An overall emission reduction target (against the 1990 benchmark) was binding under the Kyoto Protocol, but national quotas varied in different years from emission increases to reductions close to 30%. A New Entrants Reserve (NER) was established,<sup>52</sup> which allowed for the allocation of allowances to new or expanded<sup>53</sup> emission sources. The rationale behind it was an equity argument, that it was unfair to penalise new sources of emissions by forcing them to buy allowances while granting free allocations to old sources. Under Phase I and II of the ETS, the new source allocations were left to individual states. In March 2007, EU governments agreed on a very ambitious energy policy, known as the “20–20–20” target, containing a number of energy and climate change policy objectives for 2020, including:

- a reduction in the EU GHG emissions by at least 20% compared to 1990 levels;
- 20% of energy consumption to come from renewable sources; and
- a 20% reduction in energy consumption (versus business-as-usual levels) through efficiency measures.

The EU cap-and-trade programme, the cornerstone of the EU climate change policy, was launched in 2005 as the ETS,<sup>54</sup> covering 30 countries (27 EU members plus Iceland, Liechtenstein and Norway) and about 11,000 power and industrial plants (oil refineries, steel foundries, cement, glass, lime, bricks, ceramics, pulp and paper factories). As of February 2012, the installations covered under the programme account for about 50% of EU carbon emissions and about 40% of GHG emissions.<sup>55</sup> The programme has been implemented in two trading periods, 2005–07 and 2008–12, with a third period – extending through 2020 – scheduled to start in 2013. There have been important changes in the design of the ETS from one period to another, and we shall highlight some of them.

The NAPs determine the total quantity of annual emissions at the EU level. This decision-making process is decentralised, with each member state of the EU deciding on the overall numbers of allowances (each allowance is one ton of CO<sub>2</sub>) and the allocation of allowances to different installations covered under the ETS. This means that the EU level cap is the outcome of decisions taken by sovereign states. However, from 2013, the allocations will be decided at the EU level. The NAPs are subject to the scrutiny of the European Commission, which reviews the national targets against the principles governing allocations as listed in the Emissions Trading Directive (ETD).<sup>56</sup> The Commission seeks to ensure that the national plans are consistent with the commitments made under the Kyoto Protocol, achieve emission reductions and have the existing technological means to reduce emissions. The national plans can be accepted in part or in full. Partially accepted plans are not required to be re-submitted if the EU recommendations were fully implemented.

Prior to 2013, emission allowances were allocated free of charge, with a certain amount being retained for auctions. Starting in 2013, NAPs will be replaced by a centralised allocation of permits, with

about 50% of emission allowances to be auctioned. Free allocations of emission allowances to electricity producers will be discontinued (with some temporary exclusions from this rule). Free allocations will continue in the case of the installations exposed to a risk of carbon leakage, with auctions used as a primary mechanism for sources not exposed to this threat.<sup>57</sup> An auctioning regulation<sup>58</sup> envisages the creation of a common platform to auction allowances, with member states having the discretion to create their own auctions.

The allowances granted under the ETS are recorded in the national registries, which gather information about the number of allowances assigned to each state, transactions of allowances among accountholders, actual emission levels and the reconciliation of actual emissions against surrendered allowances. During 2012, the national registries were gradually replaced by a centralised registry operated at the EU level. The registry was partially activated in January 2012 to enable compliance with the ETS by aviation. The Consolidated System of European Union Registries (CSEUR) was fully activated on June 20, 2012. This decision will not only facilitate the trading of allowances across national boundaries within the EU, but will also address some irregularities plaguing the national registries.<sup>59</sup>

A covered installation (ie, a source of GHGs included in the ETS) has to surrender one allowance for each ton of CO<sub>2</sub> emitted (in plain English, this means that if you emit CO<sub>2</sub>, you have to give back a corresponding volume of allowances). An installation with a surplus of allowances may sell them in the market; an installation with a deficit has to go to the market to cover the shortfall. The allowances can be banked within a trading period. This means that an allowance issued, for example, in 2005, could be used (sold or surrendered) through 2007, and then would expire. Allowances have to be submitted by the end of April, which means that an installation can effectively borrow from its current year allocation (as new allowances are issued in February) in order to satisfy obligations related to the previous year. The failure to surrender allowances against actual emissions results in penalties of 100 euro per ton. The rules related to banking allowances<sup>60</sup> contributed to the price dynamics of allowances, characterised by persistently high levels followed by crashes to prices close to zero in some periods (as discussed in more detail below).

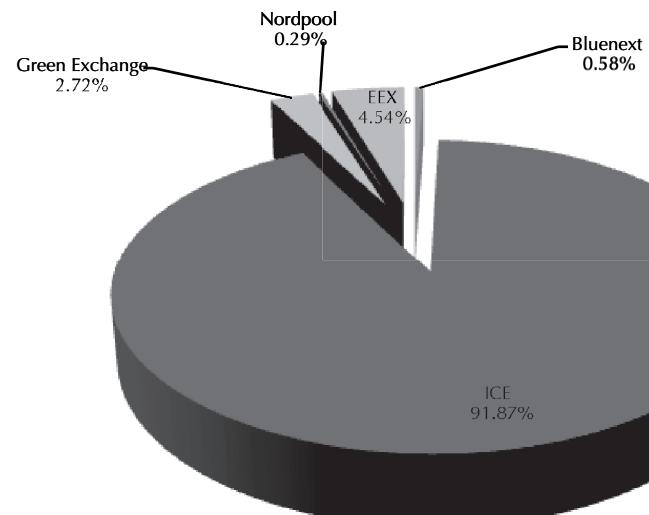
The design of ETS allows the delivery of JI and CDM credits to offset emissions, up to a ceiling defined for each NAP. Unused credits can be banked and used during the 2013–20 trading period. Starting in 2013, the EU will not accept CDM credits created through projects registered after January 1, 2013, except for projects located in the least-developed countries (LDCs).

Emission-related products are traded on a number of European exchanges, with the dominant position held by ICE (Figure 25.6).

ICE Futures Europe<sup>61</sup> lists derivative contracts on three types of carbon emission-related products: EUAs (starting in 2005 with futures, followed by options in the next year), CERs (options and futures introduced in 2008, followed by a spot contract the next year) and ERUs. EU Aviation Allowances futures (EUAAAs), introduced in February 2012, have also been added to the ICE carbon product slate.

The Green Exchange is owned by a number of financial and trading firms, including the CME Group, Constellation Energy, Credit Suisse, Evolution Markets, Goldman Sachs, ICAP Energy, JP Morgan Chase & Co, Morgan Stanley, RNK Capital, Spectron, TFS Energy, Tudor Investment and Vitol SA. The company offers a wide range of contracts covering such products as EUAs, UN-certified

**Figure 25.6 Carbon Exchanges Market Shares (June 2012)**



Source: ICE

CERs, US Regional Greenhouse Gas Initiative carbon allowances (RGGIs), US emissions allowances (SO<sub>2</sub> and NOx) and Californian carbon allowances (CCAs).<sup>62</sup>

### ETS assessment

The first ETS trading phase (the second phase will terminate at the end of 2012) can be characterised as a qualified success, and a period of experimentation during which market participants and the EU governing bodies learned many useful lessons and laid foundations for a much more efficient European carbon market than the system that has been in place previously. The ETS experience also proved very useful for other countries experimenting with emission markets (to be discussed later in this chapter). This does not mean that ETS was a flawless and seamless system.

The most obvious manifestation of flaws in the initial ETS design was excessive price volatility. An example is a price shock that took place in late April 2006, and which is still vividly remembered by many traders and debated in academic papers.<sup>63</sup> The contract for delivery in December 2007 peaked on April 19, 2006, at €31.58, dropping to €9.70 on May 11. After staging a recovery, the price of this contract dropped below €10 again towards the end of 2006, and eventually sank under €1.<sup>64</sup> Phase II allowances traded consistently at around €20 in the forward markets. The sudden drop in prices of Phase I allowances was attributed to the leakage of information (or exceptional foresight demonstrated by some analysts) about an opening gap between actual emissions and granted carbon allowances. The total cap was not really binding and, as economic theory would predict, the clearing market price of allowances should be equal to zero (with some residual noise to be expected).

The over-allocation of allowances was due to a combination of flaws in the design of ETS and inherent difficulties of predicting emission volumes under any circumstances. The level of emissions depends on the level of economic activity and weather conditions (the main drivers behind electricity production). Warm winters and mild summers reduce the demand for electricity and, therefore, CO<sub>2</sub> output from thermal generation plants. Recession and slow economic growth in the EU towards the end of the first decade reduced industrial output and emissions. Other emission sources covered under ETS included steel and cement plants, with outputs

highly sensitive to the rate of economic growth. Another important factor behind changes in emissions from power plants were abatement measures, primarily the greater utilisation of natural gas-fired units, which emit less CO<sub>2</sub> per unit of output compared to coal plants. Other abatement measures include shifting energy-intensive production processes outside the EU. The shift towards natural gas was primarily driven by the introduction of the carbon market and, in the absence of historical data and direct experience, any forecast of fuel mix would have had inevitably a large margin of error, compounded by the difficulty of predicting other drivers (thanks to the capriciousness of both weather and economic conditions).

The decentralisation of decisions regarding national allowances led inevitably to some countries being more lenient than others, and to imperfect distribution of overall allowances across different industries. The industries that were long had no urgency to sell, but those that were structurally short had to cover.<sup>65</sup> The market participants who had to cover were primarily electric utilities, the longs were often industrial companies (some of them in Eastern Europe), with limited experience in trading. This created upward market pressure and a tendency to hoard allowances as prices were increasing. Some East European countries were falling behind the schedule in creation of their national registries, and this has removed some sources of allowances from the market at the beginning of Phase I.<sup>66</sup> These countries were also undergoing economic transformation and the downsizing of heavy industries, a legacy of their communist past. This was another factor leading to excessive national allocations, with no attempt to favour particular national economies. Economic restructuring and recession resulted in an unexpected drop in emissions. Between 2006 and 2010, the drop in CO<sub>2</sub> emissions ranged from 23.5% in Romania to 3.4% in Slovakia. Poland was the only East European EU country that escaped recession and registered a small increase in emissions.<sup>67</sup>

Some of the initial design flaws were addressed going into Phase II. The European Commission pressured individual countries to be less cavalier in the determination of national allowance allocations. The economic crisis of 2007–09 caused a drop in the level of economic activity and reduced overall carbon emission levels, recreating, to a lesser extent, complications experienced in Phase I. EU annual CO<sub>2</sub> emissions in the period 2006–11 were equal, respectively, to 4520.3, 4464.3, 387.6, 4067.5, 4177.8, 4061.3 million tons.<sup>68</sup>

The prices of carbon emissions traded under the ETS have remained depressed, falling in July 2012 to new lows.<sup>69</sup> The main reason is a continued period of slow economic growth, with some EU countries sinking, after a short-lived recovery in 2010 and 2011, back into recession. One of the remedies proposed to support the sagging market is postponing as much as 40% of the carbon allowances expected to be auctioned over the following three years.<sup>70</sup> The problem with such *ad hoc* interventions is that they undermine the ability of emitting companies to implement effective mitigation measures, which require the ability to anticipate market conditions and the relative costs of different solutions.

### **US Carbon emission markets**

The US did not ratify the Kyoto Protocol.<sup>71</sup> This does not mean, however, that there is no potential for the development of vibrant regional US carbon markets or even nationwide markets. The US policy in this area is critical to the future of the green markets. At the time of writing, the US ranks second in terms of total emissions of CO<sub>2</sub> (having been recently surpassed by China)<sup>72</sup> and, without its participation, the measures countervailing global warming cannot be successful. The introduction of carbon trading at the federal level hinges on the results of elections in the autumn of 2012. The likely design of a future carbon market can be inferred with a high degree of confidence from the details of two climate change bills that failed to clear Congress a few years ago.

The American Clean Energy and Security Act of 2009 (H.R. 2454, or Waxman–Markey) was passed by the House in June 2009. A corresponding Senate bill was introduced in May 2010 (The American Power Act, or Kerry–Lieberman). The 111th Congress failed to act on these two legislative initiatives, and both bills are likely to remain in limbo unless the trend towards a more conservative House and Senate proves temporary, and the US economy fully recovers from recession. Both bills share a common approach to regulating carbon emissions and it is very likely they will be resurrected in the future. Some common features are detailed below.

- ❑ A cap-and-trade approach with tradable allowances allocated initially at no-cost to the covered sources with a growing percentage of permits being sold at auction over time. Most

allowances would be allocated at zero cost to regulated utilities until 2025, with gradual elimination of the grants between 2025 and 2030. Both bills anticipated roughly equal amounts of annual CO<sub>2</sub> allowances.

- ❑ Both bills recognise technological differences between gas-fired and coal-fired units. On average, the emissions per MWh are equal to one ton in coal units and about 0.5 ton in gas units. Assuming that gas-fired units will be at the margin, the coal units would need to receive half a ton of allowances per MWh to maintain their competitive position (the price they receive will go up as the costs of marginal generators increase, which will work to their advantage).
- ❑ The proposed legislation recognised the differences between sources of emission, classifying them into two categories:
  - large stationary sources (more than 25,000 metric tons of CO<sub>2</sub> per year); and
  - mobile sources and small stationary sources.
- ❑ Large sources of emissions would be regulated directly through an allowance system. Small sources would be regulated indirectly through allowances required from the suppliers of fuels they use. The suppliers would have to surrender a sufficient number of allowances to cover greenhouse gases emitted by the fuels they delivered to end users.
- ❑ Kerry–Lieberman bill would have introduced a collar on allowance prices to limit their volatility. The 2013 floor would be set at US\$12 per metric ton (2009 dollars), increasing annually at the rate equal to the CPI + 3 percentage points. A cost containment reserve of 4.0 billion allowances would help to contain the price spikes to the upside. Emitters would be allowed to buy up to 15% of their annual requirements from the reserve at a price starting at US\$25/ton in 2013, escalating at a rate equal to CPI + 5 percentage points.
- ❑ Both bills contained provisions seeking to address the issue of carbon leakage and protect certain energy intensive industries. The Kerry–Lieberman bill would set-up a “border adjustment” mechanism (starting in 2025), a requirement for importers of carbon-intensive materials (such as steel or cement). Imports from countries with no climate change regulation would be subject to a carbon duty, defined in terms of allowances the importers would have to deliver.

- ❑ Both bills allowed the use of offsets, both from domestic and international sources, subject to annual caps and subject to projects approval by the EPA and the Department of Agriculture.
- ❑ The Kerry–Lieberman bill limited the scope of EPA oversight of the carbon markets, and banned state and regional carbon programmes (such as those described below).

The trend in US greenhouse emissions and the expected combined outcome of the EPA regulations may help to make the topic of carbon regulations less controversial and increase the probability of a compromise. The retirement of many older and less-efficient coal plants and the expansion of gas-fired generation will slow down the growth of the US greenhouse emissions. Demographic trends (the increasing percentage of retirees in the overall US population) will have a similar impact (retired people drive less). The recession and slow economic growth has also resulted in a decrease in energy-related GHG emissions (Figure 25.7).

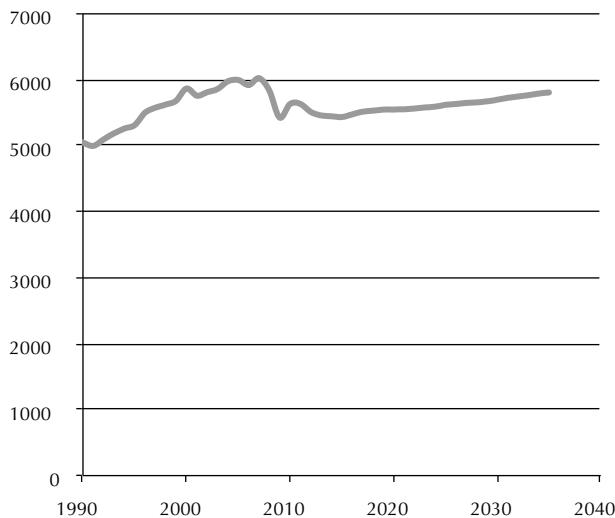
The same trends can be seen in the overall US emissions of GHGs from all sources, including energy. Energy-related carbon dioxide emissions are projected to stay below maximum historical levels between 2012 and 2035, even in the absence of policy measures targeting greenhouse gas emissions.<sup>73</sup>

Another factor that may create incentives for politicians to seek a compromise is a sense of inevitability of some form of carbon emissions regulation. In 2007, the US Supreme Court ruled that the CAA gave the EPA authority to regulate greenhouse gases if the agency declared them a threat to public welfare and that they endanger human health. The EPA followed up with an endangerment finding in December 2009 and then with a tailoring rule, which determined which emitters are subject to carbon emission regulation. The EPA actions are being contested in the courts in more than 60 suits, but one can expect that some form of regulation will be eventually put in place.

## RGGI

The RGGI is the first mandatory carbon CaT system in the US covering a number of states in the North East (Connecticut, Delaware, Maine, New Hampshire, New York, Rhode Island and Vermont). The cap, introduced in 2009, was established at 165

**Figure 25.7** Energy-related carbon dioxide emissions (million metric ton CO<sub>2</sub> equivalent)



Source: U.S. Energy Information Administration

Note: 1990–2011 actual data, 2012–35 forecast (from AEO 2012)

million short tons for 2012, and applies to fossil fuel-fired plants with capacity equal to, or greater than, 25 MW (affecting 209 facilities as of February 2012). It is expected that the cap will stabilise carbon emissions over the first implementation period (2009–14), and starting in 2015 the cap will be reduced by 2.5% annually, for a total reduction of 10% by 2018. Covered facilities have to surrender an allowance for each ton of CO<sub>2</sub> emitted. Most allowances (90%) are auctioned on a quarterly basis, with cumulative revenues equal to US\$952 million through December 2011. The revenues are spent primarily (about 80%) in programmes related to energy efficiency, renewable energy, energy bill assistance for low-income households and other projects helping to reduce greenhouse emissions.

The auctions are based on a uniform price, single round and sealed bid format, with a single bidder having the option of submitting multiple bids. Associated parties may acquire a maximum of 25% of the allowances offered at an auction. Prices have tended to generally hover under US\$2.00 (see Table 25.3 for more information). The allowances trade in the secondary markets, primarily on the

**Table 25.3** CO<sub>2</sub> RGGI allowance prices in the auction and secondary markets (volume-weighted averages)

Product	2010	2011
Futures V09/10/11	US\$2.10	US\$1.93
Clearing price – first	US\$1.93	US\$1.89
Clearing price – second	US\$1.86	US\$1.89
Physical delivery COATS	US\$2.16	US\$1.86

V refers to vintages of allowances, “first” and “second” – to the control (compliance periods) periods,<sup>74</sup> and COATS<sup>75</sup> is the registry for RGGI allowances.

Source: “Annual report on the market for RGGI CO<sub>2</sub> allowances: 2011,” Potomac Economics, May 2012

Green Exchange. The overall impact of the RGGI on retail electricity prices is very modest, and was estimated to average 46 cents per month per household.

RGGI allows the use of offsets, defined as “project-based greenhouse gas (GHG) emissions reduction or carbon sequestration achieved outside of the capped electricity sector.”<sup>76</sup> The regulated sources can use qualifying offsets to satisfy up to 3.3% of their obligation. Under certain conditions, this percentage can increase to 5% or 10% (if CO<sub>2</sub> allowance prices exceed the levels of US\$7 and US\$10, respectively). This condition has not been triggered so far, given low RGGI allowance prices. The projects that qualify for creation of offsets include:

- ❑ the capture or destruction of CH<sub>4</sub> (methane) from landfills;
- ❑ reduction of emissions of SF<sub>6</sub> from electricity transmission and distribution equipment;
- ❑ sequestration of CO<sub>2</sub> through afforestation;
- ❑ reduction of CO<sub>2</sub> emissions through non-electric, end-use energy efficiency in buildings; and
- ❑ the avoidance of CH<sub>4</sub> emissions through agricultural manure management operations.

### California cap-and-trade programme

The cap-and-trade programme in California was established under the Global Warming Solution Act (California Assembly Bill AB 32), passed in the more prosperous days of 2006. The act authorised the California Air Resources Board (CARB) to reduce greenhouse gas

emissions in California to 1990 levels by 2020. The emission sources covered by the act include electric utilities, large industrial facilities and suppliers of natural gas and transportation liquids (600 facilities owned by 350 firms). The programme starts in 2013 for electric utilities and industrial companies, and in 2015 for the distributors of fuels. The cap for 2013 was set 2% below the emission forecast for 2012, with a decline of 2% envisaged for 2014, and 3% reductions annually between 2015 and 2020. Specifically, the cap for 2013 was set at 162.8 MMtCO<sub>2</sub>e (million tons of CO<sub>2</sub> equivalent); jumping to 394.5 MMtCO<sub>2</sub>e for 2015. The target for 2020 is 334.2 MMtCO<sub>2</sub>e. The increase in 2015 is related to the extension of the programme to fuel distributors. Compliance with the cap is determined over multi-year periods: the first compliance period covers 2013 and 2014, the second extends from 2015 to 2017, the third from 2018 to 2020.

For large industrial facilities, the allowances are allocated initially free of charge with a transition to an auction-based distribution system over time. The allowance levels for each industry group are set at 90% of average emissions for a benchmark established for efficient facilities. A quarter of allowances are allocated to electric utilities, with the benefits going to ratepayers (the utilities sell allocations at auctions and use the revenues for the benefit of ratepayers).

## CONCLUSIONS

The carbon emissions market is a market of great and yet unfulfilled promise. It has a potential to become a crucial link between different geographical markets and different segments of the energy complex, such as coal, natural gas, electricity and oil. Its full implementation would create incentives to change the composition of energy supplies, with a forced shift towards such sources as solar and wind. This is not going to happen as long as the implemented carbon markets remain local and partial solutions to a global problem – a global problem that many influential opinion makers do not think exists at all. For the time being, we see this market as limping from one temporary patch to another, and from one localised solution to another.

---

<sup>1</sup> “The ideas of economists and political philosophers, both when they are right and when they are wrong, are more powerful than is commonly understood. Indeed the world is ruled by little else. Practical men, who believe themselves to be quite exempt from any intellectual influence, are usually the slaves of some defunct economist. Madmen in authority, who hear voices in the air, are distilling their frenzy from some academic scribbler of a few years back.

I am sure that the power of vested interests is vastly exaggerated compared with the gradual encroachment of ideas." John Maynard Keynes, "The General Theory of Employment, Interest and Money," Chapter 24.

- 2 See Robert N. Stavins, 2001, "Experience with market-based environmental policy instruments," *Resources for the Future*, November, discussion paper 01-58.
- 3 The firms covered under the programme could transfer reductions internally to comply with an aggregate pollution limit (this was referred to as "netting," or "bubbles"), or could sell them to other firms should they have a surplus over their internal needs. Internal transfers refer to the ability to use emission reductions achieved at one plant at the level of the entire company. In 1977, the programme was expanded by requiring new sources (ie, new plants) to comply with ambient air standards by offsetting incremental emissions with reductions of emissions elsewhere below historical levels. As far as I know, the word "offset" was used in this context for the first time.
- 4 The first CaT programme I am aware of was related to the Montreal Protocol, an international agreement targeting CFCs and halons (ozone depleting gases); see R. W. Hahn and A. M. McCartland, 1989, "Political economy of instrumental choice: An examination of the US role in implementing the Montreal Protocol," *Northwestern University Law Review*, 83, pp 592-611.
- 5 CaT programmes are associated with a jargon one has to accept. The sources of emissions are called usually covered sources or installations (as opposed to sources not included in the programme and not being a subject to a cap). Sometimes the term "budget sources" is used. Budget, in this context, refers to an emission cap.
- 6 In some countries, the courts may be involved in the process.
- 7 Any government-sponsored programme develops its own jargon. The emitters are usually referred to as "sources" or "installations."
- 8 The general tendency in cap-and-trade programmes is to expand the list of the targeted sources, both in terms of types of sources (for example, stationary versus non-stationary sources) and in terms of geographical scope. Sometimes, such changes are very controversial – as the inclusion of aviation in the EU Emission Trading Scheme demonstrates.
- 9 Source 2 would have to incur the cost  $BLL_R A$  reducing emissions through a technological solution, which is greater than the cost of buying emission credits.
- 10 Some utilities were less than prudent and sometimes treated the holdings of emission credits as a profit reserve, selling them before the end of the quarter for a quick capital gain, hoping to buy them back at lower prices. Such strategies misfired on some occasions.
- 11 A passenger on a transatlantic flight may buy a few emission credits, on a voluntary basis, to soothe a conscience troubled by their contribution to CO<sub>2</sub> and other emissions (if a few drinks are insufficient to blunt moral sensibilities).
- 12 Traders supported by competent analysts can predict such developments ahead of time and engage in highly profitable inter-temporal arbitrage.
- 13 The use of taxes to control externalities was first proposed by the British economist, Arthur Cecil Pigou. This is why such taxes are often referred to as Pigouvian taxes. They were popularised in the second half of the 20th century by William Baumol and Wallace Oates (see William J. Baumol and Wallace E. Oates, 1971, "The use of standards and prices for protection of the environment," *Swedish Journal of Economics*, (73)1, March, pp 42–54).
- 14 The first such regulation I am aware of was an act of the English Parliament banning the dumping of waste in ditches and public waterways in 1388 (see <http://www.environmentalistseveryday.org/publications-solid-waste-industry-research/information/history-of-solid-waste-management/index.php>).
- 15 "Sources of fine particles [less than 2.5 micrometres in diameter] include all types of combustion activities (motor vehicles, power plants, wood burning, etc) and certain industrial processes.[...]Particles may be formed in the air from the chemical change of gases. They are indirectly formed when gases from burning fuels react with sunlight and water vapour.

These can result from fuel combustion in motor vehicles, at power plants, and in other industrial processes.” <http://www.epa.gov/pmdesignations/faq.htm>.

- 16 To be precise, the estimate for 2008 is 18.8% (see “National summary of nitrogen oxides emissions” at [www.epa.gov](http://www.epa.gov)).
- 17 “Scrubber systems are a diverse group of air pollution control devices that can be used to remove particles and/or gases from industrial exhaust streams. Traditionally, scrubbers have referred to pollution control devices that used liquid to ‘scrub’ unwanted pollutants from a gas stream. Recently, the term scrubber is also used to describe systems that inject a dry reagent or slurry into a dirty exhaust stream to ‘scrub out’ acid gases. Scrubbers are one of the primary devices that control gaseous emissions, especially acid gases.” See [http://yosemite.epa.gov/oagps/eogtrain.nsf/ae20ef1becae534385256b4100770781/daf3098012db893e85256b6b0067b7e5/\\$FILE/si412c\\_lesson1.pdf](http://yosemite.epa.gov/oagps/eogtrain.nsf/ae20ef1becae534385256b4100770781/daf3098012db893e85256b6b0067b7e5/$FILE/si412c_lesson1.pdf).
- 18 SCR is the technology for removal of NO<sub>x</sub> from the flue gas emitted by industrial boilers.
- 19 I shall cover only the most important milestones related to implementation of the CAA that have a major impact on the current environmental policies. It is beyond the scope of this book to cover all the meanders of regulatory actions and court fights. Such details are available at <http://www.epa.gov/air/caa/>.
- 20 The “non-attainment” areas have air pollution levels persistently exceeding the national standards, as defined in the Clean Air Act Amendments of 1970. Non-attainment areas are required to develop and implement a mitigation plan. The leverage the federal government uses is denial of certain federal funds.
- 21 Alaska and Hawaii excluded (a footnote for non-North American readers).
- 22 Substitution means that a plant operator may reassign required emission reductions from one unit to other units under their control (that are not included in the emission reduction programme).
- 23 A power plant operator may apply for less stringent emission targets, called AELs.
- 24 Another regulatory initiative was creation of the Ozone Transport Assessment Group (OTAG), including over 30 eastern states.
- 25 <http://www.epa.gov/airmarkets/progsregs/nox/sip.html>.
- 26 “Testimony Of Brian Mclean Director, Office f Atmospheric Programs Office of Air and Radiation U.S. Environmental Protection Agency before the Committee On Environment and Public Works Subcommittee on Clean Air and Nuclear Safety,” United States Senate, July 29, 2008. See also [http://epa.gov/cair/charts\\_files/cair\\_final\\_presentation.pdf](http://epa.gov/cair/charts_files/cair_final_presentation.pdf).
- 27 CAIR, EPA, Office of Air and Radiation, March 2005.
- 28 United States Court of Appeal for the District of Columbia Circuit, No. 05-1244, argued March 25, 2008; decided July 11, 2008.
- 29 The same court that struck the CAIR reinstated the law on a temporary basis in December 2008.
- 30 “Fact Sheet. The Cross-State Air Pollution Rule: Reducing the Interstate Transport of Fine Particulate Matter and Ozone” (<http://epa.gov/airtransport/pdfs/CSAPRFactsheet.pdf>).
- 31 For example, technologies such as FGD or DSI used to comply with the MATS acid gas standards help to reduce SO<sub>2</sub> emissions, contributing to compliance with CSAPR. Reductions in mercury emissions required under MATS can be accomplished using a combination of FGD and SCR, which help to reduce SO<sub>2</sub> emissions as well, or by using ACI systems. FGD helps to reduce mercury output from burning bituminous coal, but this technology has to be replaced with ACI in the case of sub-bituminous coal and lignites.
- 32 <http://www.epa.gov/mats/pdfs/20111216MATSfinal.pdf>.
- 33 Section 112 contains the list of 187 Hazardous Air Pollutants (HAP) with a clear directive for the EPA to act to reduce their emissions.
- 34 Power plants account for about 50% of US mercury emissions, 62% of arsenic emissions and 82% of hydrochloric acid emissions. Other sources of mercury pollution are municipal waste incinerators and medical waste incinerators.
- 35 “In 1979, the FDA actually doubled what it considered the hazardous level of mercury in

- fish – from 0.5 parts per million to 1.0 parts per million. The EU still sets its limit at 0.5 parts per million.” See [http://topics.nytimes.com/top/news/health/diseasesconditionsand-healthtopics/foodcontaminationandpoisoning/mercury\\_in\\_tuna/index.html](http://topics.nytimes.com/top/news/health/diseasesconditionsand-healthtopics/foodcontaminationandpoisoning/mercury_in_tuna/index.html).
- 36 In the interest of full disclosure, we have stopped eating tuna a long time ago so we cannot use this excuse for any shortcomings of this book.
- 37 See “Combined National and State-level Health Benefits for the Cross-State Air Pollution Rule and Mercury and Air Toxics Standards,” Table 2 for more information. Any estimate is based on a wide range of assumptions that should be critically examined (<http://www.epa.gov/ttn/ecas/regdata/Benefits/casprmat.pdf>).
- 38 In 2005, EPA promulgated rules establishing a CaT programme, the Clean Air Mercury Rule (CAMR), a system for controlling mercury pollution. This rule was incompatible with the requirements of the CAA (which specifically required use of the MACT approach), and its implementation would not prevent formation of local areas of high concentration of mercury (hot spots). The CaT rule for mercury was vacated in 2008 by the D.C. Circuit Court of Appeals.
- 39 A sample of the reports includes North American Electric Reliability Corporation, 2011, “2011 long-term reliability assessment,” November; M. J. Bradley & Associates, 2011, “Ensuring a clean, modern electric generating fleet while maintaining electric system reliability; Fall 2011 update,” November; DOE, 2011, “Resource adequacy implications of forthcoming EPA air quality regulations,” December; Bipartisan Policy Center, 2011, “Environmental regulation and electric system reliability,” June; CERES, 2011, “New jobs – cleaner air part II: An investment in American businesses and American jobs,” November; Edison Electric Institute, 2011, “Potential impacts of environmental regulation on the US generation fleet,” January.
- 40 CFCs are organic compounds built of carbon, chlorine and fluorine. HCFCs are CFCs which additionally contain hydrogen. CFCs have many industrial uses and are known commonly as freons. However, they have another harmful effect: they contribute to the destruction of the ozone layer.
- 41 J. T. Houghton, G. J. Jenkins and J. J. Ephraums, (eds), 1990, *Climate Change. The IPCC Scientific Assessment* (Cambridge, England: Cambridge University Press) [http://www.ipcc.ch/publications\\_and\\_data/publications\\_ipcc\\_first\\_assessment\\_1990\\_wg1.shtml](http://www.ipcc.ch/publications_and_data/publications_ipcc_first_assessment_1990_wg1.shtml).
- 42 There were several efforts to quantify the level of consensus on global warming in the scientific community. Studies supporting overwhelming support for the anthropogenic hypothesis included Naomi Oreskes, 2004, “Beyond the ivory tower: The scientific consensus on climate change,” *Science*, (306)5702, December 3, p 1,686. According to Peter T. Doran and Maggie Kendall Zimmerman “the debate on the authenticity of global warming and the role played by human activity is largely nonexistent among those who understand the nuances and scientific basis of long-term climate processes.” See “Examining the scientific consensus on climate change,” *EOS*, (90)3 January 20, 2009. See R. A. Pielke, 2005, “Consensus about climate change?” *Science*, 308, pp 952–53, for a differing opinion.
- 43 They often point out to the discrepancy between the temperature readings at weather stations located in huge human agglomerations (where the urban island phenomenon, as discussed in the section on weather derivatives, is very strong) and stations located in more remote locations.
- 44 Based on <http://hyperphysics.phy-astr.gsu.edu/hbase/thermo/grnhse.html#c5>. The website contains many useful links to the actual sources and to the counter-arguments.
- 45 There are many reasons to believe that collective global warming action is an example of the “prisoner’s paradox,” an example of a suboptimal outcome if everybody acts in their best individual interest (see <http://plato.stanford.edu/entries/prisoner-dilemma/> on this topic).
- 46 <http://www.ipcc.ch/organization/organization.shtml>
- 47 <http://www.ipcc.ch/organization/organization.shtml>

- 48 [http://unfccc.int/essential\\_background/items/6031.php](http://unfccc.int/essential_background/items/6031.php).
- 49 MOP stands for Meeting of the Parties of the Kyoto Agreement.
- 50 CFCs (Chlorofluorocarbons) are covered by the Montreal Protocol (1987) on “Substances that Deplete the Ozone Layer,” signed under the Vienna Convention for the Protection of the Ozone Layer (1985).
- 51 Directive 2003/87/EC of the European Parliament and of the Council of October 13, 2003 (<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2003:275:0032:0032:EN:PDF>).
- 52 NER was established by the Article 3 of the Directive, which defined a new entrant as “any installation carrying out one or more of the activities indicated in Annex I, which has obtained a greenhouse gas emissions permit or an update of its greenhouse gas emissions permit because of a change in the nature or functioning or an extension of the installation, subsequent to the notification to the Commission of the national allocation plan.” See [http://ec.europa.eu/clima/policies/ets/docs/ecofys\\_new\\_entran\\_en.pdf](http://ec.europa.eu/clima/policies/ets/docs/ecofys_new_entran_en.pdf) for more information.
- 53 Expansion referred to increase in capacity through investments and not a higher rate of utilisation of existing capacity.
- 54 Sometimes the word “scheme” is used, but this has negative connotations in American English.
- 55 At the time of writing, only nitrous oxide emissions are included in addition to CO<sub>2</sub>. This will change from 2013.
- 56 “Directive 2009/29/EC of the European Parliament and of the Council of 23 April 2009 amending Directive 2003/87/EC so as to improve and extend the greenhouse gas emission allowance trading scheme of the Community” (<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:140:0063:0087:en:PDF>).
- 57 A sector is exposed to a risk of carbon leakage if “the extent to which the sum of direct and indirect additional costs induced by the implementation of this directive would lead to a substantial increase of production cost, calculated as a proportion of the Gross Value Added, of at least 5%; and the Non-EU Trade intensity defined as the ratio between total of value of exports to non EU + value of imports from non-EU and the total market size for the Community (annual turnover plus total imports) is above 10%.” See [http://ec.europa.eu/clima/policies/ets/leakage/index\\_en.htm](http://ec.europa.eu/clima/policies/ets/leakage/index_en.htm).
- 58 Commission Regulation No. 1031/2010 (amended on November 23, 2011) addresses the timing and administration of the auctions of the EU greenhouse gas emissions allowances.
- 59 The European Commission suspended trading of allowances in January of 2011 for a few days. “The Commission took the decision after the revelation that emissions allowances worth €7 million (US\$9.4 million) were stolen from an account in the Czech Republic. In recent days, criminals have also hacked into trading accounts in Austria, Poland, Greece and Estonia, the commission said.” Joshua Chaffin, 2011, “Cyber-theft halts EU emissions trading,” *Financial Times*, January 19.
- 60 The ETS allowed for banking and borrowing of allowances (see Julien Chevallier, 2012, “Banking and borrowing in the EU ETS: A review of economic modelling, current provisions and prospects for future design,” *Journal of Economic Surveys*, February, (26)1, pp 157–76, for more details).
- 61 The ICE acquired the Chicago Climate Exchange, Chicago Climate Futures Exchange and European Climate Exchange on April 30, 2010.
- 62 The full list of contracts is available at <http://www.thegreenx.com/products/index.html>.
- 63 See United States Government Accountability Office, 2008, “International climate change programs: Lessons learned from the European Union’s Emissions Trading Scheme and the Kyoto Protocol’s Clean Development Mechanism,” November (<http://www.gao.gov/assets/290/283397.pdf>).
- 64 PointCarbon, 2007, “Carbon 2007. A new climate for carbon trading,” March 13 (<http://www.pointcarbon.com/research/promo/research/>).
- 65 A. Denny Ellerman and Paul L. Joskow, 2008, “The European Union’s emissions trading

- system in perspective," MIT, May. This paper contains an in-depth analysis of the price dynamics of CO<sub>2</sub> allowances in Phase I.
- 66 At the beginning of August 2005, only 11 out of 25 countries had operational registries (see <http://www.icis.com/heren/articles/2005/08/03/9271861/official-spanish-emissions-registry-launch-delayed-until-after-summer.html>).
- 67 BP, "Statistical Review of World Energy 2012," June.
- 68 Ibid.
- 69 Pilita Clark and Jack Farchy, 2012, "Carbon prices tumble to record low," *Financial Times*, July 18.
- 70 Pilita Clark, 2012, "EU moves to tighten carbon trading," *Financial Times*, July 25.
- 71 President George W. Bush announced in March 2001 that the US would not ratify the Kyoto Protocol, citing potential harm to the US economy as the reason.
- 72 This happened in 2006 with, respectively, 6412 and 6415.5 million tons of CO<sub>2</sub> emitted in the US and China.
- 73 EIA, "Annual Energy Outlook 2012," p 2.
- 74 The first control period: January 1, 2009 to December 31, 2011; the second: January 1, 2012 to December 31, 2014.
- 75 CO<sub>2</sub> Allowance Tracking System.
- 76 RGGI Fact Sheet.



## *Coal Markets*

The coal markets can be considered a good laboratory for anybody interested in studying the emergence and development of a new market. Coal has historically been supplied under long-term bilateral contracts (in the US, often as long as 20–30 years), with a significant component of the delivered price determined by the cost of transportation. In the US (and elsewhere), the primary reason for reliance on long-term contracts was the high capital cost of new mining projects and related transportation infrastructure, which required many years of reliable revenues to justify the risks involved. On the buy-side, regulated utilities were in a position to enter into such contracts as they could transfer the cost to ratepayers with a high degree of certainty. The operators of coal-fired power plants with high capital costs and a long expected life had similar objectives to the coal miners: they required long-term arrangements with regular deliveries of coal to their sites to avoid costly shutdowns. Since around 2000 we have seen the emergence of a more active spot market, combined with the development of financial derivatives. This has happened at the time when the coal industry has come under pressure from many directions, including:

- ❑ environmental concerns related to the high level of CO<sub>2</sub> and other emissions (see Chapter 25), and the permanent impact on the earth's landscape of coal mining operations (especially in the case of open-pit mining and mountain top removal); and
- ❑ competition from natural gas as a fuel of choice, given the falling prices of natural gas in the US, which favour gas-fired generation units over coal.

This chapter will examine the physical properties of coal, review current production, exports and imports trends, as well as available coal price benchmarks and emerging financial markets for this fuel.

Coal versus natural gas substitution has become an important trend in the US, and an important driver of prices of both fuels.

### **PHYSICAL PROPERTIES OF COAL**

Coal is a rock containing over 50% of carbon. Coal was created, like other fossil fuels, through the accumulation, compression (and, additionally, induration (hardening)) of organic matter (trees and other plants). As explained by the World Coal Association:<sup>1</sup>

Coal formation began during the Carboniferous Period – known as the first coal age – which spanned 360 million to 290 million years ago. The build-up of silt and other sediments, together with movements in the earth's crust – known as tectonic movements – buried swamps and peat bogs, often to great depths. With burial, the plant material was subjected to high temperatures and pressures. This caused physical and chemical changes in the vegetation, transforming it into peat and then into coal.

Anthracite is the coal of the highest degree of maturation and carbon content (over 85%), and is used for space heating, power generation and as a metallurgical fuel. Bituminous coal is less mature than anthracite, and contains 45–85% carbon. Bituminous coal of a high quality (to be defined below) is used for the production of coke,<sup>2</sup> a fuel used in iron smelters. Sub-bituminous coal has carbon content of 35–45% and a higher moisture content than bituminous coal. Lignite (known as brown coal in Europe) has low-carbon content (25–35%) and high-moisture content. One has to bear in mind that there are many competing classifications of coals and the numerical ranges provided here should be treated as indicative information.

Coal is a highly heterogeneous commodity with properties varying widely from one source to another. Combinations of different properties of coal, unit transportation costs and distance travelled, environmental regulations and different end-uses explain why many countries can be simultaneously big producers and big users – exporters and importers – of this fuel. The characteristics of coal produced by a given country or region may be different than the structure of final demand, creating a need for imports and a surplus of output that may be exported. The main physical and chemical properties of coal include:<sup>3</sup>

- coal rank;
- sulphur content; and
- mercury content.

Coal rank is a classification based on a combination of the two factors: calorific value and moisture content. Lignite coals through to high-volatile bituminous coals are classified based on a moist, mineral-matter-free basis, whereas medium-volatile bituminous through to anthracite coals are classified based on a dry mineral-matter-free basis. The difference between the two types of coal is moisture content, which is important because transportation cost is an important component of the delivered price. Transportation of high-rank coal (lower moisture level) is less expensive.

Another classification used extensively in the industry is based on a distinction between thermal and metallurgical coal. Thermal coals (also called steam coals) range from lignite coal through to bituminous coals. They are used primarily for power generation, in industrial boilers installed in many industrial plants (chemical plants, cement plants, etc) and as transportation fuel (on a decreasing scale). Most Americans now only see coal locomotives in Westerns.

Sulphur content is an important property of coal. High-sulphur content leads to the corrosion of boilers and pipes in industrial and power generation units. Sulphur content is also an important factor in the selection of coals for the production of coke: low-sulphur content is required for this. Burning coal produces sulphur compounds (such as SO<sub>2</sub>) with adverse environmental impacts (the very vigorous measures taken in the US to reduce sulphur emissions were discussed in the previous chapter). According to the EIA 1993 classification,<sup>4</sup> low-sulphur coal has less than 0.6 pounds of sulphur per million Btus, whereas medium-sulphur has between 0.61 and 1.67, while high-sulphur has over 1.68. In general, US coal produced in the western part of the US has low-sulphur content, coal from the central US (Illinois Basin and Western Interior Basin) has high-sulphur content, and Appalachia is somewhere in between.

The mercury content of US coals ranges from “a low of 2.04 pounds of mercury per trillion Btu for low-sulphur subbituminous coal originating from mines in the Rocky Mountain supply region, to a high of 63.90 pounds of mercury per trillion Btu for waste coal.”<sup>5</sup> Mercury content is important, as this trace element has been recently targeted by US environmental regulations (as was discussed in detail in Chapter 25).

---

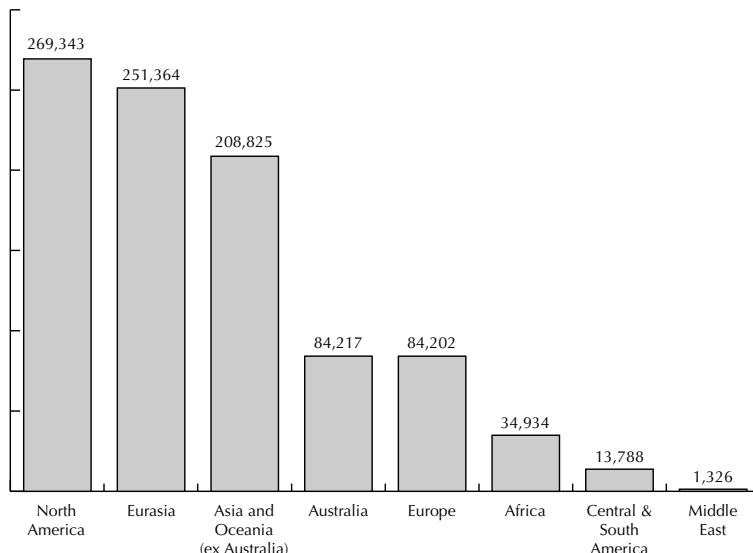
*Ash content* refers to the non-combustible residue left after coal is burned. Ash content varies for anthracite coals from 9.7–20.2% of weight, 3.3–11% for bituminous coals and around 4.2% for lignite. Ash content depends of the volume of foreign material trapped in coal during the process of its formation. Ash produced during burning of coal is classified either as fly ash or bottom ash (ie, ash that is not emitted into the air with flue gases but remains at the bottom of the furnace). Environmental regulations require that fly ash be trapped using electrostatic devices and removed, together with bottom ash, to landfills.<sup>6</sup> Some ash is used in production of cement.

## RESERVES AND PRODUCTION

One of the main advantages of coal is that it is the most abundant fossil fuel, and widely distributed across the globe. Current proven coal reserves are about 948 billion tons worldwide (based on EIA statistics for 2008),<sup>7</sup> translating into a coverage rate of over 100 years at current consumption levels. Figure 26.1 shows the distribution of coal reserves around the world.

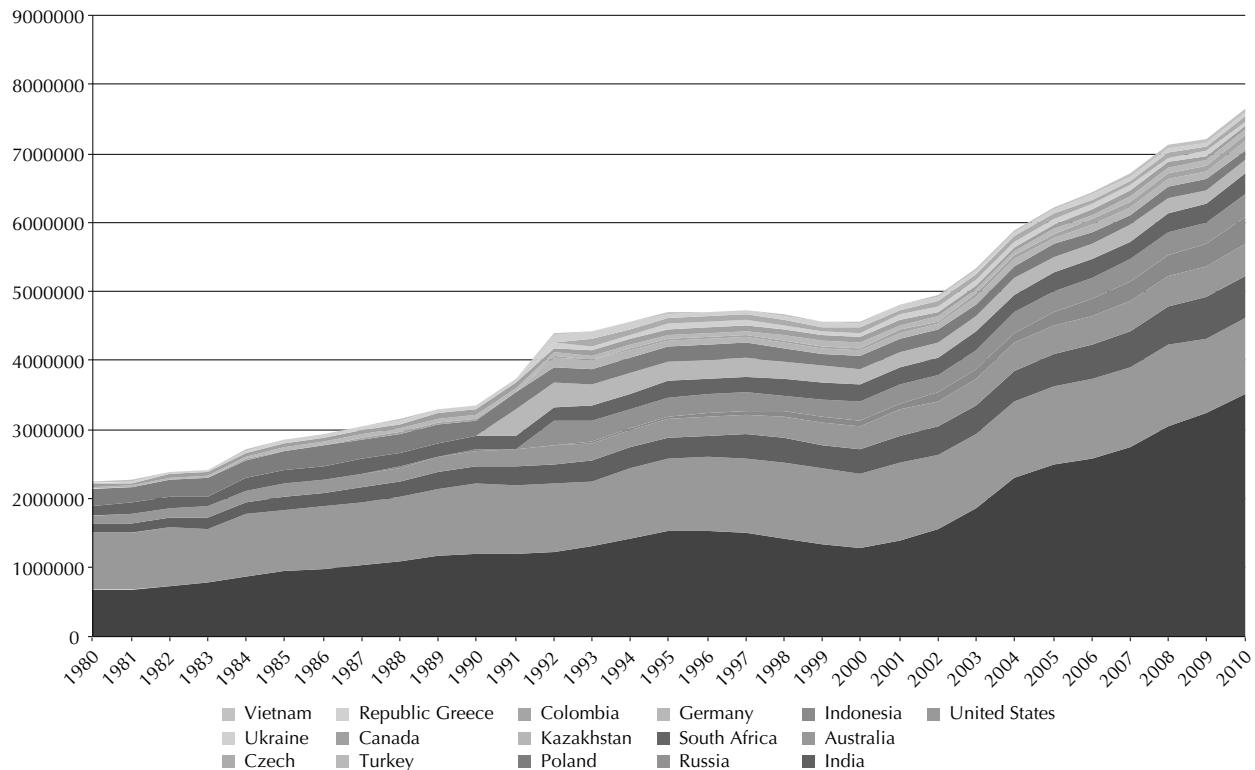
Figures 26.2 to 26.5 illustrate the level of production, exports,

**Figure 26.1** Total recoverable coal (million short tons)



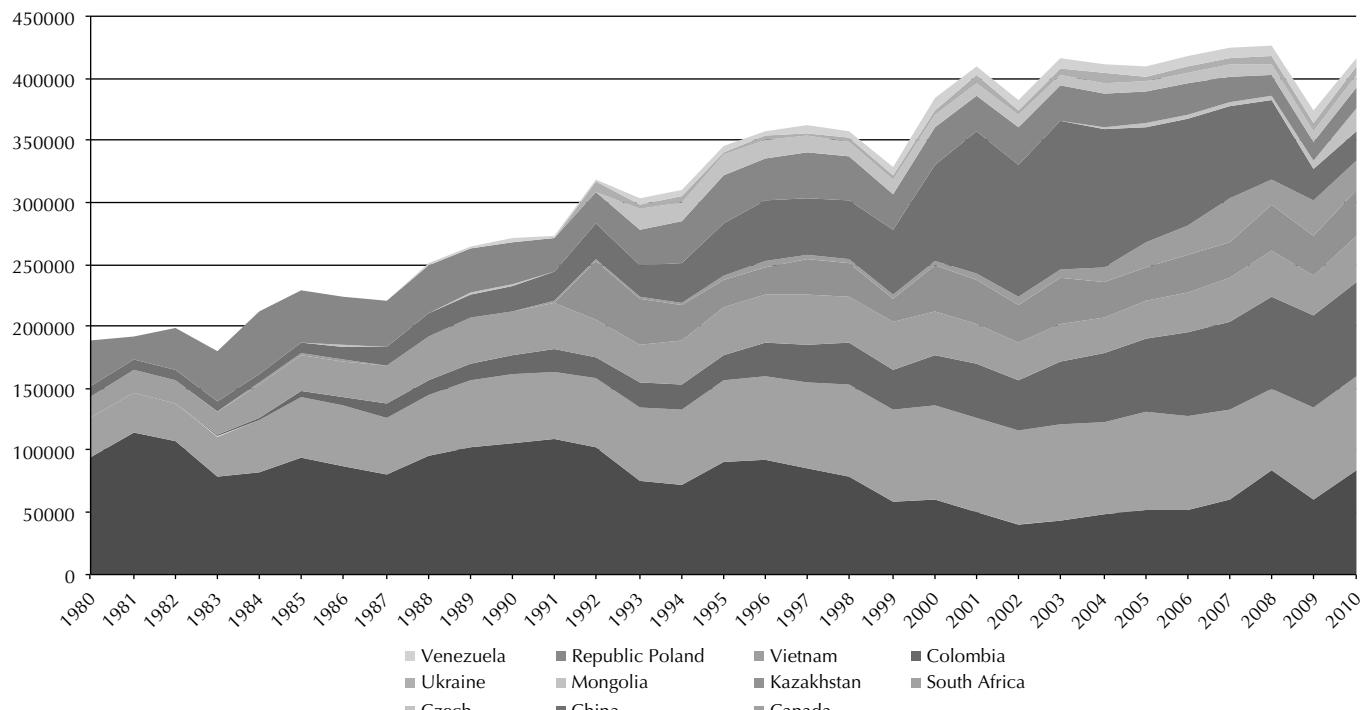
Source: U.S. Energy Information Administration

**Figure 26.2** Production of coal by selected countries (thousand short tons)



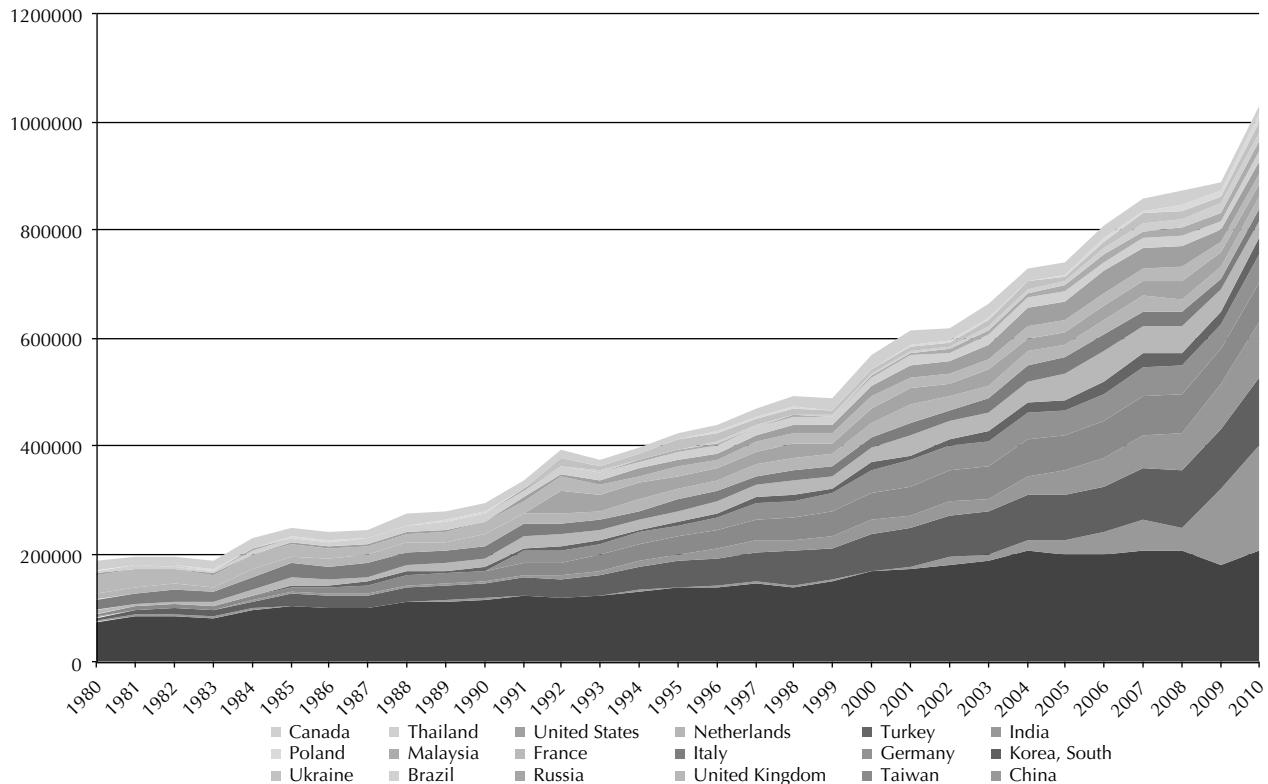
Source: U.S. Energy Information Administration

**Figure 26.3** Coal exports by selected countries (thousand short tons)



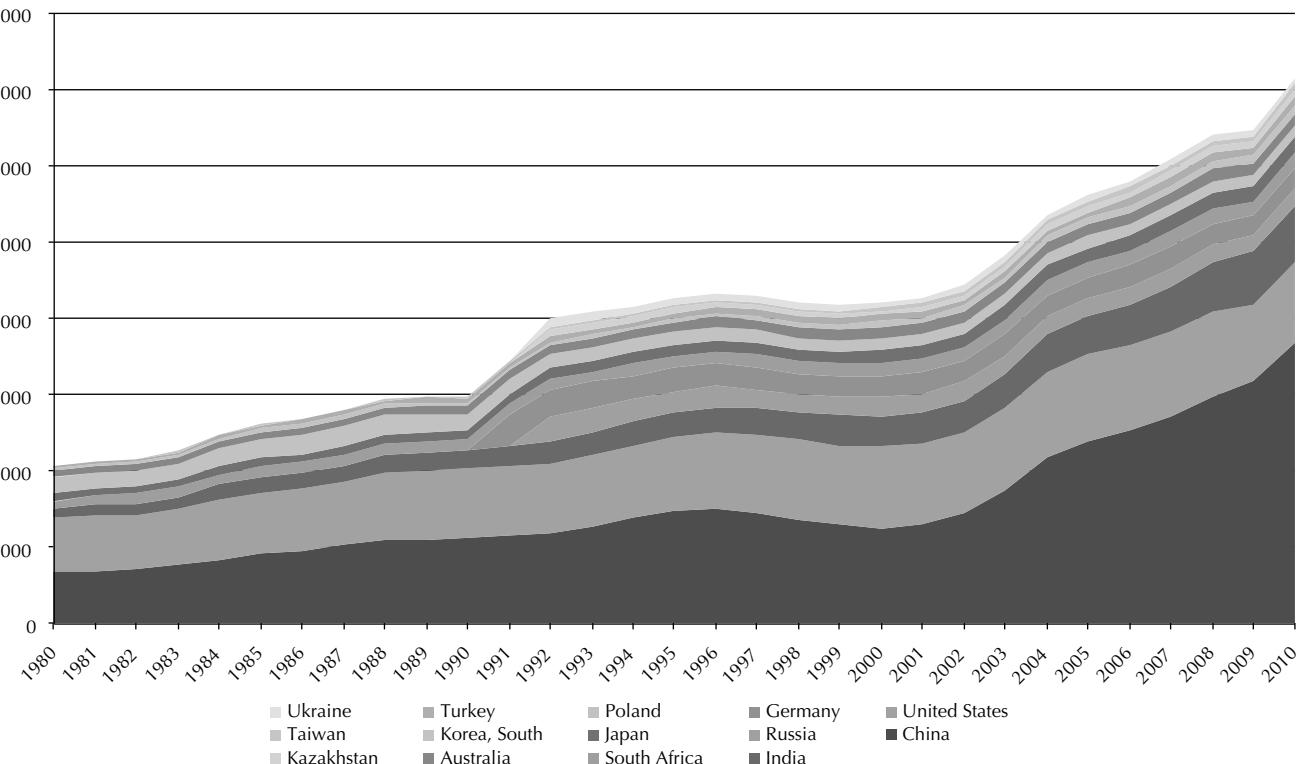
Source: U.S. Energy Information Administration

**Figure 26.4** Coal imports by selected countries (thousand short tons)



Source: U.S. Energy Information Administration

Figure 26.5 Coal consumption by selected countries (thousand short tons)



Source: U.S. Energy Information Administration

imports and consumption of coal for the largest producers and consumers.

The countries included in Figure 26.2 accounted for over 95% of the world production, with China, the US and India alone accounting for about 65%.

As in the case of other commodities, recent trade flows and price dynamics have been to a large extent a manifestation of the economic and social transformation of China. A jolt to the market happened in 2009, when China was transformed almost overnight from a net exporter of coal to an importer (139 million tons, 114 million tons net).<sup>8</sup> Chinese imports represent about 15% of all globally traded coal. Richard Morse and Gang He argued in the paper quoted above that this shift does not represent shortages of coal in China, but rather an arbitrage between domestic and international prices, especially as the latter weakened in the aftermath of the 2008 economic crisis. This is another manifestation of the growing sophistication of international energy markets, and the need to monitor energy-related developments across the globe.

The US is endowed with very significant coal resources concentrated in several major basins: CAPP (Central Appalachian), NAPP (Northern Appalachian), ILB (Illinois Basin), PRB (Powder River Basin), UB (Uinta Basin) and GCL (Gulf Coast Lignite). Summary information about these regions is provided in Table 26.1.

Coal from the PRB in Wyoming and Montana region has low calorific value, but this is offset by a low cost of production (coal is located close to the surface and can be produced through surface mining), and low-sulphur content also more than compensates for it. PRB is the fastest-growing region of production in the US.

**Table 26.1** US coal reserves

Basin	Reserves (BB tons)	Heat content (Btu/lb)	SO <sub>2</sub> content (lb/MMBtu)	2008 transport cost (US\$/ton)	2008 mining cost (US\$/ton)
Powder River	115	8,800	0.8	25	8–12
CAPP	20	12,500	1.2	15	45–65
NAPP	15	13,000	<3	15	35–50
Illinois	51	11,800	5.0	15	30–40
Uinta	12	11,700	0.8	20	20–30

Source: Tudor, Pickering, Holt & Co

## COAL TRADING

### Price indexes

This section will discuss the most important price indexes used in the coal markets. More information can be found in the methodology section of Platts<sup>9</sup> and Argus.<sup>10</sup> The main price benchmarks used for coal are available from Argus/McCloskey<sup>11</sup> and Platts. The most important McCloskey indexes are shown below.

API 2 index is the price for coal imported into Western Europe. This index represents an average of the Argus CIF Rotterdam assessment and McCloskey's northwest European steam coal marker. As explained by McCloskey:

The NW Europe Steam Coal marker reflects market value for any origin of standard bituminous material that is delivered into North West Europe. Bituminous material from all origins is included in the NW Europe marker as long as the material's specification reaches the general European standard, established by McCloskey in 1991, of under 1% sulphur, with prices c.v.<sup>12</sup> adjusted to a 6,000kc NAR<sup>13</sup> basis. The NW Europe marker is based on delivery into the ARA hub of the most economic vessel in the current freight market from the sourcing country.

The Argus component is published daily in *Argus Coal Daily International*. McCloskey Fax and the *McCloskey Coal Report* publish on Fridays the averages of the week's daily markers, which are available during the week from the McCloskey newswire service.

API 4 index<sup>14</sup> is the price benchmark for all coal exported out of Richards Bay, South Africa. The API 4 index is calculated as an average of the Argus FOB Richards Bay assessment and McCloskey's FOB Richards Bay marker.

The API 2 and API 4 indexes are used as benchmarks for about 90% of coal-related derivatives. The difference between the two indexes is often called implied freight. This calculation should be used with caution because the dynamics of FOB and CIF prices is often very different, and the difference can be negative on some occasions (a combination of weak markets in Europe with strong markets in Asia, an alternative destination for South African coal).

API 5 index is the price benchmark used for exports of 5,500 kcal/kg net as received (NAR), high-ash coal from Australia. The index is

calculated as an average of the Argus FOB<sup>15</sup> Newcastle 5,500 kcal/kg assessment and the equivalent from IHS McCloskey.

API 6 index represents 6,000 kcal/kg NAR coal exported from Australia. The API 6 index is calculated as an average of the Argus FOB Newcastle 6,000 kcal/kg assessment and the equivalent from IHS McCloskey.

API 8 index is the benchmark for 5,500 kcal/kg NAR coal delivered to south China. It is calculated as an average of the Argus 5,500 kcal/kg CFR south China price assessment, and the IHS McCloskey/Xinhua Infolink South China marker.

The OTC Broker Index has been published by Platts since 2003. It is based on the market-on-close approach and covers the following contracts:

- CAPP, barge, 12,000 BTU/lb;
- CAPP, rail (CSX),<sup>16</sup> 12,500 Btu/lb;
- Powder River Basin, rail, 8,800 Btu/lb; and
- Powder River Basin, rail, 8,400 Btu/lb.

Global coal platform was created in 2001 as a joint venture of major coal producers (BHP Billiton, Anglo American, Glencore and Rio Tinto) and end users (EON, Enel and J-Power), plus a ship broking group, SSY. Global Coal publishes the Newcastle price index, which is based on transactions executed on this platform as well as bid and offers.

The ICI (Indonesian Coal Index) is a spot price of four grades of Indonesian coal – 6,500, 5,800, 5,000 and 4,200 kcal/kg “gross as received” (GAR) calculated by Argus.

The definitions of price indexes include a number of technical terms that have not been introduced earlier, and which are explained in Panel 26.1.

### **Coal trading worldwide**

Coal in the US and worldwide was historically traded under long-term contracts, which benefited both producers and end-users from

### PANEL 26.1 CALORIFIC VALUE OF COAL

The industry distinguishes between gross and net calorific values of coal. The gross value is the amount of heat liberated under laboratory conditions. The net calorific value refers to the amount of heat during the actual combustion process under normal industrial conditions. Some energy is lost – for example, in water vapour.

$$\text{Net CV} = \text{Gross CV} - 91.2H - 10.5M \text{ (in Btu/lb)}$$

where H stands for hydrogen content (in percentages) and M is moisture content (in percentages). A rule of thumb is that, for bituminous coals, the difference is equal to about 470 Btu/lb. Calorific value is measured in the US in Btu/lb, while gigajoule/tonne is used in the UK; South Africa and Australia use mega joule/kg. The rest of the world uses kilocalories per kilogram.

Based on Larry Thomas, 2002, *Coal Geology* (Chichester, England: Wiley).

the point of view of price stability, reliability of deliveries and quality of coal, and guarantees regarding demand levels. The capital intensity of coal production and coal-fired power plants favoured long-term contracts that allowed amortisation of initial mine and plant investments. A significant portion of the coal price represented transportation costs over which both the producers and the buyers had limited control, an additional reason for the perpetuation of long-term contractual arrangements.<sup>17</sup>

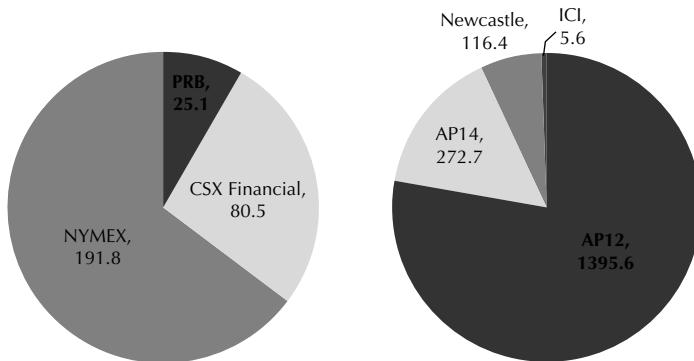
Internationally, starting in the late 1980s, international coal trade was based on Japan's benchmark prices, negotiated between Japanese utilities, steel companies and BHP<sup>18</sup> for steel and thermal coal imported into Japan. Other grades of coal were priced at a discount or a premium to benchmark. In practice, coal was traded at a discount to the benchmark, which was eventually abandoned and replaced with a "fair treatment system,"<sup>19</sup> with prices being bilaterally negotiated.

The years after the late 1990s witnessed a rapid growth of physical and financial swaps settling on the indexes listed above. The information about market size is spotty and based on estimates. A summary of information about coal trading in the US and in the international markets is shown in Figure 26.6.

## CME and ICE coal contracts

The late 1990s and early years of the 21st century were a period of rapid development for OTC and exchange-based coal trading. Nymex led the charge, developing a proposed design of futures contracts. The industry started trading bilateral contracts based on the Nymex contract specification even before the futures had been officially listed. The most important contracts offered by the CME include Central Appalachian coal futures (Globex, ClearPort) and Powder River Basin Coal (Platts OTC Broker Index) swap futures and Coal (API 2) CIF ARA (Argus/McCloskey) swap futures offered on ClearPort. The Central Appalachian contract quality specification includes: heat content (minimum 12,000 Btu/lb gross calorific value, with a tolerance of 250 Btu/lb below), ash content (maximum 13.50%), sulphur content (maximum 1.00%, with a tolerance of 0.050% above), moisture (maximum 10.00%), volatile matter (minimum 30.00%), grindability<sup>20</sup> (minimum 41 Hardgrove Index, HGI, with three-point analysis tolerance below) and sizing (three inches top size). Deliveries are made to the buyer's barge at the seller's delivery facility on the Ohio River between Mileposts 306 and 317 or on the Big Sandy River. The contract size is 1,550 tons (a typical barge). The ClearPort contracts mentioned above are cash-settled. The list of the ClearPort coal contracts can be found at the CME website.

**Figure 26.6** Coal trading in 2011, US versus international (millions metric tons)



Source: Matthew Moore, 2011, "Tools of Trade: 'User's Perspective,'" CTA Coal Trading Conference, December 6.

ICE offers an equally diversified portfolio of coal-related contracts covering the Europe, South Africa and North America and Australia axes of the global coal market:

- ❑ Rotterdam coal futures settled against API#2 coal index;
- ❑ Richards Bay coal futures settled against the API#4 coal index;
- ❑ Newcastle coal futures settled against the globalCOAL NEWC Index;
- ❑ FOB Indo sub-bituminous coal futures;
- ❑ CSX coal futures;
- ❑ Central Appalachian coal futures; and
- ❑ Powder River Basin coal futures.

### **OTC contracts**

The Coal Trading Association (CTA), established in 1999, is a coalition of market participants which has made a significant contribution to the development and growth of the US coal market. The CTA developed the Master Coal Purchase and Sales Agreement (MCPSA), which was launched in 2000, and helped to streamline the process of negotiating coal-related deals and reduce the contractual risk related to potential non-performance by counterparties. Members could use the agreement as a template for negotiating company-specific master agreements. A revised version of the MCPSA was ratified at the 2006 CTA meeting.<sup>21</sup> The revised 2006 version contained the quantity variance adjustment (QVA) clause, which has been embraced by the industry and is now a standard feature of most contracts.

The QVA is a provision formulated to address the volumetric risk present in coal contracts. This risk is related to the operational difficulties of forming coal trains with volumes corresponding exactly to contractual quantities (or loading barges exactly to the contract specification). This is not a minor problem. For example, the US drought of 2012 lowered the levels of water in many US rivers and forced barge operators to load them only to 50% or less of capacity (and to suspend barge traffic completely on some rivers). Of course, under- or over-loading may be intentional and related to the current price levels. As in the case of oil tankers, if spot prices exceed contract prices, producers will have all the incentive to under-load. The opposite would happen should the spot prices fall below the contract price. In

the past, when coal was traded under long-term contracts, and with relationship between buyers and sellers cemented over many years of mutual dependence, this did not represent a serious problem. Discrepancies with respect to weight and quality would be smoothed over time and resolved in an amicable way. However, with short-term contracts and anonymous markets, a different solution was required. The QVA defines a deadband (typically +/-2%) around the contractual quantity. The volumes outside the band would be priced at current market prices, and not at the contract price.

The most popular US OTC contracts include:<sup>22</sup>

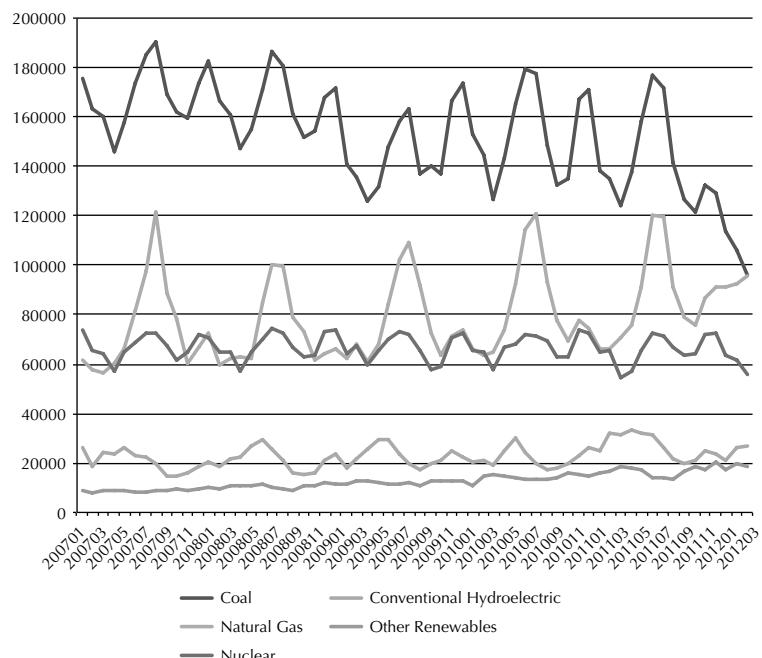
- Nymex Look-Alike Coal Contract Specifications (the standard size is five times the Nymex contract, ie, 7,750 tons):
  - origin: Central Appalachia
- Southern Powder River Basin, 8,800 Btus/lb and 8,400 Btus/lb:
  - origin: Southern Powder River Basin
  - size: unit trains ~15,000 tons (Over/under rule +/-2% monthly contract volume dead band)<sup>23</sup>
  - delivery point: FOB Railcar, mine, jointly served by both UP and BN railroads
- CSX compliance and 1% (12,500 Btu/lb):
  - origin: CSX Kanawha and/or Big Sandy Freight Districts
  - size: unit trains ~11,000 tons
- NS<sup>24</sup> compliance and 1% (12,500 Btu/lb):
  - origin: NS Kenova and/or Thacker (I) Freight Districts
  - size: unit trains ~11,000 tons.

For the last two contracts, 1% refers to the sulphur content. Compliance is a property of coal or a blend of coals. Compliance in this context means that coal meets "sulphur dioxide emission standards for air quality without the need for flue gas desulphurisation."<sup>25</sup> All the contracts, except for NS Compliance, have an over/under rule of +/-2% monthly contract volume deadband. The contracts contain standardised provisions for price adjustments in case of sulphur content and/or heat content diverging from the levels envisaged in the contract.

### Gas for coal substitution

The US energy industry reached a critical milestone in April 2012: the net production of electricity from coal-fired and natural gas-fired power plants was equal for the first time (see Figure 26.7). There are several factors behind this development, for which one can use the term “historical” without running the risk of putting one’s foot in one’s mouth and chewing on it. The obvious reason is the falling price of natural gas, which makes gas-fired generation competitive in comparison with coal units. A traditional generation merit order with coal units in the US being dispatched ahead of gas units, with the latter setting electricity prices at the margin in most regions of the US, has been changed. Coal-fired power plants at some locations have become marginal units from the point of view of price setting. We show below a simplified example of the calculation of cross-over

**Figure 26.7** Net generation of electricity in the US by source (thousand megawatt hours)



Source: U.S. Energy Information Administration

natural gas prices – ie, prices at which natural gas units become less expensive to run than coal plants.

There are, however, other important reasons why the industry could increase the production of electricity from natural gas. The first was the development of mature energy markets and energy trading that made possible the physical substitution of one fuel for another. In the case of natural gas and coal, this happens through shutting down coal units and switching generation to gas units. Without efficient energy markets, this substitution would be limited to the generation fleet owned by one operator, who can optimise the use of their assets. Even if a local utility or independent power producer has no spare gas-fired capacity, the solution is to shut down coal units and acquire electricity in the market. This option would not have been available on the same scale back in the early 1990s. Many years of experience of participating in increasingly sophisticated energy markets moved the industry to a point on the learning curve that makes trading around assets and asset optimisation a routine activity.

The second reason is the availability of cost-efficient spare gas-fired generation capacity. This is a result of large-scale investments in gas generation units in the late 1990s and in the first years of this century. At the time, these investments were seen as a mistake, a costly bet on a fuel in short supply and likely to become increasingly expensive. In some cases, the companies that led the charge suffered huge losses and sold some units at about 20 cents on a dollar. The industry is now reaping the benefits of these decisions, further proof that the energy markets have a tendency to humiliate even the smartest person and that luck can be quite important in business.

Of course, as with any seismic shift in the markets, gas versus coal substitution requires the modification of once-profitable business strategies. Many independent power producers built their business model around their fleets of coal units. With gas-fired units at the margin, coal plants were earning attractive economic rents. Low prices not only eliminate or reduce such rents, but also wreak havoc on the hedging strategies implemented in the past by merchant and financial players. Historically, natural gas was treated as a convenient proxy for hedging electricity positions, given the strong relationship between the prices of both commodities. We have witnessed many times that a standard procedure for hedging long-

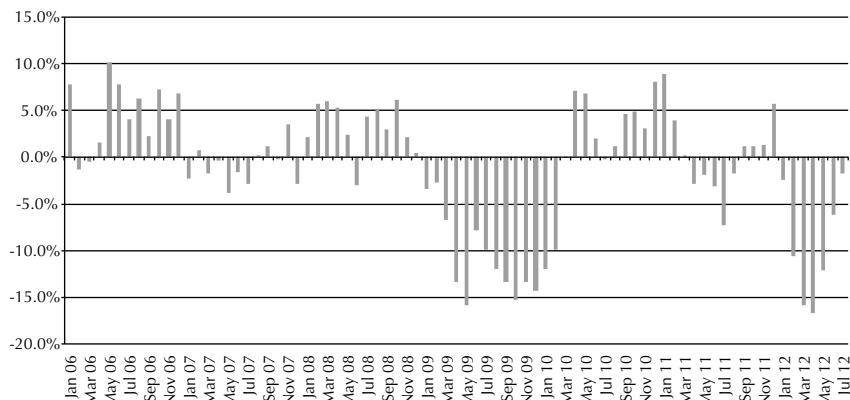
---

term electricity contracts was transacting in Henry Hub natural gas futures (assuming a certain level of market heat rates). Such strategies were often dangerous in the past (electricity prices in the US Northwest may decouple from prices of natural gas, given the importance of hydropower). In the current climate, one should be even more cautious in hedging electricity with gas.

The low prices of natural gas spilled into the coal markets, depressing prices and reducing rail shipments during the first months of 2012 for which the data is available at the time of writing. As a matter of fact, US rail carloads statistics are a useful canary in the mine (no pun intended), data that offer an early warning about unfolding economic trends.

A simplified example<sup>26</sup> will demonstrate how economic decisions regarding dispatch of different units are made. We assume that the heat rate for gas units is equal to 7.5 MMBtu/MWh, and for coal units to 10,500 MMBtu/MWh. We shall explain the calculations for the case of CAPP (Central Appalachian) coal. Its heat content is assumed to be equal to 12,500 Btu/lb, and this translates into 25 MMBtu/ton (with 2,000 lb per ton). Sulphur content is assumed to be equal to 2 lb of SO<sub>2</sub> per MMBtu, or 50 lb per ton. The price of coal is equal to US\$100/tons (US\$80/ton at the mine mouth plus transportation). With SO<sub>2</sub> allowance prices assumed to be equal to US\$2/ton of sulphur dioxide, the cost of SO<sub>2</sub> allowances per MMBtu

**Figure 26.8** % change in US rail carloads of coal from same month previous year: Jan. 2006–July 2012



Source: AAR

**Table 26.2** Natural Gas vs. Coal Crossover Prices

	<b>NAPP</b>	<b>CAPP</b>	<b>Illinois</b>	<b>Rockies</b>	<b>PRB</b>	
Btu/lb	13000	12500	12000	11500	9000	
MMBTU/ton	26	25	24	23	18	
SO2 (lb/MMBTU)	3.1	2	5.1	1	0.8	Pounds per ton
SO2 (lb/ton)	80.6	50	122.4	23	14.4	Coal Heat Rate
Price (\$/ton)	76.75	80	52.43	39.1	13.28	NG HR
Shipping cost (\$/lb)	15	20	9	20	22	
Total Price	91.75	100	61.43	59.1	35.28	
SO2 Allowance (\$/ton)	2	2	2	2	2	
SO2 (\$/ton of coal)	0.0806	0.05	0.1224	0.023	0.0144	
SO2 (\$/MMBTU)	0.0031	0.002	0.0051	0.001	0.0008	
NOx (lb/MMBTU)	0.15	0.15	0.15	0.15	0.15	
NOx allowance (\$/ton)	17.5	17.5	17.5	17.5	17.5	
NOx (\$/ton of coal)	0.034125	0.0328125	0.0315	0.0301875	0.023625	
NOx (\$/MMBTU)	0.0013125	0.0013125	0.0013125	0.0013125	0.0013125	
Cost of Coal (\$/ton)	91.864725	100.0828125	61.5839	59.1531875	35.318025	
Cost of Coal (\$/MMBTU)	3.533258654	4.0033125	2.565995833	2.571877717	1.9621125	
O&M	2	2	2	2	2	
Cost of electricity coal (\$/MWh)	39.09921587	44.03478125	28.94295625	29.00471603	22.60218125	
	1	1	1	1	1	
Breakeven Price of NG (\$/MMBTU)	5.079895449	5.737970833	3.7257275	3.733962138	2.880290833	

Source: Based on SNL calculations

is equal to 0.002, and to US\$0.05 per ton of coal. A similar calculation is carried out for the cost of NOx allowances. This translates into the total cost of coal (coal price plus the cost of allowances per tons) of US\$100.0828/ton and US\$4.0033/MMBtu. The cost of emission allowances is so low in this example that they can be ignored (we include them in the calculations because good times do not often last forever). Assuming a variable O&M cost of US\$2.00/MWh, we arrive at an electricity cost of US\$44.035/MWh (US\$4.0033/MMBtu  $\times$  10.5 MMBtu/MWh + US\$2.00/MWh). Assuming a heat rate of 7.5 MMBtu/MWh and O&M of US\$1/MWh for gas-fired units, we arrive at a parity price of natural gas of US\$5.74/MMBtu. This is quite shocking, given that the prices of natural gas in June 2012 hovered around US\$2.50/MMBtu.

## CONCLUSIONS

The coal market – given its size, its importance to a number of rapidly expanding economies (such as India and China) and its transition from traditional long-term contracts to more flexible and shorter-duration transactions – offers great opportunities to trading organisations with sufficient scale and sophistication to address the logistical challenges of this business. It is not an easy market, given the complicated operational aspects of coal storage and transportation, and quality differentials. The interactions between coal and natural gas markets are a wild card creating the potential for very high price volatility for both commodities. This may be a perfect opportunity for the next generation of fearless energy traders.

1 <http://www.worldcoal.org/coal/what-is-coal/>.

2 Coke is a product obtained through heating coal without contact with air. This process removes volatile components of coal and leaves carbon with small amounts of other substances, such as hydrogen, nitrogen and sulphur, and minerals originally present in coal.

3 “Meeting projected coal production demands In the USA: Upstream issues, challenges, and strategies,” The Virginia Center for Coal and Energy Research Virginia Polytechnic Institute and State University, December 2008.

4 EIA, 1993, “U.S. coal reserves: An update by heat and sulfur content,” February, DOE/EIA-0529(92), Washington, DC.

5 <http://205.254.135.7/oiaf/analysispaper/stb/index.html>.

6 Ash represents a challenge, not only because of its volume, but also because it contains many highly toxic substances and may be mildly radioactive (see Mara Hvistendahl, 2007, “Coal ash is more radioactive than nuclear waste,” *Scientific American*, December 13).

7 860 billion tons according to BP 2012, “Statistical review of world energy,” June.

- 8 Richard K. Morse and Gang He, 2010, "The world's greatest coal arbitrage: China's coal import behavior and implications for the global coal market," working paper #94, Stanford University, Program on Energy and Sustainable Development, August.
- 9 <http://www.platts.com/IM.Platts.Content/methodologyreferences/methodologyspecs/coalmethodology.pdf>.
- 10 [http://www.argusmedia.com/~/media/Files/PDFs/Meth/argus\\_coal\\_dailyint.ashx](http://www.argusmedia.com/~/media/Files/PDFs/Meth/argus_coal_dailyint.ashx).
- 11 McCloskey (a unit of IHS) is a company specialising in the dissemination of news, analysis and data on the international coal industry.
- 12 Calorific value.
- 13 Net as received.
- 14 Information about the API 4, 5, 6 and 8 indexes is taken verbatim from the Argus website.
- 15 In FOB, the supplier absorbs the costs of delivering the commodity to the ship (transportation to the port, storage at the port and loading). FAS stands for "free alongside ship" and does not include the cost of loading.
- 16 CSX Corporation is a railway company operating coal-loading terminals.
- 17 In the US railways would set the tariff rates to prevent competition from coal delivered through other transportation modes (such as barges).
- 18 BHP and Billiton merged in June 2001.
- 19 The end of trade sanctions against South Africa in 1991–92 created a glut of coal (especially metallurgical coal) and increased bargaining power of the Japanese importers. They succeeded in imposing in 1995 the "fair treatment system" based on bilateral negotiations, with confidentiality of commercial terms. See, Stephen G. Bunker and Paul S. Ciccantell, 2007, *East Asia and the Global Economy: Japan's Ascent, with Implications for China's Future* (Baltimore, MD: Johns Hopkins University Press).
- 20 Grindability is a coal property measuring the difficulty of grinding coal to a specific size for boiler combustion. HGI was developed in the 1930 to measure empirically this property.
- 21 The 2010 version is available at <http://www.coaltrade.org/industry-standards/>.
- 22 See <http://www.icapenergy.com/US/markets/coal.aspx> for more information.
- 23 Pro rata quantity variation adjustment will be made outside the monthly contractual dead-band range based on average monthly Platts OTC Broker Index.
- 24 NS stands for Norfolk Southern, a railway company.
- 25 <http://www.eia.gov/tools/glossary/?id=coal>.
- 26 Based on analysis published by SNL.



## *Conclusions*

As we completed the first part of the book and started working on these last few pages of conclusions, we thought briefly about another author, who had put down his plume one June night in 1787 and took a short walk under the acacias of his garden, from which he could see the countryside, bordered by a lake and mountains. "The air was temperate," he wrote, "the sky was serene, the silver orb of the moon was reflected from the waters, and all nature was silent."<sup>1</sup> Gibbon is this author's favourite writer, and every day we listen to a few pages of his books on *LibriVox* whilst jogging. However, we realised that the completion of our work allowed for none of the satisfaction and pride that he must have experienced. As a matter of fact, we could feel nothing but envy. He was writing about a topic for which he had at his fingertips all the information available to the mankind at that time. No new facts would become available any time soon, barring the discovery of other written sources or archeological research. When this happens, it is over decades, not days. His books, controversial at the time and even a source of a scandal, have survived the test of time. He had an incomparable command of language, which allowed him not only to convey in a few words the crystalline beauty of the night, but also to bring alive a great historical drama. He did not have to keep up with new developments every single day, and the ending of his story has been known for centuries. He was not afraid of taking an uncompromising stand, untroubled by the objections and criticism of some of his readers: he knew how his story would end and he believed he could explain it.

Our situation was quite different. We faced the challenge of writing about a field in which the ground is constantly shifting under our feet, and whatever was written could be obsolete by the next day. We do not have the luxury of long years of reflection and careful

evaluation of facts: every day we step into a different river. Every decision and every conclusion is difficult: almost nothing is black and white, there are only many shades of grey. If everything was certain in this business, it would be very boring indeed. We wrote in a language that was not the language we grew up with, and this inevitably shows. We did not even have the consolation of taking a leisurely walk to reflect on the last few pages remaining to be written. Instead, we stepped out into the oppressive heat and humidity of a Houston summer evening, with all sorts of crawling and flying creatures ready to torment us. The reason we were able to continue writing that night was air conditioning made possible by energy, which is still plentiful and relatively cheap. Only one thing has not changed over the centuries: writers always expect more recognition than they get.<sup>2</sup>

The question we were trying to answer was whether the message of this book would survive the furious pace of change that is one of the defining features of the industry. We think the answer is quite obvious, but still worth repeating. Energy is critical to our standard of living and national security and, by extension, our future. No responsible government can fail in its responsibility to secure safe energy supplies and take precautionary measures to avoid costly interruptions in energy flows that would not only inflict direct cost on the economy, but would also undermine confidence, critical to economic growth in the long run.

We shall remain dependent for a long time on sources of energy that are finite (or, at best, from resources that cannot be replenished in geological time). In principle, as a society we have not progressed too far from a caveman satisfying his needs for heat and light. When he was cold or scared of dark, he would go out, find “stuff” and set it on fire. We do exactly the same, except that we do it on a larger scale, using sophisticated technology and in more invasive ways (as anybody who travelled in the Gulf of Mexico in 2010 or has driven through the hills of West Virginia that have been harvested for coal could testify). We may not experience absolute shortages of energy, but we do face limitations that can be compared to an elastic band. We can stretch it through the ingenuity of scientists and engineers, but there is no way to escape the physical limitations of our planet. Periods of oversupply should not fool us: the conflict between finite resources and the exponential growth of population and our needs

will become apparent sooner or later, probably sooner than most people think. After all, “the greatest shortcoming of the human race is our inability to understand the exponential function.”<sup>3</sup>

This is why efficient and unobstructed energy markets are critical to the future of our society. We need an effective way to transmit information about the scarcity and abundance of different energy alternatives, both to producers and consumers. This is the best way to induce waste-reducing behaviour and promote energy-saving innovations. This book was written in the hope that it would make a small contribution to the development of the energy trading business – not by offering another panegyric about the virtues of a free market (this battle was won a long time ago), but by providing a road map for young people trying to break into the profession. It is an information- and skill-intensive business, but learning by working in a trading-related job is comparable to the transmission of folk wisdom: it happens primarily through immersion in daily business activities and by finding mentors who can train the new hands. The road map offered here may be imperfect, but there are few comprehensive alternatives available. Hopefully, the book will help shorten their initiation rites and be of use to a new generation of professionals entering this important industry.

However, the time has now also come to embark on the next leg of this journey. The follow-up book written by the author, and published by Risk Books, will cover option valuation and risk management topics with respect to the energy industry. The ability to identify and manage energy-related risks will become increasingly important, and will be even more critical if our view of the growing volatility of the energy markets is correct. Understanding and managing energy risks requires a combination of many diverse skills and a familiarity with multiple and interacting layers of this industry: the physical and institutional layer we have covered here and the financial engineering layer that remains to be covered.

Mastering these two layers is a difficult task, requiring many years of hard work – and, as with cleaning your house, the work is never finished. However, while apologising for any imperfections in this book, the one message we want to leave the reader with is that the emotional and financial rewards are enormous.

- 1 Edward Gibbon, 1827, *Memoirs of my Life and Writings* (London, England: Hunt & Clarke), available at <http://www.gutenberg.org/ebooks/6031>.
- 2 "Another damned fat book, Mr. Gibbon? Scribble, scribble, scribble, eh Mr. Gibbon?" Attributed to a number of powerful people alive at the time.
- 3 Albert Allen Bartlett ([http://www.albartlett.org/presentations/arithmetic\\_population\\_energy.html](http://www.albartlett.org/presentations/arithmetic_population_energy.html)).

# *Index*

(page numbers in italic type refer to figures and tables)

## **A**

- abiotic oil theory 559–60
  - see also* oil: non-conventional
- Alabi, Adebayo 584
- Al-Bisharah, Mohammed 553
- American Gas Association 400
- analytical tools 801–36, 803, 807, 808, 810, 817
- data sources 819–34
  - bilateral transactions: cash markets 828–9
  - forward prices 826–7
  - generation units 821–5
  - hydro conditions 832–4
  - power plant outage information 830–2
  - power/fuel price spreads 829–30
- heat rate and thermal efficiency 805–19
- spark spread, alternative
  - definitions 810–12
- spark spread models,
  - shortcomings in 812–15
- spark spread and its uses in valuation models 809
- supply stack 816–19

- and interactions between natural gas and electricity industries 815–16
- technical characteristics of generation and load 802–4
- capacity factors 802–4
- Argus 277
- aromatics 505
- Aronofski, J. 584
- assets, trading around 68–70

## **B**

- back office 60–1
- Bakken Marketlink 649
- Baku–Tbilisi–Ceyhan pipeline 601
- Bardi, Ugo 552, 558
- Barnett Shale Basin 330, 332
- BentekEnergy 76, 77, 293, 314, 344–5, 348, 351, 388, 491, 745
- Berman, Art 541
- bitumen, defined 534–5
- Blas, Javier 634
- Bourgeois Gentilhomme, Le* (Molière) 152
- Brent Crude 629–41, 630
  - Brent complex 632–3
- Contract for Difference (CFD) 634–8, 637

- dated Brent 632, 633  
forward Brent 632  
futures contracts 633  
OTC Brent-related derivative contracts 633  
other Brent-related contacts 638–40  
25-day BFOE 633–4  
WTI's divergence from 647  
*see also* oil pricing  
British Gas 466–8  
privatisation of 466–7  
British Thermal Unit:  
conversion table for 13  
explained 12–13  
*see also* natural gas: units of measurement  
Brown, Jeffrey 553
- C**
- California cap-and-trade programme 933–4  
*see also* emission markets
- Campbell, John 592–3
- Canadian Enerdata 277–8
- Canadian Natural Gas Storage Survey 278
- Canadian tar sands 534–8  
*see also* oil
- Castro, Jordi 584
- Central Intelligence Agency, study by 75
- Centrica 467
- channels of transmission, commodity markets as 86–9
- classification of non-conventional natural gas deposits 326–8  
*see also* natural gas: non-conventional
- clearinghouse design 220–6  
and role of clearing members 222–6
- Climate Prediction Center 94
- CME and ICE coal contracts 953–5  
*see also* coal
- CME Nymex 201–7
- coal 941–61, 944, 945, 946, 947, 948, 949, 953  
calorific value of 952  
physical properties of 942–4  
reserves and production 944–9  
trading 950–60, 953  
CME and ICE contracts 953–5  
gas-for-coal substitution 956–60  
OTC contracts 954–5  
price indexes 950–1  
worldwide 951–2
- coal-fired power plants 728–30  
*see also* electricity: generation of
- Committee of Chief Risk Officers (CCRO) 124
- commodities as asset class 33–7  
hedging opportunities 34–7  
quest for high returns 33–4
- Commodity Exchange Act (CEA) 201, 213–14, 262–3, 267, 888  
most important change made to 215  
Section 4a, quoted 213
- Commodity Futures Modernization Act 23, 262, 263, 440
- Commodity Futures Trading Commission (CFTC) (US) 40, 200–1, 209, 212–18  
*passim*, 231–3, 441, 886, 887

- COT reports of 231  
 “distinguishing characteristics” formulated by 259  
 fifteen-second rule of 266  
 “interim final rule” of 260  
 and participants and regulatory developments, *see*  
 participants and regulatory developments  
 two tests of 261
- commodity indexes 180–7, 182, 183  
 Dow Jones–UBS 183–4, 185  
 Goldman Sachs 181  
 mutual funds 193–4  
 performance issues concerning 192–3  
 S&P GSCI 183  
 Thomson Reuters/Jefferies 184–6, 186
- commodity markets as channels of transmissions 86–9
- competitive intelligence 74–80  
 and event-based disruptions 77–80  
 Gulf Coast Refiner Impacts 79
- compressed natural gas (CNG) 299–300
- convenience yield 134–7
- Cushing 641, 643
- D**
- data sources, and data quality issues 308–16, 387–90; *see also* natural gas: consumption and production data for
- DCMs, *see* designated contract markets; direct clearing members
- de Jager, Jan 329
- Deas, Thomas C. Jr 241
- decline curves and depletion analysis 319
- demand and supply in energy markets 6–7, 7
- Dennis, Richard J. 420–1
- deregulation 22–3
- designated contract markets (DCMs) 200–1, 214, 262
- direct clearing members (DCMs) 222
- directional trading 65–8
- disruptions, event-based 77–80
- Dodd–Frank Act 23, 65, 67, 123, 125, 199, 200, 201, 210, 211, 214, 215, 220, 238, 245, 258–9, 265, 273, 887, 888
- Dunning–Kruger Syndrome 51
- E**
- ECMs, *see* exempt commercial markets
- Edison, Thomas 698–9
- Edison Electric Institute (EEI) 124
- EIA US energy flows 13–17
- EIA Weekly* 618
- El Niño 85, 93, 94, 96–8, 833  
*see also* weather
- electricity:
- alternating current 700–12, 700
  - and complex numbers 705
  - definitions 700–4
  - power 704–6, 713
- basics of 695–721, 713, 715, 716, 718
- definitions 697–700
- current 697–8
- Ohm’s Law 698

- power 698–9  
unit of charge 697  
generation of 723–53, 725, 726,  
730, 735, 736, 738, 739, 740,  
749  
hydro 746–8  
plants: coal 728–30  
plants: gas-fired 731–4  
plants: thermal 727–8  
poor understanding of 724  
solar 748–50  
in US: basic facts 724–7, 725, 726  
wind 740–6; *see also* wind  
    power  
interactions between natural gas  
    industry and 815–16  
loads 760–7, 762, 763, 764, 765  
    and capacitors 761  
    inductive 761  
    profile 763–7  
    resistive 761
- hockey stick bidding strategy  
    878–80  
manipulation of settlement  
    prices 880–1  
and market rules, taking  
    advantage of 881–3  
physical and economic

- European exchanges 773–81  
 power pool model 781–96, 790,  
     791, 795, 796  
 transmission and distribution  
     756–60, 757, 759  
     loop-flow problem of 759–60  
 units of measurement 717–19  
 US residential consumption 718  
 emission markets 895–939, 899,  
     910, 917, 926, 933  
 carbon 916–34, 926, 932  
 California cap-and-trade  
     programme 933–4  
 ETS assessment 927–9  
 EU emissions trading 923–7  
 legal framework 919–22  
 RGGI 931–3  
 US emissions markets 929–31  
 environment policy measures  
     896–904, 910  
     and carbon leakage 901–2  
     command and control 904  
     and efficiency 898–900  
     and emission levels, forecast of  
         902  
     emission taxes 903–4  
     and hot spots 902  
     and offsets 901  
     and price volatility 900–1  
     tradable permits 896–902  
     US regulations 2012–20 910–12  
 $\text{SO}_2$  and  $\text{NO}_x$  emission  
     regulation 904–16  
 CSAPR 912–13  
     market impact 914–16  
     utility MATS 913–14  
 Enbridge Energy Company 649  
 energy flows and consumption  
     12–21  
 British Thermal Unit 12–13  
     conversion table for 13  
 EIA US 13–17, 14, 16, 17, 20  
 Lawrence Livermore National  
     Laboratory: US energy  
     sources and uses 17–21  
 Energy Information  
     Administration (EIA) (US)  
         12, 13–16, 301, 330, 336–7,  
         353  
     energy flow 20  
     energy flows 14, 16, 17  
     natural-gas data of 304–5, 305,  
         306, 308–16, 312, 314  
     projections of 337  
     storage reports of 400–6  
     *see also* natural gas:  
         transportation and storage  
         of  
 energy markets:  
     coal 944, 945, 946, 947, 948, 949,  
         953  
     calorific value of 952  
     physical properties of 942–4  
     reserves and production 944–9  
     trading 950–60, 953  
 commodity swaps and options  
     139–52, 143, 146; *see also*  
         natural gas: US markets for  
     options 147–52  
 physical, financial and basis  
     swaps 142–4, 143  
     swaps, customised 144–7  
     swaps, mechanics of 139–42  
 common features of 4–12  
     complexity 11–12  
     contracts, prices, markets 4–5  
     market manipulation 8–10  
     price discovery 8

- price formation process 5–6  
price reporting agencies 10–11  
supply and demand 6–7, 7  
transactions for spreads 7–8  
complexity in 11–12  
and contracts and prices 4–5  
electricity, analytical tools for  
801–36, 803, 807, 808, 810,  
817; *see also* analytical tools  
data sources 819–34  
heat rate and thermal  
efficiency 805–19, 807, 808,  
810  
and interactions between  
natural gas and electricity  
industries 815–16  
technical characteristics of  
generation and load 802–4  
electricity, power pools and  
exchanges 767–96, 853–7  
electricity market design  
767–99  
electricity markets as auctions  
769–73  
European exchanges 773–81  
power pool model 781–96, 790,  
791, 795, 796  
electricity, transactions 842, 847,  
853–7, 856  
ancillary services 838–43  
capacity payments 843–9, 847  
capacity payments, and  
market-design issues 845–9  
FTR market 849–53  
power market transactions  
853–7  
structured transactions 857–62  
structured transactions:  
contract risks 858–60  
structured transactions: full  
requirements deals 857  
structured transactions:  
services provided 857–8  
structured transactions: tolling  
transactions 860–2  
structured transactions:  
volumetric commitments  
858  
US power pools 853–7  
emissions 895–939, 899, 910, 917,  
926, 933; *see also* emission  
markets  
carbon 916–34, 926, 932  
environment policy measures  
896–904, 910–12, 910  
 $\text{SO}_2$  and  $\text{NO}_x$  emission  
regulation 904–16  
and energy flows and  
consumption 12–21  
British Thermal Unit 12–13  
EIA US 13–17, 14, 16, 17, 20  
Lawrence Livermore National  
Laboratory: US energy  
sources and uses 17–21  
and environmental issues 42–4  
exchanges 199–251, 204, 208, 209,  
213, 219, 229, 234, 235, 239,  
240, 256  
clearinghouse design 220–6  
CME Nymex 201–7  
and commitments of traders  
and related reports 231–8  
Intercontinental 207–11  
Natural Gas 211–12  
position limits in futures  
markets 212–18  
position limits and margins:  
current practice 218–20

- reviewed 201–12  
 and role of clearing members 222–6  
 safeguards 222–6  
 SPAN system 226–31, 229  
 whether or not to clear 238–45  
 financial innovation in 37–44, 39, 40, 41, 42  
 commodity futures, bonds, notes and index funds 38–9  
 and environmental issues 42–4  
 exchange-traded funds 39  
 instruments, spot, forward and futures markets 110–38, 116, 120, 121  
 and interactions between natural gas and electricity industries 815–16  
 intertwined nature of 3  
 manipulation and gaming in 8–10, 865–93, 870, 878; *see also* manipulation and gaming in electricity 875–85  
 Enron strategies in California 883–5, 883  
 hockey stick bidding strategy 878–80  
 manipulation of settlement prices 880–1  
 market behaviour rules 886–90  
 and market rules, taking advantage of 881–3  
 physical and economic withholding 877–8, 878  
 by price: the mechanics 865–75  
 of settlement prices 880–1  
 natural gas, Europe 464–87  
 natural gas, international 463–500, 496  
 European 464–87, 467, 471, 483, 488, 489, 492  
 European: Continental 473–5  
 European: hubs 482–7, 488  
 European: long-term supply contracts 475–9  
 European: regulatory developments 479–82  
 European: UK 466–73  
 liquefied natural gas (LNG) markets and transactions 487–95  
 US liquefaction plants 496  
 natural gas, US 418, 425, 427, 430, 433, 436, 441, 447, 454, 455, 457, 458, 471, 480  
 balancing market 418–19  
 demand curve 416–18  
 essential characteristics 414–18  
 futures 436–7  
 index prices, importance of 424–6  
 monthly and daily transactions 422–4  
 municipal prepay transactions 448–9  
 physicals, exchange for 437–40  
 processing plant contracts 449–54  
 spot, forward and futures 414–58  
 supply curve 418  
 swap valuation, basis model 430–4  
 swaps, basis 429–30  
 swaps, floating-for-floating 434

- swaps, ICE 440–2  
swaps, mechanics and uses of 434–5  
swaps, natural gas 426–8  
swaps and index transactions 435–6  
trader's perspective 419–21  
volumetric production payments (VPPs) 442–7  
volumetric risk management 454–8
- oil:  
crude, classifications 507  
crude, and its properties 504–10  
output and demand statistics 512–19, 517, 518  
processing 565–600  
properties, production and reserves 503–32, 504, 507, 514, 521, 522, 523, 526, 527, 528  
units of measurement 510–11  
oil pricing 625–61, 630, 637, 642, 644, 647, 648, 650, 653  
Brent Crude 629–41, 630, 637, 647  
and government selling prices (GSPs) 627  
Oman–Dubai 653–6, 653  
posted price 626–7  
regimes for 625–9  
West Texas Intermediate (WTI) 641–53, 647, 648
- oil, transactions in 663–91, 669, 675, 676, 689  
and basis risk 673  
collars 668–72, 669
- crude and refined products 664–8  
examples of 668–79  
freight risk management 679–88; *see also* freight risk management  
participating swaps 677–9  
and refinery configuration risk 674  
3:2:1 crack spread 672–7, 674  
and volumetric risk 673
- participants and regulatory developments, *see* participants and regulatory developments
- and price discovery 8  
price formation process 5–6  
and price reporting agencies 10–11
- and regulatory developments, *see* participants and regulatory developments
- spot, forward and futures 110–38, 116, 120, 121, 127; *see also* analytical tools: data sources
- and bid–offer spreads 133–4  
and convenience yield 134–7  
credit risk 123–4  
definitions 110–22  
differences: futures and forward 122–7  
forward prices versus futures: numerical differences 126–7
- and large individual transactions 126
- and limited market activity 126
- and limited participation 126

- market participation 125–6  
 and normal backwardation 128–31  
 and practitioners, what they do 137–8  
 price standardisation 125  
 relationship: spot and forward 127–8  
 standardisation 122–3  
 and storage theory 131–3  
 and transaction costs and access to infrastructure 134  
 US natural gas market 414–58;  
*see also* natural gas: US markets for  
 and spreads, transactions for 7–8  
 structured transactions in 157–98, 160, 161, 162, 163, 164, 165, 166, 174, 175, 176, 185, 186, 187  
 explained and discussed 158–67  
 passive investment instruments 179–94, 182, 183; *see also* passive investment instruments  
 weather derivatives 167–79  
 and supply and demand 6–7, 7  
 technical progress in 44–5  
 trading, *see* energy trading  
 US natural gas, *see* natural gas: US markets for  
*see also* energy trading; *individual commodities*  
 energy return on investment (EROI) 536–7  
 energy trading:  
 brief history of 21–37  
 and commodities as asset class 33–7  
 deregulation 22–3  
 financialisation 23–7  
 globalisation 27–33  
 and commodities as asset class:  
 hedging opportunities 34–7  
 quest for high returns 33–4  
 organisation of 51–83  
 back office 60–1  
 competitive intelligence 74–80  
 described 52–61  
 front office 57–60  
 fundamental analysis 71–4  
 middle office 60  
 and sources of profits 61–71  
 and profits, sources of:  
 asset management 70–1  
 directional trading 65–8  
 market making 63–5  
 trading around assets 68–70  
 weather information in 85–105  
 anomalies 93–8  
 and commodity markets:  
 channels of transmission 86–9  
 forecasting models 89–93  
*see also* energy markets  
 Enron, California strategies of 883–5  
 Environment Canada 89  
 ethane 290  
*see also* natural gas: gathering and processing of; natural gas: liquids  
 ethanol 513, 518, 544–9 *passim*  
 energy required to produce 546  
 production costs in main production countries 548  
*see also* oil: non-conventional

- European Center for Medium-Range Weather Forecasts (ECMWF) 89, 91–2
- European Union, natural-gas consumption in 465; *see also* natural gas: international markets for
- event-based disruptions 77–80
- Gulf Coast Refiner Impacts 79
- Hurricane Katrina 78, 288, 306, 614
- Hurricane Rita 78, 288, 306
- Excel Solver 346
- exchanges 199–251, 204, 208, 209, 219, 229
- clearinghouse design 220–6
- and commitments of traders and related reports 231–8
- Intercontinental 207–11
- market data 211
- OTC markets 209–11
- Natural Gas 211–12
- position limits and margins: current practice 218–20
- and role of clearing members 222–6
- whether or not to clear 238–45
- see also* energy markets
- exchange-traded weather derivatives 170–1
- see also* structured transactions; weather: derivatives
- exempt commercial markets (ECMs) 201, 211, 262–3, 264, 441
- F**
- Federal Agency Regulation Commission (FERC) 273, 276, 370, 372, 373–4, 379–80, 383, 387, 389, 407–9, 479, 782–4, 820, 838, 839, 841, 843, 886
- and intrastate pipelines 383–4
- fifteen-second rule 266
- financialisation 23–7
- firm transmission rights 849–53
- Fischer, Franz 549
- freight risk management 679–88
- and Baltic Clean Tanker Index 680–1
- and Baltic Dirty Tanker Index 681–4
- and listed tanker voyage FFA products 685–8
- US\$/mt 687–8
- Worldscale 685–7
- see also* oil markets
- front office 57–60
- fuel oils 585–8
- physical and chemical properties of 586–8
- see also* oil
- futures commission merchants 266–7
- G**
- gas-fired power plants 731–4
- see also* electricity: generation of
- gas-for-coal substitution 956–60
- see also* coal
- gasoil 588
- gasoline/petrol 588–97 *passim*
- see also* oil
- Genscape 77, 388
- Gibbon, Edward 963
- Global Forecast System 91
- globalisation 27–33

- Goldman Sachs Commodity Index 181
- Gottschalck, Jon 94
- government selling prices (GSPs) 627
- Graeber, David 167–8
- Gulf of Mexico, 2012 spill in 87
- Gulf Coast Refiner Impacts 79
- H**
- Hall, Charles 537
- Hicks, John Richard 128, 129
- Höök, Mikael 559
- Houston Ship Canal (HSC) 51, 139, 145, 389, 422
- Howard, Frank 549
- Hubbert, Marion King 551–6, 552, 553
- Hurrel, James W. 93
- Hurricane Katrina 78, 288, 306, 614  
*see also* weather
- hurricane prediction and impact assessment 99–102  
*see also* weather
- Hurricane Rita 78, 288, 306  
*see also* weather
- hydrocarbons, classification of 504
- hydropower 476–8  
*see also* electricity: generation of
- I**
- Ibbotson Associates, study by 36
- ICIS Heron 277
- “Inside FERC Gas Market Report” 423, 424
- Intelligence Press (IP) 278
- International Petroleum Monthly* 512
- International Swap and Derivatives Association (ISDA) 124, 260, 640  
suit filed by 214
- ISDA, *see* International Swap and Derivatives Association
- isobutane 291  
*see also* natural gas: gathering and processing of; natural gas: liquids
- J**
- Johnson, Lyndon 837
- K**
- Katy 389
- Kay, John 26
- kerosene 588
- Keynes, John Maynard 128, 129, 130
- Kirchoff, Gustav 699–700, 700  
Current Law of 701  
Voltage Law/Loop Law of 700
- Kudryavtsev, Nikolai Alexandrovich 559
- Kunzig, Robert 535
- L**
- La Niña 85, 93, 94, 96–8, 833  
*see also* weather
- Lavacchi, A. 552
- Lawrence Livermore National Laboratory: US energy sources and uses 17–21  
*see also* EIA US energy flows
- LCI Energy Insight 388
- Libya, and oil posted price 626–8
- liquefied natural gas (LNG) 390–8  
markets and transactions 487–95, 488, 489, 492, 493

- supply chain for, cost of 396–8  
technology and markets 392–6,  
394, 395
- liquefied petroleum gas (LPG) 289,  
291, 571, 597  
*see also* oil
- LNG, *see* liquefied natural gas
- loop-flow problem 759–60  
*see also* electricity: transmission  
and distribution
- Lorentz, Edward 90
- M**
- McGlade, Christopher 524
- Madden–Julian Oscillation (MJO)  
93, 94–5, 833
- major swap participant, defined  
259
- Malallah, Adel 553
- manipulation and gaming 8–10,  
865–93, 878
- in electricity markets 875–85, 883  
hockey stick bidding strategy  
878–80, 879  
and market rules, taking  
advantage of 881–3  
physical and economic  
withholding 877–8, 878  
of settlement prices 880–1
- Enron strategies in California  
883–5  
hockey stick bidding strategy  
878–80
- market behaviour rules 886–90  
physical and economic  
withholding 877–8, 878
- by price: the mechanics 865–75  
of indexes 868–72, 870  
motivation 871–2
- two parallel markets of  
different levels of  
transparency 872–5
- of settlement prices of  
outstanding contracts 880–1
- market making 63–5
- market manipulation, *see*  
manipulation and gaming
- middle office 60
- Moliere 152
- monitoring pipeline flows 384–90  
data sources 387–90  
physical flows 384–7
- “Monthly Gross Natural Gas  
Production Report” 313
- “Monthly Oil Market Report” 518
- Motor Octane Number 593
- N**
- naphtha 571, 572, 575  
straight-run 597  
*see also* oil
- naphthalenes 505
- Nashawi, Ibrahim Sami 552–3
- National Center for Environmental  
Prediction 89, 91
- National Oceanic and Atmospheric  
Administration (NOAA) 91
- natural gas:  
compressed (CNG) 299–300  
consumption and production  
data for 300–16, 301, 302,  
303, 304, 305, 306, 307, 309,  
310, 312, 314  
data sources and data quality  
issues 308–16  
US trends 300–7, 302, 303, 304,  
305, 306, 307  
world production 308

- decline curves and depletion analysis 316–21, 318, 319
- distribution companies, as market participants 256
- and electricity industries, interactions between 815–16
- energy equivalent conversions 299
- gathering and processing of 287–97, 292, 295
- gathering systems 287–9
- liquids (NGLs) 289–93
- plants 293–7
- international markets for 463–500, 496
- European 464–87, 467, 471, 483, 488, 489, 492
- European: Continental 473–5
- European: hubs 482–7, 488
- European: long-term supply contracts 475–9
- European: regulatory developments 479–82
- European: UK 466–73
- liquefied natural gas (LNG) markets and transactions 487–95
- US liquefaction plants 496
- liquefied (LNG) 390–8
- in international markets 487–95
- supply chain for, cost of 396–8, 398
- technology and markets 392–6, 394, 395
- liquids (NGLs) 289–93, 455
- ethane 290
- isobutane 291
- normal butane 291
- pentane 292
- propane 290–1
- midstream, *see* natural gas: transportation and storage of
- non-conventional, *see* non-conventional natural gas
- origins and chemical properties of 284–6, 285
- power plants fired by 731–4; *see also* electricity: generation of
- SI prefixes concerning 298
- transportation and storage of 365–412
- capacity estimates 408
- liquefied gas (LNG) 390–8, 394, 395
- nominal process 374–9, 375, 376
- pipeline designs 368
- pipeline flows, monitoring 384–90
- pipeline regulation 379–84
- pipeline transportation contracts 368–74, 371, 372
- storage contracts 406–9
- storage facilities 398–400
- storage forecasts 386–7
- storage reports 400–6
- US pipeline grid 366–8
- unit conversion for 301
- units of measurement 297–300
- upstream 283–324
- consumption and production data for 300–16, 301, 302, 303, 304, 305, 306, 307, 309, 310, 312, 314
- decline curves and depletion analysis 316–21, 318, 319

- gathering and processing of  
287–97
- liquids (NGLs) 289–93
- origins and chemical  
properties of 284–6, 285
- US liquefaction plants 496
- US markets for 413–61, 418, 425,  
427, 430, 433, 436, 441, 447,  
454, 455, 457, 458
- balancing market 418–19
- demand curve 416–18
- essential characteristics  
414–18
- futures 436–7
- index prices, importance of  
424–6
- monthly and daily  
transactions 422–4
- municipal prepay transactions  
448–9
- physicals, exchange for  
437–40
- processing plant contracts  
449–54
- spot, forward and futures  
414–58
- supply curve 418
- swap valuation, basis model  
430–4
- swaps, basis 429–30
- swaps, floating-for-floating  
434
- swaps, ICE 440–2
- swaps, mechanics and uses of  
434–5
- swaps, natural gas 426–8
- swaps and index transactions  
435–6
- trader’s perspective 419–21
- volumetric production  
payments (VPPs) 442–7
- volumetric risk management  
454–8
- Nelson Complexity Index 577
- NGLs, *see* natural gas: gathering  
and processing of
- Nichols, Jeff 420
- NIST Handbook* 299
- non-conventional natural gas  
325–63
- coal-bed methane 327, 328
- deposits, classification of 326–8
- “greenness” of 353–9
- shale gas 327, 328–53, 331, 332,  
333, 337, 342, 347, 349, 350;  
*see also* shale gas
- tight gas 327; *see also* natural gas
- non-conventional oil 533–64, 539,  
542, 546
- abiotic 559–60
- biofuels 543–51
- Canadian tar sands 534–8
- Orinoco Belt Heavy 538–9
- and peak oil theory 551–60; *see*  
*also* oil: peak, theory of; peak  
oil theory
- shale 539–43
- see also* oil
- normal backwardation 128–31
- normal butane 291
- see also* natural gas: gathering  
and processing of; natural  
gas: liquids
- North American Natural Gas  
Methodology and  
Specifications Guide* 423
- Nuclear Regulatory Commission  
(NRC) 737–9

- nuclear-powered power plants 734–9, 735, 736, 738, 739  
*see also* electricity: generation of
- O**
- Oasis 389
- oil:
- biofuels 543–51
  - advanced 544
  - bio-alcohols 544–5
  - production, 2011 545
- crude, definitions 525
- crude, and its properties 504–10
- physical properties and
  - classification of 506–10, 507
- and importance of inventories 615–22
- classifications 620–2
- markets, *see* oil markets
- non-conventional 533–64, 539, 542, 546
- abiotic 559–60
  - biofuels 543–51
  - Canadian tar sands 534–8
  - Orinoco Belt Heavy 538–9
  - and peak oil theory 551–60; *see also* oil: peak, theory of; peak oil theory
  - shale 539–43
- output and demand statistics 512–19
- production trends 516–19
  - US supply 518
  - world supply 517
- peak, theory of 551–60, 552, 553
- and abiotic oil 559–60
  - and future of oil 556–9
  - Hubbert’s model, extension to 552–6
- and pipeline operations 609–14
- and market disruptions 613–14
- US oil and product pipelines 609–13
- pricing of, *see* oil pricing
- processing of 565–600
- refined products 585–97
  - refining 566–85
- properties, production and
- reserves 503–32, 504, 507, 514, 521, 522, 523, 528
  - crude 504–10
- output and demand statistics 512–19
- reserves 519–28, 521
  - reserves data 526–8, 527, 528
  - and SEC rules 526
  - units of measurement 510–11
- refined products 585–97
- fuel oils 585–8
  - gasoil 588
  - gasoline/petrol 588–97 *passim*, 593, 596
- industry terms 589
- kerosene 588
- refining 566–85, 576, 581, 583
- capacity 578–80, 578, 580
  - conversion and upgrading 573–6
  - desulphurisation 572–3
  - distillation 569–72
  - main processes, listed 568
  - and Nelson Complexity Index 577
  - optimisation 580–5
  - preprocessing 569
  - refined products, listed 567
  - US yields 576

- shale 539–43  
US resources 539  
in tankers 602–9, 604  
charter contracts 605–7  
Worldscale 607–9  
transportation and storage of  
601–24, 602, 604, 605, 609,  
610, 611, 612, 615, 619, 620  
charter contracts 605–7  
and oil inventories,  
importance of 615–22  
and pipeline operations  
609–14  
tankers 602–9, 604  
Worldscale 607–9  
units of measurement 510–11  
world capacity 543  
oil markets:  
transactions in 663–91, 669, 675,  
676, 689  
and basis risk 673  
collars 668–72, 669  
crude and refined products  
664–8  
examples of 668–79  
freight risk management  
679–88; *see also* freight risk  
management  
participating swaps 677–9  
and refinery configuration risk  
674  
3:2:1 crack spread 672–7, 674  
and volumetric risk 673  
Oil Price Information Service  
(OPIS) 278  
oil pricing 625–61, 630, 637, 642,  
644, 647, 648, 650, 653  
Brent Crude 629–41, 630  
Brent complex 632–3  
Contract for Difference (CFD)  
634–8, 637  
dated Brent 633  
forward Brent 632  
futures contracts 633  
OTC Brent-related derivative  
contracts 633  
other Brent-related contracts  
638–40  
25-day BFOE 633–4  
WTI's divergence from 647  
and government selling prices  
(GSPs) 627  
Oman–Dubai 653–6, 653  
posted price 626–7  
regimes for 625–9  
and posted price 626–7  
“seven sisters” 626  
West Texas Intermediate (WTI)  
641–53, 648  
Basin, Centurion, Spearhead  
and Seaway pipelines 643–4  
Brent's divergence from 647  
new pipeline projects 648–50  
physical trading 651–3  
WTI CME futures contract  
650–1  
olefins 505–6  
Oman–Dubai 653–6, 653  
*see also* oil pricing  
Oneok Partners 651  
Organization of the Petroleum  
Exporting Countries:  
Middle East reserves of 521  
and posted-price increase 626–7  
Organization of the Petroleum  
Exporting Countries (OPEC)  
519  
*see also* oil pricing: regimes for

- Orinoco Belt Heavy Oils 538–9  
*see also* oil
- OTC Brent-related derivative contracts 633  
*see also* Brent Crude; oil pricing
- over-the-counter (OTC) voice brokers 264–7  
 futures commission merchants 266–7
- over-the-counter (OTC) weather derivatives 172–6  
*see also* structured transactions; weather: derivatives
- P**
- Pacific Decadal Oscillation (PDO) 93, 95–6, 95, 96
- paraffins 505
- participants and regulatory developments 253–80
- over-the-counter (OTC) voice brokers 264–7  
 futures commission merchants 266–7
- participants 254–8  
 brokers 256  
 electricity producers 255–6  
 exchanges 256  
 financial institutions 256  
 fossil-fuel producers 255  
 major swap participant 259–60  
 natural gas distribution companies 256  
 processors 255  
 swap dealers 259  
 transporters 255
- price reporting agencies (PRAs) 268–78, 270
- business model and services 268–74  
 regulatory scrutiny of 274–6  
 reviewed 277–8
- regulatory developments 258–64
- passive investment instruments 179–94
- commodity indexes 180–7, 182, 183  
 Dow Jones–UBS 183–4, 185  
 Goldman Sachs 181  
 mutual funds 193–4  
 S&P GSCI 183  
 Thomson Reuters/Jeffries 184–6, 186
- exchange-traded funds and notes 187–92
- mutual funds 193–4
- performance issues concerning 192–3
- Patzek, Tad 548
- peak oil theory 551–60  
 and future of oil 556–9  
 Hubbert's model, extension to 552–6  
*see also* oil
- pentane 292  
*see also* natural gas: gathering and processing of; natural gas: liquids
- petrol/gasoline 588–97 *passim*  
*see also* oil
- Pichler, Helmut 549
- Pineapple Express 94
- pipeline flows, monitoring 384–90  
 data sources 387–90  
 physical flows 384–7

- pipeline regulation 379–84, 380  
  interstate 380–4  
  intrastate 383–4
- pipeline transportation contracts  
  368–74, 371, 372  
  authorised overrun 370  
  firm service 369  
  interruptible service 369  
  no-notice 369–70  
  other primary firm 370  
  secondary firm 370
- Platts 277, 422, 632, 633, 638, 640  
Crude Oil Marketwire 638  
E-Window 640–1  
posted price 626–7  
  *see also* oil pricing
- Potential Gas Committee 335
- power generation 723–53, 725, 726,  
  730, 735, 736, 738, 739, 740,  
  749  
  hydro 746–8  
  plants: coal 728–30  
  plants: gas-fired 731–4  
  plants: thermal 727–8  
  poor understanding of 724  
  solar 748–50  
  in US: basic facts 724–7, 725,  
  726  
  wind 740–6; *see also* wind power  
  *see also* electricity
- power pool model 781–96  
  locational marginal price (LMP)  
  785–96  
  mathematical formulation and  
  an example 787–96  
  *see also* electricity: power pools  
  and exchanges
- price discovery process 8
- price reporting agencies 10–11
- price reporting agencies (PRAs)  
  268–78
- business model and services  
  268–74
- index construction 271–2  
manipulation 272–4
- regulatory scrutiny of 274–6
- reviewed 277–8  
  Argus 277  
  Canadian Enerdata 277–8  
  Intelligence Press (IP) 278  
  OPIS 278  
  Platts 277
- processing plants 293–7
- contracts:  
  fixed-fee arrangements 449  
  “keep whole” contract 450  
  percentage-of-index  
    arrangements 450  
  percentage-of-proceeds  
    arrangements 449–50  
  *see also* natural gas: gathering  
  and processing of
- profits, sources of 61–71
- propane 290–1  
  *see also* natural gas: gathering  
  and processing of; natural  
  gas: liquids
- propanol 544  
  *see also* oil: non-conventional
- Province of Alberta Energy  
  Resource Conservation  
  Board 535
- R**
- Research Octane Number 593
- S**
- Sabatier, P. 549

- Securities Industry & Financial Management Association, suit filed by 214
- Senderens, J. D. 549
- Seneca 558
- “seven sisters” 626  
*see also* oil pricing
- Shakespeare, William 646
- shale gas 328–53, 331, 332, 333, 337, 342, 347, 349, 350
- controversy over 338–53
- environmental impact of 338–40
- geopolitical repercussions of 352–3
- half-cycle breakeven cost 345–50, 347
- industry countermeasures concerning 351–2
- prices, depressed, reason for 343–5
- prices and production costs 340–2
- and asset prices 342
- and cashflow and balance-sheet stress 342
- dynamics of 340–1
- resources 335–8
- technological breakthrough in 334
- SIFMA, *see* Securities Industry & Financial Management Association
- Skrebowski, Chris 553
- SO<sub>2</sub> and NO<sub>x</sub> emission regulation 904–16
- CSAPR 912–13
- market impact 914–16
- utility MATS 913–14
- see also* emission markets
- Society of Petroleum Engineers (SPE) 317, 520
- solar energy 748–50
- see also* electricity: generation of
- Sommers, Jill 267
- sources of profits 61–71
- SPAN system 226–31, 229
- spreads, transactions for 7–8
- Standard Portfolio Analysis of Risk (SPAN) system 226–31, 229
- storage and transportation, *see* transportation and storage
- storage forecasts, natural gas 386–7
- storage theory 131–3
- straight-run naphtha 579
- structured transactions 157–98, 160, 161, 162, 163, 164, 165, 166, 174, 175, 176, 185, 186, 187
- electricity markets 857–62
- explained and discussed 158–67
- passive investment instruments 179–94, 182, 183; *see also* passive investment instruments
- weather derivatives 167–79
- design of 169–70
- examples of 170–2
- over-the-counter (OTC) 172–6
- pros and cons of 176–9
- and weather’s impact on prices 176
- Stulz, René 124
- supply and demand in energy markets 6–7, 7
- swap data repositories (SDRs) 125, 254, 273
- swap dealer, defined 259

**T**

- Teagle, Walter 549–50  
technical progress in energy markets 44–5  
Tesla, Nikola 699  
thermal power plants 727–8  
*see also* electricity: generation of  
Tolkien, J. R. R. 535  
trading around assets 68–70  
transaction costs and access to infrastructure 134  
transactions for spreads 7–8  
transactions in oil markets 669, 675, 676, 689  
and basis risk 673  
crude and refined products 664–8  
examples of 668–79  
collars 668–72, 669  
participating swaps 677–9  
3:2:1 crack spread 672–7  
freight risk management 679–88  
participating swaps 677–9  
and refinery configuration risk 674  
and volumetric risk 673  
*see also* freight risk management  
Trans-Alaska Pipeline 509  
transportation and storage:  
of natural gas 365–412; *see also* natural gas  
capacity estimates 408  
liquefied gas (LNG) 390–8, 394, 395  
nominal process 374–9, 375, 376  
pipeline designs 368  
pipeline flows, monitoring 384–90

- pipeline regulation 379–84  
pipeline transportation contracts 368–74, 371, 372  
storage contracts 406–9  
storage facilities 398–400  
storage forecasts 386–7  
storage reports 400–6  
US pipeline grid 366–8  
of oil 601–24, 602, 604, 605, 609, 610, 611, 612, 615, 619, 620;  
*see also* oil  
charter contracts 605–7  
and oil inventories, importance of 615–22  
and pipeline operations 609–14  
tankers 602–9, 604  
Worldscale 607–9

Tropsch, Hans 549

**U**

- units of measurement:  
electricity 717–19  
natural gas 297–300  
oil 510–11  
US pipeline grid 366–8

**V**

- Ventyx 76  
Voltaire 413  
volumetric production payments (VPPs) 442–7  
volumetric risk management 454–8

**W**

- weather:  
anomalies 93–8, 95, 96  
Southern Oscillation Index 98  
derivatives 167–79

- design of 169–70  
examples of 170–2  
exchange-traded 170–1  
pros and cons of 176–9  
and energy trading 85–105  
anomalies 93–8, 95, 96  
and commodity markets:  
    channels of transmission  
    86–9  
forecasting models:  
    American/Canadian 92  
    European 91–2  
    Global Forecast System 91  
and hurricane prediction and  
    impact assessment 99–102  
impact of, on energy prices 176  
West Texas Intermediate (WTI)  
    641–53, 648  
Basin, Centurion, Spearhead and  
    Seaway pipelines 643–4  
Brent's divergence from 647  
new pipeline projects 648–50  
physical trading 651–3  
WTI CME futures contract 650–1;  
    *see also* oil pricing  
Westinghouse, George 699  
wind power 740–6, 740  
    controversy over 743–6  
    and reactive power 745–6  
    turbines 742  
    turbulence caused by 744  
    *see also* electricity: generation of  
Woolley Victoria 205  
World Coal Association 942  
World Petroleum Congress 520  
*World without Oil* 268  
WTI, *see* West Texas Intermediate

## Y

- Yamani, Ahmed Zaki 557  
Yergin, Dan 557

