

Tadahiro Nakajima
Shigeyuki Hamori

Energy Trading and Risk Management

Commentary on Arbitrage, Risk
Measurement, and Hedging Strategy



Kobe University Monograph Series in Social Science Research

Series Editors

Yunfang Hu, Kobe University Graduate School of Economics, Kobe, Japan

Shigeyuki Hamori, Kobe University Graduate School of Economics, Kobe, Japan

Editorial Board

Masahiro Enomoto, Kobe University RIEB, Kobe, Japan

Yoshihide Fujioka, Kobe University Graduate School of Economics, Kobe, Japan

Yuka Kaneko, Kobe University Center for Social Systems Innovation, Kobe, Japan

Kazumi Suzuki, Kobe University Graduate School of Business Administration, Kobe, Japan

Kenji Yamamoto, Kobe University Graduate School of Law, Kobe, Japan

The Kobe University Monograph Series in Social Science Research is an exciting interdisciplinary collection of monographs, both authored and edited, that encompass scholarly research not only in the economics but also in law, political science, business and management, accounting, international relations, and other sub-disciplines within the social sciences. As a national university with a special strength in the social sciences, Kobe University actively promotes interdisciplinary research. This series is not limited only to research emerging from Kobe University's faculties of social sciences but also welcomes cross-disciplinary research that integrates studies in the arts and sciences.

Kobe University, founded in 1902, is the second oldest national higher education institution for commerce in Japan and is now a preeminent institution for social science research and education in the country. Currently, the social sciences section includes four faculties—Law, Economics, Business Administration, and International Cooperation Studies—and the Research Institute for Economics and Business Administration (RIEB). There are some 230-plus researchers who belong to these faculties and conduct joint research through the Center for Social Systems Innovation and the Organization for Advanced and Integrated Research, Kobe University. This book series comprises academic works by researchers in the social sciences at Kobe University as well as their collaborators at affiliated institutions, Kobe University alumni and their colleagues, and renowned scholars from around the world who have worked with academic staff at Kobe University. Although traditionally the research of Japanese scholars has been publicized mainly in the Japanese language, Kobe University strives to promote publication and dissemination of works in English in order to further contribute to the global academic community.

Tadahiro Nakajima · Shigeyuki Hamori

Energy Trading and Risk Management

Commentary on Arbitrage, Risk Measurement, and Hedging Strategy



Tadahiro Nakajima
The Kansai Electric Power Company,
Incorporated (Japan)
Osaka, Japan

Shigeyuki Hamori 
Graduate School of Economics
Kobe University
Kobe, Japan

ISSN 2524-504X

ISSN 2524-5058 (electronic)

Kobe University Monograph Series in Social Science Research

ISBN 978-981-19-5602-7

ISBN 978-981-19-5603-4 (eBook)

<https://doi.org/10.1007/978-981-19-5603-4>

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721,
Singapore

Acknowledgements

In publishing this book, we have benefited greatly from various people. First of all, we would like to express our special thanks to Ms. Juno Kawakami for excellent editorial guidance. Next, we would like to thank the officers and co-workers at the Kansai Electric Power Company and the faculty members and graduate students at Kobe University for many helpful discussions and appropriate comments. Finally, we would like to thank our family members, Shie, Aiki, and Naoko. Without their heartwarming supports, we could not have completed this work. This work was supported by JSPS KAKENHI Grant Number 22K01424.

Osaka, Japan
Kobe, Japan

Tadahiro Nakajima
Shigeyuki Hamori

Contents

1	Preface	1
	References	3
2	Arbitrage Trading in Energy Markets and Measuring Its Risk	5
2.1	Introduction	5
2.2	Data and Preliminary Analyses	7
2.2.1	Descriptive Statistics	7
2.2.2	Stationarity and Unit Root Test	12
2.2.3	Cointegration Test	16
2.2.4	Long-Term Equilibrium Estimation	20
2.3	Trading Strategies	20
2.3.1	Arbitrage Between Own Spot Spread and Future Spread	22
2.3.2	Statistical Arbitrage	23
2.4	Simulation Results	24
2.5	Risk Measurement in Statistical Arbitrage	28
2.5.1	Value-At-Risk and Expected Shortfall	29
2.5.2	Copula	31
2.5.3	Copula Estimation and Risk Measurement	41
2.6	Concluding Remarks	44
	References	49
3	Fuel Market Connectedness and Fuel Portfolio Risk	53
3.1	Introduction	53
3.2	Data	54
3.2.1	Crude Oil	55
3.2.2	Natural Gas	56
3.3	Methodology	57
3.3.1	Connectedness Index (Diebold and Yilmaz [4])	58
3.3.2	Spectral Decomposition (Baruník and Křehlík [1])	60
3.3.3	EGARCH Volatility Series Estimation	61
3.4	Analysis Results	62

3.4.1 Crude Oil	63
3.4.2 Natural Gas	71
3.5 Concluding Remarks	81
References	83
4 Hedging Strategy with Futures Contracts	85
4.1 Introduction	85
4.2 Data	87
4.3 Optimal Hedge Ratio and Hedge Effectiveness	89
4.4 Multivariate GARCH Model	90
4.4.1 Diagonal VECM Model	91
4.4.2 Diagonal BEKK Model	92
4.4.3 CCC Model	93
4.5 Analysis Results	94
4.5.1 HH Market	94
4.5.2 NBP Market	98
4.6 Concluding Remarks	102
References	104
5 Market Risk of a Power Generation Business	105
5.1 Introduction	105
5.2 Methodology	107
5.3 Data and Preliminary Analyses	107
5.3.1 Price Series	107
5.3.2 Return Series	108
5.3.3 Volatility Series	110
5.4 Analysis Results	114
5.4.1 Return Series	114
5.4.2 Volatility Series	115
5.4.3 Risk Measurement	119
5.5 Concluding Remarks	119
References	122
6 Alternative to Postface: Market Risk Transfer in Power Companies	123
References	129
Index	131

About the Authors

Dr. Tadahiro Nakajima is Senior Researcher at the Kansai Electric Power Company, Incorporated. He received a Ph.D. from Kobe University. He received the Highly Commended Paper of the *Studies in Economics and Finance*—Literati Award (Emerald Publishing). His main research interests are applied time series analysis, energy economics, and energy markets. He has published about 20 articles in international peer-reviewed journals and a book from Springer.

Dr. Shigeyuki Hamori is Professor of Economics at Kobe University in Japan. He received a Ph.D. from Duke University. He is President of the International Research Institute for Economics and Management, Distinguished Fellow of the International Engineering and Technology Institute (DFIETI), and Distinguished Fellow of the Institute of Data Science and Artificial Intelligence (DFIDSAI). His main research interests are applied time series analysis, empirical finance, data science, and international finance. He is Co-editor of the *Singapore Economic Review*, Associate Editor of the *International Review of Financial Analysis*, and Associated Editor of the *Eurasian Economic Review*. He served as Editor of special issues of various journals such as *Frontiers in Environmental Science*, *Energies*, *Emerging Market Finance and Trade* and *Journal of Risk and Financial Management*. He has published about 250 articles in international peer-reviewed journals and 20 books from Springer, Routledge, World Scientific, etc.

List of Figures

Fig. 2.1	Mean and median illustration	8
Fig. 2.2	Distribution examples and their skewness	10
Fig. 2.3	Distribution examples and their kurtosis	11
Fig. 2.4	Natural gas and electricity futures prices	12
Fig. 2.5	Natural gas and electricity spot prices	13
Fig. 2.6	Spurious regression example (series plot)	15
Fig. 2.7	Spurious regression example (scatter plot)	15
Fig. 2.8	Trading model	21
Fig. 2.9	Power generation cost and long-term equilibrium relationship in the spot market	25
Fig. 2.10	Coefficients of the equilibrium equation for statistical arbitrage	27
Fig. 2.11	Time-series plots of the profit from each trade	28
Fig. 2.12	Time-series plots of the profit from each combination of the two trading strategies	29
Fig. 2.13	One-day 99% VaR and the expected shortfall in a normal distribution	30
Fig. 2.14	One-day 99% VaR and the expected shortfall in a normal distribution with a deformed left tail	31
Fig. 2.15	Procedure to measure portfolio risk using various copulas	42
Fig. 2.16	Degree of freedom and log-likelihood in estimating the t copula	43
Fig. 2.17	Scatter plots of random numbers following the bivariate Gaussian copula	45
Fig. 2.18	Scatter plots of random numbers following the bivariate t copula	45
Fig. 2.19	Scatter plots of random numbers following the bivariate Clayton copula	46
Fig. 2.20	Scatter plots of random numbers following the bivariate Gumbel copula	46

Fig. 2.21	Scatter plots of random numbers following the bivariate Frank copula	47
Fig. 2.22	Simplified diagram of an electric power business model	49
Fig. 3.1	Crude oil prices	56
Fig. 3.2	Natural gas prices	57
Fig. 3.3	Crude oil return series	64
Fig. 3.4	Spillover effects between return series (crude oil)	65
Fig. 3.5	Crude oil volatility series	68
Fig. 3.6	Spillover effects between volatility series (crude oil)	69
Fig. 3.7	Degree of freedom and log-likelihood, t copula (crude oil portfolio)	71
Fig. 3.8	Natural gas return series	74
Fig. 3.9	Spillover effects between return series (natural gas)	75
Fig. 3.10	Natural gas volatility series	77
Fig. 3.11	Spillover effects between volatility series (natural gas)	78
Fig. 3.12	Degree of freedom and log-likelihood in estimating the t copula (natural gas portfolio)	80
Fig. 4.1	HH spot and futures prices	88
Fig. 4.2	NBP spot and futures prices	89
Fig. 4.3	HH spot and futures return series	95
Fig. 4.4	Covariance series between futures and spot prices (HH)	96
Fig. 4.5	OHR calculated using the estimated diagonal VECH model for the HH	97
Fig. 4.6	OHR calculated using the estimated diagonal BEKK model for the HH	97
Fig. 4.7	OHR calculated using the estimated CCC model for the HH	98
Fig. 4.8	NBP spot and futures return series	100
Fig. 4.9	Covariance series between futures and spot prices (NBP)	101
Fig. 4.10	OHR calculated using the estimated diagonal VECH model for the NBP	101
Fig. 4.11	OHR calculated using the estimated diagonal BEKK model for the NBP	102
Fig. 4.12	OHR calculated using the estimated CCC model for the NBP	103
Fig. 5.1	Commodity price series, Europe	109
Fig. 5.2	Commodity return series, Europe	111
Fig. 5.3	Commodity volatility series, Europe	113
Fig. 5.4	Spillover effects between return series	115
Fig. 5.5	Spillover effects between volatility series	117
Fig. 5.6	Degree of freedom and log-likelihood in the t copula estimates (European commodity portfolio)	120
Fig. 6.1	LNG price formulas	125

Fig. 6.2	Impact of LNG procurement with a lower threshold price on the business balance	125
Fig. 6.3	Power price formulas	127
Fig. 6.4	Impact of power sales with an upper threshold price on the business balance	127

List of Tables

Table 2.1	Descriptive statistics	7
Table 2.2	ADF unit root tests	17
Table 2.3	Johansen cointegration test: Futures market	19
Table 2.4	Johansen cointegration test: Spot market	19
Table 2.5	Power generation business specifications	24
Table 2.6	Long-term equilibrium estimation	26
Table 2.7	Profit from each trade	27
Table 2.8	Profit from each combination of the two trading strategies	28
Table 2.9	Maximum unrealized loss and total profit from statistical arbitrage	29
Table 2.10	Parameters of each marginal distribution function	42
Table 2.11	Risk measurement using the estimated copulas	44
Table 3.1	Descriptive statistics (crude oil prices)	55
Table 3.2	Descriptive statistics (natural gas prices)	57
Table 3.3	Descriptive statistics (crude oil return series)	63
Table 3.4	ADF unit-root test results (crude oil return series)	64
Table 3.5	VAR model (crude oil return series)	65
Table 3.6	Residual variance–covariance matrix in the VAR model (crude oil return series)	65
Table 3.7	Spillover index and spectral analysis (crude oil return series)	66
Table 3.8	Estimated AR-EGARCH model (crude oil return series)	67
Table 3.9	Descriptive statistics (crude oil volatility series)	67
Table 3.10	ADF unit-root test results (crude oil volatility series)	68
Table 3.11	VAR model (crude oil volatility series)	68
Table 3.12	Residual variance–covariance matrix in the VAR model (crude oil volatility series)	69
Table 3.13	Spillover index and spectral analysis (crude oil volatilities) ...	70
Table 3.14	Parameters of each marginal distribution function (crude oil returns)	70

Table 3.15	Risk measurement using the estimated copulas (crude oil portfolio)	72
Table 3.16	Risk of each crude oil market	72
Table 3.17	Descriptive statistics (natural gas return series)	73
Table 3.18	ADF unit-root test results (natural gas return series)	74
Table 3.19	VAR model (natural gas returns)	74
Table 3.20	Residual variance–covariance matrix in the VAR model (natural gas returns)	74
Table 3.21	Spillover index and spectral analysis (natural gas returns)	75
Table 3.22	Estimated AR-EGARCH model (Natural gas return series)	76
Table 3.23	Descriptive statistics (natural gas volatility series)	77
Table 3.24	ADF unit-root test results (natural gas volatility series)	77
Table 3.25	VAR model (natural gas volatility series)	78
Table 3.26	Residual variance–covariance matrix in the VAR model (natural gas volatility series)	78
Table 3.27	Spillover index and spectral analysis (natural gas volatility series)	79
Table 3.28	Parameters of each marginal distribution function (natural gas return series)	79
Table 3.29	Risk measurement using the estimated copulas (natural gas portfolio)	81
Table 3.30	Risk of each natural gas market	81
Table 4.1	Descriptive statistics	88
Table 4.2	Descriptive statistics (HH return series)	94
Table 4.3	Estimated multivariate GARCH model (HH)	95
Table 4.4	Average OHR and HE (HH)	98
Table 4.5	Descriptive statistics (NBP return series)	99
Table 4.6	Estimated multivariate GARCH model (NBP)	100
Table 4.7	Average OHR and HE (NBP)	103
Table 5.1	Descriptive statistics (price series)	108
Table 5.2	Descriptive statistics (return series)	109
Table 5.3	ADF unit root test results (return series)	112
Table 5.4	Estimated AR-EGARCH model	112
Table 5.5	Descriptive statistics (volatility series)	112
Table 5.6	ADF unit root test results (volatility series)	114
Table 5.7	VAR model (return series)	114
Table 5.8	Residual variance–covariance matrix, VAR model (return series)	115
Table 5.9	Spillover index and spectral analysis (return series)	116
Table 5.10	VAR model (volatility series)	116
Table 5.11	Residual variance–covariance matrix, VAR model (volatility series)	117

Table 5.12	Spillover index and spectral analysis (volatility series)	118
Table 5.13	Parameters of each marginal distribution function (return series)	119
Table 5.14	Risk measurement using estimated copulas (European commodity portfolio)	121

Chapter 1

Preface



This book introduces empirical methods for researchers and graduate students to analyze energy markets empirically. In addition, it is helpful for those aiming to be energy traders. To provide beginners and practitioners with the tools necessary to analyze energy markets appropriately, it focuses on how to use analytical methods and interpret the results rather than a comprehensive explanation of theory. Moreover, it begins with basic statistics and econometrics, so it is understandable for aspiring undergraduate students. We next review the latest academic and discussion papers. We often see that explanations are omitted or insufficient, despite the analysis of relatively new studies, and that the established knowledge is used without mentioning the applicable conditions. Therefore, in many cases, we cannot keep up with the logic of the papers, and thus cannot correctly evaluate and utilize the findings. To prevent such problems, this book provides brief explanations of the procedures adopted.

Most of the latest empirical analysis approaches have been, and will continue to be, developed in studies of traditional financial markets. They made not only academic contributions to clarifying the truth, but also practical contributions to decision-making by governments and central banks, as well as the risk and fund management of financial institutions. Energy commodities include crude oil, natural gas, coal, and electricity. These have typical commodity characteristics. They are economic goods with full or partial, but substantial, fungibilities. They are treated as equivalents regardless of who produces them. Each price is determined by the function of the market as a whole. Their spot and derivative markets are well-established and highly liquid, and energy markets are already financialized.

Energy traders sell and/or buy non-differentiated commodities in highly liquid markets with the same aims as traditional financial traders. Power generation companies procure various fuels in highly liquid fuel markets, convert them into electricity, and sell the output in highly liquid electricity markets. Petroleum refiners buy crude oil in highly liquid crude oil markets, convert it to petroleum products such as gasoline, heating oil, and kerosene, and then sell these in highly liquid markets. Then, to maximize profits and minimize risk, they trade derivatives whose underlying assets

are the commodities that they buy and sell in highly liquid energy derivatives markets. Their operations are very similar to those of financial traders who raise funds in the market and manage financial securities on their own accounts and those of forex traders exchanging currencies in the forex market.

Unlike traditional financial securities, each type of energy certainly has its own unique properties (e.g., essentiality for the real economy, high storage costs, deterioration during storage, and a constant perfect match between supply and demand). However, if we consider these properties in the analysis or measurement, or when interpreting the results, it is possible to apply approaches and techniques from traditional financial markets to energy markets.

Just as empirical studies in traditional financial markets contribute both academically and to practical and industrial applications, studies of energy markets offer significant contributions. Certainly, many empirical studies already contributed to a wide range of fields. Therefore, we assume that neither lazy regulation nor loopy business management, which relies on thoughtless intuition and shallow experience, exist in the energy industry.

This book offers the opportunity to acquire advanced empirical approaches and techniques, and discusses some examples of their application to actual energy markets. Specifically, Chap. 2 presents two potential trading strategies that utilize the long-term equilibrium relationship between natural gas prices and electricity prices via simulation using historical data. Moreover, it measures the risk of a portfolio consisting of natural gas futures short positions and wholesale electricity futures long positions, which are held during the statistical arbitrage between gas and power futures. In this context, this book introduces descriptive statistics and the augmented Dickey-Fuller (ADF) unit root test developed using the approach proposed by Dickey and Fuller [4], the cointegration test proposed by Johansen [7], the dynamic ordinary least squares method to estimate the long-term equilibrium equation, arbitrage simulation, and the concept of copula first published by Sklar [9], which functionally expresses the connection of stochastic variables, value-at-risk (VaR), and expected shortfall as risk indicators. Chapter 3 measures the connectedness in the return and volatility for crude oil and natural gas between North American, European, and Asian markets. Furthermore, it spectrally decomposes the connectedness indicators. These results reveal the spillover effects between these regions and suggest the approximate potential of each fuel portfolio consisting of the three regional markets. Moreover, it estimates some types of copulas for each fuel and calculates the VaR and expected shortfall for each fuel portfolio. In this context, this book explains the vector autoregressive (VAR) model, vector moving average (VMA) model, the Diebold and Yilmaz [5] approach to measuring spillover effects, the Baruník and Křehlík [1] technique for spectral decomposition of Diebold and Yilmaz's [5] indicators based on the Fourier transform, and the exponential generalized autoregressive conditional heteroscedasticity (EGARCH) model by Nelson [8] to generate each volatility series. Chapter 4 estimates the dynamic correlation between spot returns and future returns in the natural gas markets in the United States and the United Kingdom. It then calculates the optimal hedge ratio, which is the future position ratio to the spot position,

to minimize portfolio risk. In this context, we introduce three multivariate generalized autoregressive conditional heteroscedasticity (GARCH) models: the constant conditional correlation (CCC) model by Bollerslev [2], diagonal VECM model by Bollerslev et al. [3], and diagonal BEKK model by Engle and Kroner [6]. Chapter 5 examines and considers the relationship between the returns and volatility of crude oil, natural gas, coal, electricity, and carbon credits in the European futures market. It then measures the market risk of a power generation business by interpreting that business as a portfolio consisting of long positions in fuel futures and short positions in electricity futures. Chapter 6, as an alternative to postface, introduces two transaction cases that include market risk transfers. We break down the energy prices into the energy and risk values by interpreting the energy transaction contract from the business balance.

References

1. Baruník, J., & Křehlík, T. (2018). Measuring the frequency dynamics of financial connectedness and systemic risk. *Journal of Financial Econometrics*, 16(2), 271–296.
2. Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model. *Review of Economics and Statistics*, 72(3), 498–505.
3. Bollerslev, T., Engle, R. F., & Wooldridge, J. M. (1998). A capital asset pricing model with time-varying covariances. *Journal of Political Economy*, 96(1), 116–131.
4. Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366a), 427–431.
5. Diebold, F. X., & Yilmaz, K. (2012). Better to give than to receive: Predictive directional measurement of volatility spillovers. *International Journal of Forecasting*, 28(1), 57–66.
6. Engle, R. F., & Kroner, K. F. (1995). Multivariate simultaneous generalized ARCH. *Econometric Theory*, 11(1), 122–150.
7. Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12(2–3), 231–254.
8. Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59(2), 347–370.
9. Sklar, A. (1959). *Distribution functions of n dimensions and margins* (pp. 229–231). Publications of the Institute of Statistics at the University of Paris.

Chapter 2

Arbitrage Trading in Energy Markets and Measuring Its Risk



2.1 Introduction

Power generation companies procure fuel, process it into electricity, and sell the electricity unless they use fuel-free power generation methods such as hydropower, solar power, and wind power.

Typical fuels include coal, natural gas, petroleum gas, crude and refined products, and uranium. Many of their spot and financial derivatives are bought and sold through commodity exchanges. Furthermore, even unlisted commodities and derivatives related to their listed commodities and securities are traded at prices linked to their spot prices. Fuel spots are often traded on long-term fixed-price contracts to promote investment in each upstream fuel business. However, we can interpret these transaction prices as being linked to their related markets from a long-term perspective. On the other hand, various types of electricity in each region are traded, while many derivatives whose underlying assets are electric power are listed on exchanges. Like fuel, electricity is also traded as long-term fixed-price contracts; however, it is extremely common to buy and sell electricity at market prices or their linked prices.

The various fuels power generation companies procure and the electricity they generate are commodities that cannot be differentiated in terms besides their prices. Moreover, these markets, including their derivatives markets, are liquid. Therefore, both fuel and power prices can be considered efficiently formed. In other words, no one can buy fuel continuously at a lower price than others can, regardless of how much they investigate the fuel market. Similarly, no one can sell electricity continuously at a higher price than others can, regardless of how much they examine the power market. Of course, a power generation company invests in a fuel production and/or distribution business to ensure superior fuel procurement relative to other companies. However, the investment is not considered to be part of the power generation business.

The power generation business can only be differentiated by how efficiently it converts fuel to electricity. However, the efficiency of a power generation facility depends on the plant manufacturer and not on the power generation operator.

Certainly, if a firm sells so much electricity that it requires many generators, it can aim to build a generator portfolio that minimizes the expected future cash output by considering the fixed and variable costs of various power generation methods. Regretfully, few businesses in one market can take such measures. In addition, rational maintenance and an efficient organizational structure can reduce operating costs depending on the reliability of equipment. However, the effect is not large and other firms easily catch up. Unfortunately, the price of commoditized goods generally approaches their marginal cost, so the power generation business tends to have low profits, despite the burden of large-scale investment risk. If a power generation company has only inferior production equipment relative to the market, it could go out of business. Given these conditions, it is difficult for power companies that earn gross profit only by the difference between the procurement price of fuel and the selling price of electricity to improve their profit margins unless they take advantage of this price difference.

Therefore, in this chapter, we propose arbitrage between the spot and futures price differences and statistical arbitrage between fuel and power prices as trading strategies that utilize the unit price difference between natural gas and electricity. Moreover, our simulation with actual historical data shows that even if a company owns an inferior power generator, it might be able to profit by adopting these arbitrage strategies. We then estimate the five types of copulas to grasp the joint distribution of gas and power price returns. The simulation with random numbers that follow those copulas measures the risk of a portfolio consisting of gas futures short positions and power futures long positions, which are held during statistical arbitrage.

Many studies analyze arbitrage. Alexakis [1] discusses the implications of statistical arbitrage strategies based on the cointegration relationship between global stock indexes. Mayordomo et al. [25] study statistical arbitrage strategies in credit derivative markets. Focardi et al. [10] introduce a new arbitrage approach based on dynamic factor models of prices rather than returns, and demonstrate their performance in the stock market. Hain et al. [15] examine the profits of a cointegration-based statistical arbitrage strategy in European energy markets and confirm the statistically significant risk-adjusted excess returns. Baviera and Baldi [2] introduce a new statistical arbitrage strategy with stop-loss and leverage in high-frequency trading and apply the method to the spread on heating oil and gas oil futures. Liu and Su [22] analyze the dynamic causality between gold and silver returns and provide the implications for statistical arbitrage strategies. Nakajima [26] investigates statistical arbitrage using the cointegration relationship between wholesale electricity futures and natural gas futures. Stübinger and Schneider [35] propose an integrated statistical arbitrage strategy based on overnight price gaps and applied it to high-frequency representative stock data. Sánchez-Granero et al. [32] propose a novel approach based on a statistical arbitrage technique to test the efficiency of the Latin American stock markets. Keilbar and Zhang [19] result a great performance by a simple statistical arbitrage trading strategy using the cointegration spreads between cryptocurrencies.

The remainder of this chapter is organized as follows. Section 2.2 explains the data we use in this chapter. Section 2.3 proposes the two trading strategies. Section 2.4

provides the simulation results. Section 2.5 measures the risk in statistical arbitrage. Section 2.6 provides the overall concluding remarks and considerations.

2.2 Data and Preliminary Analyses

Here, we employ the PJM Western Hub Real-Time Peak to represent wholesale power and Henry Hub Natural Gas to represent natural gas. We use futures prices from January 2, 2015 to December 30, 2020 and spot prices from January 2, 2015 to January 29, 2021. Henry Hub and PJM prices are expressed in USD per mmbtu and per MWh, respectively. We obtain these daily data from Bloomberg. Henry Hub and PJM are among the most representative natural gas and wholesale electricity sources in the United States. In addition, we use the most recent daily data.

2.2.1 Descriptive Statistics

Before conducting various analyses and simulations, it is extremely important to interpret the representative statistics of the data. Table 2.1 provides the summary statistics of the Henry Hub and the PJM.

Considering that each future has a maturity of one month, we set each spot price to January 29, 2021 and each future to December 30, 2020 to simulate the spot-future arbitrage described later in Sect. 2.3.1. Because we extract only the days when both

Table 2.1 Descriptive statistics

		Henry Hub		PJM Western Hub peak	
Period	From	Futures	Spot	Futures	Spot
	To	12/30/2020	1/29/2021	12/30/2020	1/29/2021
Observations		1511	1477	1511	1477
Mean		2.65	2.64	36.49	35.18
Median		2.72	2.71	35.10	32.50
Maximum		4.84	7.13	87.60	366.91
Minimum		1.48	1.33	19.20	15.63
Standard deviation		0.50	0.56	8.75	17.24
Skewness		0.26	0.93	1.75	8.48
Kurtosis		4.26	8.84	8.95	124.26
Jarque–Bera		118 (0.00)	2313 (0.00)	3002 (0.00)	922,552 (0.00)

Note Values in parentheses indicate p -values

the Henry Hub and PJM data are available, we have 1511 and 1477 observations for the futures and spot prices, respectively.

The mean and median are numerical values located in the center of the economic variables. The mean \bar{x} of the series $(x_i | i = 1, 2, \dots, N)$ is calculated as

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i. \quad (2.1)$$

On the other hand, the median is a value located in the center of each series arranged in descending order. The medians of these futures and spot series are at the 756th and 739th values, respectively. If the number of observations is even, then the median is the average of the two data points in the center. Thus, the median is a more stable index expressing the middle than the mean because outlier values have less effect. Figure 2.1 shows three distribution examples with the same mean, but different medians. Table 2.1 indicates that both the mean and median of each future are higher than those of each spot. In other words, both Henry Hub and PJM tend to be contango. We can infer that the supply and demand are not very tight during this period. We can express the relationship between the future price p_f and its spot price p_s as

$$p_f = p_s e^{C_c \Delta T}, \quad (2.2)$$

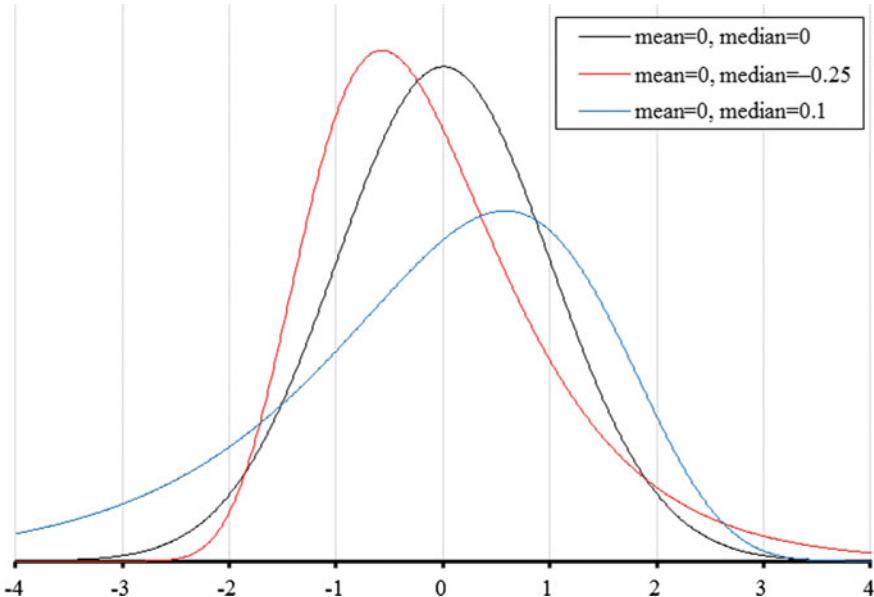


Fig. 2.1 Mean and median illustration

where C_c is the cost of carry expressed in terms of yield and ΔT is the period from the present to maturity. The cost of carry is the sum of the risk-free interest rate and holding cost, expressed as yield minus the convenience yield. Therefore, if their supply and demand remained tight during the period, then the utility of holding their spots would be increasing. Thus, their costs of carry should become negative, and their futures should become lower than their spots. In addition, the medians of both the Henry Hub future and spot prices are higher than their respective means. Therefore, we can expect to find many outliers in the left tail of each distribution. On the contrary, the medians of both the PJM future and spot prices are lower than their respective means. Therefore, we can expect to find many outliers in the right tail of each distribution.

The next topic is the dispersion of each variable, which indicates the spread of values from the mean. The simplest indicator is the range, calculated by subtracting the minimum value from the maximum value. The ranges of the PJM and Henry Hub futures are narrower than the range of each spot price. Furthermore, for all four variables, the difference between the maximum and mean is much larger than that between the minimum and mean. This indicates a long distribution in the right tail.

As mentioned above, the median provides an interpretation of dispersion by comparing it with the mean. The percentile is a statistic developed based on the median concept. The α -th percentile is the value located at the first α percentage in ascending order. For example, the 65th percentile is the value located at 65% of the sample size, counting from the minimum value. The 25th, 50th, and 75th percentiles are called the lower quartile, median, and upper quartile, respectively. The percentile is an important concept in risk measurement; however, we omit this statistic in Table 2.1, because we discuss risk measurement in Sect. 2.5.

One statistic that can express the dispersion of a distribution is the mean absolute deviation, calculated as

$$\text{Mean absolute deviation} = \frac{1}{N} \sum_{i=1}^N |x_i - \bar{x}|. \quad (2.3)$$

However, this statistic is rarely used, so we omit it from Table 2.1.

The most representative statistic for dispersion is the standard deviation σ , calculated using the following formula as the corrected sample standard deviation:

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}. \quad (2.4)$$

Table 2.1 shows that the standard deviation of each spot is larger than that of each future. It turns out that the dispersion for their spots are larger than that of their futures. The main reason is that it is technically difficult or too expensive to store natural gas and power.

The skewness μ_3 is a measure of the asymmetry of the distribution about its mean, defined as

$$\mu_3 = \frac{1}{N} \sum_{i=1}^N \left(\frac{x_i - \bar{x}}{\hat{\sigma}} \right)^3, \quad (2.5)$$

where $\hat{\sigma}$ is the standard deviation of the observations, calculated using the following equation:

$$\hat{\sigma} = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2}. \quad (2.6)$$

The $\hat{\sigma}$ includes bias because it is calculated using the estimated mean \bar{x} . The standard deviation corrected for this bias is the σ defined in Eq. (2.4). However, there is little difference between σ and $\hat{\sigma}$ in practice because we often have an extremely large set of observations. If a distribution is symmetric, its skewness is 0. Distributions with a long right tail have positive skewness, whereas those with a long left tail have negative skewness. Figure 2.2 plots three distribution samples with the same mean but different skewness values. Table 2.1 indicates that all variables have positive skewness, meaning that the distribution has a long right tail. In particular, the distortion to the right of the PJM spot distribution is clearly large.

The kurtosis μ_4 is a measure of the sharpness and flatness of the distribution, defined as

$$\mu_4 = \frac{1}{N} \sum_{i=1}^N \left(\frac{x_i - \bar{x}}{\hat{\sigma}} \right)^4. \quad (2.7)$$

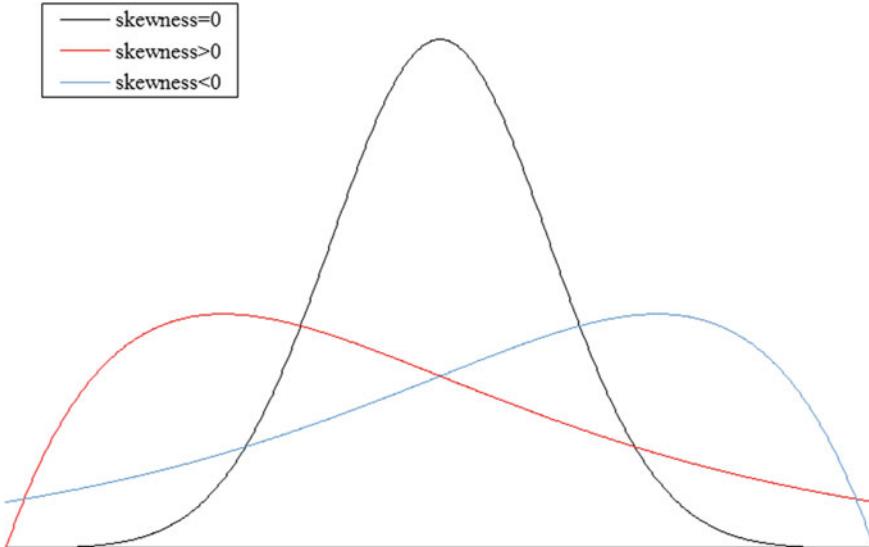


Fig. 2.2 Distribution examples and their skewness

The kurtosis of the normal distribution is 3. If a distribution has a sharp peak and long fat-tails, then its kurtosis exceeds 3. Conversely, if a distribution has a rounded peak and short lean tails, its kurtosis is less than 3. Figure 2.3 plots three distribution samples with the same mean but different kurtosis values. Table 2.1 indicates that the kurtosis of each series is larger than 3. Each variable had its own long fat-tail and a peak distribution. In particular, the PJM spot price series is remarkable.

Jarque and Bera's [16] test is a goodness-of-fit test for whether the skewness and kurtosis of a distribution follow a normal distribution. In the null hypothesis that observations are normally distributed, the Jarque–Bera statistic, defined by the following equation, asymptotically has an χ^2 distribution with degree of freedom 2.

$$\text{Jarque-Bera} = \frac{N}{6} \left\{ \mu_3^2 + \frac{1}{4} (\mu_4 - 3)^2 \right\}. \quad (2.8)$$

The Jarque–Bera statistics in Table 2.1 reject the normal distribution hypothesis for all the variables. Electric power prices are more fat-tailed than gas prices and their spot prices are more fat-tailed than their futures prices.

Finally, Figs. 2.4 and 2.5 provide the time plots of these future prices and spot prices, respectively. Each variable has properties estimated from Table 2.1. We can see that both Henry Hub and PJM spot prices fluctuate in the same way as their respective future prices. However, these spot prices spike frequently, and this tendency is remarkable, especially for the PJM. In energy spot markets, price spikes and jumps often occur because of a sudden increase in demand and shortage in supply capacity. This is because equal amounts of supply and demand are simultaneously required.

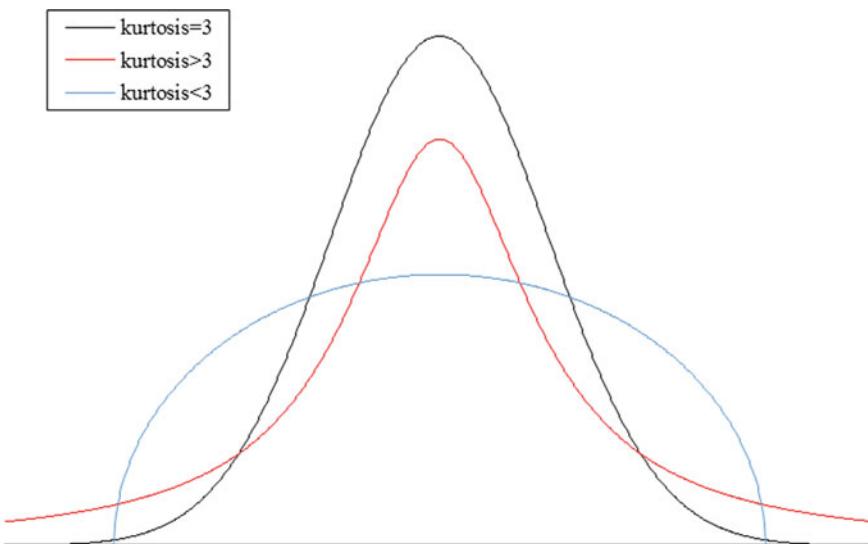


Fig. 2.3 Distribution examples and their kurtosis

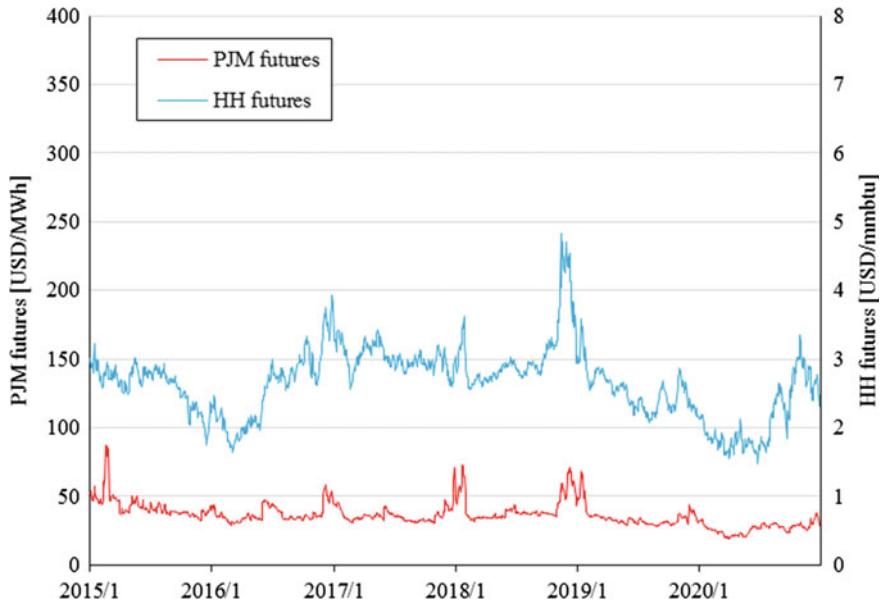


Fig. 2.4 Natural gas and electricity futures prices

Unfortunately, it is extremely difficult or virtually impossible to store natural gas and electricity. Moreover, we note that there may be a certain relationship between the Henry Hub prices and the PJM prices. We can observe that the prices of gas and power are synchronized, regardless of their spots and futures.

2.2.2 *Stationarity and Unit Root Test*

We must not forget the concept of “stationarity” in analyzing the time-series data of economic variables. A stochastic process $\{x_t\}$ is defined as a stationary process if it satisfies the following three conditions:

Condition 1. The expected value $E(x_t)$ does not depend on time t . In other words, all expected values are equal at any t .

Condition 2. The variance $V(x_t)$ does not depend on time t . In other words, all variances are equal at any t .

Condition 3. The autocovariance $Cov(x_t, x_{t-i})$ does not depend on time t , but only on time difference $i (> 0)$.

As the most representative stationary process, we introduce white noise, which satisfies the following three conditions:

Condition 1. The expected value $E(x_t)$ is 0 at any t .

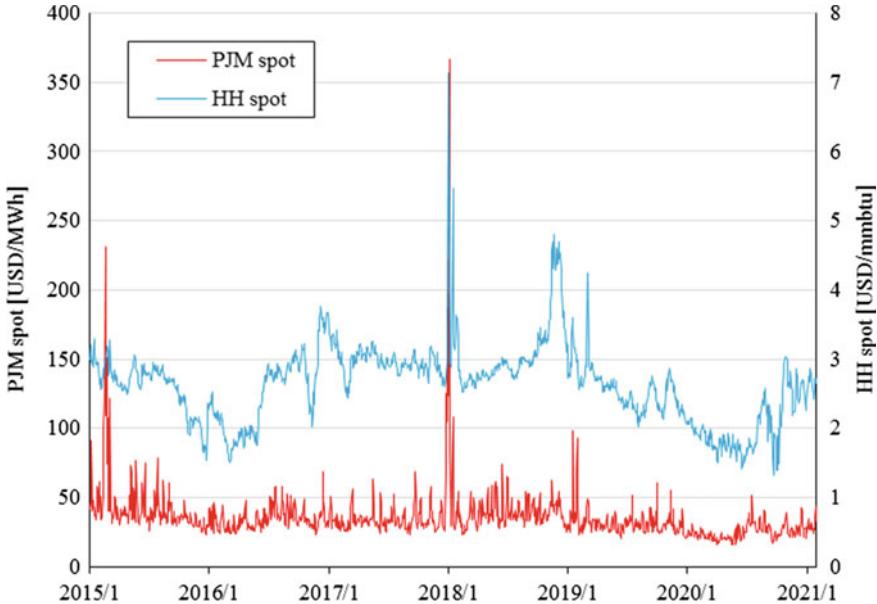


Fig. 2.5 Natural gas and electricity spot prices

Condition 2. The variance $V(x_t)$ is a constant at any t .

Condition 3. The autocovariance $\text{Cov}(x_t, x_{t-i})$ is 0 if i is not 0.

Most market price series that energy market analysts frequently deal with are nonstationary processes, like many other economic variable series.

To consider the characteristics of stochastic processes, suppose the p -th order autoregression process of the following equation.

$$x_t = \sum_{i=1}^p \theta_i x_{t-i} + u_t, \quad (2.9)$$

where u_t is the white noise. We can express the characteristic polynomial of Eq. (2.9) as

$$m^p - m^{p-1}\theta_1 - m^{p-2}\theta_2 - \cdots - \theta_p = 0. \quad (2.10)$$

If one of the solutions of this equation is less than one, then this $\{x_t\}$ is a stationary process. If one of the roots is one, then $\{x_t\}$ is a unit root process. If one root is larger than 1, then $\{x_t\}$ is called an explosive process. Both unit root and explosive processes are nonstationary processes.

Next, we discuss spurious regressions as one of the most important points to keep in mind when analyzing a unit root process. Suppose two unit root processes x_t and y_t that are independent of each other. We estimate the following regression model

using ordinary least squares (OLS).

$$y_t = \alpha + \beta x_t + \varepsilon_t. \quad (2.11)$$

Theoretically, x_t and y_t are irrelevant. Therefore, β should be 0. However, the estimated model has the following characteristics.

Characteristic 1. The correlation coefficient is large, meaning that the model fits well.

Characteristic 2. The t -value is large, meaning that β is significant.

Characteristic 3. The Durbin-Watson ratio is small, meaning that ε_t has a positive autocorrelation.

Granger and Newbold [13] discovered the phenomenon called “spurious regression” using Monte Carlo simulations. Phillips [29] proves this analytically. We must study economic variables with great care because most of them have unit roots. We provide an example below. We generate 100,000 data points for each series using the following formula:

$$\begin{aligned} x_t &= x_{t-1} + \varepsilon_{xt} \\ y_t &= y_{t-1} + \varepsilon_{yt}, \end{aligned} \quad (2.12)$$

where $x_1 = y_1 = 0$, and ε_{xt} and ε_{yt} are random numbers that follow an independent standard normal distribution. Figures 2.6 and 2.7 show the time series and scatter plots, respectively. Although x_t and y_t have no theoretical relationship, they seem to be in harmony. The estimated result of Eq. (2.11) using OLS is

$$\begin{aligned} \text{Correlation coefficient} &= 0.903 \\ \beta &= -144 (t\text{-value} = -518, p\text{-value} = 0) \\ \text{Durbin-Watson ratio} &= 0.000 \end{aligned} \quad (2.13)$$

This result thus represents a spurious regression.

The augmented Dickey–Fuller (ADF) unit root test is based on the approach proposed by Dickey and Fuller [5]. The ADF test is one of the most representative unit root tests, with a unit root null hypothesis and stationary alternative hypothesis. We transform Eq. (2.9) into

$$\Delta x_t = \xi x_{t-1} + \sum_{i=1}^p \eta_i \Delta x_{t-i} + u_t. \quad (2.14)$$

The null hypothesis is that the unit root process is $\xi = 0$, and the alternative hypothesis is that the stationary process is $\xi < 0$. In the null hypothesis, the t -value of ξ does not have a Student’s t distribution, but has a specific distribution known as the Dickey–Fuller (DF) table. This table tests the unit root process. A common

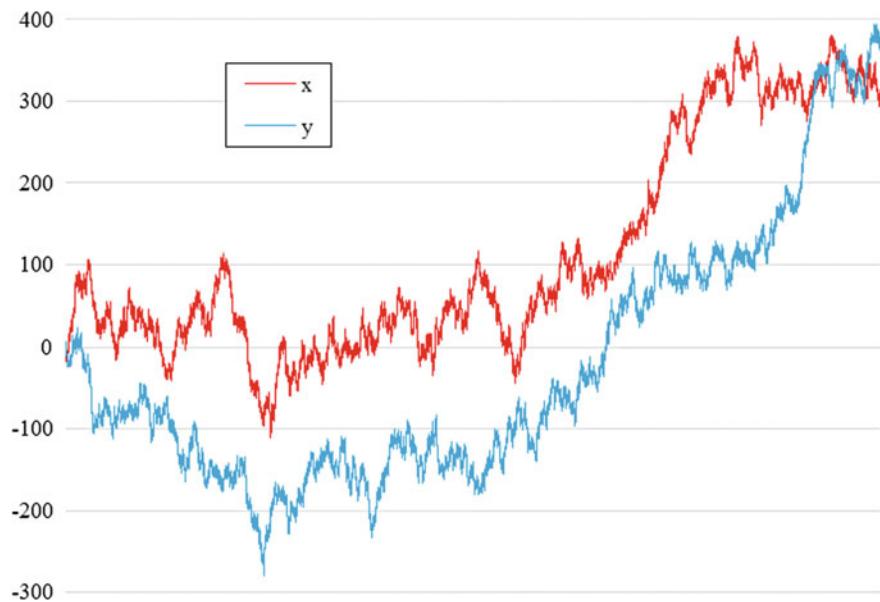


Fig. 2.6 Spurious regression example (series plot)

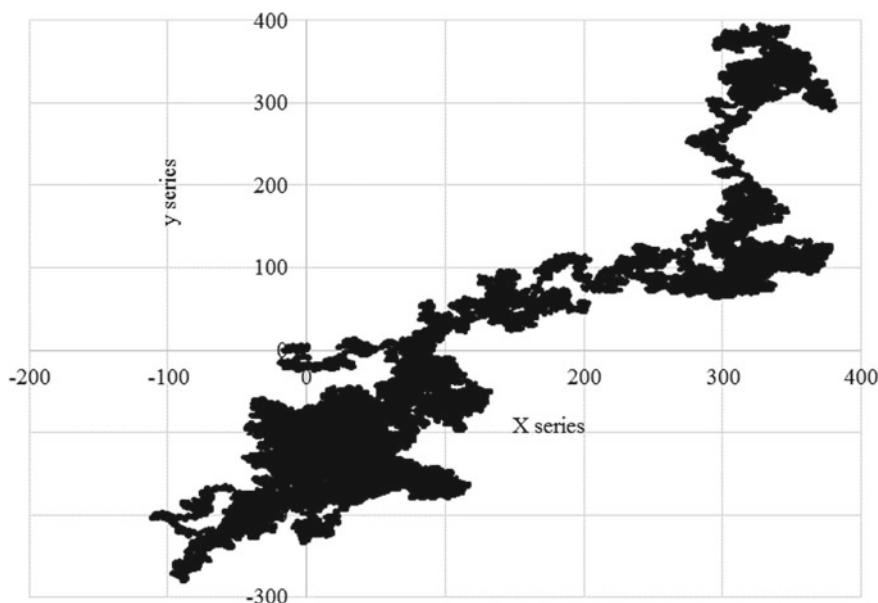


Fig. 2.7 Spurious regression example (scatter plot)

assumption is that a constant term is added to Eq. (2.9). Furthermore, it is also frequently supposed that a time trend term is added. Each model has the following equations:

$$\Delta x_t = \delta + \xi x_{t-1} + \sum_{i=1}^p \eta_i \Delta x_{t-i} + u_t \quad (2.15)$$

$$\Delta x_t = \delta + \zeta t + \xi x_{t-1} + \sum_{i=1}^p \eta_i \Delta x_{t-i} + u_t, \quad (2.16)$$

where δ is a constant term and ζt is a time trend term. Regarding the selection of autoregression lag order p , it is reasonable to extend the order until the test for the residual term no longer accepts autocorrelation. The possibility of autocorrelation of u_t extends the model. However, we often simply select the order based on information criteria such as the Akaike information criterion (AIC), Schwarz Bayesian information criterion (SBIC), and Hannan-Quinn information criterion. In addition, we can select the order by referring to previous studies on the same economic variable. The Dickey and Fuller [5] test is an ADF test that uses a model without lagged terms.

Phillips et al. [31] propose an ADF test for rational bubbles using the right tail of the DF distribution. In other words, the unit root process null hypothesis remains $\xi = 0$, whereas the explosive process alternative hypothesis is $\xi > 0$.

Other typical unit root tests include the PP test developed by Phillips and Perron [30], and the KPSS test by Kwiatkowski et al. [20]. The PP test, similar to the ADF test, is based on a unit root null hypothesis and a stationary alternative hypothesis, whereas the KPSS test presupposes a stationary null hypothesis and a unit root alternative hypothesis. We should select the test type in consideration of the power and purpose of the unit root test.

Table 2.2 shows the ADF unit root test results for each energy price and its first difference. Equation (2.14) is the model adopted. For all variables, their levels accept the unit root hypothesis and their first difference rejects the null hypothesis. In general, most price series are unit root processes and their first difference series are stationary processes. These results are consistent with previous studies.

2.2.3 Cointegration Test

Figures 2.1 and 2.2 bring to mind the long-term equilibrium relationship between Henry Hub and the PJM in both futures and spot markets. However, as all four variables accept the unit root hypothesis, we must suspect a spurious regression.

Engle and Granger [7] introduced the concept of “cointegration,” which connects multiple unpredictable stochastic variables with a unit root. If a linear combination of multiple unit root processes is stationary, then these variables have a cointegrated

Table 2.2 ADF unit root tests

Commodity	Variable	ADF- <i>t</i> statistic
Henry Hub futures	Level	-0.80 (0.37)
	First difference	-30.48* (0.00)
PJM Western Hub peak futures	Level	-1.34 (0.17)
	First difference	-18.83* (0.00)
Henry Hub spot	Level	-0.88 (0.34)
	First difference	-13.75* (0.00)
PJM Western Hub peak spot	Level	-2.46 (0.01)
	First difference	-19.43* (0.00)

Note * indicates that the unit root hypothesis is rejected at the 1% significance level. The values in parentheses indicate *p*-values

relationship. In other words: suppose that the following vector consists of v variables in a unit root process:

$$\mathbf{X}_t = {}^T(x_{1t}, x_{2t}, \dots, x_{vt}). \quad (2.17)$$

The following linear combination is derived from the inner product of the v dimensional coefficient vector and \mathbf{X}_t :

$$\boldsymbol{\beta} \mathbf{X}_t = (\beta_1, \beta_2, \dots, \beta_v) {}^T(x_{1t}, x_{2t}, \dots, x_{vt}). \quad (2.18)$$

If $\boldsymbol{\beta} \mathbf{X}_t$ is a stationary process, then $x_{1t}, x_{2t}, \dots, x_{vt}$ have a cointegrated relationship. Additionally,

$$\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_v) \quad (2.19)$$

is the cointegrating vector. If there is cointegration between some variables, then the deviation of the observed values from their long-term equilibrium is a stable stochastic process. Because many economic variables have unit roots, this concept is very often applied in a wide range of fields to examine the relationships between economic variables.

Therefore, we test whether the Henry Hub and PJM prices are cointegrated and expect to use this cointegrated relationship in the trading strategies.

Engle and Granger's [7] proposed test for cointegration has limitations. First, it does not expect a system with three or more variables to have two or more cointegration relationships. Second, the test results may change when the variables are interchanged.

Section 2.2.3 adopts the test proposed by Johansen [17] and Johansen and Juselius [18]. We assume the following model:

$$\Delta \mathbf{X}_t = \mathbf{A} \mathbf{X}_{t-1} + \sum_{i=1}^{p-1} \mathbf{B}_i \Delta \mathbf{X}_{t-i} + \mathbf{u}_t. \quad (2.20)$$

Because the cointegration condition is that Eq. (2.18) is a stationary process, we can add a constant term vector or time trend vector to Eq. (2.20). The rank of \mathbf{A} represents the number of independent cointegrating vectors; therefore, we have the following three cases:

Case 1. $\text{rank}(\mathbf{A}) = v$: All elements of \mathbf{X}_t are a stationary process.

Case 2. $\text{rank}(\mathbf{A}) = \rho (0 < \rho < v)$: There are ρ cointegrating vectors.

Case 3. $\text{rank}(\mathbf{A}) = 0$: $\beta \mathbf{X}_t$ is a nonstationary process.

Therefore, after estimating Eq. (2.20), we perform the cointegration test based on the eigenvalues of \mathbf{A} .

Section 2.2.3 tests for the relationship between Henry Hub prices and PJM prices in the futures and spot markets. Let the eigenvalues of \mathbf{A} be $k_1, k_2 (k_1 > k_2)$. The acceptance of the unit root hypotheses should not hold $\text{rank}(\mathbf{A}) = 2$. Therefore, if there is a cointegrating vector; that is, if $\text{rank}(\mathbf{A}) = 1$, then the following equations hold:

$$\begin{aligned} \log(1 - k_1) &< 0 \\ \log(1 - k_2) &= 0. \end{aligned} \quad (2.21)$$

Alternatively, if there is no cointegrating vector; that is, if $\text{rank}(\mathbf{A}) = 0$, then

$$\begin{aligned} \log(1 - k_1) &= 0 \\ \log(1 - k_2) &= 0. \end{aligned} \quad (2.22)$$

In general, the null hypothesis of the trace test is the existence of ρ or fewer cointegrating vectors, and the alternative hypothesis is that there are more than ρ cointegrating vectors. The test statistic k_{trace} is

$$k_{trace} = -T \sum_{i=\rho+1}^v \log(1 - \hat{k}_i), \quad (2.23)$$

where \hat{k}_i is the estimated eigenvalue and T is the number of observations. We should test the relationship for each value of $\rho = 0, 1, \dots, v - 1$. Since Sect.2.2.3 adopts a two-variable model, the null hypothesis is one or fewer cointegrating vectors, and the alternative hypothesis is two or more cointegrating vectors. The statistics $k_{trace,v=2,\rho=0}$ and $k_{trace,v=2,\rho=1}$ are

$$k_{trace,v=2,\rho=0} = -T \log(1 - \hat{k}_1) - T \log(1 - \hat{k}_2)$$

Table 2.3 Johansen cointegration test: Futures market

Hypothesized number of coefficients	Trace test	Maximum eigenvalue test
None	31.18* (0.00)	30.72* (0.00)
One	0.47 (0.56)	0.47 (0.56)

Note * indicates rejection of the hypothesis at the 1% significance level. The values in parentheses indicate p -values

$$k_{trace,v=2,\rho=1} = -T \log(1 - \hat{k}_2). \quad (2.24)$$

In general, the null hypothesis of the maximum eigenvalue test is the existence of ρ cointegrating vectors, and the alternative hypothesis is that $\rho + 1$ cointegrating vectors exist. The test statistic k_{max} is

$$k_{max} = -T \log(1 - \hat{k}_{\rho+1}). \quad (2.25)$$

Since Sect. 2.2.3 uses a two-variable model, the null hypothesis is one cointegrating vector, and the alternative hypothesis is two cointegrating vectors. The statistics $k_{max,v=2,\rho=0}$ and $k_{max,v=2,\rho=1}$ are

$$\begin{aligned} k_{max,v=2,\rho=0} &= -T \log(1 - \hat{k}_1) \\ k_{max,v=2,\rho=1} &= -T \log(1 - \hat{k}_2). \end{aligned} \quad (2.26)$$

Tables 2.3 and 2.4 present the Johansen cointegration test results for Henry Hub prices and PJM prices in the futures and spot markets, respectively. Both the trace test and the maximum eigenvalue test reject the hypothesis of no cointegrating vector and accept the hypothesis of one cointegrating vector. We have the same result in both the futures and spot markets. We can conclude that there is a cointegration relationship between Henry Hub prices and PJM prices.

Table 2.4 Johansen cointegration test: Spot market

Hypothesized number of coefficients	Trace test	Maximum eigenvalue test
None	147.97* (0.00)	146.79* (0.00)
One	1.18 (0.32)	1.18 (0.32)

Note * indicates rejection of the hypothesis at the 1% significance level. The values in parentheses indicate p -values

2.2.4 Long-Term Equilibrium Estimation

We can estimate the cointegrating vectors by using dynamic OLS (DOLS). OLS estimates the following equation with lag terms for the explanatory variables to eliminate autocorrelation:

$$x_{v,t} = \varphi_0 + \sum_{i=1}^{v-1} \left(\beta_i \varphi_{i,t} + \sum_{j=-K}^K \phi_{i,j} \Delta x_{i,t-j} \right). \quad (2.27)$$

Since Sect. 2.2.4 utilizes a two-variable model, the model for estimating the long-term equilibrium is

$$PJM_t = \varphi_0 + \varphi_1 HenryHub_t + \sum_{j=-K}^K \phi_j \Delta HenryHub_{t-j}. \quad (2.28)$$

The lag order K was determined using SBIC. The long-term equilibrium equation for future prices is

$$PJM_{future} = 10.673 \times HenryHub_{future} + 8.159. \quad (2.29)$$

The long-term equilibrium equation for the spot prices is

$$PJM_{spot} = 11.142 \times HenryHub_{spot} + 5.732. \quad (2.30)$$

2.3 Trading Strategies

The only way to profit by trading goods is to “buy at a lower price and sell at a higher price.” If we trade only one item, then price forecasting is the most important matter. Is this realistically possible? A market is efficient if the information that affects the market price is comprehensive, constant, and has a timely effect on the price. Markets for securities and commodities listed on exchanges are almost efficient and depend on liquidity. In other words, we cannot forecast the price because the price already reflects all the currently available information, and any information that affects the price will occur independently of the price. Unfortunately, it is impossible for market participants to earn returns above the market average. Certainly, a “fully efficient market” is theoretical or virtual. Therefore, some investors and speculators try to collect information before it is reflected in the price. However, these actions make the market more efficient. Because the stationary hypothesis for most energy prices is rejected by the unit root test using daily data, energy companies should consider energy markets as efficient, and energy prices as unpredictable.

In general, power companies procure various types of fuels from various markets, produce electricity using various power generation methods, and sell the power through various sales channels. Section 2.3 assumes a simple model of purchasing natural gas at the Henry Hub price and selling electricity at the PJM price, as Fig. 2.8 illustrates. We propose two trading strategies. Section 2.4 will simulate these methods using actual historical data. Both focus not on these prices but on the price difference between Henry Hub prices and PJM prices. We cannot expect profit owing to market efficiency, even if we analyze each price in detail. On the other hand, we demonstrate the potential to make a profit by investigating price differences, which is a stationary process. When buying the gas required to produce one unit of electricity and selling it, the gross margin is often called the spark spread.

The trading strategy introduced in Sect. 2.3.1 is the arbitrage between the futures market spreads and a company's spreads expected from its power generation efficiency. This takes advantage of the spread of futures as a stochastic process. All we have to do is take the Henry Hub long position and the PJM short position to secure profits when a favorable futures spread occurs stochastically. The strategy proposed in Sect. 2.3.2 is statistical arbitrage utilizing the cointegration relationship between Henry Hub prices and PJM prices in the futures market. Making use of the long-term equilibrium equation in the futures market that expresses the futures spread, the lower PJM long positions and the higher Henry Hub short positions are expected to yield profit in the narrower spreads than the market when the spread approaches the long-term equilibrium. Of course, the higher PJM short positions and lower Henry Hub long positions, which are taken in wider spreads than the market, are expected

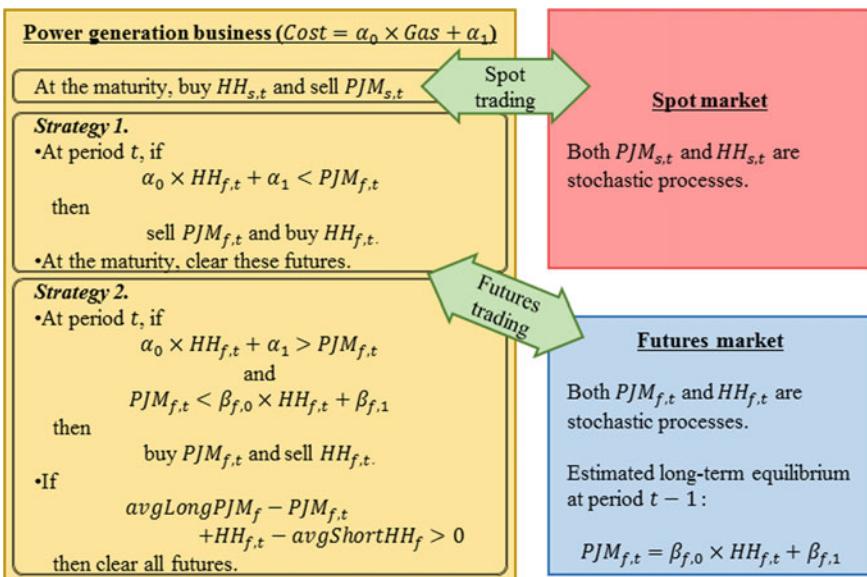


Fig. 2.8 Trading model

to make a profit. However, this is almost always included in the arbitrage between futures spreads and the company's spot spreads. Therefore, we did not simulate this method.

2.3.1 Arbitrage Between Own Spot Spread and Future Spread

Power generation costs consist of fixed costs (e.g., equipment depreciation costs, labor costs, and maintenance costs) and variable costs (e.g., fuel costs and exhaust processing costs). However, the variable cost is almost the fuel cost. Therefore, the unit cost of gas-fired generation expressed as *Cost* and the corresponding natural gas procurement cost expressed as *Gas* satisfy the following equation:

$$\text{Cost} = \alpha_0 \times \text{Gas} + \alpha_1, \quad (2.31)$$

where α_0 and α_1 are the coefficients. Although the Henry Hub futures price $\text{HenryHub}_{f,t}$ and the PJM futures price $\text{PJM}_{f,t}$, both of which are unit root processes, are cointegrated, the long-term equilibrium equation, which is a stochastic process, can have large outliers. Then, when the own spot spread, that is, the difference between the power generation unit cost and the corresponding gas procurement unit price, is smaller than the future spread, that is, the future price difference between the PJM and Henry Hub, we swap the spot spread, Spread_s and the future spread, Spread_f , which we express as

$$\text{Spread}_s = \alpha_0 \times \text{HenryHub}_{f,t} + \alpha_1 - \text{HenryHub}_{f,t} \quad (2.32)$$

$$\text{Spread}_f = \text{PJM}_{f,t} - \text{HenryHub}_{f,t}. \quad (2.33)$$

Therefore, the difference between these spreads is

$$\text{Spread}_s - \text{Spread}_f = \alpha_0 \times \text{HenryHub}_{f,t} + \alpha_1 - \text{PJM}_{f,t}. \quad (2.34)$$

In the following equation:

$$\alpha_0 \times \text{HenryHub}_{f,t} + \alpha_1 - \text{PJM}_{f,t} < 0. \quad (2.35)$$

If we take the Henry Hub long position and the PJM short position corresponding to the electric energy planned for generation, we can lock in profit.

2.3.2 Statistical Arbitrage

By estimating the long-term equilibrium equation of $\text{HenryHub}_{f,t}$ and $\text{PJM}_{f,t}$ in a cointegration relationship, we can determine whether the futures spread on a candidate trading date is wider or narrower than the expected spread. This determination enables statistical arbitrage trading between Henry Hub and PJM.

Since the prices in period t are not available for trading in period t , we estimate the following long-term equilibrium equation using the price series up to period $t - 1$:

$$\text{PJM}_{f,t} = \beta_{f,0} \times \text{HenryHub}_{f,t} + \beta_{f,1}. \quad (2.36)$$

If the futures spread is higher than the expected value, then we express it as

$$\text{PJM}_{f,t} > \beta_{f,0} \times \text{HenryHub}_{f,t} + \beta_{f,1}. \quad (2.37)$$

We can consider that the PJM price is higher and the Henry Hub price is lower; therefore, we take the PJM short position and Henry Hub long position. Then, the condition for closing these arbitrage positions is

$$\begin{aligned} \text{PJM}_{f,t} - \text{avgShortPJM}_f + \text{avgLongHenryHub}_f \\ - \text{HenryHub}_{f,t} > 0, \end{aligned} \quad (2.38)$$

where avgShortPJM_f is the average price of the PJM futures short positions taken, and avgLongHenryHub_f is the average price of the Henry Hub futures long positions taken. The clearance of all these futures positions under this condition leads to profit.

Conversely, if the futures spread is below the expected value, then we express it as

$$\text{PJM}_{f,t} < \beta_{f,0} \times \text{HenryHub}_{f,t} + \beta_{f,1}. \quad (2.39)$$

We determine that the PJM price is lower and the Henry Hub price is higher; therefore, we take the PJM long position and Henry Hub short position. Then, the condition for closing these arbitrage positions is

$$\begin{aligned} \text{avgLongPJM}_f - \text{PJM}_{f,t} + \text{HenryHub}_{f,t} \\ - \text{avgShortHenryHub}_f > 0, \end{aligned} \quad (2.40)$$

where avgLongPJM_f is the average price of the PJM futures long positions and $\text{avgShortHenryHub}_f$ is the average price of the Henry Hub futures short positions. If this condition is satisfied, we can lock in profits by clearing all these positions.

As mentioned above, there can be two patterns of statistical arbitrage trading: one is when the spread is larger than the expected value, and the other when the spread is smaller than the expected value. However, Sect. 2.4 simulates only the case

of Eq. (2.39) where the spread is narrower because the case of Eq. (2.37), which indicates a wider spread, often includes the case of Eq. (2.35).

Statistical arbitrage trading can take each position of any size in a single transaction. In Sect. 2.4, we take the Henry Hub short position and the PJM long position corresponding to the planned electric energy generation.

2.4 Simulation Results

This section simulates the trading strategy in Sect. 2.3 using one year of historical Henry Hub and PJM futures and spot data. We adopt futures with a one-month maturity. Thus, the futures trading period is from January 1, 2020 to December 31, 2020, and the spot trading period is from February 1, 2020 to January 31, 2021.

Table 2.5 presents the specifications of the assumed power generation business used in the simulation. We calculate the fixed unit cost by allocating the annual fixed cost to the peak hours, and dividing it by the amount of electric energy sold in one day. Calculating each coefficient of Eq. (2.31) from these yields the following equation:

$$\text{Cost}[\text{USD/MWh}] = 7.590 \times \text{Gas}[\text{USD/mmbtu}] + 10.464. \quad (2.41)$$

As a reference, we estimate the cointegrating vector by the DOLS using the historical data of these spot prices from February 1, 2020 to January 31, 2021. Consequently, the long-term equilibrium equation between the Henry Hub spot price $\text{HenryHub}_{s,t}$ and the PJM spot price $\text{PJM}_{s,t}$ during this period is

$$\begin{aligned} \text{PJM}_{s,t}[\text{USD/MWh}] &= 6.551 \\ &\times \text{HenryHub}_{s,t}[\text{USD/mmbtu}] + 11.576 \end{aligned} \quad (2.42)$$

Figure 2.9 plots Eqs. (2.41) and (2.42), which express the relationship between the unit prices of natural gas and electricity. Since the lowest price of the Henry Hub

Table 2.5 Power generation business specifications

(A) Electric Power (Power transmission end)	600	MW
(B) Power selling time	19	h/day
(C) Electric energy to sell ($= A \times B$)	11,400	MWh/day
(D) Thermal efficiency	45	%
(E) Natural gas energy required ($= C/D$)	86,528	mmbtu/day
(F) Fixed cost	55,000,000	USD/year
(G) $\alpha_0 (= E/C)$	7.590	mmbtu/MWh
(H) $\alpha_1 (= F/C/365 \times B/24)$	10.464	USD/MWh

spot during the trading period is 1.33 [USD/mmbtu], the power generation cost is always higher than the expected PJM spot price. In other words, the competitiveness of the power generation company is less than the spot market. Regrettably, the expected value of profit from spot trading alone is negative. Substituting the mean value of the Henry Hub spot prices during the simulation period, which is about 2.05 [USD/mmbtu], in Eq. (2.41), the mean value of costs is about 26.05 [USD/MWh]. On the other hand, the mean value of the PJM spot prices is about 25.36 [USD/MWh]. Thus, the expected loss in spot trading is

$$\begin{aligned} \text{ExpectedLoss} &= (26.05 - 25.36) \times 11400 \times 365 \\ &= 2,873,803 \left[\frac{\text{USD}}{\text{year}} \right]. \end{aligned} \quad (2.43)$$

As we can see, this simulation sets conditions that are disadvantageous to the business operator.

To reflect changes in the market environment (e.g., economic conditions, seasonality, climate), this statistical arbitrage simulation estimates the long-term equilibrium equation of these futures prices monthly. For example, statistical arbitrage in January 2020 uses the equilibrium equation estimated by DOLS using observations from January 2015 to December 2019. We estimate the long-term equilibrium equations used in the subsequent trading months by the moving window method. Table 2.6 summarizes the results.

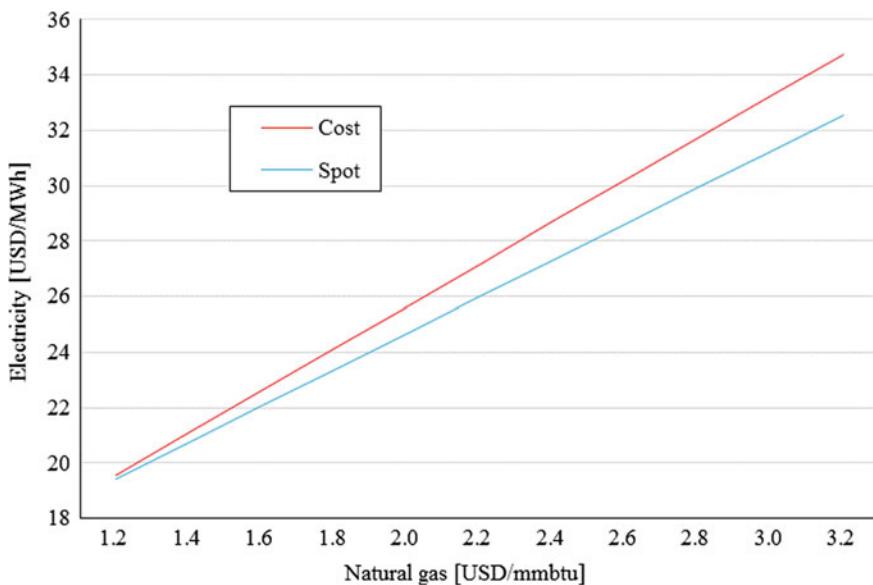


Fig. 2.9 Power generation cost and long-term equilibrium relationship in the spot market

Table 2.6 Long-term equilibrium estimation

Trading period	Long-term equilibrium $PJM_{f,t} = \beta_{f,0} \times HH_{f,t} + \beta_{f,1}$		Observation period
	$\beta_{f,0}$	$\beta_{f,1}$	
January 2020	9.18	13.12	January 2015 to December 2019
February 2020	9.16	13.00	February 2015 to January 2020
March 2020	8.81	13.44	March 2015 to February 2020
April 2020	9.31	11.78	April 2015 to March 2020
May 2020	9.76	10.40	May 2015 to April 2020
June 2020	10.40	8.50	June 2015 to May 2020
July 2020	10.31	8.67	July 2015 to June 2020
August 2020	10.18	8.98	August 2015 to July 2020
September 2020	10.20	8.84	September 2015 to August 2020
October 2020	10.32	8.38	October 2015 to September 2020
November 2020	10.27	8.27	November 2015 to October 2020
December 2020	10.21	8.20	December 2015 to November 2020

Figure 2.10 plots the coefficients of the monthly long-term equilibrium equations. These coefficients capture changes in the surrounding market conditions. This technique, which is able to recognize changes in the market structure, can be expected to generate profits as long as we do not hold a position for a long time.

The simulation trades these spots from February 2020 to January 2021, and futures from January 2020 to December 2020. The spot-future arbitrage and statistical arbitrage methods take positions corresponding to the electric energy planned to be sold.

The simulation results are as follows. Table 2.7 indicates the profit from each trade. The unrealized loss is due to futures that are not cleared on the last trading day. Spot-future arbitrage makes a profit of 6,446,848 [USD], which can cover the spot trading loss of 2,881,676 [USD]. The realized profit and unrealized loss from statistical arbitrage are 1,086,396 [USD] and 211,463 [USD], respectively. Therefore, the total profit is 874,933 [USD].

Figure 2.11 provides the time plots of each trading strategy's profit. The black line representing spot trading started in February 2020, because there was no spot trading in January 2020. The red, blue, and green lines express each arbitrage ending in December 2020 because the simulation does not trade futures in January 2021.

Table 2.8 provides the results for each combination of spot and futures trades at the end of the simulation period. If we trade only spots, we make a loss of 2,881,676 [USD], which is almost the same as the loss of 2,873,803 [USD] estimated from Eq. (2.43). By adopting both trading strategies, we earn a profit of 4,440,105 [USD]. Applying only the spot-future arbitrage results in a profit of 3,565,172 [USD], while

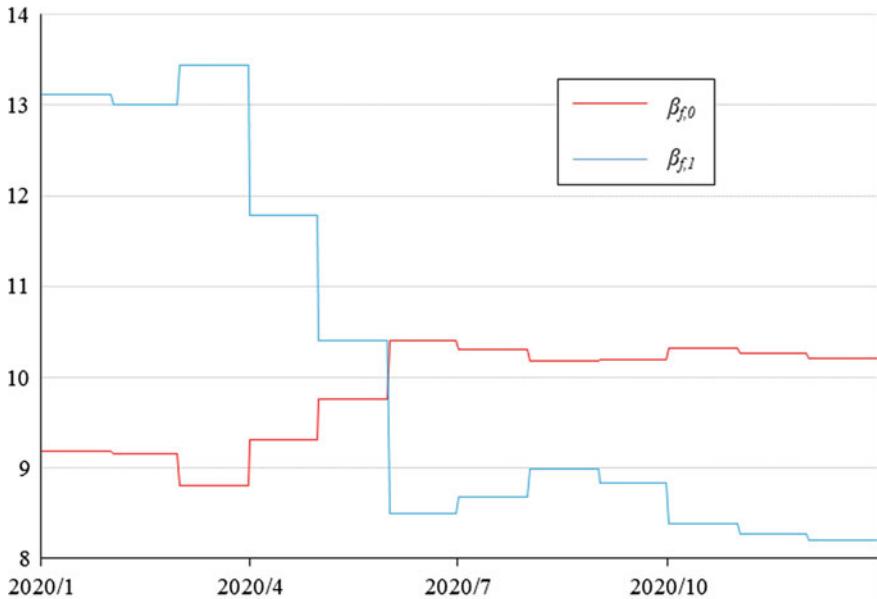


Fig. 2.10 Coefficients of the equilibrium equation for statistical arbitrage

Table 2.7 Profit from each trade

	Realized profit	Unrealized profit	Total profit
Spot trade	-2,881,676	0	-2,881,676
Spot-future arbitrage	6,446,848	0	6,446,848
Statistical arbitrage	1,086,396	-211,463	874,933

applying only the statistical arbitrage results in a loss of 2,006,743 [USD] which is less than the loss of 2,881,676 [USD] when only spot trading occurs.

Figure 2.12 presents the time-series plots of the profit from each combination of trading strategies. We can observe that the effectiveness of spot-future arbitrage is larger than that of statistical arbitrage throughout the period. However, the larger the leverage of statistical arbitrage, which is defined as the ratio of futures positions to spot positions, the greater the effectiveness of statistical arbitrage.

Therefore, to increase profits from statistical arbitrage trading, we consider increasing leverage. Table 2.9 shows the maximum unrealized loss during the trading period and total profit at the end of the trading period by leverage size. Clearly, both losses and profits are proportional to leverage. Therefore, if we take futures positions three times as great as spots, then statistical arbitrage can cover the loss from spot trading. Companies want to maximize their profits, but the risk they can take depends

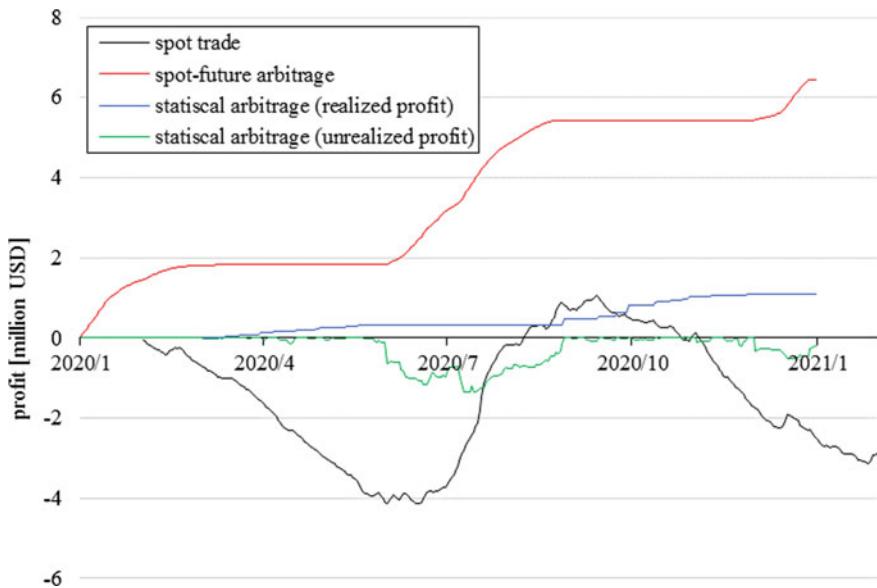


Fig. 2.11 Time-series plots of the profit from each trade

Table 2.8 Profit from each combination of the two trading strategies

	Statistical arbitrage	No statistical arbitrage
Spot-future arbitrage	4,440,105	3,565,172
No spot-future arbitrage	-2,006,743	-2,881,676

mainly on their capital. We must rationally manage the maximum unrealized loss during the trading period to maximize the profit from statistical arbitrage.

2.5 Risk Measurement in Statistical Arbitrage

If the expected return of a statistical arbitrage is positive, then we should maximize the futures positions to maximize profits. However, the trading technique expects a long-term equilibrium in the future, so we should manage risk, which means a possible future loss. This is because excessive unrealized losses may hinder business operations. Risks that occur during trading mainly include credit risk caused by the default of business partners, operational risk caused by failure in daily business activities and procedures, liquidity risk caused by an extreme decrease in trading volume that forces us to settle at a significantly unfavorable price, and market risk caused

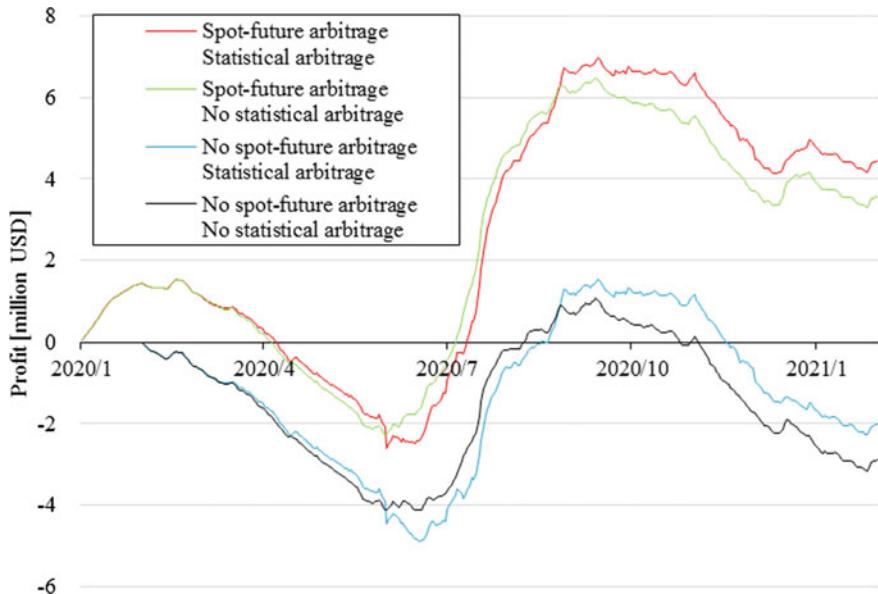


Fig. 2.12 Time-series plots of the profit from each combination of the two trading strategies

Table 2.9 Maximum unrealized loss and total profit from statistical arbitrage

Leverage	Maximum unrealized loss during the trading period	Total profit at the end of the trading period
1	1,362,795	874,933
2	2,725,589	1,749,866
3.3	4,497,223	2,887,279

by fluctuations in market prices. Chapter 2 covers only the market risk. We quantitatively measure the risk of a portfolio held during statistical arbitrage consisting of Henry Hub futures short positions and PJM futures long positions. This risk is due to fluctuations in futures prices.

2.5.1 Value-At-Risk and Expected Shortfall

This section introduces value-at-risk (VaR) and expected shortfall, which are the two most representative risk indicators that summarize portfolio risk into one numerical value.

VaR is the predicted maximum loss of a portfolio currently held in a given period within a given confidence interval, and is expressed in monetary amounts or yields. For example, if the maximum loss rate of a portfolio that can occur with a 99%

probability after one day is 10%, then the VaR at a confidence level of 99% with a one-day holding period is 10%, or the 1-day 99% VaR is called 10%. We can also interpret this condition as a 1% chance of a return worse than -10% after one day. Figure 2.13 expresses the 1-day 99% VaR when the return on the portfolio is normally distributed.

VaR is common and is one of the most attractive risk indicators because it is relatively easy to measure and helps us understand the worst risk. However, VaR can be misleading about risk, depending on the shape of the left tail of the return distribution. Supposing the distribution of the return of a portfolio in Fig. 2.14, which is a normal distribution with the left tail less than -10% partially inverted, we can understand that the VaR in Fig. 2.14 is same as in Fig. 2.13, while the risk in Fig. 2.14 is greater than in Fig. 2.13. VaR is not appropriate as a single numerical indicator of portfolio risk, depending on the shape of the left tail. Therefore, expected shortfall is the conditional expected loss in the case of loss exceeding VaR; in other words, expected shortfall can be understood as the mean value of loss in the worst case, and is consistent with intuition. We see that although the VaRs in Figs. 2.13 and 2.14 are at the same point, the expected shortfall in Fig. 2.14 is to the left of the expected shortfall in Fig. 2.13.

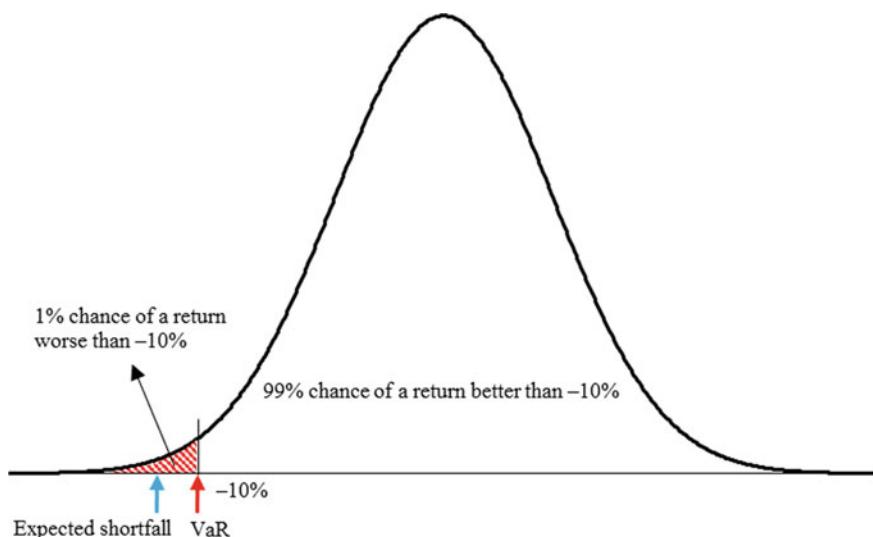


Fig. 2.13 One-day 99% VaR and the expected shortfall in a normal distribution

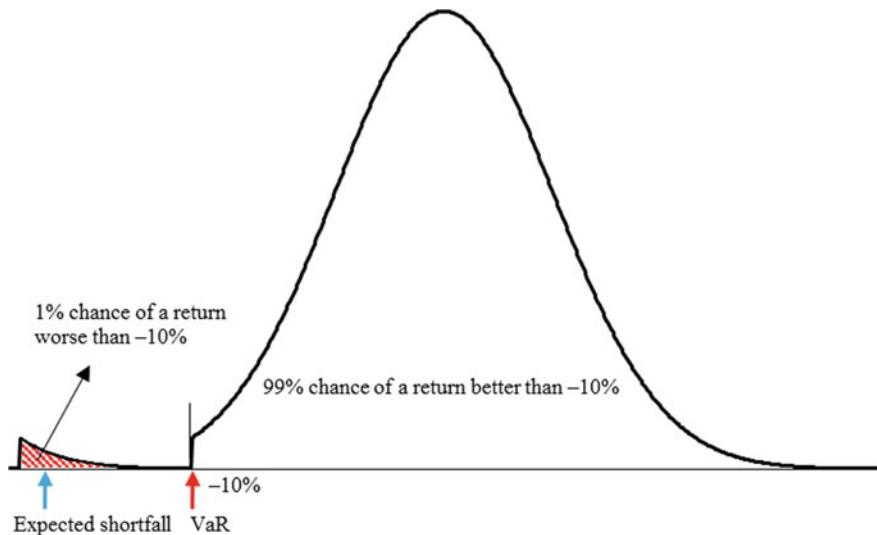


Fig. 2.14 One-day 99% VaR and the expected shortfall in a normal distribution with a deformed left tail

2.5.2 Copula

There are historical and parametric methods for measuring VaR and expected shortfalls. The historical method does not mathematically assume the distribution of value fluctuations but finds the percentile or conditional mean from the actual historical data as VaR or expected shortfall. Although this method can easily measure these indicators even in a complicated distribution, the measurement results are extremely dependent on the observations. The parametric method models the distribution of value fluctuations to calculate VaR or expected shortfall. In some cases, we may mathematically calculate the risk values easily by assuming a normal or lognormal distribution. In other cases, Monte Carlo simulations often derive risk values by assuming a complicated distribution. This measurement result depends on the assumed distribution model and estimated parameters. Therefore, the selection of a distribution model and estimation of its parameters are very important. As a parametric method, this section introduces a method using copulas, which convert from multiple marginal distributions into joint distributions. Sklar [33] proposed the copula concept.

Assuming m returns as risk factors, which are expressed by the stochastic variables R_1, R_2, \dots, R_m , let their distribution functions be F_1, F_2, \dots, F_m . To capture the stochastic value fluctuations of the portfolio consisting of m securities, we must estimate the joint distribution function F of the stochastic variables R_1, R_2, \dots, R_m from the marginal distribution functions F_1, F_2, \dots, F_m . In this case, a unique function C exists that satisfies the following equation:

$$\text{Probability}(R_1 \leq \rho_1, R_2 \leq \rho_2, \dots, R_m \leq \rho_m)$$

$$= F(\rho_1, \rho_2, \dots, \rho_m) = C(F_1(\rho_1), F_2(\rho_2), \dots, F_m(\rho_m)). \quad (2.44)$$

This is known as Sklar's [33] theorem. Thus, for any $\mathbf{u} = (u_1, u_2, \dots, u_m)$, the following equation holds:

$$C(u_1, u_2, \dots, u_m) = F(F_1^{-1}(u_1), F_2^{-1}(u_2), \dots, F_m^{-1}(u_m)), \quad (2.45)$$

where $0 \leq u_1, u_2, \dots, u_m \leq 1$ is satisfied. Further, let $c, f_1, f_2, \dots, f_m, f$ be the density function of $C, F_1, F_2, \dots, F_m, F$, respectively. We obtain the following equation from Eq. (2.44):

$$f(r_1, r_2, \dots, r_m) = c(F_1(r_1), F_2(r_2), \dots, F_m(r_m)) \prod_{i=1}^m f_i(r_i). \quad (2.46)$$

Therefore, the log-likelihood function $ll(\boldsymbol{\pi}_1, \boldsymbol{\pi}_2, \dots, \boldsymbol{\pi}_m; \boldsymbol{\pi})$ is

$$\begin{aligned} & ll(\boldsymbol{\pi}_1, \boldsymbol{\pi}_2, \dots, \boldsymbol{\pi}_m; \boldsymbol{\pi}) \\ &= \sum_{j=1}^N \log c(F_1(r_{1,j}; \boldsymbol{\pi}_1), F_2(r_{2,j}; \boldsymbol{\pi}_2), \dots, F_m(r_{m,j}; \boldsymbol{\pi}_m); \boldsymbol{\pi}) \\ &\quad + \sum_{j=1}^N \sum_{i=1}^m \log f_i(r_{i,j}; \boldsymbol{\pi}_i), \end{aligned} \quad (2.47)$$

where $r_{i,j}$ is the j -th observation of the i -th stochastic variable ($i = 1, 2, \dots, m$; $j = 1, 2, \dots, N$), $\boldsymbol{\pi}_i$ is the parameters of the marginal distribution function $F_i(\cdot; \boldsymbol{\pi}_i)$ ($i = 1, 2, \dots, m$), and $\boldsymbol{\pi}$ is the copula parameters. Estimating Eq. (2.47) by the maximum likelihood means estimating all parameters, namely copula and marginal distribution function parameters, simultaneously. The copula is estimated after estimating the marginal distributions considering the computational load. We adopt a double exponential distribution, expressed as follows, for the probability density function f_i of the marginal distribution function F_i :

$$f_i(x) = \frac{1}{\sqrt{2}\sigma_i} \exp\left(-\left|\frac{x - \mu_i}{\sigma_i/\sqrt{2}}\right|\right), \quad (2.48)$$

where μ_i and σ_i are the mean and standard deviation, respectively. Integrating Eq. (2.48), we obtain the following equation for the marginal distribution function F_i :

$$\begin{aligned} F_i(x) &= 1 - \frac{1}{2} \exp\left(-\frac{x - \mu_i}{\sigma_i/\sqrt{2}}\right), \text{ if } x > \mu \\ F_i(x) &= \frac{1}{2} \exp\left(\frac{x - \mu_i}{\sigma_i/\sqrt{2}}\right), \text{ if } x \leq \mu \end{aligned} \quad (2.49)$$

From Eq. (2.49), we obtain the inverse function:

$$\begin{aligned} x &= \mu_i - \frac{\sigma_i}{\sqrt{2}} \log(2 - 2F_i(x)), \text{ if } F_i(x) > \frac{1}{2} \\ x &= \mu_i + \frac{\sigma_i}{\sqrt{2}} \log(2F_i(x)), \text{ if } F_i(x) \leq \frac{1}{2} \end{aligned} \quad (2.50)$$

The most common copulas are broadly divided into two families: the elliptical copula and Archimedean copula. First, the two most common elliptical copula functions are the Gaussian copula, introduced by Lee [21] and generalized by Ophem [28], and the t copula, proposed by Embrechts et al. [6] and Fang et al. [8]. These copulas have elliptical distributions. These represent the correlation structure between the stochastic variables as a matrix. The three most popular Archimedean copulas are Gumbel [14], Clayton [4], and Frank [11] copulas. These copulas express a correlation structure with a single parameter. Here, we provide a brief description of the Gaussian, t , Gumbel, Clayton, and Frank copulas, although we refer the reader to Nelsen [27] for the theory and precise properties.

Gaussian Copula

Given that m stochastic variables follow the m variate standard normal distribution Φ_m with the correlation matrix Σ , Gaussian copula $C_{Gaussian}(u_1, u_2, \dots, u_m; \Sigma)$ is

$$\begin{aligned} C_{Gaussian}(u_1, u_2, \dots, u_m; \Sigma) \\ = \Phi_m(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \dots, \Phi^{-1}(u_m); \Sigma), \end{aligned} \quad (2.51)$$

where Φ^{-1} is the inverse function of the univariate standard normal distribution. The density function $c_{Gaussian}(u_1, u_2, \dots, u_m; \Sigma)$ is

$$\begin{aligned} c_{Gaussian}(u_1, u_2, \dots, u_m; \Sigma) &= \frac{\partial C_{Gaussian}^m(u_1, u_2, \dots, u_m; \Sigma)}{\partial u_1 \partial u_2 \dots \partial u_m} \\ &= \frac{1}{\sqrt{|\Sigma|}} \exp\left(-\frac{1}{2} {}^T \boldsymbol{\omega} (\Sigma^{-1} - E) \boldsymbol{\omega}\right), \end{aligned} \quad (2.52)$$

where E denotes the identity matrix, and

$$\begin{aligned} \boldsymbol{\omega} &= {}^T (\omega_1, \omega_2, \dots, \omega_m) \\ &= {}^T (\Phi^{-1}(u_1), \Phi^{-1}(u_2), \dots, \Phi^{-1}(u_m)) \\ &= \Phi^{-1}(\mathbf{u}). \end{aligned} \quad (2.53)$$

Thus, using the N observations of each stochastic variable, the log-likelihood $ll_{Gaussian}$ is

$$\begin{aligned}
ll_{Gaussian}(\boldsymbol{\Sigma}) &= \log \prod_{j=1}^N \frac{\partial C_{Gaussian}^m(u_1, u_2, \dots, u_m; \boldsymbol{\Sigma})}{\partial u_1 \partial u_2 \dots \partial u_m} \\
&= -\frac{N}{2} \log |\boldsymbol{\Sigma}| - \frac{1}{2} \sum_{j=1}^N {}^T \boldsymbol{\omega}_j (\boldsymbol{\Sigma}^{-1} - \mathbf{E}) \boldsymbol{\omega}_j
\end{aligned} \quad (2.54)$$

where

$$\boldsymbol{\omega}_j = \Phi^{-1}(\mathbf{u}_j). \quad (2.55)$$

Therefore, the maximum likelihood estimator $\widehat{\boldsymbol{\Sigma}}$ is given by the following equation:

$$\widehat{\boldsymbol{\Sigma}} = -\frac{1}{N} \sum_{j=1}^N \boldsymbol{\omega}_j {}^T \boldsymbol{\omega}_j. \quad (2.56)$$

t Copula

Given that m stochastic variables follow the m variate t distribution $t_{df, \Sigma}^m$ with degree of freedom $df (\geq 3)$ and the correlation matrix $\boldsymbol{\Sigma}$, the t copula $C_t(u_1, u_2, \dots, u_m)$ is

$$C_t(u_1, u_2, \dots, u_m; \boldsymbol{\Sigma}, df) = t_{df, \Sigma}^m(t_{df}^{-1}(u_1), t_{df}^{-1}(u_2), \dots, t_{df}^{-1}(u_m)), \quad (2.57)$$

where t_{df}^{-1} is the inverse function of the univariate t distribution with degree of freedom df . The density function $c_t(u_1, u_2, \dots, u_m; \boldsymbol{\Sigma}, df)$ is

$$\begin{aligned}
c_t(u_1, u_2, \dots, u_m; \boldsymbol{\Sigma}, df) &= \frac{\partial C_t^m(u_1, u_2, \dots, u_m; \boldsymbol{\Sigma}, df)}{\partial u_1 \partial u_2, \dots, \partial u_m} \\
&= \frac{\Gamma\left(\frac{df+m}{2}\right)\left(\Gamma\left(\frac{df}{2}\right)\right)^m \left(1 + \frac{{}^T \boldsymbol{\omega} \boldsymbol{\Sigma}^{-1} \boldsymbol{\omega}}{df}\right)^{-\frac{df+m}{2}}}{\sqrt{|\boldsymbol{\Sigma}|} \Gamma\left(\frac{df}{2}\right) \left(\Gamma\left(\frac{df+1}{2}\right)\right)^m \prod_{j=1}^N \left(1 + \frac{\omega_j^2}{df}\right)^{-\frac{df+1}{2}}},
\end{aligned} \quad (2.58)$$

where

$$\begin{aligned}
\boldsymbol{\omega} &= {}^T (\omega_1, \omega_2, \dots, \omega_m) \\
&= {}^T (t_{df}^{-1}(u_1), t_{df}^{-1}(u_2), \dots, t_{df}^{-1}(u_m)).
\end{aligned} \quad (2.59)$$

Thus, using the N observations of each stochastic variable, the log-likelihood ll_t is

$$\begin{aligned}
ll_t(\boldsymbol{\Sigma}, df) &= \log \prod_{j=1}^N \frac{\partial C_t^m(u_1, u_2, \dots, u_m; \boldsymbol{\Sigma}, df)}{\partial u_1 \partial u_2 \dots \partial u_m} \\
&= N \log \left(\frac{\Gamma\left(\frac{df+m}{2}\right)}{\Gamma\left(\frac{df}{2}\right)} \right) + mN \log \left(\frac{\Gamma\left(\frac{df}{2}\right)}{\Gamma\left(\frac{df+1}{2}\right)} \right) - \frac{N}{2} \log |\boldsymbol{\Sigma}| \\
&\quad - \frac{df+m}{2} \sum_{j=1}^N \log \left(1 + \frac{^T \boldsymbol{\omega}_j \boldsymbol{\Sigma}^{-1} \boldsymbol{\omega}_j}{df} \right) \\
&\quad + \frac{df+1}{2} \sum_{j=1}^N \sum_{i=1}^m \log \left(1 + (\boldsymbol{\omega}_j)_i^2 \right), \tag{2.60}
\end{aligned}$$

where

$$\boldsymbol{\omega}_j = t_{df}^{-1}(\mathbf{u}_j). \tag{2.61}$$

Therefore, the maximum likelihood estimator $\widehat{\boldsymbol{\Sigma}}$ satisfies

$$\widehat{\boldsymbol{\Sigma}} = \frac{df+m}{N} \sum_{j=1}^N \frac{{\boldsymbol{\omega}_j}^T \boldsymbol{\omega}_j}{df + {^T \boldsymbol{\omega}_j \widehat{\boldsymbol{\Sigma}}^{-1} \boldsymbol{\omega}_j}}. \tag{2.62}$$

Given various degrees of freedom df , the maximum likelihood estimator $\widehat{\boldsymbol{\Sigma}}$ is obtained by iterative convergence calculation. Then, we calculate the log-likelihood $ll_t(\boldsymbol{\Sigma}, df)$. Let df and $\widehat{\boldsymbol{\Sigma}}$, which maximize $ll_t(\boldsymbol{\Sigma}, df)$, be the parameters of the t copula.

Clayton Copula

We can express the Clayton copula $C_{Clayton}(u_1, u_2, \dots, u_m; \pi_c)$ with the following equation using a parameter $\pi_c (> 0)$:

$$C_{Clayton}(u_1, u_2, \dots, u_m; \pi_c) = \left(\sum_{i=1}^m u_i^{-\pi_c} - m + 1 \right)^{-\frac{1}{\pi_c}}. \tag{2.63}$$

The density function $c_{Clayton}(u_1, u_2, \dots, u_m; \pi_c)$ is then

$$\begin{aligned}
c_{Clayton}(u_1, u_2, \dots, u_m; \pi_c) &= \frac{\partial C_{Clayton}^m(u_1, u_2, \dots, u_m; \pi_c)}{\partial u_1 \partial u_2 \dots \partial u_m} \\
&= \left(\prod_{i=1}^{m-1} (1 + i\pi_c) \right) \left(\prod_{i=1}^m u_i^{-\pi_c - 1} \right) \left(\prod_{i=1}^m u_i^{-\pi_c} - m + 1 \right)^{-\frac{1}{\pi_c} - m}. \tag{2.64}
\end{aligned}$$

Thus, using the N observations of each stochastic variable, the log-likelihood $ll_{Clayton}$ is

$$\begin{aligned} ll_{Clayton}(\pi_c) &= \log \prod_{j=1}^N \frac{\partial C_{Clayton}^m(u_1, u_2, \dots, u_m; \pi_c)}{\partial u_1 \partial u_2 \dots \partial u_m} \\ &= N \sum_{i=1}^{m-1} \log(1 + i\pi_c) \\ &\quad - \sum_{j=1}^N \left(\begin{array}{l} (1 + \pi_c) \log \left(\prod_{i=1}^m u_i^j \right) \\ + \left(\frac{1}{\pi_c} + m \right) \log \left(\sum_{i=1}^m (u_i^j)^{-\pi_c} - m + 1 \right) \end{array} \right) \end{aligned} \quad (2.65)$$

We may calculate the π_c that maximizes the log-likelihood $ll_{Clayton}$ numerically.

Gumbel Copula

We can express the Gumbel copula $C_{Gumbel}(u_1, u_2, \dots, u_m; \pi_g)$ with the following equation using a parameter $\pi_g (> 1)$:

$$C_{Gumbel}(u_1, u_2, \dots, u_m; \pi_g) = \exp \left(- \left(\sum_{j=1}^m (-\log u_j)^{\pi_g} \right)^{\frac{1}{\pi_g}} \right). \quad (2.66)$$

The density function $c_{Gumbel}(u_1, u_2, \dots, u_m; \pi_g)$ is

$$c_{Gumbel}(u_1, u_2, \dots, u_m; \pi_g) = \frac{\partial C_{Gumbel}^m(u_1, u_2, \dots, u_m; \pi_g)}{\partial u_1 \partial u_2 \dots \partial u_m}. \quad (2.67)$$

Thus, using the N observations of each stochastic variable, the log-likelihood ll_{Gumbel} is

$$ll_{Gumbel}(\pi_g) = \log \prod_{j=1}^N \frac{\partial C_{Gumbel}^m(u_1, u_2, \dots, u_m; \pi_g)}{\partial u_1 \partial u_2 \dots \partial u_m}. \quad (2.68)$$

We can numerically calculate the π_g that maximizes the log-likelihood ll_{Gumbel} . For simplicity of notation, we define the following equations:

$$\begin{aligned} v_i &= -\log u_i (i = 1, 2, \dots, m) \\ \psi_i^j &= \left(\sum_{i=1}^m (v_i^j)^{\pi_g} \right)^{\frac{1}{\pi_g}} \end{aligned} \quad (2.69)$$

Then, the log-likelihoods ${}^m ll_{Gumbel}$ of the $m (= 2, 3, 4, 5)$ variate Gumbel copula can be described by Eqs. (2.70), (2.71), (2.72), and (2.73).

$$\begin{aligned} {}^2 ll_{Gumbel}(\pi_g) &= \log \prod_{j=1}^N \frac{\partial C_{Gumbel}^2(u_1, u_2; \pi_g)}{\partial u_1 \partial u_2} \\ &= \sum_{j=1}^N \left(\begin{array}{l} -\left((v_1^j)^{\pi_g} + (v_2^j)^{\pi_g} \right)^{\frac{1}{\pi_g}} \\ + \log \left(\frac{(v_1^j v_2^j)^{\pi_g-1}}{u_1^j u_2^j} \left((v_1^j)^{\pi_g} + (v_2^j)^{\pi_g} \right)^{\frac{1}{\pi_g}-2} \right) \\ + \log \left(\pi_g - 1 + \left((v_1^j)^{\pi_g} + (v_2^j)^{\pi_g} \right)^{\frac{1}{\pi_g}} \right) \end{array} \right) \\ &= \sum_{j=1}^N \left(\begin{array}{l} -\psi_2^j \\ + \log \left(\frac{(v_1^j v_2^j)^{\pi_g-1}}{u_1^j u_2^j} \left(\psi_2^j \right)^{1-2\pi_g} \right) \\ + \log \left(\pi_g - 1 + \psi_2^j \right) \end{array} \right) \end{aligned} \quad (2.70)$$

$$\begin{aligned} {}^3 ll_{Gumbel}(\pi_g) &= \log \prod_{j=1}^N \frac{\partial C_{Gumbel}^3(u_1, u_2, u_3; \pi_g)}{\partial u_1 \partial u_2 \partial u_3} \\ &= \sum_{j=1}^N \left(\begin{array}{l} -\psi_3^j \\ + \log \left(\frac{(v_1^j v_2^j v_3^j)^{\pi_g-1}}{u_1^j u_2^j u_3^j} \left(\psi_3^j \right)^{1-3\pi_g} \right) \\ + \log \left(\begin{array}{l} (2\pi_g - 1)(\pi_g - 1) \\ + 3(\pi_g - 1)\psi_3^j + (\psi_3^j)^2 \end{array} \right) \end{array} \right) \end{aligned} \quad (2.71)$$

$$\begin{aligned} {}^4 ll_{Gumbel}(\pi_g) &= \log \prod_{j=1}^N \frac{\partial C_{Gumbel}^4(u_1, u_2, u_3, u_4; \pi_g)}{\partial u_1 \partial u_2 \partial u_3 \partial u_4} \\ &= \sum_{j=1}^N \left(\begin{array}{l} -\psi_4^j \\ + \log \left(\frac{(v_1^j v_2^j v_3^j v_4^j)^{\pi_g-1}}{u_1^j u_2^j u_3^j u_4^j} \left(\psi_4^j \right)^{1-4\pi_g} \right) \\ + \log \left(\begin{array}{l} (3\pi_g - 1)(2\pi_g - 1)(\pi_g - 1) \\ + (11\pi_g - 7)(\pi_g - 1)\psi_4^j \\ + 6(\pi_g - 1)\left(\psi_4^j \right)^2 + \left(\psi_4^j \right)^3 \end{array} \right) \end{array} \right) \end{aligned} \quad (2.72)$$

$${}^5 ll_{Gumbel}(\pi_g) = \log \prod_{j=1}^N \frac{\partial C_{Gumbel}^5(u_1, u_2, u_3, u_4, u_5; \pi_g)}{\partial u_1 \partial u_2 \partial u_3 \partial u_4 \partial u_5}$$

$$\begin{aligned}
&= \sum_{j=1}^N \left(\begin{array}{l} -\psi_5^j \\ + \log \left(\frac{(v_1^j v_2^j v_3^j v_4^j v_5^j)^{\pi_g-1}}{u_1^j u_2^j u_3^j u_4^j u_5^j} (\psi_5^j)^{1-5\pi_g} \right) \\ + \log \left(\begin{array}{l} (4\pi_g-1)(3\pi_g-1)(2\pi_g-1)(\pi_g-1) \\ + 5(5\pi_g-3)(2\pi_g-1)(\pi_g-1)\psi_5^j \\ + 5(7\pi_g-5)(\pi_g-1)(\psi_5^j)^2 \\ + 10(\pi_g-1)(\psi_5^j)^3 + (\psi_5^j)^4 \end{array} \right) \end{array} \right) \quad (2.73)
\end{aligned}$$

Frank Copula

We can express the Frank copula $C_{Frank}(u_1, u_2, \dots, u_m; \pi_f)$ with the following equation using parameter $\pi_f (> 0)$:

$$C_{Frank}(u_1, u_2, \dots, u_m; \pi_f) = -\frac{1}{\pi_f} \log \left(1 + \frac{\prod_{i=1}^m (e^{-\pi_f u_i} - 1)}{(e^{-\pi_f} - 1)^{m-1}} \right). \quad (2.74)$$

The density function $c_{Frank}(u_1, u_2, \dots, u_m; \pi_f)$ is

$$c_{Frank}(u_1, u_2, \dots, u_m; \pi_f) = \frac{\partial C_{Frank}^m(u_1, u_2, \dots, u_m; \pi_f)}{\partial u_1 \partial u_2 \dots \partial u_m}. \quad (2.75)$$

Thus, using the N observations of each stochastic variable, the log-likelihood ll_{Frank} is

$$ll_{Frank}(\pi_f) = \log \prod_{j=1}^N \frac{\partial C_{Frank}^m(u_1, u_2, \dots, u_m; \pi_f)}{\partial u_1 \partial u_2 \dots \partial u_m}. \quad (2.76)$$

We can numerically calculate the π_f that maximizes the log-likelihood ll_{Frank} . For simplicity of notation, we define the following equation:

$$w_i = e^{-\delta u_i} - 1 (i = 1, 2, \dots, m). \quad (2.77)$$

Then, the log-likelihoods $^m ll_{Frank}$ of the $m (= 2, 3, 4, 5)$ variate Frank copula can be described by Eqs. (2.78), (2.79), (2.80), and (2.81), respectively.

$$^2 ll_{Frank}(\pi_f) = \log \prod_{j=1}^N \frac{\partial C_{Frank}^2(u_1, u_2; \pi_f)}{\partial u_1 \partial u_2}$$

$$= \sum_{j=1}^N \log \frac{-\delta(w_1^j + 1)(w_2^j + 1)(e^{-\delta} - 1)}{(e^{-\delta} - 1 + w_1^j w_2^j)^2} \quad (2.78)$$

$$\begin{aligned} {}^3ll_{Frank}(\pi_f) &= \log \prod_{j=1}^N \frac{\partial C_{Frank}^3(u_1, u_2, u_3; \pi_f)}{\partial u_1 \partial u_2 \partial u_3} \\ &= \sum_{j=1}^N \log \left(\frac{-\delta^2(w_1^j + 1)(w_2^j + 1)(w_3^j + 1)(e^{-\delta} - 1)^2}{\left((e^{-\delta} - 1)^2 + w_1^j w_2^j w_3^j \right)^3} \right) \end{aligned} \quad (2.79)$$

$$\begin{aligned} {}^4ll_{Frank}(\pi_f) &= \log \prod_{j=1}^N \frac{\partial C_{Frank}^4(u_1, u_2, u_3, u_4; \pi_f)}{\partial u_1 \partial u_2 \partial u_3 \partial u_4} \\ &= \sum_{j=1}^N \log \left(\frac{-\delta^3(w_1^j + 1)(w_2^j + 1)(w_3^j + 1)(w_4^j + 1)(e^{-\delta} - 1)^3}{\left((e^{-\delta} - 1)^6 - 4(e^{-\delta} - 1)^3 w_1^j w_2^j w_3^j w_4^j \right.} \right. \\ &\quad \times \left. \left. \left. + (w_1^j)^2 (w_2^j)^2 (w_3^j)^2 (w_4^j)^2 \right) \over \left((e^{-\delta} - 1)^3 + w_1^j w_2^j w_3^j w_4^j \right)^4 \right) \right) \end{aligned} \quad (2.80)$$

$$\begin{aligned} {}^5ll_{Frank}(\pi_f) &= \log \prod_{j=1}^N \frac{\partial C_{Frank}^5(u_1, u_2, u_3, u_4, u_5; \pi_f)}{\partial u_1 \partial u_2 \partial u_3 \partial u_4 \partial u_5} \\ &= \sum_{j=1}^N \log \left(\frac{-\delta^4(w_1^j + 1)(w_2^j + 1)(w_3^j + 1)(w_4^j + 1)(w_5^j + 1)}{\left((e^{-\delta} - 1)^4 \right.} \right. \\ &\quad \times \left. \left. \left. \left. \left. + (e^{-\delta} - 1)^4 \right) \over \left((e^{-\delta} - 1)^{12} - 11(e^{-\delta} - 1)^8 w_1^j w_2^j w_3^j w_4^j w_5^j \right. \right. \right. \right. \right. \\ &\quad \left. \left. \left. \left. \left. \left. + 11(e^{-\delta} - 1)^8 (w_1^j)^2 (w_2^j)^2 (w_3^j)^2 (w_4^j)^2 (w_5^j)^2 \right) \right. \right. \right. \right. \right. \\ &\quad \left. \left. \left. \left. \left. \left. - (w_1^j)^3 (w_2^j)^3 (w_3^j)^3 (w_4^j)^3 (w_5^j)^3 \right) \right) \over \left((e^{-\delta} - 1)^4 + w_1^j w_2^j w_3^j w_4^j w_5^j \right)^5 \right) \right) \end{aligned} \quad (2.81)$$

After estimating all the parameters using the above, we should select the copula to measure the risk indicators. Although it is desirable to adopt a copula that minimizes the difference between the joint distribution obtained from historical data and the joint distribution estimated by each copula, it is rational to select the copula based on the information criterion, considering that the number of parameters and the degrees of freedom differ for each type of copula. As in Breymann et al. [3], we select the copula based on the Akaike information criterion (AIC).

We now generate the uniform random numbers U_1, U_2, \dots, U_m , which follow the estimated copula $C(u_1, u_2, \dots, u_m)$. Then, we can find the VaR and expected

shortfall from $\tilde{R}_1 = F_1^{-1}(U_1)$, $\tilde{R}_2 = F_2^{-1}(U_2)$, ..., $\tilde{R}_m = F_m^{-1}(U_m)$, where $F_1^{-1}, F_2^{-1}, \dots, F_m^{-1}$ are the inverse functions of the marginal distribution functions. These allow us to assess the portfolio risk. The generation of random numbers based on each copula type is as follows.

Gaussian Copula

We obtain the lower triangular matrix using the Cholesky decomposition of the estimated $\widehat{\Sigma}$. By multiplying the matrix by a vector $(s_1, s_2, \dots, s_m)^T$ whose elements are random numbers that follow an independent standard normal distribution, we may generate random numbers S_1, S_2, \dots, S_m that follow a multivariate normal distribution with a correlation matrix $\widehat{\Sigma}$. Using the standard normal distribution function Φ , we can generate random numbers $U_1 = \Phi(S_1)$, $U_2 = \Phi(S_2)$, ..., $U_m = \Phi(S_m)$ that follow the Gaussian copula.

t Copula

We obtain the lower triangular matrix using the Cholesky decomposition of the estimated $\widehat{\Sigma}$. By multiplying the matrix by a vector $(s_1, s_2, \dots, s_m)^T$ whose elements are random numbers that follow an independent standard normal distribution, we may generate random numbers S_1, S_2, \dots, S_m that follow a multivariate normal distribution with a correlation matrix $\widehat{\Sigma}$. Generating random numbers Z_1, Z_2, \dots, Z_m that follow an independent standard normal distribution, which follows the χ^2 distribution with degree of freedom df , we may calculate $T_1 = \sqrt{df}S_1/Z_1$, $T_2 = \sqrt{df}S_2/Z_2, \dots, T_m = \sqrt{df}S_m/Z_m$, which itself follows a multivariate *t* distribution with degree of freedom df and a correlation matrix $\widehat{\Sigma}$. Using the *t* distribution function t_{df} with degree of freedom df , we can generate random numbers $U_1 = t_{df}(T_1)$, $U_2 = t_{df}(T_2)$, ..., $U_m = t_{df}(T_m)$ that follow the *t* copula.

Clayton Copula

Generating uniform random numbers $I_1, I_2, \dots, I_m (0 \leq I_1, I_2, \dots, I_m \leq 1)$ and random numbers γ that follow the standard gamma distribution $G(1/\pi_c)$, we can calculate random numbers U_1, U_2, \dots, U_m that follow the Clayton copula, as follows:

$$\begin{aligned} U_1 &= (1 - (1/\gamma) \log I_1)^{-1/\pi_c}, U_2 = (1 - (1/\gamma) \log I_2)^{-1/\pi_c}, \\ &\dots, U_m = (1 - (1/\gamma) \log I_m)^{-1/\pi_c}. \end{aligned} \quad (2.82)$$

Gumbel Copula

Generating uniform random numbers $V (0 \leq V \leq \text{circle ratio } \pi)$ and random numbers W that follow the standard exponential distribution, we calculate the following equation:

$$\left(\frac{\sin((\pi_g - 1)V/\pi_g)}{W} \right)^{\pi_g - 1} \equiv \theta. \quad (2.83)$$

Using uniform random numbers $I_1, I_2, \dots, I_m (0 \leq I_1, I_2, \dots, I_m \leq 1)$, we can calculate random numbers U_1, U_2, \dots, U_m that follow the Gumbel copula, as follows:

$$\begin{aligned} U_1 &= \exp\left(-(-(1/\theta) \log I_1)^{1/\pi_g}\right), \\ U_2 &= (-(1/\theta) \log I_2)^{1/\pi_g}, \dots, U_m = (-(1/\theta) \log I_m)^{1/\pi_g}. \end{aligned} \quad (2.84)$$

Frank Copula

We use only a bivariate Frank copula in this book, and thus omit the generation of random numbers that follow trivariate or more Frank copulas. Generating uniform random numbers $I_1, I_2 (0 \leq I_1, I_2 \leq 1)$, we can set random numbers U_1 and U_2 that follow the Frank copula, as follows:

$$\begin{aligned} U_1 &= I_1 \\ U_2 &= -\frac{1}{\pi_f} \log \left(\frac{I_2(1 - e^{-\pi_f})}{I_2(e^{-\pi_f}U_1 - 1) - e^{-\pi_f}U_1} + 1 \right). \end{aligned} \quad (2.85)$$

While countless studies apply copulas to financial markets, other studies analyze energy markets using copulas. Ghorbel and Trabelsi [12] adopt the concept of copulas to study the dependence structure between the crude oil, natural gas, and heating oil markets. Lu et al. [23] construct the conditional joint distribution using copula functions with GARCH-type models to calculate the VaR of a portfolio composed of crude oil futures and natural gas futures. Zhang [36] examines the impact of changes in crude oil prices on bunker prices and tanker freight rates using copula models. Soliman and Nasir [34] employ a time-varying copula connection function to analyze the risk dependency relationship between the energy and carbon markets in Europe. Fernandes et al. [9] propose a stochastic-copula approach that considers seasonality, mean reversion, and jumps to examine the dependence between European electricity and natural gas day-ahead prices. Ly et al. [24] reveal the dependencies and co-movements between European electricity markets based on copula models.

2.5.3 *Copula Estimation and Risk Measurement*

In this section, we aim to measure the risk of the spread between the Henry Hub futures prices and the PJM futures prices because the statistical arbitrage simulated in Sect. 2.4 holds a portfolio consisting of Henry Hub futures short positions and PJM futures long positions.

First, we calculate each return series of the Henry Hub futures prices and PJM futures prices. We also estimate the mean and standard deviation of each return

Table 2.10 Parameters of each marginal distribution function

	Henry Hub futures	PJM futures
Mean	3.52×10^{-4}	6.52×10^{-4}
Standard deviation	3.16×10^{-2}	4.64×10^{-2}

series. Table 2.10 provides these values, which are the parameters of each marginal distribution function.

We then estimate each copula and generate random numbers to determine the VaR and expected shortfall according to the procedure in Sect. 2.5.2. However, we should note the following. For long positions, the smaller the return, the larger the loss. Therefore, we should measure the risk in the range of the left tail of the distribution. However, for short positions, the larger the return, the larger the loss; therefore, we should measure the risk in the range of the right tail of the distribution.

Figure 2.15 presents a simplified diagram of the procedure in this section. Figure 2.16 plots the relationship between the degree of freedom and the log-likelihood in the estimated t copula. The t copula parameter is the $\widehat{\Sigma}$ estimated with the degree of freedom 3 that maximizes the log-likelihood. We generate the distribution of the daily rate of return by simulating each estimated copula 10,000 times to calculate the VaR and expected shortfall. Table 2.11 provides the estimated copulas and the VaRs and expected shortfalls calculated based on these copulas.

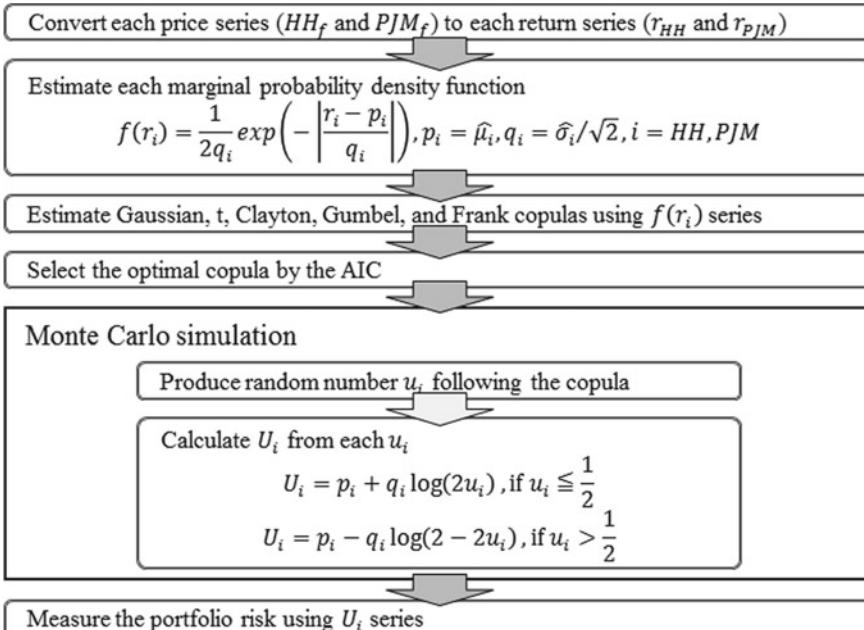


Fig. 2.15 Procedure to measure portfolio risk using various copulas

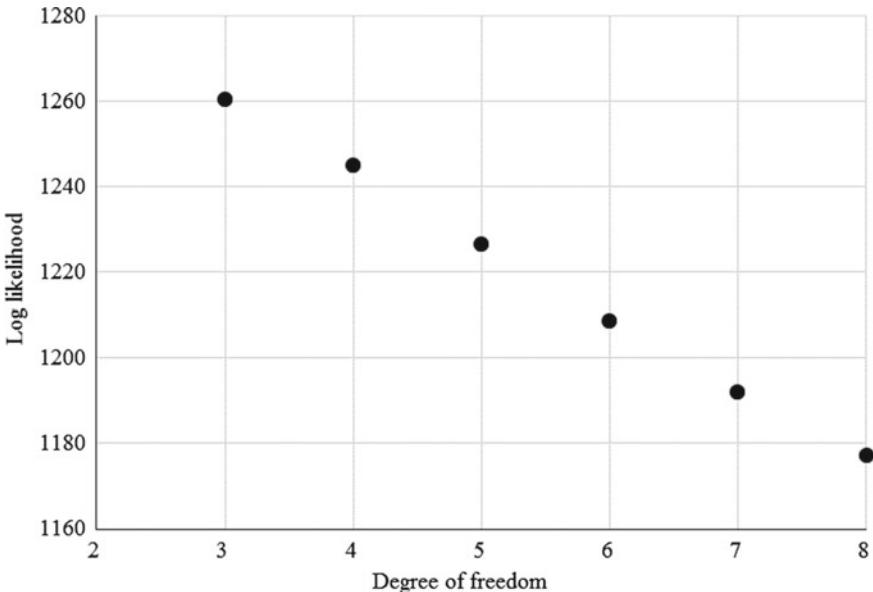


Fig. 2.16 Degree of freedom and log-likelihood in estimating the t copula

The VaR and expected shortfall measured based on the Frank copula, which has the smallest AIC, can be regarded as the risk of statistical arbitrage. Accordingly, the possibility of losing more than 4.79% over a day is less than 1%, and the expected loss is 6.53% if the loss is more than 4.79%. These values aid in risk management, such as portfolio restructuring, position reductions, clearance before maturity, increasing the allowance, and capital expansion.

Figures 2.17, 2.18, 2.19, 2.20, and 2.21 provide scatter plots of the random numbers u_{PJM} and u_{HH} that follow the Gaussian, t , Clayton, Gumbel, and Frank copulas, which we estimate in this section

In adapting to the risk measurement of the portfolio consisting of the Henry Hub futures short positions and the PJM futures long positions, the generation of random numbers is reversed. In other words, $u_{PJM} \rightarrow 0$, $u_{HH} \rightarrow 1$ expresses the left tail of each marginal distribution function, and $u_{PJM} \rightarrow 1$, $u_{HH} \rightarrow 0$ expresses the right tail of each marginal distribution function.

We can observe a difference in the dependency structure between these variables for each type of copula. Although the dots are relatively concentrated in backslash-shaped copulas, these variations are larger in the Gaussian and t copulas. From another point of view, the dots were also generated near the lower left and upper right. This indicates that when one price fluctuates significantly, the other price may fluctuate in the opposite direction. In the Clayton copula, we can see that mutual dependence is extremely strong when the PJM return is positive and the Henry Hub return is negative. In the Gumbel copula, the interconnection is very strong when the PJM return is negative and the Henry Hub return is positive. The Frank copula

Table 2.11 Risk measurement using the estimated copulas

Gaussian	AIC	-7.687	
	$\widehat{\Sigma}$	1.000	0.493
		0.493	1.000
	Value-at-risk ^a	5.46%	
	Expected shortfall ^a	6.64%	
<i>t</i>	AIC	- 6.279	
	Degree of freedom	3	
	$\widehat{\Sigma}$	1.000	0.583
		0.583	1.000
	Value-at-risk ^a	5.56%	
	Expected shortfall ^a	6.83%	
Clayton	AIC	- 10.372	
	π_c	3.163	
	Value-at-risk ^a	3.24%	
	Expected shortfall ^a	4.05%	
Gumbel	AIC	- 10.230	
	π_g	2.689	
	Value-at-risk ^a	3.37%	
	Expected shortfall ^a	4.40%	
Frank	AIC	-10.651	
	π_f	9.342	
	Value-at-risk ^a	4.79%	
	Expected shortfall ^a	6.53%	

Note ^a indicates the value at the 99% confidence level

Bold indicates the minimum AIC and the risk measures at that time

has no distinctive features in the strength of the interdependence. If we know the dependency of each return before estimating some types of copulas, it is rational to adopt the copula that fits the facts instead of selecting a copula type by the AIC.

2.6 Concluding Remarks

Many power generation companies purchase primary energy as fuel and produce and sell electricity as secondary energy. Energy is generally a good that cannot be differentiated except for price. Moreover, both fuel and power are traded in highly liquid, efficient markets. Therefore, we do not interpret power generation companies as electric power manufacturers, but reinterpret them as business operators that add value to energy economics by converting fuel currency to power currency. Like traditional financial firms, they make efficient use of energy derivatives that correspond to their physical position.

Fig. 2.17 Scatter plots of random numbers following the bivariate Gaussian copula

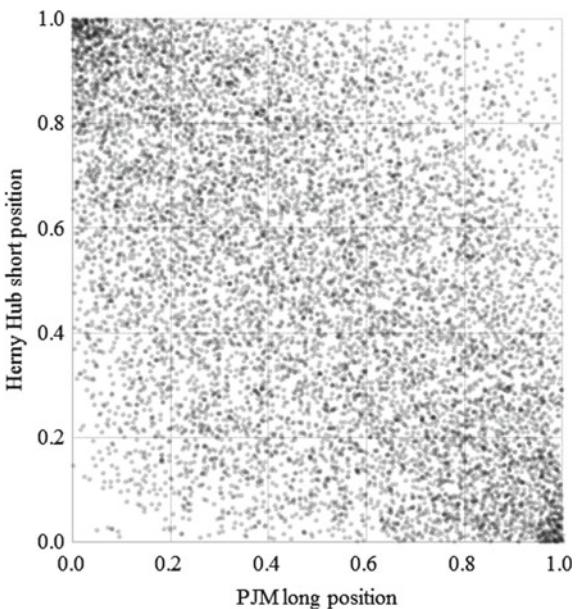


Fig. 2.18 Scatter plots of random numbers following the bivariate t copula

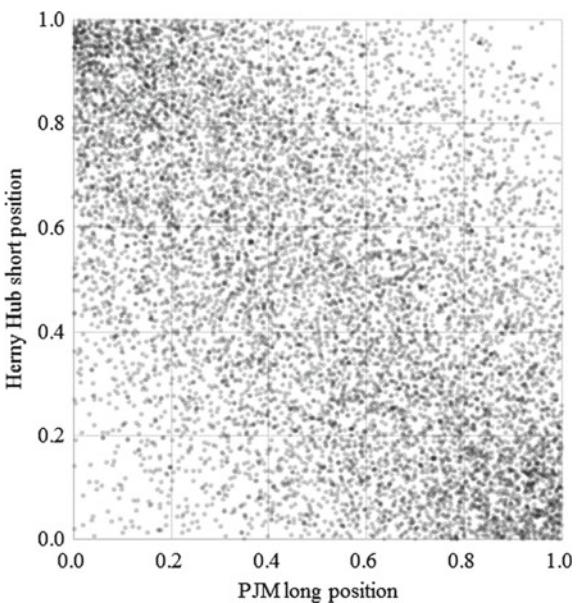


Fig. 2.19 Scatter plots of random numbers following the bivariate Clayton copula

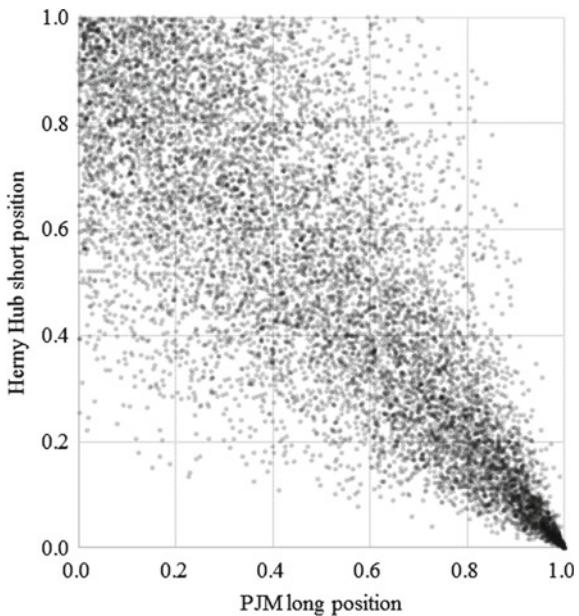


Fig. 2.20 Scatter plots of random numbers following the bivariate Gumbel copula

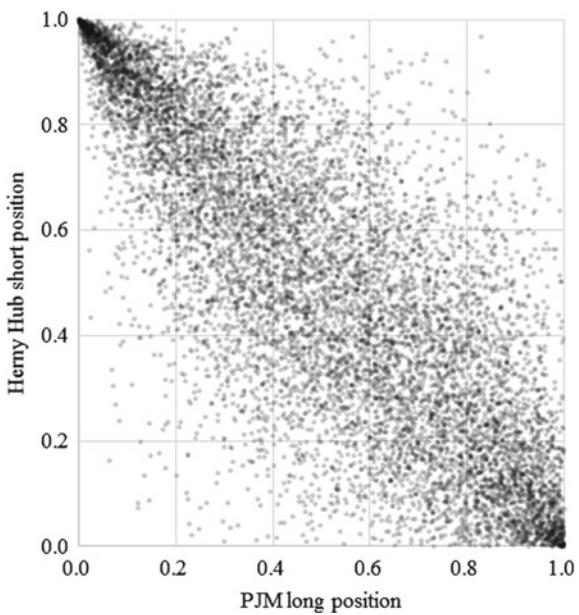
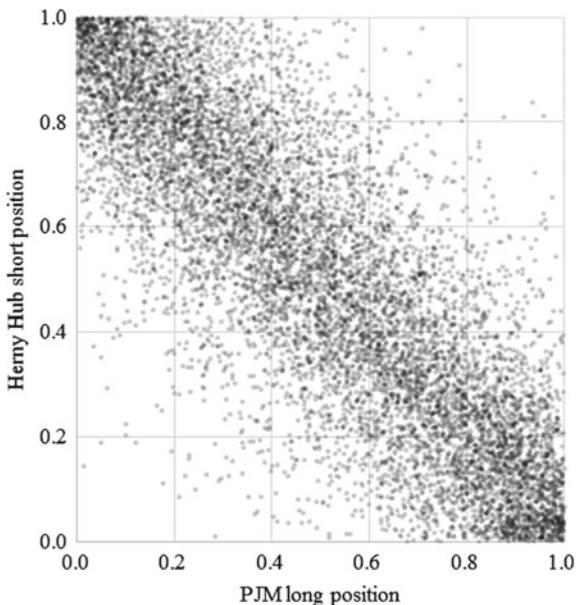


Fig. 2.21 Scatter plots of random numbers following the bivariate Frank copula



This chapter proposes arbitrage between the spot and futures of the price difference between fuel and electricity, and the statistical arbitrage between fuel futures and power futures. In the simulation and risk measurement, we adopt the most primitive model that buys the Henry Hub, one of the most representative natural gas price indicators, and sells the PJM, one of the most representative wholesale electricity indicators.

First, we clarify the characteristics of futures and spot prices by examining their descriptive statistics. The mean and median for each future were higher than those for each spot. Because supply and demand are not so tight during the period, their medians are higher than their respective means in the Henry Hub futures and spot markets. Many outliers are present in the left tail. By contrast, the medians are lower than their respective means in the PJM futures and spot markets. Several outliers may be present in the right tail. Additionally, the ranges of both the Henry Hub futures and PJM futures are narrower than the ranges of each spot market. Moreover, the standard deviation of each spot is larger than that of each future. This is because it is technically difficult or too expensive to store natural gas and electricity. For all four prices, the difference between the maximum and mean is much larger than the difference between the minimum and mean. Moreover, all price series have positive skewness. These distributions have a long right tail. Each variable's kurtosis is larger than 3. All distributions have long fat-tails and a sharp peak.

The time plots of these prices indicate that each spot fluctuates in the same manner as its respective future. However, the spots spike frequently because a simultaneous, equal amount of the supply and demand is needed. Moreover, we can see that the Henry Hub and PJM are synchronized, regardless of their spots and futures.

The Jarque-Berra test rejects the hypothesis that each price series has kurtosis and skewness following a normal distribution. According to the ADF test results, all price series have a unit root and the first difference series does not. The Johansen tests between the Henry Hub and PJM both in the futures and spot markets accept the cointegration hypothesis. Therefore, we expect various trading strategies to utilize their long-term equilibrium relationships.

In this chapter, we propose two trading strategies. The first is arbitrage between spot and futures spreads, selling the PJM futures and buying the Henry Hub futures if the PJM futures price is higher than the power generation cost estimated from the Henry Hub futures price. The other is statistical arbitrage between the Henry Hub futures and PJM futures, buying the PJM futures and selling the Henry Hub futures if the PJM price is estimated by substituting the Henry Hub futures price into the long-term equilibrium equation between the Henry Hub futures price and the PJM futures price is higher than the PJM futures price.

To demonstrate their effectiveness, we simulate both trading strategies using historical data, assuming that a power generation company is inferior to the spot market. According to the simulation, although spot trading alone results in an expected loss of its market subordinated level, adopting both trading strategies can make a large profit instead of only offsetting the expected loss.

Moreover, we measure the risk of a portfolio consisting of Henry Hub future short positions and PJM future long positions held in statistical arbitrage. We estimated five types of copulas that represent the mutual dependency of these two return series after estimating their marginal distributions. Using Monte Carlo simulations with the Frank copula, which is the most suitable for historical data, we calculate the VaR and expected shortfall.

This chapter provides knowledge on the interpretation and handling of time-series data, proposes trading strategies with confirmation of their effectiveness, and teaches how to measure portfolio risk. However, the trading simulation in this section ignores the market impact. Usually, a buy order raises the market price and a sell order lowers the market price. In addition, the impact of energy companies on commodity exchanges is often significant. Therefore, they may not be able to execute trades as expected. It is worth re-examining trading strategies by simulating artificial markets or using a market impact model.

More importantly, there is no limit to the development of trading algorithms aimed at maximizing profits and minimizing risk. The following are five examples of development directions:

- Add futures with different contract months to our portfolio in the spot-future arbitrage.
- Adjust the leverage of statistical arbitrage considering the size of equity capital.
- Testing the cointegration on a moving window sample, which is similar to estimating the cointegrating vector and incorporating the test statistics into the leverage in statistical arbitrage.
- Estimate the volatility of the long-term equilibrium and utilize it for statistical arbitrage and/or clearance conditions.

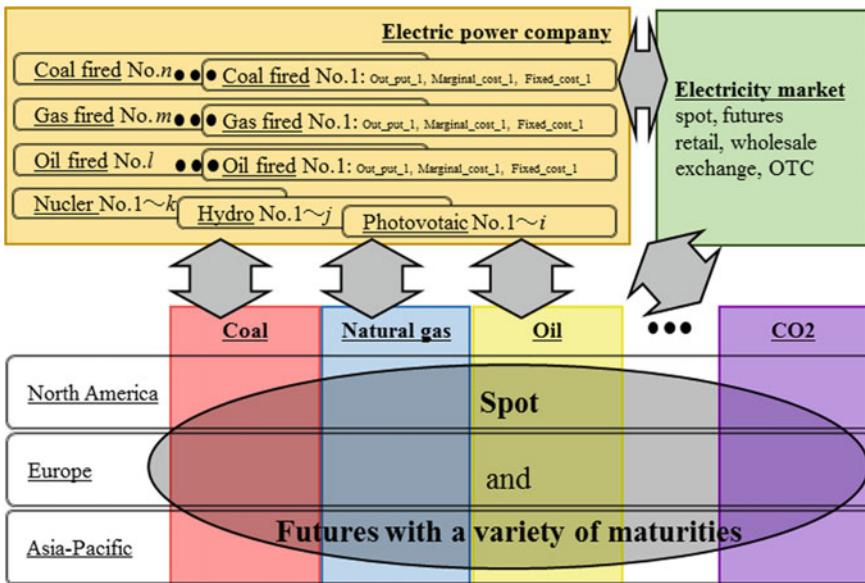


Fig. 2.22 Simplified diagram of an electric power business model

- Change the update frequency of the cointegrating vector estimation in the statistical arbitrage strategy.

Moreover, establishing a more realistic business model is challenging. Figure 2.22 presents a simple and realistic model, but it is still not simple in practice. Actual utilities hold a variety of fuel-type power plants with various power generation efficiencies and trade fuel and power in spot markets and futures with varying maturity and various indicators, including special bilateral contracts. Therefore, we can expect many opportunities for market distortion. In other words, there should be many opportunities to make money at low risk. To maximize profits and minimize risk, we must aim to build rational models and develop algorithms.

References

1. Alexakis, C. (2010). Long-run relations among equity indices under different market conditions: Implications on the implementation of statistical arbitrage strategies. *Journal of International Financial Markets, Institutions and Money*, 20(4), 389–403.
2. Baviera, R., & Baldi, T. S. (2019). Stop-loss and leverage in optimal statistical arbitrage with an application to energy market. *Energy Economics*, 79, 130–143.
3. Bremann, W., Dias, A., & Embrechts, P. (2003). Dependence structures for multivariate high-frequency data in finance. *Quantitative Finance*, 3(1), 1–14.
4. Clayton, D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*, 65(1), 141–151.

5. Dickey, A. D., & Fuller, W. A. (1979). Distribution of the estimates for autoregressive time series with a unit root. *Journal of the American Statistical Society*, 74, 427–431.
6. Embrechts, P., McNeil, A. J., & Straumann, D. (2002). Correlation and dependence in risk management: Properties and pitfalls. In M. A. H. Dempster (Ed.), *Value at risk and beyond* (pp. 176–223). Cambridge University Press.
7. Engle, R. F., & Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, 55(2), 251–276.
8. Fang, H. B., Fang, K. T., & Kotz, S. (2002). The meta-elliptical distributions with given marginals. *Journal of Multivariate Analysis*, 82(1), 1–16.
9. Fernandes, M. C., Diasa, J. C., & Nunes, J. P. V. (2021). Modeling energy prices under energy transition: A novel stochastic-copula approach. *Economic Modelling*, 105, 105671.
10. Focardi, S. M., Fabozzi, F. J., & Mitov, I. K. (2016). A new approach to statistical arbitrage: Strategies based on dynamic factor models of prices and their performance. *Journal of Banking & Finance*, 65, 134–155.
11. Frank, M. J. (1979). On the simultaneous associativity of $F(x, y)$ and $x+y=F(x, y)$. *Aequationes Mathematicae*, 19, 194–226.
12. Ghorbel, A., & Trabelsi, A. (2014). Energy portfolio risk management using time-varying extreme value copula methods. *Economic Modelling*, 38, 470–485.
13. Granger, C. W., & Newbold, P. (1974). Spurious regressions in econometrics. *Journal of Econometrics*, 14, 114–120.
14. Gumbel, E. J. (1960). Bivariate exponential distributions. *Journal of the American Statistical Association*, 55(292), 698–707.
15. Hain, M., Hess, J., & Uhrig-Homburg, M. (2018) Relative value arbitrage in European commodity markets. *Energy Economics*, 69, 140–154.
16. Jarque, C. M., & Bera, A. K. (1987). A test for normality of observations and regression residuals. *International Statistical Review*, 55(2), 163–172.
17. Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica*, 59(6), 1551–1580.
18. Johansen, S., & Juselius, K. (1990). Maximum likelihood estimation and inferences on cointegration with application to the demand for money. *Oxford Bulletin of Economics and Statistics*, 52(2), 169–210.
19. Keilbar, G., & Zhang, Y. (2021). On cointegration and cryptocurrency dynamics. *Digital Finance*, 3, 1–23.
20. Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 54, 159–178.
21. Lee, L. F. (1983). Generalized econometric models with selectivity. *Econometrica*, 51(2), 507–512.
22. Liu, G. D., & Su, C. W. (2019). The dynamic causality between gold and silver prices in China market: A rolling window bootstrap approach. *Finance Research Letters*, 28, 101–106.
23. Lu, X. F., Lai, K. K., & Liang, L. (2014). Portfolio value-at-risk estimation in energy futures markets with time-varying copula-GARCH model. *Annals of Operations Research*, 219(1), 333–357.
24. Ly, S., Sriboonchitta, S., Tang, J., & Wong, W. K. (2022). Exploring dependence structures among European electricity markets: Static and dynamic copula-GARCH and dynamic state-space approaches. *Energy Reports*, 8, 3827–3846.
25. Mayordomo, S., Peña, J. I., & Romo, J. (2014). Testing for statistical arbitrage in credit derivatives markets. *Journal of Empirical Finance*, 26, 59–75.
26. Nakajima, T. (2019). Expectations for statistical arbitrage in energy futures markets. *Journal of Risk and Financial Management*, 12(1), 14.
27. Nelsen, R. B. (2006). *An introduction to copulas*. Springer Science & Business Media.
28. Ophem, H. V. (1999). A general method to estimate correlated discrete random variables. *Econometric Theory*, 15(2), 228–237.
29. Phillips, P. C. B. (1986). Understanding spurious regressions in econometrics. *Journal of Econometrics*, 33(3), 311–340.

30. Phillips, P. C. B., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, 75(2), 335–346.
31. Phillips, P. C. B., Wu, Y., & Yu, J. (2011). Explosive behavior in the 1990s NASDAQ: When did exuberance escalate asset values? *International Economic Review*, 52(1), 201–226.
32. Sánchez-Granero, M. A., Balladares, K. A., Ramos-Requena, J. P., & Trinidad-Segovia, J. E. (2020). Testing the efficient market hypothesis in Latin American stock markets. *Physica A: Statistical Mechanics and its Applications*, 540, 123082.
33. Sklar, A. (1959). Fonctions de Répartition à n Dimensions et Leurs Marges. *Publications de l'Institut Statistique de l'Université de Paris*, 8, 229–231.
34. Soliman, A. M., & Nasir, M. A. (2019). Association between the energy and emission prices: An analysis of EU emission trading system. *Resources Policy*, 61, 369–374.
35. Stübinger, J., & Schneider, L. (2019). Statistical arbitrage with mean-reverting overnight price gaps on high-frequency data of the S&P 500. *Journal of Risk and Financial Management*, 12(2), 51.
36. Zhang, Y. (2018). Investigating dependencies among oil price and tanker market variables by copula-based multivariate models. *Energy*, 161, 435–446.

Chapter 3

Fuel Market Connectedness and Fuel Portfolio Risk



3.1 Introduction

In energy importing countries, many trading companies, oil refineries, electric power companies, and city gas companies procure crude oil and natural gas globally. They must construct and adjust their energy portfolios while considering the spillover effect between each price index because they adopt various price indexes and special pricing formulas. Furthermore, they must measure and manage portfolio risk.

Even for portfolios with many components, its market integration neutralizes the diversified procurement. If the total spillover effect of returns is large, then the portfolio may not bring fuel procurement to stable price levels. The strong connectivity of volatility may insulate a portfolio against any risks incurred in one component. Even under market integration, if the rate of spillover between markets is slow, then monitoring each market may allow for some action.

As these facts illustrate, it is extremely meaningful to analyze the relationships between portfolio component candidates from multiple perspectives. Specifically, measuring the spillover effect of volatility and returns, the connectedness of an entire portfolio as well as between its components, and the spectrum of the spillover effect can give us a rough idea of the portfolio's potential. It is reasonable to narrow down a few probable portfolio candidates from a large number of candidates by measuring the spillover effect with the lowest calculation cost possible; while measuring risk using Monte Carlo simulation is highly accurate, it is very computationally intensive.

A considerable number of studies discuss whether oil markets are integrated. Gülen [6,7] accepts the global crude oil market integration hypothesis using cointegration analysis. Kleit [11] adopts an approach based on arbitrage theory and demonstrates the convergence of the global crude oil market. Reboredo [19] accepts the hypothesis that the global crude oil market is “one great pool” by estimating several copula models with different conditional dependence structures and time-varying dependence parameters. Ji and Fan [9] confirm market integration by constructing a minimal spanning tree for the global crude oil market. Using a threshold unit-root

approach, Fattouh [5] argues that oil markets are not necessarily integrated in every period, despite the generally integrated global crude oil market. Using a threshold cointegration analysis, Hammoudeh et al. [8] find an asymmetric adjustment process, despite the long-run equilibrium relationships between different crude oil benchmarks. Jia et al. [10] analyze the evolutionary features of global crude oil market integration using a proposed novel wavelet-based complex network. Zhang [23] examines the integration between the Chinese and North American crude oil markets by testing for return and volatility spillovers. Nakajima [12] examines the risk transmission between crude oil and petroleum product prices in Japan's oil futures market using the realized volatility and exponential generalized autoregressive conditional heteroscedasticity (EGARCH) model.

Several other studies analyze information efficiency in natural gas markets. Silverstovs et al. [22] and Nakajima and Toyoshima [13] examine the global natural gas market integration between North America, Europe, and Japan. Chai et al. [3] study the relationship between the Chinese and global natural gas markets. Olsen et al. [16], Scarcioffolo and Etienne [20], and Ren et al. [18] investigate integration in the North American natural gas market. Nick [15], Osička et al. [17], and Bastianin et al. [2] discuss natural gas market integration in Europe. Shi et al. [21] analyze the interrelationships of liquefied natural gas (LNG) markets in Asia.

This chapter examines portfolios of crude oil and natural gas as case studies. We cover three representative price indices for each energy type in the North American, European, and Asian markets. First, we provide an overview of each energy market using descriptive statistics and time plots for price, returns, and volatility. Second, we measure the total connectedness of the crude oil and natural gas markets, and the spillover effect between each price index in each market. We then spectrally analyzed the result. In this context, this chapter describes the vector autoregression (VAR) model, vector moving average (VMA) model, connectedness index proposed by Diebold and Yilmaz [4], spectral analysis approach proposed by Baruník and Křehlík [1], and EGARCH model by Nelson [14]. Finally, we measure the risk of each energy portfolio using the copulas introduced in Chap. 2.

The remainder of this chapter is organized as follows. Section 3.2 describes the data used in this chapter. Section 3.3 introduces and explains the methodologies. Section 3.4 provides the results of the analyses and measurements. Section 3.5 discusses the overall concluding remarks and considerations.

3.2 Data

In this chapter, we analyze data representing the European, North American, and Asian markets. For crude oil, we use Brent futures in Europe, West Texas Intermediate (WTI) futures in North America, and Dubai-Oman futures in Asia. For natural gas, we utilize the Title Transfer Facility (TTF) futures in Europe, Henry Hub (HH) futures in North America, and Japan Korea Marker (JKM) futures in Asia. We take advantage of the daily data from January 1, 2018 to December 31, 2019, though we

extract only days when all three crude oil (natural gas) prices are available. The Brent and WTI prices are expressed in USD per barrel. Dubai-Oman prices are given in USD per kiloliter. All natural gas prices were in USD per mmbtu. We obtain all the data from Bloomberg.

3.2.1 Crude Oil

First, we must interpret the descriptive statistics to review the data. Table 3.1 reports the summary statistics of the Brent, WTI, and Dubai-Oman futures prices. Since the mean values for these price series are larger than their respective medians and all price series have positive skewness, we can assume that these distributions have a long right tail, which includes many outliers. The mean ratio of the ranges is 50% for Brent, 52% for WTI, and 45% for Dubai-Oman, with standard deviations of 9.9%, 11%, and 9.6%, respectively. Therefore, while the WTI prices are the most volatile and the Dubai-Oman prices are the most stable, we can conclude that these three price series have almost the same level of dispersion. Because the kurtosis for all variables is less than 3, these distributions have rounded peaks and short, thin tails. The Jarque–Bera test rejects the normal distribution hypothesis for these three variables.

Figure 3.1 shows time plots of these prices. These three prices seem to be completely synchronized and fluctuating, as the descriptive statistics suggest. These three markets can be integrated. Therefore, a portfolio comprising these commodities may be worthless. However, this is merely a prejudice that relies on intuition. Mathematical verification is essential for understanding these facts.

Table 3.1 Descriptive statistics (crude oil prices)

	Brent	WTI	Dubai-Oman
Period	From 1/4/2018 to 12/30/2019		
Observations	470	470	470
Mean	68.01	61.06	417.88
Median	67.02	60.95	409.99
Maximum	86.29	76.41	532.16
Minimum	52.16	44.61	346.15
Standard deviation	6.75	6.59	40.15
Skewness	0.33	0.10	0.55
Kurtosis	2.32	2.20	2.48
Jarque–Bera	17.89 (0.00)	13.47 (0.00)	28.98 (0.00)

Note p -values are in parentheses

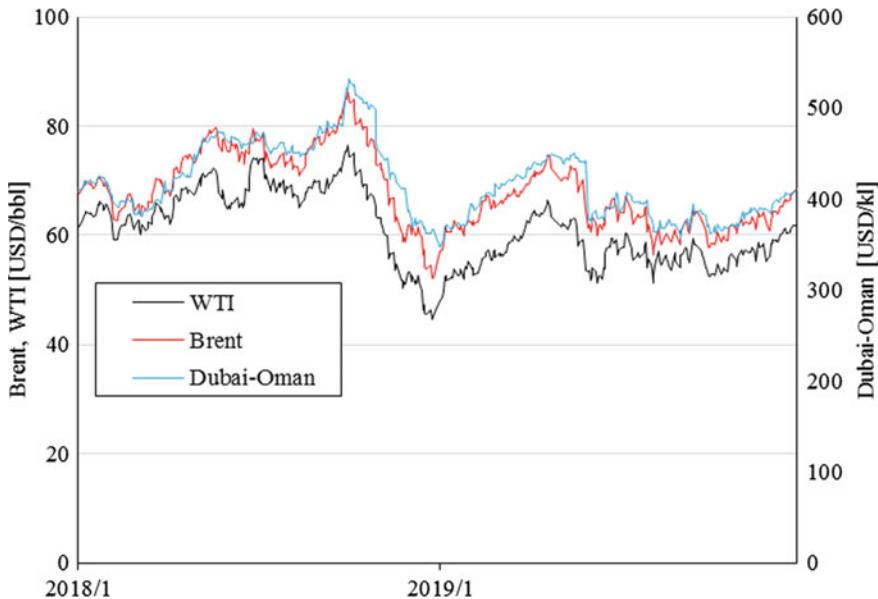


Fig. 3.1 Crude oil prices

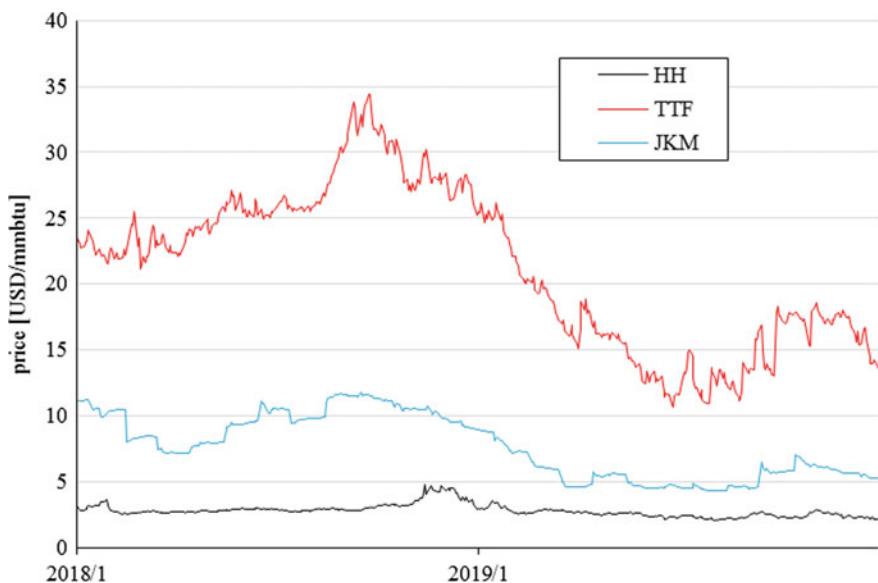
3.2.2 Natural Gas

As in Sect. 3.2.1, we calculate the representative statistics to provide an overview of the data. Table 3.2 lists the summary statistics of the TTF, HH, and JKM futures prices. Because the TTF's mean price is smaller than its median and its skewness is 0.00, the distribution has many outliers in the left tail, while its tails have nearly symmetrical lengths. For the HH and JKM prices, because the means are larger than their medians and their skewness is positive, the distribution has many outliers in the right tail, which is longer than the left tail. The ranges have a mean ratio of 112% for the TTF, 99% for the HH, and 98% for the JKM, with standard deviations of 28, 17, and 31%, respectively. Therefore, although the TTF prices fluctuate the most widely, the TTF and JKM prices have the same level of dispersion. Because the kurtosis for the TTF and JKM is less than 3, their distributions have a rounded peak and short, thin tails. However, the HH price kurtosis is larger than 3, indicating that it has a sharp peak and a thick, long-tailed distribution. The Jarque–Bera test did not support a normal distribution for any variable. Unlike the distribution of the three crude oil price indices, we assume that each index has a characteristic distribution.

Figure 3.2 shows the time plots of the natural gas prices. It can be observed that the price of the TTF swings greatly, while the price of the HH hardly swings. Although these three variables seem to be temporarily synchronized, they seem to fluctuate in three ways. This is consistent with the interpretation of the descriptive

Table 3.2 Descriptive statistics (natural gas prices)

	TTF	HH	JKM
Period	From 1/2/2018 to 12/31/2019		
Observations	504	504	504
Mean	21.29	2.80	7.68
Median	22.29	2.74	7.49
Maximum	34.41	4.84	11.81
Minimum	10.66	2.07	4.28
Standard deviation	5.97	0.48	2.42
Skewness	0.00	1.67	0.11
Kurtosis	1.92	6.78	1.56
Jarque–Bera	24.31 (0.00)	535.86 (0.00)	44.85 (0.00)

**Fig. 3.2** Natural gas prices

statistics. Because these markets do not appear to be integrated, the potential value of a portfolio consisting of these three natural gas indices might be expected.

3.3 Methodology

In this chapter, we employ the connectedness index proposed by Diebold and Yilmaz [4] to examine the spillover effects between the three markets and their integration.

Moreover, to capture the spillover speed between these markets, we apply the spectral analysis approach proposed by Baruník and Křehlík [1] to the connectedness index. We also measure the spillover effect of market risk for crude oil and natural gas by analyzing these volatility series, which we estimate using Nelson's [14] EGARCH model. Finally, we measure the risk of the crude oil and natural gas portfolios using four types of copulas (i.e., Gaussian, t , Clayton, and Gumbel). We calculate the value-at-risk (VaR) and expected shortfall using the daily rate of return distribution of the portfolio generated from 100,000 random numbers for each estimated copula. In this section, we introduce the connectedness index (Diebold and Yilmaz [4]), spectral analysis approach (Baruník and Křehlík [1]), and EGARCH model (Nelson [14]). For the copulas, we refer the reader to Sect. 2.5.2.

3.3.1 Connectedness Index (Diebold and Yilmaz [4])

If \mathbf{x}_t is defined as a stochastic process of N variables, we can describe a general VAR(p) model containing values up to the past time p as follows:

$$\mathbf{x}_t = \Phi_0 + \sum_{j=1}^p \Phi_j \mathbf{x}_{t-j} + \boldsymbol{\varepsilon}_t, \quad (3.1)$$

where Φ_0 is an N -dimensional constant vector, Φ_j is an $N \times N$ coefficient matrix, p is the lag length, and $\boldsymbol{\varepsilon}_t$ is an N -dimensional vector with the following characteristics:

$$\begin{aligned} E(\boldsymbol{\varepsilon}_t) &= \mathbf{O} \\ V(\boldsymbol{\varepsilon}_t) &= E(\boldsymbol{\varepsilon}_t^T \boldsymbol{\varepsilon}_t) = \boldsymbol{\Sigma} \\ E(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_{t-s}) &= \mathbf{O}, \text{ if } s > 0, \end{aligned} \quad (3.2)$$

where $\boldsymbol{\varepsilon}_t N$ is the independently and identically distributed sequence of dimensional random vectors with zero mean and covariance matrix $\boldsymbol{\Sigma}$. \mathbf{x}_t is a covariance stationary process if an N -variable stochastic process satisfies the following three conditions:

Condition 1. The expected value $E(\mathbf{x}_t)$ does not depend on time t .

Condition 2. The variance $V(\mathbf{x}_t) = E((\mathbf{x}_t - E(\mathbf{x}_t))^T (\mathbf{x}_t - E(\mathbf{x}_t)))$ does not depend on time t .

Condition 3. The autocovariance $Cov(\mathbf{x}_t, \mathbf{x}_{t-s})$ does not depend on time t , but only on the time difference $s (> 0)$.

We can estimate the coefficient matrices of the VAR(p) model using ordinary least squares (OLS) as a consistent estimator. It is theoretically correct to select lag order p under the condition that the autocorrelation of the estimated residuals is rejected. However, considering the calculation cost, it is more reasonable to select lag order p based on the information criterion.

Diebold and Yilmaz [4] assumes $\Phi_0 = \mathbf{O}$. We set the lag length p based on the Schwarz Bayesian information criterion (SBIC).

$$\mathbf{x}_t = \sum_{j=1}^p \Phi_j \mathbf{x}_{t-j} + \boldsymbol{\epsilon}_t. \quad (3.3)$$

If \mathbf{x}_t is a covariance stationary vector, then we can represent Eq. (3.3) in the following VMA model to examine the dynamic interdependencies between the vector elements:

$$\mathbf{x}_t = \sum_{j=0}^{\infty} \mathbf{A}_j \boldsymbol{\epsilon}_{t-j}, \quad (3.4)$$

where $\mathbf{A}_i = \sum_{j=1}^p \Phi_j \mathbf{A}_{t-j}$, \mathbf{A}_0 is the $N \times N$ identity matrix and $\mathbf{A}_j = \mathbf{O}$, if $j < 0$.

Using the H -step-ahead forecast error variance decompositions, the spillover effect from the l -th to the k -th variable up to H -step-ahead is

$$\theta_{kl} = \frac{1}{\sigma_{ll}} \frac{\sum_{h=0}^{H-1} ({}^T \mathbf{e}_k \mathbf{A}_h \mathbf{e}_l)^2}{\sum_{h=0}^{H-1} {}^T \mathbf{e}_k \mathbf{A}_h \boldsymbol{\Sigma}^T \mathbf{A}_h \mathbf{e}_k}, \quad (3.5)$$

where σ_{ll} is the standard deviation of the error term for the l -th equation, and \mathbf{e}_k and \mathbf{e}_l are the selection vectors whose k -th and l -th elements are 1, and the others are 0, respectively. Each element of the variance decomposition matrix is normalized by the row sum N as pairwise connectedness $\tilde{\theta}_{kl}$:

$$\tilde{\theta}_{kl} = \frac{\theta_{kl}}{N}. \quad (3.6)$$

Total connectedness S is

$$S = \frac{\sum_{k=1}^N \sum_{l=1, k \neq l}^N \tilde{\theta}_{kl}}{N}. \quad (3.7)$$

The numerator of Eq. (3.7) is the sum of the spillover effects, excluding its effect on itself. The total connectedness indicates the sum of the relative proportion of the portfolio's response to the shock of a variable. Furthermore, we can measure the directional spillover effects S_k received by the k -th variable from all other variables as

$$S_k = \frac{\sum_{k=1, k \neq l}^N \tilde{\theta}_{kl}}{N}. \quad (3.8)$$

Similarly, we can measure the directional spillover effects $S_{.l}$ transmitted by the l -th variable to all other variables as

$$S_{.l} = \frac{\sum_{l=1, k \neq l}^N \tilde{\theta}_{kl}}{N}. \quad (3.9)$$

3.3.2 Spectral Decomposition (Baruník and Křehlík [1])

Adopting the approach introduced by Baruník and Křehlík [1], we spectrally decompose the connectedness indices by Diebold and Yilmaz [4] to grasp when the spillover effect factors occur.

We define the map from the time-domain function $f(t)$ to the frequency domain function $F(\omega)$ by the following equation, which is called the discrete Fourier transform (DFT):

$$F(\omega) = \sum_{t=0}^{N-1} e^{-i\omega t} f(t), \text{ for } 0 \leq \omega < 2\pi. \quad (3.10)$$

The DFT of Eq. (3.5) is

$$\varphi(\omega)_{jk} = \frac{1}{\sigma_{kk}} \frac{\left(\left(\left(\sum_{h=0}^{H-1} e^{-i\omega h} \mathbf{A}_h \right) \boldsymbol{\Sigma} \right)_{jk} \right)^2}{\left(\left(\sum_{h=0}^{H-1} e^{-i\omega h} \mathbf{A}_h \right) \boldsymbol{\Sigma} \left(\sum_{h=0}^{H-1} e^{i\omega h} \mathbf{A}_h \right)' \right)_{jj}}. \quad (3.11)$$

This equation indicates the spillover effect from the k -th variable to the j -th variable up to H -steps-ahead, expressed by the angular frequency ω . The ω component ratio to all the frequency components concerning the spillover to the j -th variable, which is defined as the weighting function $B_j(\omega)$, is calculated as

$$B_j(\omega) = \frac{\left(\left(\sum_{h=0}^{H-1} e^{-i\omega h} \mathbf{A}_h \right) \boldsymbol{\Sigma} \left(\sum_{h=0}^{H-1} e^{i\omega h} \mathbf{A}_h \right)' \right)_{jj}}{\frac{1}{2\pi} \int_0^{2\pi} \left(\left(\sum_{h=0}^{H-1} e^{-i\lambda h} \mathbf{A}_h \right) \boldsymbol{\Sigma} \left(\sum_{h=0}^{H-1} e^{i\lambda h} \mathbf{A}_h \right)' \right)_{jj} d\lambda}. \quad (3.12)$$

We express the spillover effect from the k -th variable to the j -th variable in all the bands as

$$(\theta_\infty)_{jk} = \frac{1}{2\pi} \int_0^{2\pi} B_j(\omega) \varphi(\omega)_{jk} d\omega. \quad (3.13)$$

We calculate this spillover effect in band D is as

$$(\theta_D)_{jk} = \frac{1}{2\pi} \int_D B_j(\omega) \varphi(\omega)_{jk} d\omega. \quad (3.14)$$

We convert $(\theta_D)_{jk}$ to the relative contribution as follows:

$$\left(\tilde{\theta}_D\right)_{jk} = \frac{(\theta_D)_{jk}}{\sum_{j=1}^N (\theta_\infty)_{jk}}. \quad (3.15)$$

The connectedness in band D is

$$C_D = \frac{\sum_{j=1}^N \sum_{k=1, j \neq k}^N \left(\tilde{\theta}_D\right)_{jk}}{\sum_{j=1}^N \sum_{k=1}^N \left(\tilde{\theta}_D\right)_{jk}}. \quad (3.16)$$

We decompose connectedness into short-term (1–5 business days), medium-term (6–20 business days), and long-term (over 21 business days) factors. In other words, Eq. (3.14) becomes

$$\begin{aligned} (\theta_{Short})_{jk} &= \frac{1}{2\pi} \int_{\frac{\pi}{5}}^{\frac{9\pi}{5}} B_j(\omega) \varphi(\omega)_{jk} d\omega \\ (\theta_{Medium})_{jk} &= \frac{1}{2\pi} \int_{\frac{\pi}{20}}^{\frac{\pi}{5}} B_j(\omega) \varphi(\omega)_{jk} d\omega + \frac{1}{2\pi} \int_{\frac{9\pi}{5}}^{\frac{39\pi}{20}} B_j(\omega) \varphi(\omega)_{jk} d\omega \\ (\theta_{Long})_{jk} &= \frac{1}{2\pi} \int_0^{\frac{\pi}{20}} B_j(\omega) \varphi(\omega)_{jk} d\omega + \frac{1}{2\pi} \int_{\frac{39\pi}{20}}^{2\pi} B_j(\omega) \varphi(\omega)_{jk} d\omega. \end{aligned} \quad (3.17)$$

3.3.3 EGARCH Volatility Series Estimation

To measure the risk spillover between markets, we examine the volatility series that represents the second moment of the return series. We can categorize volatility series estimations into three types: autoregressive conditional heteroskedasticity (ARCH), stochastic volatility (SV), and realized volatility (RV) models. The ARCH model is a simultaneous model of the return and latent volatility series. The SV model is a formulation of the logarithmic fluctuation of volatility using an autoregressive

moving average model. The RV series is not estimated by a certain model but is the sum of the squares of the returns in a very short time calculated using high-frequency data over the measurement period. We adopt Nelson's [14] EGARCH model, an ARCH model, here. The EGARCH model is described by the autoregressive (AR)-EGARCH model:

$$\begin{aligned}
 R_t &= c_0 + \sum_{i=1}^k c_i R_{t-i} + \varepsilon_t \\
 \varepsilon_t &= \sqrt{h_t} u_t \\
 u_t &= \frac{\nu}{\lambda 2^{1+1/\nu} \Gamma(1/\nu)} \exp\left(-\frac{1}{2} \left|\frac{z}{\lambda}\right|^\nu\right) \\
 \lambda &= \sqrt{\frac{2^{-2/\nu} \Gamma(1/\nu)}{\Gamma(3/\nu)}} \\
 \ln(h_t) &= \theta + \sum_{i=1}^p \alpha_i \left| \frac{\varepsilon_{t-i}}{\sqrt{h_{t-i}}} \right| + \sum_{i=1}^q \beta_i \ln(h_{t-i}) + \sum_{i=1}^r \gamma_i \frac{\varepsilon_{t-i}}{\sqrt{h_{t-i}}}, \quad (3.18)
 \end{aligned}$$

where R_t is the conditional mean of the return series; h_t is the conditional variance of the return series; ε_t is the prediction error series; $c_i, \theta, \alpha_i, \beta_i$, and γ_i are the AR-EGARCH model parameters; u_t is the generalized error distribution (GED); $\Gamma(\cdot)$ is the Gamma function; and $\nu (> 0)$ is the shape parameter. We can estimate all parameters by the maximum likelihood method. We selected the lag length of each term based on the SBIC. After determining the lag order of the AR term, ARCH term, generalized autoregressive conditional heteroscedasticity (GARCH) term, and asymmetric term, we have the order of the terms. This model offers two advantages. First, the estimated volatility always meets the non-negative constraint regardless of the estimated parameters because we use the logarithmic value of volatility as the dependent variable. Second, this model captures the asymmetry of volatility fluctuations due to positive and negative returns.

3.4 Analysis Results

This section reports the results of the analysis of the crude oil and natural gas portfolios consisting of representative indexes in Europe, North America, and Asia. After the necessary pre-analysis and model estimation, we measure the connectedness indices and perform the spectral decomposition. We examine the spillover effects of both the return and volatility series between these three markets. Finally, we measure the VaR and expected shortfall for each portfolio.

3.4.1 Crude Oil

Diebold and Yilmaz's [4] method requires VMA representation. That is, all variables must be stationary. As described in Chap. 2, ordinary price series are often non-stationary processes. Thus, we apply this technique to the crude oil return series rather than their price series to examine the price spillover effect between crude oil markets. Here, we obtain the return series using the following equation:

$$(Return)_t = \frac{(Price)_t - (Price)_{t-1}}{(Price)_{t-1}}. \quad (3.19)$$

However, in general, the following equation is also common:

$$(Return)_t = \log(Price)_t - \log(Price)_{t-1}. \quad (3.20)$$

Table 3.3 shows the basic statistics for the crude oil return series. The mean and median values for Brent and WTI are significantly higher than those for Dubai-Oman. For all three variables, the median value was greater than the mean. Thus, each distribution appeared to have a considerable number of outliers in the left tail. These characteristics are very strong, especially for Brent and WTI. The range of the Brent return series, which is approximately the same as the range of the WTI return series, is narrower than that of the Dubai-Oman return series. Conversely, the standard deviation for Brent, which is the same level as that for WTI, is larger than that for Dubai-Oman. The dispersions for Brent and WTI tend to be similar. In summary, the Dubai-Oman return series are less variable than the Brent and WTI, although the range is wider because of the extremely small minimum value. As all three variables have negative skewness, their distributions have a long left tail. Because all three return series have kurtosis over 3, their distributions have a sharp peak and long, fat tails. The Jarque–Bera test rejects the hypothesis of normally distributed skewness and kurtosis for each crude oil return series.

Table 3.3 Descriptive statistics (crude oil return series)

	Brent	WTI	Dubai-Oman
Observations	469	469	469
Mean	0.019%	0.018%	0.011%
Median	0.190%	0.152%	0.021%
Maximum	9.31%	8.19%	7.76%
Minimum	– 7.17%	– 7.90%	– 14.41%
Standard deviation	0.0186	0.0197	0.0154
Skewness	– 0.26	– 0.32	– 1.39
Kurtosis	5.79	4.83	23.16
Jarque–Bera	158 (0.00)	73 (0.00)	8094 (0.00)

Note p-values are in parentheses

Figure 3.3 plots the return series. As Table 3.3 implies, this figure shows that the Dubai-Oman return series has some outliers that differ from the other two variables. However, we observe that all three variables tended to be similar.

Table 3.4 provides the results of the unit-root test for the crude oil return series. As the augmented Dickey-Fuller (ADF) test without a constant term and time-trend term rejects the unit-root hypothesis for all variables, we can confirm that all return series are stationary processes. In other words, we can represent the VAR model in the VMA.

We estimate the VAR model of the return series by selecting the lag order based on the SBIC. Table 3.5 summarizes the results. The lag order is 1. We present the variance–covariance matrix of the residuals of the VAR model in Table 3.6.

We depict the spillover indices obtained from Tables 3.5 and 3.6 in Fig. 3.4. It is no exaggeration to say that the Brent and WTI return series are almost linked because their mutual spillover effects are extremely strong, at about 45%. Both Brent and WTI

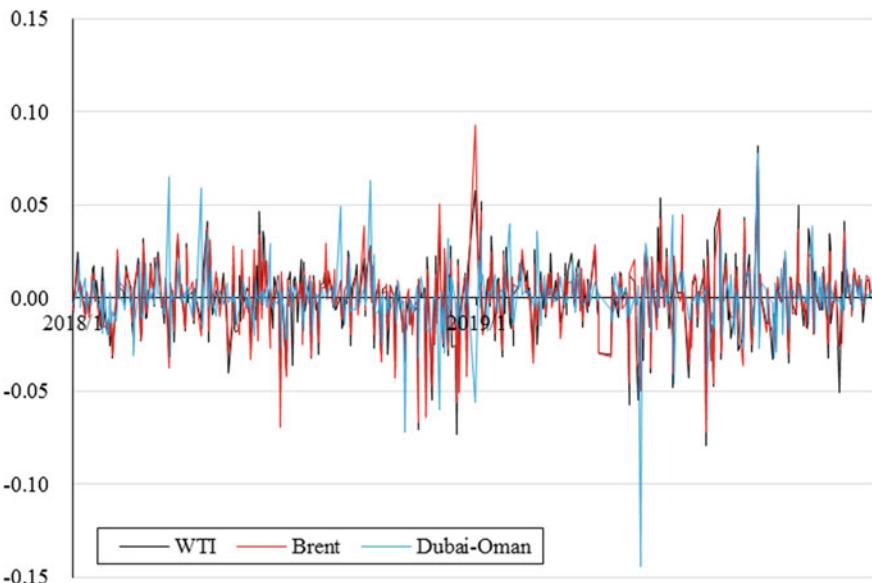


Fig. 3.3 Crude oil return series

Table 3.4 ADF unit-root test results (crude oil return series)

Return series	ADF- <i>t</i> statistics
Brent	-22.55* (0.00)
WTI	-22.48* (0.00)
Dubai-Oman	-23.49* (0.00)

Note * indicates rejection of the unit-root hypothesis at the 1% significance level. *p*-values are in parentheses

Table 3.5 VAR model (crude oil return series)

	Brent (t)	WTI (t)	Dubai-Oman (t)
Brent ($t - 1$)	-0.231	-0.105	0.207
WTI ($t - 1$)	0.203	0.056	0.188
Dubai-Oman ($t - 1$)	-0.090	-0.070	-0.130

Table 3.6 Residual variance–covariance matrix in the VAR model (crude oil return series)

	Brent	WTI	Dubai-Oman
Brent	3.44×10^{-4}	3.34×10^{-4}	3.16×10^{-5}
WTI	3.34×10^{-4}	3.89×10^{-4}	2.97×10^{-5}
Dubai-Oman	3.16×10^{-5}	2.97×10^{-5}	1.83×10^{-4}

have a moderately strong impact on Dubai-Oman, at around 18%. On the other hand, the spillover effect from Dubai-Oman to Brent and WTI is barely observable. The total connectedness for these crude oil markets was 42.95%. We can thus conclude that market integration has progressed to some extent. In terms of returns, the value of a portfolio consisting of these three crude oil indexes might not be so high and might not be much different from a portfolio consisting of Dubai-Oman plus Brent or Dubai-Oman plus WTI.

We present the spectral analysis of the connectedness indices in Table 3.7. The total connectedness from 1 to 5 business days, from 6 to 20 business days, and over 21 business days is 34.97, 6.01, and 1.97%, respectively. Considering that the total

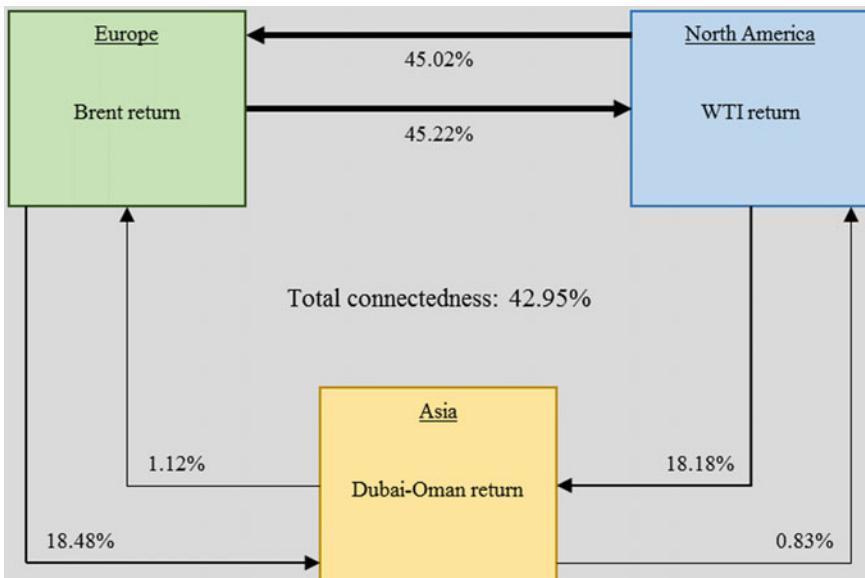
**Fig. 3.4** Spillover effects between return series (crude oil)

Table 3.7 Spillover index and spectral analysis (crude oil return series)

To	From				Bandwidth
	Brent (%)	WTI (%)	Dubai-Oman (%)	Others (%)	
Brent	44.38	36.60	1.07	12.56	$0 \leq \text{date} \leq 5$
WTI	37.44	44.31	0.78	12.74	
Dubai-Oman	14.64	14.37	52.56	9.67	
Others	17.36	16.99	0.62	34.97	
Brent	7.14	6.34	0.04	2.13	$6 \leq \text{date} \leq 20$
WTI	5.87	7.26	0.04	1.97	
Dubai-Oman	2.89	2.86	8.12	1.92	
Others	2.92	3.07	0.03	6.01	
Brent	2.33	2.08	0.01	0.70	$21 \leq \text{date}$
WTI	1.92	2.38	0.01	0.64	
Dubai-Oman	0.95	0.94	2.65	0.63	
Others	0.96	1.01	0.01	1.97	
Brent	53.85	45.02	1.12	15.38	Total
WTI	45.22	53.95	0.83	15.35	
Dubai-Oman	18.48	18.18	63.34	12.22	
Others	21.24	21.07	0.65	42.95	

connectedness for the entire period is 42.95%, the short-term factors contribute the most to the return spillover. The spillover effect depends largely on events occurring within approximately one week. Each spillover effect between these variables tends to be almost the same as total connectedness. The short-, medium- and long-term factors had the largest, second largest, and smallest effects, respectively.

To investigate risk spillover, we measure volatility connectedness. Table 3.8 shows the estimated AR-EGARCH model for generating volatility series. We determine the lag length of each term based on the SBIC. We select the lag order of the AR term, ARCH term, GARCH term, and asymmetric term in this order. Comparing the estimated coefficients, each model for the three variables is similar. The ARCH term (α_1) is statistically significant at the 1% level for Dubai-Oman, but not for Brent and WTI. The GARCH term (β_1) and asymmetric parameter (γ_1) are statistically significant at the 1% level for all variables. The GED parameter is also statistically significant for all variables and is less than 2. This means that these error terms have a fat-tailed distribution.

Table 3.9 lists the descriptive statistics of the crude oil volatility series generated by the estimated EGARCH model. For these crude oil volatility series, the mean is larger than the median; that is, the right tail has a large number of outliers. The Dubai-Oman volatility series has a smaller mean, median, maximum, and minimum than do the Brent and WTI volatility series. This is consistent with the standard deviation of the return series in Table 3.3. Moreover, the standard deviation of the Dubai-Oman

Table 3.8 Estimated AR-EGARCH model (crude oil return series)

		Brent	WTI	Dubai-Oman
Mean equation	c_0	0.001 (0.03)	0.001 (0.11)	0.000 (0.90)
	c_1	-0.061 (0.10)	-0.049 (0.23)	-0.057 (0.00)
Variance equation	θ	-0.320 (0.00)	-0.213 (0.03)	-0.267 (0.00)
	α_1	-0.065 (0.05)	-0.022 (0.47)	-0.148 (0.00)
	β_1	0.956 (0.00)	0.972 (0.00)	0.958 (0.00)
	γ_1	-0.152 (0.00)	-0.116 (0.00)	-0.121 (0.00)
GED	v	1.173 (0.00)	1.273 (0.00)	0.813 (0.00)

Note p -values are in parentheses

volatility series is smaller than the other two series. Skewness is positive for the volatility series for all these indexes. That is, each distribution has a long right tail. Kurtosis is over 3 for Brent and WTI. That is, each distribution has a sharp peak and long, fat tails. The Dubai-Oman kurtosis is less than 3. Its distribution has a rounded peak and short, thin tails. According to the Jarque–Bera statistics, no volatility series is significantly normally distributed in terms of skewness and kurtosis.

Figure 3.5 provides the time plot of the volatility series for the three crude oil indices. The Brent and WTI volatilities fluctuate synchronously at approximately the same level. The Dubai-Oman volatility level is lower than that of the other two series. However, these three volatility series appear synchronized.

Table 3.10 presents the ADF test results for these volatility series. All tests adopted a model without a time-trend term and with a constant term, as in Eq. (2.15). We cannot reject that the WTI and Dubai-Oman volatility series have a unit root at the 1% significance level. However, we reject the unit-root hypothesis for all volatility series at the 10% significance level. These volatility series are considered to be stationary processes. Thus, we apply Diebold and Yilmaz’s [4] approach.

Table 3.9 Descriptive statistics (crude oil volatility series)

	Brent	WTI	Dubai-Oman
Observations	468	468	468
Mean	3.45×10^{-4}	3.96×10^{-4}	1.94×10^{-4}
Median	2.98×10^{-4}	3.52×10^{-4}	1.76×10^{-4}
Maximum	1.10×10^{-3}	1.16×10^{-3}	5.27×10^{-4}
Minimum	8.28×10^{-5}	1.50×10^{-4}	1.45×10^{-5}
Standard deviation	1.78×10^{-4}	1.93×10^{-4}	1.13×10^{-4}
Skewness	1.69	1.81	0.55
Kurtosis	6.35	6.47	2.57
Jarque–Bera	440 (0.00)	491 (0.00)	27 (0.00)

Note p -values are in parentheses

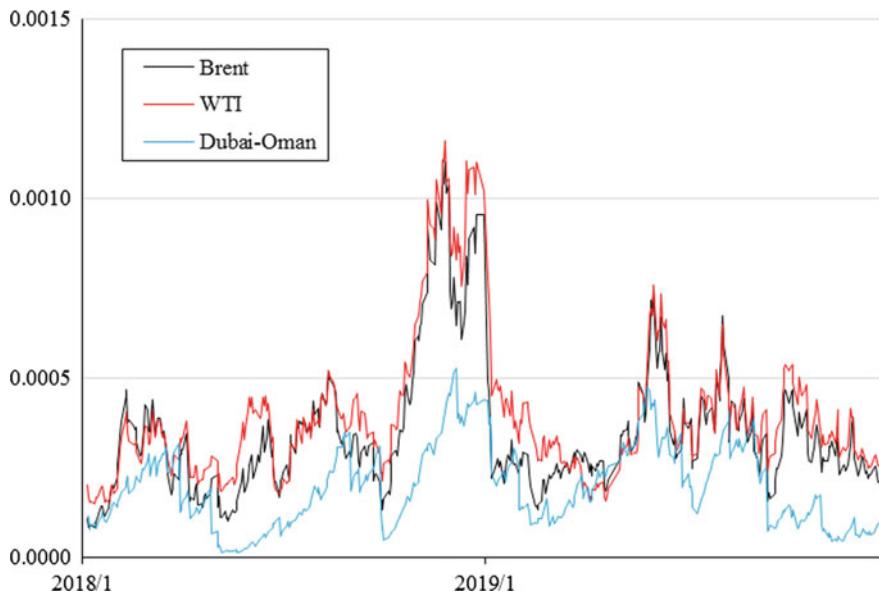


Fig. 3.5 Crude oil volatility series

Table 3.10 ADF unit-root test results (crude oil volatility series)

Volatility series	ADF- <i>t</i> statistics
Brent	-3.44* (0.01)
WTI	-2.79*** (0.06)
Dubai-Oman	-2.88** (0.05)

Note *, **, and *** indicate rejection of the unit-root hypothesis at the 1, 5, and 10% significance level, respectively. *p*-values are in parentheses

We estimate the VAR model for the volatility series after selecting the lag order using the SBIC. We provide the estimates in Table 3.11 and the variance–covariance matrix of the residuals of this VAR model in Table 3.12. We require this matrix to apply Diebold and Yilmaz's [4] approach.

We illustrate the spillover effects calculated using Tables 3.11 and 3.12, in Fig. 3.6. Similar to return spillover, Brent and WTI have a very strong mutual spillover effect. The spillover effect on Dubai-Oman volatility is extremely strong from both Brent

Table 3.11 VAR model (crude oil volatility series)

	Brent (<i>t</i>)	WTI (<i>t</i>)	Dubai-Oman (<i>t</i>)
Brent (<i>t</i> - 1)	0.904	-0.015	0.126
WTI (<i>t</i> - 1)	0.068	0.999	-0.062
Dubai-Oman (<i>t</i> - 1)	0.019	0.018	0.901

Table 3.12 Residual variance–covariance matrix in the VAR model (crude oil volatility series)

	Brent	WTI	Dubai-Oman
Brent	2.97×10^{-9}	2.42×10^{-9}	1.53×10^{-10}
WTI	2.42×10^{-9}	2.39×10^{-9}	1.37×10^{-10}
Dubai-Oman	1.53×10^{-10}	1.37×10^{-10}	8.15×10^{-10}

and WTI volatility. On the other hand, the spillover from Dubai-Oman to the other markets is less than that from either Brent or WTI. The total connectedness of these volatilities is 62.00%. This result implies that these markets are integrated at a considerable level in terms of risk. However, it does not completely deny the value of this portfolio, which consists of the three indexes of risk diversification.

Table 3.13 reports the results of the spectral decomposition of the connectedness indexes shown in Fig. 3.6. The total connectedness of the crude oil volatility series is 62.00%, which is 0.21% for short-term, 0.69% for medium-term, and 61.10% for long-term factors. In contrast to the return series, long-term factors contribute the most to volatility spillover; that is, most of the spillover effect is caused by events that occurred more than one month previously. Each spillover effect between these variables tends to be almost the same as total connectedness. Events within a week have the least impact, events that occurred between a month and two weeks prior have the second lowest impact, and events that occurred more than one month prior had the largest impact.

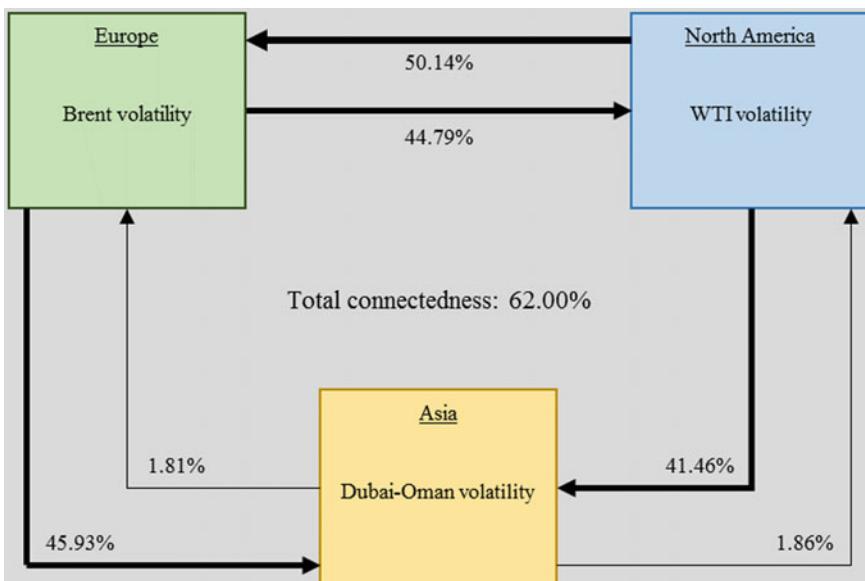


Fig. 3.6 Spillover effects between volatility series (crude oil)

Table 3.13 Spillover index and spectral analysis (crude oil volatilities)

To	From				Bandwidth
	Brent (%)	WTI (%)	Dubai-Oman (%)	Others (%)	
Brent	0.27	0.16	0.00	0.05	$0 \leq \text{date} \leq 5$
WTI	0.05	0.07	0.00	0.02	
Dubai-Oman	0.20	0.21	0.69	0.14	
Others	0.09	0.12	0.00	0.21	
Brent	0.82	0.49	0.00	0.16	$6 \leq \text{date} \leq 20$
WTI	0.17	0.25	0.01	0.06	
Dubai-Oman	0.76	0.63	2.03	0.47	
Others	0.31	0.37	0.00	0.69	
Brent	46.95	49.49	1.81	17.10	21 \leq date
WTI	44.56	53.03	1.86	15.47	
Dubai-Oman	44.96	40.62	9.89	28.53	
Others	29.84	30.04	1.22	61.10	
Brent	48.05	50.14	1.81	17.32	Total
WTI	44.79	53.35	1.86	15.55	
Dubai-Oman	45.93	41.46	12.61	29.13	
Others	30.24	30.53	1.23	62.00	

Various petroleum products and fuel oils used to generate power are often traded at prices linked to the international crude oil index. Therefore, we measure the risk of a portfolio consisting of these crude oil indexes (Brent, WTI, and Dubai-Oman). First, we estimate the Gaussian, t , Clayton, and Gumbel copulas introduced in Chap. 2. Second, we select the most appropriate copula. Finally, we determine the VaR and expected shortfall using random numbers following the copula.

Table 3.14 presents the estimated means and standard deviations, which are the parameters of the marginal distribution function of each return series.

We estimate each copula according to the procedure described in Sect. 2.5.2. While Chap. 2 estimates bivariate copulas, in this chapter, we estimate trivariate copulas.

Figure 3.7 plots the relationship between the degree of freedom and the log-likelihood when estimating the t copula. The maximum log-likelihood occurs when the degree of freedom is 4. Therefore, we estimate the t copula parameters; that is, $\widehat{\Sigma}$, with degree of freedom 4. We report all estimated copulas in Table 3.15.

Table 3.14 Parameters of each marginal distribution function (crude oil returns)

	Brent	WTI	Dubai-Oman
Mean	1.85×10^{-4}	1.84×10^{-4}	1.09×10^{-4}
Standard deviation	1.86×10^{-2}	1.97×10^{-2}	1.54×10^{-2}

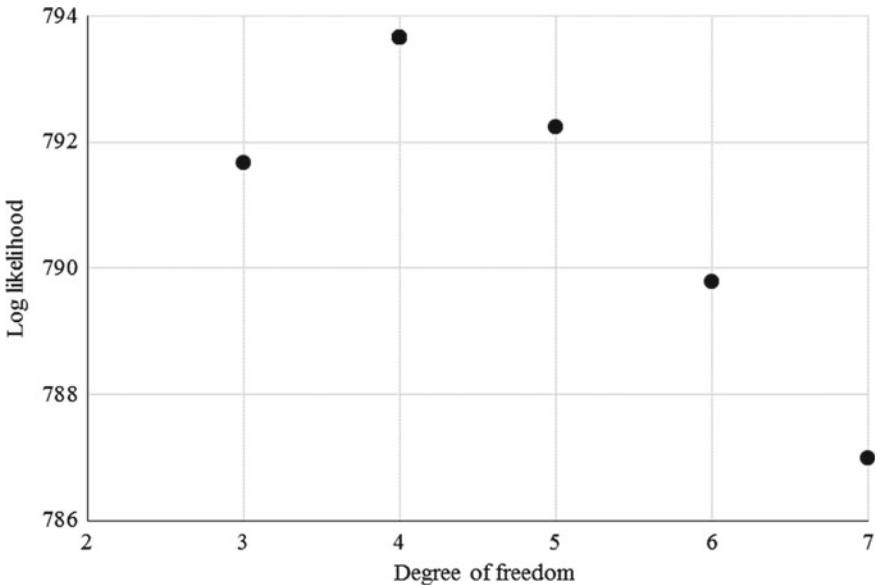


Fig. 3.7 Degree of freedom and log-likelihood, t copula (crude oil portfolio)

We generated the distribution of the daily rate of return for the portfolio by simulating each copula 100,000 times. We then calculate the VaR and expected shortfall. Table 3.15 provides the measured VaRs and expected shortfalls based on these copulas. According to the Gumbel copula with the minimum Akaike information criterion (AIC), the VaR and expected shortfall are 4.81 and 6.07%, respectively. This portfolio has a 1% chance of losing more than 4.81% one day later, with an average loss of 6.07% and a probability of 1%.

Table 3.16 lists the risk for each crude oil index alone. There is no diversification effect that decreases the VaR and expected shortfall of the portfolio to below that of Dubai-Oman; its volatility is much lower (about 50%) than that of Brent and WTI, as Table 3.9 indicates. However, the risk is smaller when we include it in the portfolio than when procuring fuel oil at the price linked to Brent or WTI alone. Firms that have only long positions in Brent or WTI, should consider the risk of such a portfolio.

3.4.2 Natural Gas

Similar to the analysis of the crude oil markets in Sect. 3.4.1, we examine the spillover effect between three representative natural gas price indices in the European, North American, and Asian markets. Moreover, we measure the risk of such a natural gas portfolio.

Table 3.15 Risk measurement using the estimated copulas (crude oil portfolio)

Gaussian	AIC	-1.144		
	$\widehat{\Sigma}$	1.000	0.914	0.078
		0.914	1.000	0.094
		0.078	0.094	1.000
	Value-at-risk*	4.41%		
	Expected shortfall*	5.47%		
<i>t</i>	AIC	0.647		
	Degree of freedom	4		
	$\widehat{\Sigma}$	1.000	0.927	0.073
		0.927	1.000	0.079
		0.073	0.079	1.000
	Value-at-risk*	3.61%		
	Expected shortfall*	4.21%		
Clayton	AIC	- 8.455		
	π_c	1.954		
	Value-at-risk*	4.79%		
	Expected shortfall*	6.03%		
Gumbel	AIC	-9.044		
	π_g	2.326		
	Value-at-risk*	4.81%		
	Expected shortfall*	6.07%		

Note * indicates the value at the 99% confidence level

Bold indicates the minimum AIC and the risk measures at that time

Table 3.16 Risk of each crude oil market

	Brent (%)	WTI (%)	Dubai-Oman (%)
Value-at-risk*	6.34	5.44	4.29
Expected shortfall*	7.87	6.82	5.39

Note * indicates the value at the 99% confidence level

Table 3.17 shows the descriptive statistics for the TTF return series in Europe, the HH return series in North America, and the JKM return series in Asia. All mean values are negative, and all medians are negative, except for JKM, which had a median of 0.00%. We can conclude that all prices have a downward trend during this period. The descending order of the mean value is -0.023% for HH, -0.049% for TTF, and -0.100% for JKM; the descending order of the median value is 0.000% for JKM, -0.070% for HH, and -0.204% for TTF; the descending order of the maximum values is 42.5% for JKM, 36.8% for TTF, and 17.9% for HH; the descending order of the minimum values is -11.8% for TTF, -16.5% for HH, and -23.4% for JKM; and the descending order of the standard deviation is 0.0294 for TTF, 0.0322 for JKM, and

0.0356 for HH. We assume that these three return series have different distributions. Skewness had a positive value for all three variables, indicating a long right tail of each distribution. Kurtosis exceeded 3 for all variables. Their distributions have sharp peaks and long, fat tails. The Jarque–Bera test rejects the normality hypothesis for each distribution.

We plot each return series in Fig. 3.8. Similar to the impression Table 3.17 gives, we can observe that these variables fluctuate without synchronization.

Table 3.18 shows the ADF unit-root test results for the natural gas volatility series. All tests adopt a model that has neither a constant term nor a time-trend term. We can reject the unit-root hypothesis at the 1% significance level in all tests. We can apply the Diebold and Yilmaz [4] approach to these natural gas return series because we confirm the stationarity of all series.

Table 3.19 presents the estimated coefficients of the VAR model for these natural gas return series. The lag order is 1, which we select based on SBIC. Table 3.20 shows the variance–covariance matrix.

Figure 3.9 illustrates the spillover effects, calculated by applying Tables 3.19 and 3.20 to Eq. (3.5). The strongest spillover effect is 2.91% from the JKM return series to the TTF return series, and the second-strongest spillover effect is 1.50% from the TTF return series to the JKM return series. All the other spillover effects are less than 1%. Moreover, total connectedness has a small value, of 1.72%. These markets are not integrated in terms of returns. The value of a portfolio consisting of these three natural gas indices may be more attractive than holding each index alone. This is because we can expect a diversification effect from this procurement approach.

Table 3.21 lists the spectral analysis results. The total connectedness is 1.72%, which consists of 1.56% for short-term factors, 0.12% for medium-term factors, and 0.04% for long-term factors. Hence, the short-term factors explain total connectedness. Each spillover effect has the same tendency. The largest is due to short-term factors, and the weakest is due to long-term factors. However, the significance of the spectral decomposition might be low because the total connectedness is extremely small.

Table 3.17 Descriptive statistics (natural gas return series)

	TTF	HH	JKM
Observations	503	503	503
Mean	– 0.049%	– 0.023%	– 0.100%
Median	– 0.204%	– 0.070%	0.000%
Maximum	36.8%	17.9%	42.5%
Minimum	– 11.8%	– 16.5%	– 23.4%
Standard deviation	0.0356	0.0294	0.0322
Skewness	2.96	0.13	4.86
Kurtosis	29.33	10.92	76.63
Jarque–Bera	15,264 (0.00)	1316 (0.00)	115,608 (0.00)

Note p -values are in parentheses

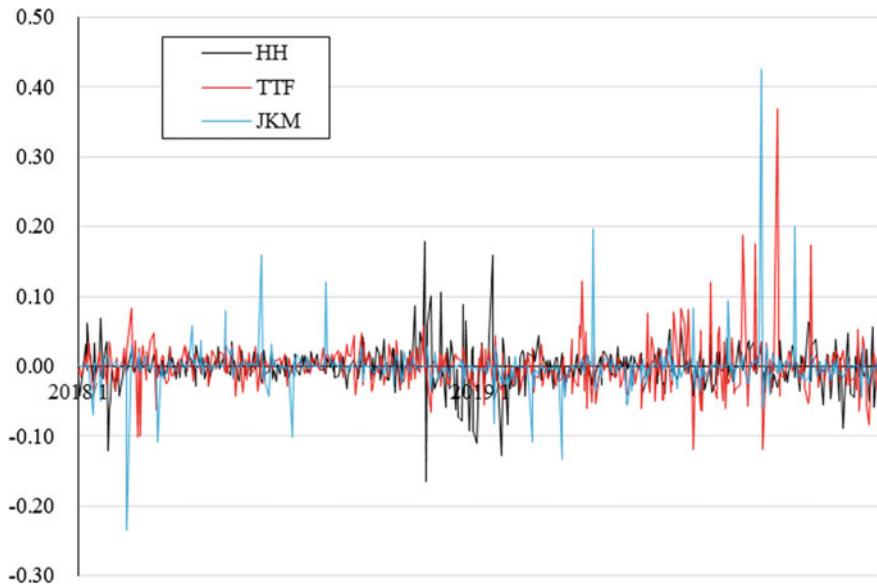


Fig. 3.8 Natural gas return series

Table 3.18 ADF unit-root test results (natural gas return series)

Return	ADF- <i>t</i> statistics
TTF	-21.53* (0.00)
HH	-12.32* (0.00)
JKM	-23.28* (0.00)

Note * indicates rejection of the unit-root hypothesis at the 1% significance level. *p*-values are in parentheses

Table 3.19 VAR model (natural gas returns)

	TTF (<i>t</i>)	HH (<i>t</i>)	JKM (<i>t</i>)
TTF (<i>t</i> - 1)	0.055	0.056	0.008
HH (<i>t</i> - 1)	0.059	-0.049	0.025
JKM (<i>t</i> - 1)	-0.141	-0.010	-0.040

Table 3.20 Residual variance–covariance matrix in the VAR model (natural gas returns)

	TTF	HH	JKM
TTF	1.25×10^{-3}	-9.95×10^{-6}	1.41×10^{-4}
HH	-9.95×10^{-6}	8.64×10^{-4}	-1.16×10^{-7}
JKM	1.41×10^{-4}	-1.16×10^{-7}	1.04×10^{-3}

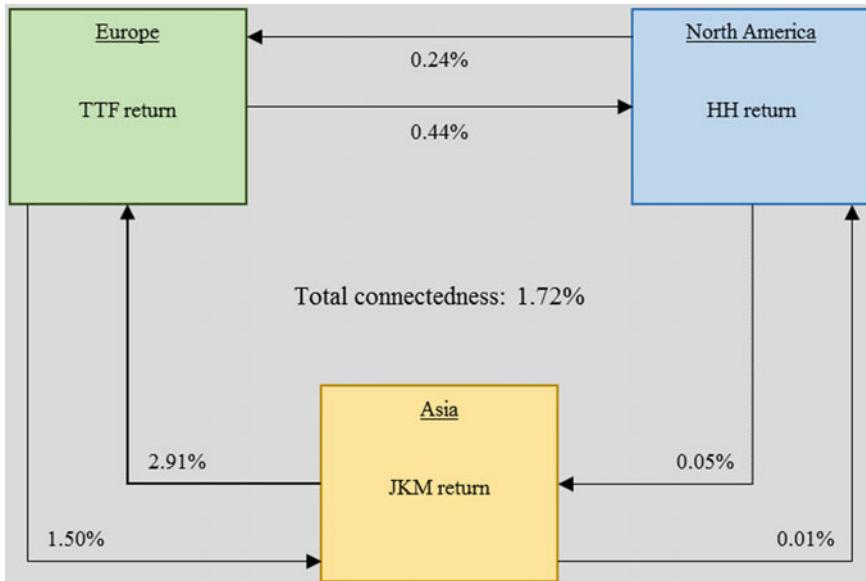


Fig. 3.9 Spillover effects between return series (natural gas)

Table 3.21 Spillover index and spectral analysis (natural gas returns)

To	From				Bandwidth
	TTF (%)	HH (%)	JKM (%)	Others (%)	
TTF	75.92	0.21	2.87	1.03	$0 \leq \text{date} \leq 5$
HH	0.38	81.35	0.00	0.13	
JKM	1.18	0.04	80.19	0.41	
Others	0.52	0.08	0.96	1.56	
TTF	15.66	0.02	0.04	0.02	$6 \leq \text{date} \leq 20$
HH	0.05	13.67	0.00	0.02	
JKM	0.24	0.01	13.72	0.08	
Others	0.10	0.01	0.01	0.12	
TTF	5.27	0.01	0.00	0.00	$21 \leq \text{date}$
HH	0.02	4.53	0.00	0.01	
JKM	0.08	0.00	4.54	0.03	
Others	0.03	0.00	0.00	0.04	
TTF	96.85	0.24	2.91	1.05	Total
HH	0.44	99.55	0.01	0.15	
JKM	1.50	0.05	98.45	0.52	
Others	0.65	0.10	0.97	1.72	

Table 3.22 Estimated AR-EGARCH model (Natural gas return series)

		TTF	HH	JKM
Mean equation	c_0	-0.001 (0.51)	0.000 (0.65)	0.000 (1.00)
	c_1	0.017 (0.67)	-0.043 (0.35)	0.000 (0.21)
	c_2		-0.059 (0.18)	
	c_3		0.043 (0.34)	
Variance equation	θ	-0.277 (0.00)	-0.247 (0.01)	-4.655 (0.02)
	α_1	0.200 (0.00)	0.200 (0.00)	2.051 (0.25)
	β_1	0.981 (0.00)	0.988 (0.00)	-0.207 (0.63)
	β_2			-0.229 (0.45)
	γ_1	-0.020 (0.59)	0.078 (0.00)	-0.509 (0.75)
GED	v	1.154 (0.00)	1.447 (0.00)	0.103 (0.00)

Note p -values are in parentheses

Next, we examine the spread of risk. We generate the volatility series for each natural gas index, which we require to measure the connectedness of the volatility. We estimate the AR-EGARCH model of natural gas returns and select the lag order of the model based on the SBIC. We select the lag order of the AR term, ARCH term, GARCH term, and asymmetric term in this order. The estimated GED parameter is less than 2. Table 3.22 presents the estimated parameters of the AR-EGARCH model.

Table 3.23 provides the basic statistics of these natural gas volatility series, which we generated using the estimated EGARCH models. The mean, median, maximum, and minimum for the JKM is two orders of magnitude larger than those of the TTF and HH. These results express the instability of the JKM return series. Skewness for all variables is positive. That means that all distributions have a long right tail. Kurtosis for all variables is over 3, indicating that all distributions have a sharp peak and long, fat tails. According to the Jarque–Bera test, none of the volatility series are significantly normally distributed.

Figure 3.10 displays a time plot of the volatility series. The TTF series had small peaks in March 2018 and April 2019, and increased from June to October 2019. The HH series exhibited a large peak at the end of 2018. The JKM series has some very large spikes, whereas it barely fluctuated from approximately 0.04. All three variables have different, though characteristic, fluctuations.

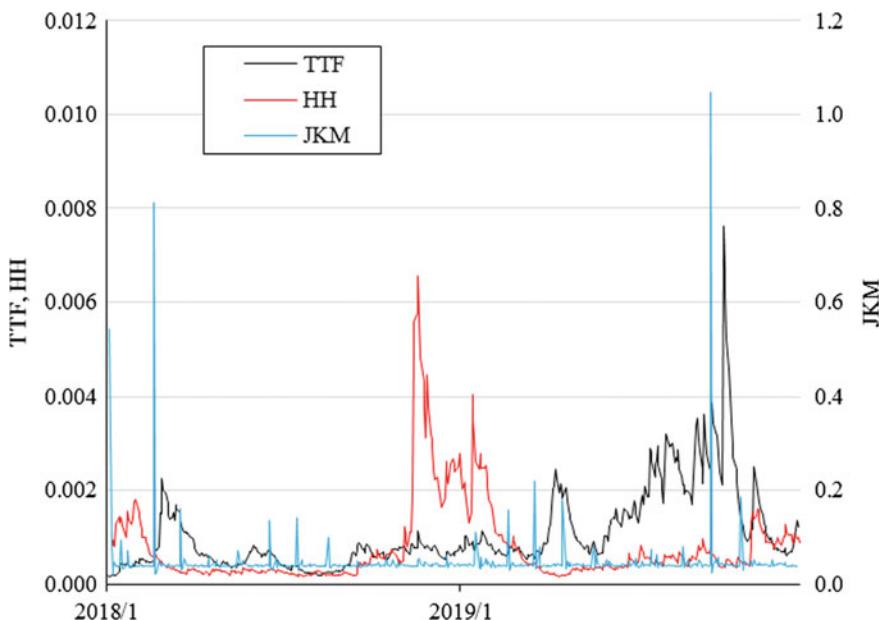
We test whether Diebold and Yilmaz's [4] approach is applicable to these volatility series. The ADF test, which uses a model with a constant term and without a time-trend term, rejects the unit-root hypothesis for all volatility series. Table 3.24 shows the ADF unit-root test results. The vector of these volatility series is a stationary covariance process. Therefore, we can represent the VAR model as a VMA model.

We estimate a VAR model for the natural gas volatility series using the lag order chosen based on the SBIC. We summarize the estimated coefficients in Table 3.25 and provide the variance–covariance matrix of the residuals of this VAR model in Table 3.26. We require these residuals to apply the Diebold and Yilmaz [4] approach.

Table 3.23 Descriptive statistics (natural gas volatility series)

	TTF	HH	JKM
Observations	502	500	502
Mean	1.10×10^{-3}	7.87×10^{-4}	4.81×10^{-2}
Median	7.65×10^{-4}	4.61×10^{-4}	4.00×10^{-2}
Maximum	7.61×10^{-3}	6.56×10^{-3}	1.05
Minimum	1.67×10^{-4}	1.63×10^{-4}	1.83×10^{-2}
Standard deviation	9.13×10^{-4}	8.83×10^{-4}	6.32×10^{-2}
Skewness	2.54	3.08	12.63
Kurtosis	13.07	14.81	177.74
Jarque–Bera	2662 (0.00)	3694 (0.00)	6.52×10^5 (0.00)

Note *p*-values are in parentheses

**Fig. 3.10** Natural gas volatility series**Table 3.24** ADF unit-root test results (natural gas volatility series)

Volatility series	ADF- <i>t</i> statistics
TTF	-4.12* (0.00)
HH	-3.64* (0.01)
JKM	-24.47* (0.00)

Note * indicates rejection of the unit-root hypothesis at the 1% significance level. *p*-values are in parentheses

Table 3.25 VAR model (natural gas volatility series)

	TTF (t)	HH (t)	JKM (t)
TTF ($t - 1$)	0.948	0.010	18.547
HH ($t - 1$)	0.008	0.972	12.675
JKM ($t - 1$)	0.001	0.000	0.109

Table 3.26 Residual variance–covariance matrix in the VAR model (natural gas volatility series)

	TTF	HH	JKM
TTF	1.03×10^{-7}	1.36×10^{-9}	5.55×10^{-7}
HH	1.36×10^{-9}	6.10×10^{-8}	5.42×10^{-7}
JKM	5.55×10^{-7}	5.42×10^{-7}	4.04×10^{-3}

Figure 3.11 depicts the connectedness indexes calculated from Tables 3.25 and 3.26. The total connectedness is 16.90%, which is stronger than the total connectedness of the return series. The risk impact of the HH on the other two indexes is significant, at 10.66% for the TTF and 10.17% for the JKM. The volatility spillover from the TTF to the JKM is the strongest, at 16.59%, while the volatility spillover from the TTF to the HH is not large, at 7.49%. The volatility spillover from the JKM to the TTF is 4.85%, and the volatility spillover from the JKM to the HH is 0.96%. The JKM does not have a large risk impact on the natural gas market.

Table 3.27 reports the spectral analysis of the connectedness indexes of the natural gas volatility series. The total connectedness during the full period is 16.90%, while

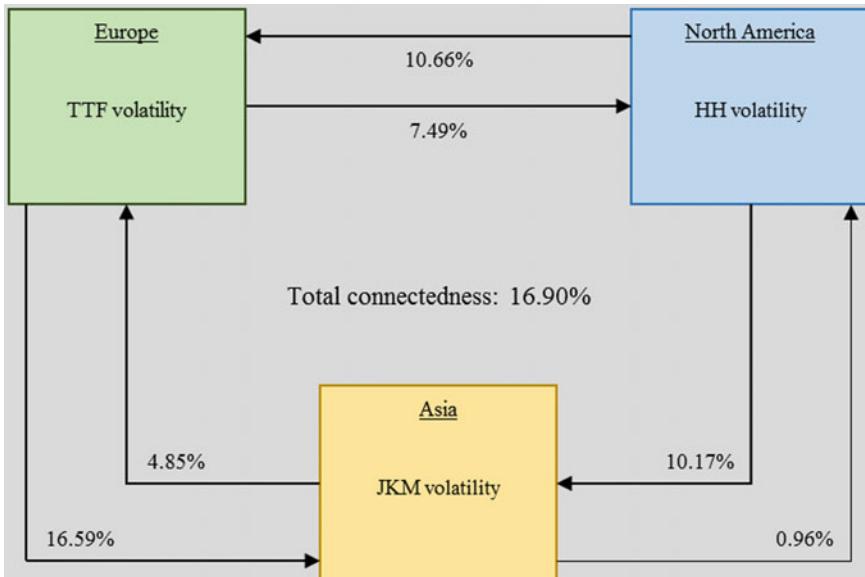


Fig. 3.11 Spillover effects between volatility series (natural gas)

the total connectedness from 1 to 5 business days, from 6 to 20 business day, and over 21 business days is 0.14, 0.56, and 16.21%, respectively. The volatility spillover effect depends mostly on events before 21 business days. The spectral decomposition of each spillover tends to be the same as of total connectedness. The short-term factor is the smallest, followed by the medium- and long-term factors.

LNG transported by ships and natural gas supplied in pipelines is often traded at prices linked to the international natural gas index. Similar to how we measured the risk of the crude oil portfolio in Sect. 3.4.1, we measure the risk of a portfolio consisting of the TTF, HH, and JKM. We obtain the VaR and expected shortfall of this natural gas portfolio by simulation using random numbers following the Gaussian, t , Clayton, and Gumbel copulas. We select the VaR and expected shortfall obtained by simulation using the copula with the smallest AIC.

Table 3.28 presents the means and standard deviations, which are the parameters of the marginal distribution function of each return series.

Table 3.27 Spillover index and spectral analysis (natural gas volatility series)

To	From				Bandwidth
	TTF (%)	HH (%)	JKM (%)	Others (%)	
TTF	1.70	0.07	0.09	0.05	$0 \leq \text{date} \leq 5$
HH	0.05	1.36	0.00	0.02	
JKM	0.19	0.02	53.02	0.07	
Others	0.08	0.03	0.03	0.14	
TTF	6.23	0.30	0.30	0.20	$6 \leq \text{date} \leq 20$
HH	0.19	5.19	0.00	0.07	
JKM	0.79	0.10	12.39	0.30	
Others	0.33	0.13	0.10	0.56	
TTF	76.56	10.29	4.46	4.92	$21 \leq \text{date}$
HH	7.25	85.01	0.95	2.73	
JKM	15.61	10.06	7.83	8.56	
Others	7.62	6.78	1.80	16.21	
TTF	84.49	10.66	4.85	5.17	Total
HH	7.49	91.56	0.96	2.81	
JKM	16.59	10.17	73.24	8.92	
Others	8.02	6.94	1.93	16.90	

Table 3.28 Parameters of each marginal distribution function (natural gas return series)

	TTF	HH	JKM
Mean	-4.92×10^{-4}	-2.30×10^{-4}	-1.00×10^{-3}
Standard deviation	3.56×10^{-2}	2.94×10^{-2}	3.22×10^{-2}

First, we estimate four trivariate copulas: the Gaussian, t , Clayton, and Gumbel copulas. Figure 3.12 provides a graph of the relationship between the degree of freedom and the log-likelihood for the t copula estimate. The maximum log-likelihood occurs when the degree of freedom is 3. Therefore, the t copula parameters are the elements of the correlation matrix $\hat{\Sigma}$ estimated with degree of freedom 3. We summarize all estimated copulas and measured risk values in Table 3.29. The VaR and expected shortfall are 8.39 and 10.65%, respectively. We measure these using random numbers following the Gumbel copula with the smallest AIC. This natural gas portfolio has a 1% chance of losing more than 8.39% one day later, with an average loss of 10.65% and a probability of 1%. Compared to the VaR of 4.81% and expected shortfall of 6.87% of the crude oil portfolio, this natural gas portfolio carries higher risk. The risk of each natural gas index in Table 3.30 is much higher than the risk of each crude oil index in Table 3.16. The reason that the natural gas portfolio is of higher risk than the crude oil portfolio does not seem to be that the diversification of natural gas procurement is ineffective. Rather, although the risk of TTF or JKM alone is high, it is possible to reduce the HH risk level by forming this natural gas portfolio.

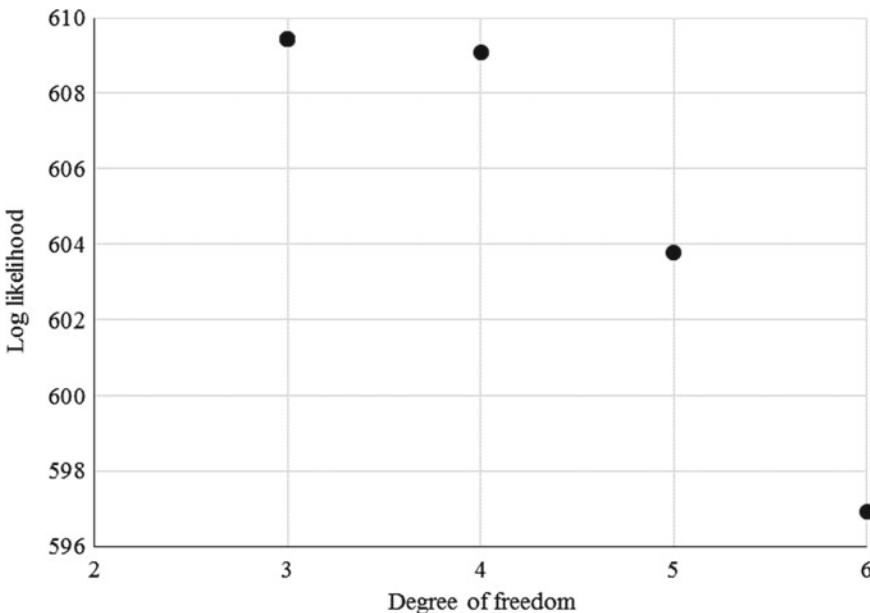


Fig. 3.12 Degree of freedom and log-likelihood in estimating the t copula (natural gas portfolio)

Table 3.29 Risk measurement using the estimated copulas (natural gas portfolio)

Gaussian	AIC	– 0.180		
	$\widehat{\Sigma}$	1.000	0.004	–0.002
		0.004	1.000	0.218
		–0.002	0.218	1.000
	Value-at-risk*	5.56%		
	Expected shortfall*	6.71%		
<i>t</i>	AIC	1.175		
	Degree of freedom	3		
	$\widehat{\Sigma}$	1.000	0.021	0.037
		0.021	1.000	0.285
		0.037	0.285	1.000
	Value-at-risk*	5.28%		
	Expected shortfall*	6.34%		
Clayton	AIC	– 7.301		
	π_c	1.530		
	Value-at-risk*	8.64%		
	Expected shortfall*	10.91%		
Gumbel	AIC	–8.106		
	π_g	2.090		
	Value-at-risk*	8.39%		
	Expected shortfall*	10.65%		

Note * indicates the value at the 99% confidence level

Bold indicates the minimum AIC and the risk measures at that time

Table 3.30 Risk of each natural gas market

	TTF (%)	HH (%)	JKM (%)
Value-at-risk*	9.88	8.15	9.21
Expected shortfall*	12.28	10.24	11.51

Note * indicates the value at the 99% confidence level

3.5 Concluding Remarks

When we procure different fuels at prices linked to different price indicators, we should understand the spillover effect of the price level and the risk between these indicators. We should determine the diversification effect of procurement based on the total connectedness of an energy portfolio. We must monitor the connectedness of the portfolio, including potential price indicators and grasp the impact of excluding components already in our portfolio. Quantitative measures of risk of such a potential portfolio is extremely important. It is beneficial to consider what components we

should have and how much we should hold to reconstruct the portfolio. Moreover, the measurement results may affect the financial strategy.

This chapter describes the connectedness index proposed by Diebold and Yilmaz [4], which can capture the spillover effect between markets. We then introduce Baruník and Křehlík's [1] spectral decomposition method, which determines when spillover factors occur. Moreover, we explain the EGARCH model, which can generate each volatility series needed to examine the risk spillover effects.

We illustrate two cases: a crude oil portfolio and a natural gas portfolio. We examine each portfolio, which consists of representative markets for Europe, North America, and Asia; specifically, Brent, WTI, and Dubai-Oman to represent crude oil markets and the TTF, HH, and JKM to represent natural gas markets, respectively.

First, we describe the crude oil markets. While Dubai-Oman's returns and volatility are smaller than those of Brent or WTI, the price series, return series, and volatility series of these three indicators appear to fluctuate synchronously. The total connectedness of the return and volatilities series are 42.95 and 62.00%, respectively. These crude oil markets appear to be integrated at a relatively high level. The spillover effects of returns and volatility between Brent and WTI are mutually strong. Dubai-Oman receives considerable risk from both Brent and WTI.

We next describe the natural gas markets. The prices, returns, and volatility for the TTF, HH, and JKM fluctuate in three different ways. Although the total connectedness of volatility is 16.90%, that of returns is 1.72%. This result indicates that intercontinental natural gas market liquidity may still be low. Any spillover effect of returns is less than 1%, except between TTF and JKM. The volatility spillover effects larger than 10%, are from HH to TTF, TTF to JKM, and HH to JKM.

Third, we describe the results of the spectral decomposition of connectedness. Both the crude oil and natural gas analysis results indicate that the spillover effects of the return series depend mostly on short-term factors (i.e., events within 5 business days), while the spillover effects of the volatility series depend mostly on long-term factors (i.e., events more than a month previously).

Finally, we describe the risk measurement results. The crude oil portfolio has a VaR of 4.81% and an expected shortfall of 6.07%, whereas the natural gas portfolio has a VaR of 8.39% and an expected shortfall of 10.65%. The risk of the natural gas portfolio is higher than that of the crude oil portfolio because the procurement diversification effect of natural gas is less than that of crude oil as natural gas has much higher risk than does crude oil.

Readers should understand how to monitor markets and measure portfolio risk when designing a portfolio. In practice, we should handle more variables and monitor connectedness more frequently, that is, with the observation period as a moving window. While we analyze only a fuel procurement portfolio, we can also measure the risk of an electricity sales portfolio by keeping in mind that the right tail of the return distribution is a risk. Electricity retailers often sell power using a variety of price formulas, including fixed prices. The approaches and exercises introduced in this chapter are meaningful for such practitioners.

References

1. Baruník, J., & Křehlík, T. (2018). Measuring the frequency dynamics of financial connectedness and systemic risk. *Journal of Financial Econometrics*, 16(2), 271–296.
2. Bastianin, A., Galeotti, M., & Polo, M. (2019). Convergence of European natural gas prices. *Energy Economics*, 81, 793–811.
3. Chai, J., Wei, Z., Hu, Y., Su, S., & Zhang, Z. G. (2019). Is China's natural gas market globally connected? *Energy Policy*, 132, 940–949.
4. Diebold, F. X., & Yilmaz, K. (2012). Better to give than to receive: Predictive directional measurement of volatility spillovers. *International Journal of Forecasting*, 28(1), 57–66.
5. Fattouh, B. (2010). The dynamics of crude oil price differentials. *Energy Economics*, 32(2), 334–342.
6. Gülen, S. G. (1997). Regionalization in the world crude oil market. *Energy Journal*, 18(2), 109–126.
7. Gülen, S. G. (1999). Regionalization in the world crude oil market: Further results. *Energy Journal*, 20(1), 125–139.
8. Hammoudeh, S., Thompson, M., & Ewing, B. (2008). Threshold cointegration analysis of crude oil benchmarks. *Energy Journal*, 29(4), 79–95.
9. Ji, Q., & Fan, Y. (2016). Evolution of world crude oil market integration: A graph theory analysis. *Energy Economics*, 53, 90–100.
10. Jia, X., An, H., Sun, X., Huang, X., & Wang, L. (2017). Evolution of world crude oil market integration and diversification: A wavelet-based complex network perspective. *Applied Energy*, 185(2), 1788–1798.
11. Kleit, A. N. (2001). Are regional oil markets growing closer together? An arbitrage cost approach. *Energy Journal*, 22(2), 1–15.
12. Nakajima, T. (2019). Test for volatility spillover effects in Japan's oil futures markets by a realized variance approach. *Studies in Economics and Finance*, 36(2), 224–239.
13. Nakajima, T., & Toyoshima, Y. (2019). Measurement of connectedness and frequency dynamics in global natural gas markets. *Energies*, 12(20), 3927.
14. Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59(2), 347–370.
15. Nick, S. (2016). The informational efficiency of European natural gas hubs: Price formation and intertemporal arbitrage. *Energy Journal*, 37(2), 1–30.
16. Olsen, K. K., Mjelde, J. W., & Bessler, D. A. (2015). Price formulation and the law of one price in internationally linked markets: An examination of the natural gas markets in the USA and Canada. *Annals of Regional Science*, 54(1), 117–142.
17. Osička, J., Lehotský, L., Zapletalová, V., Černoch, F., & Dančák, B. (2018). Natural gas market integration in the Visegrad 4 region: An example to follow or to avoid? *Energy Policy*, 112, 184–197.
18. Ren X, Lu Z, Cheng C, Shi Y, Shen J (2019) On dynamic linkages of the state natural gas markets in the USA: Evidence from an empirical spatio-temporal network quantile analysis. *Energy Economics*, 80, 234–245.
19. Reboredo, J. C. (2011). How do crude oil prices co-move? A copula approach. *Energy Economics*, 33(5), 948–955.
20. Scarcioffolo, A. R., & Etienne, X. L. (2019). How connected are the U.S. regional natural gas markets in the post-deregulation era? Evidence from time-varying connectedness analysis. *Journal of Commodity Markets*, 15, 100076.
21. Shi, X., Shen, Y., & Wu, Y. (2019). Energy market financialization: Empirical evidence and implications from East Asian LNG markets. *Finance Research Letters*, 30, 414–419.
22. Silverstovs, B., L'Hégaret, G., Neumann, A., & Hirschhausen, C. (2005). International market integration for natural gas? A cointegration analysis of prices in Europe, North America and Japan. *Energy Economics*, 27(4), 603–615.
23. Zhang, B. (2019). Are Chinese and international oil markets integrated? *International Review of Economics and Finance*, 62, 41–52.

Chapter 4

Hedging Strategy with Futures Contracts



4.1 Introduction

A futures contract is a promise to buy or sell a security at a currently agreed upon price at a predetermined time in the future. Futures are one of the most representative derivatives along with options and swaps. Commodity futures are those whose underlying assets are specific standardized commodities (e.g., precious metals, agricultural products, energy, etc.) and are often listed on commodity exchanges. Although they can be sold and bought within their maturity, they will automatically settle at maturity. The main purposes of trading commodity futures are risk hedging, speculation (e.g., diversified investment beyond traditional financial securities and trading based on market forecasts), arbitrage (e.g., pair trading between highly correlated securities and trading between securities with different maturities considering risks and interest rates), and procurement only in the case of physical settlement.

For many non-financial companies, commodity futures represent a means to hedge risk. Trading futures can eliminate uncertainties arising from price fluctuations because they determine future cash flows. We can not only determine profits in advance but also avoid unacceptable losses by trading futures. For example, firms often buy crude oil futures to avoid losses caused by future spot price increases when procuring crude oil. If the crude oil spot price actually rises, the profit obtained by liquidating the futures can cover the loss from the price increase in the spot market. As a completely opposite example, when selling electricity, firms often sell electricity futures to avoid losses due to future spot price declines. If the electricity spot price actually falls, the profit gained by counter-trading futures can cover the loss from the actual price drop. Many futures trades have the advantage of lowering hedging costs because of margin trading, which itself often reduces hedging costs in futures trades.

If we trade futures to hedge the risk of spot price fluctuations, we must be careful to curb the diversification of the portfolio return consisting of a spot and its futures. The ratio of future positions to the spot position that minimizes this variance is the

optimal hedge ratio (OHR). The OHR is obtained by dividing the covariance of the spot and futures returns by the variance of the futures return. Various multivariate generalized autoregressive conditional heteroscedasticity (GARCH) models have been proposed to capture the conditional covariance between multiple asset returns and the conditional variance of each return. Traders can monitor the OHRs calculated from the estimated multivariate GARCH model to design an optimal portfolio.

Many studies investigate hedging portfolio strategies in the oil and/or gas markets. Knill et al. [13] indicate that oligopolists with superior knowledge of supply and demand, such as oil and gas companies, hedge only when they expect unfavorable events. Ripple and Moosa [17] examine the difference in the hedging effect of crude oil futures on maturity and reveal that futures hedging is more effective in near-month contracts. Chang et al. [5] evaluate the OHR and optimal portfolio weights of crude oil portfolios using a multivariate GARCH model. Chang et al. [6] compare the performance of five types of multivariate volatility models for the crude oil spot and futures return series by calculating the OHR and hedging effectiveness (HE) index. Toyoshima et al. [18] estimate three types of multivariate GARCH models for crude oil spot and futures markets. The strategy built by the asymmetric dynamic conditional correlation (ADCC) model has the best variance reduction performance. Wang et al. [20] examine the performance of an unhedged strategy and the minimum variance hedging strategies, for which they estimate required parameters from five types of models that generate constant hedge ratios, six types of bivariate GARCH models, two types of copula models, and five types of regime-switching models for 18 commodities (e.g., crude oil and natural gas), three currencies, and three stock indices. Ghoddusi and Emamzadehfard [10] test multiple features of hedging performance using upstream and downstream prices in the United States (US) natural gas market. Hanly [11] reports significant differences between the OHRs and hedging performance from an analysis of crude oil, petroleum product, and natural gas benchmarks. They argue that the hedging performance of natural gas futures is less effective than that of crude oil and heating oil. Wang et al. [19] evaluate hedging performance not by the ability to minimize the variance of the hedged portfolio but by the ability to minimize risk. They employ three constant and seven time-varying hedging models for the crude oil index. Lv et al. [15] investigate the performance of portfolios consisting of Chinese petrochemical-related stocks and crude oil indices in terms of risk and return. Furió and Torró [9] design optimal hedging strategies with natural gas futures based on the expected utility maximization approach. Their study reveals that United Kingdom (UK) natural gas futures have a significant seasonally variable futures bias. Li et al. [14] reveal that the hedging effect is high due to the strong correlations between Chinese crude oil spots and futures, and confirm that the Chinese crude oil futures market, launched in 2018, realizes its main function and launching aim.

This chapter considers the procurement of natural gas in the US and the UK. We adopt the Henry Hub (HH) in the US and the national balancing point (NBP) in the UK. We estimate the three types of bivariate GARCH models of spot returns and future returns for each natural gas market and calculate the OHR and HE indices. We use the diagonal VECM model proposed by Bollerslev et al. [3], diagonal BEKK

model proposed by Baba et al. [1] and Engle and Kroner [8], and constant conditional correlation (CCC) model proposed by Bollerslev [2], which are the basic representative multivariate GARCH models. A comparison of each HE shows that the portfolio constructed on the OHR calculated by the diagonal BEKK model has the largest hedging effect for both the HH and NBP.

The remainder of this chapter is organized as follows. Section 4.2 provides the price series of natural gas spots and futures in the US and UK markets. Section 4.3 describes the OHR and HE. Section 4.4 explains the multivariate GARCH model. Section 4.5 presents the results of the analysis. Section 4.6 expresses the overall concluding remarks and considerations.

4.2 Data

We use the HH and NBP as natural gas price indices to represent the US and UK markets, respectively. We use the spot and futures prices from January 2, 2015 to December 30, 2020. HH and NBP prices are expressed in USD and GBP per mmbtu, respectively. We obtain these daily data from Bloomberg.

Before the analysis, we confirm the representative statistics to provide an overview of the data. Table 4.1 lists the descriptive statistics of the price series. Because we extract only the days for which both futures and spot prices are available, we have 1484 and 1517 observations for the HH and NBP, respectively. Both the mean and median of each future are higher than those of each spot. Contangos occur in both the HH and NBP markets. We can infer that supply and demand were not very tight during this period. The medians of both the HH futures and spot were higher than their respective means. Therefore, we can expect many outliers in the left tail of each distribution. On the other hand, the means of both the NBP futures and spot are almost the same as their respective medians, indicating that we can expect a roughly equivalent number of outliers on the left and right tails. Although their spot price series means are lower than their futures price series means, their spot maximum values are much larger than their futures maximum values. We can see common price spikes in the energy spot markets. For both the HH and NBP, the futures price series have a narrower range and a smaller standard deviation than the spot price series. That is, their futures price series are more scattered than their spot price series are. All the variables have positive skewness and a distribution with a long right tail. All variables had kurtosis greater than 3 and had a distribution with a sharp peak and long fat-tails. The Jarque–Bera test rejects the normal distribution null hypothesis for the HH future, HH spot, and NBP spot price series, while accepting the hypothesis for the NBP future price series at the 10% significance level.

Figure 4.1 shows the time plots of the HH futures and spot prices. There is almost no separation between the futures and spot prices. However, only the spot price series has large price spikes that occurred in January 2018 and March 2019. In November 2016 and September 2020, spot prices were slightly lower than futures prices.

Table 4.1 Descriptive statistics

	HH		NBP	
	Futures	Spot	Futures	Spot
Period	From 1/2/2015 to 12/30/2020			
Observations	1484	1484	1517	1517
Mean	2.656	2.638	0.410	0.404
Median	2.718	2.717	0.410	0.402
Maximum	4.837	7.135	0.777	2.300
Minimum	1.482	1.330	0.083	0.073
Standard deviation	0.501	0.562	0.132	0.146
Skewness	0.253	0.904	0.127	1.652
Kurtosis	4.280	8.639	3.078	21.777
Jarque–Bera	117 (0.00)	2168 (0.00)	4.428 (0.12)	22,976 (0.00)

Note p -values are in parentheses

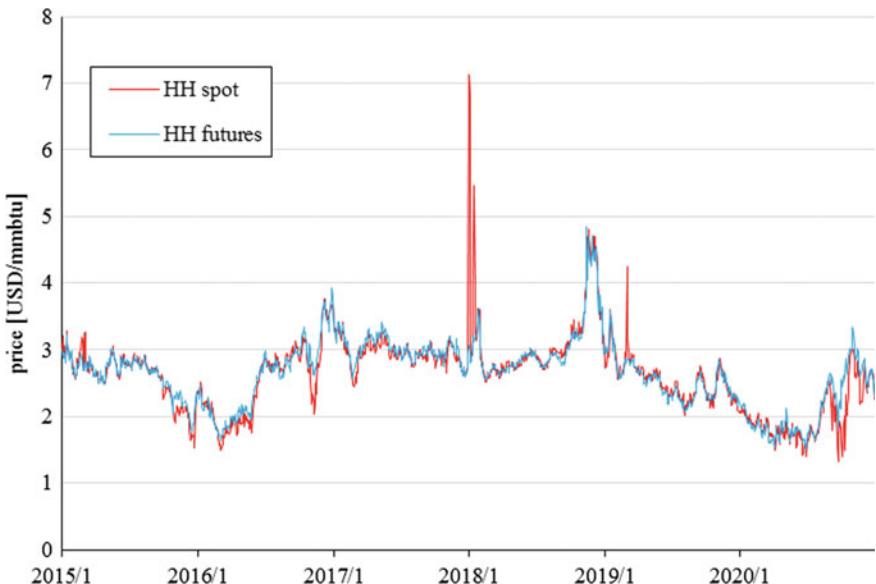
**Fig. 4.1** HH spot and futures prices

Figure 4.2 shows the time plots of the NBP futures and spot prices. The price difference between the futures and its spot tends to be smaller than that of the HH. We observe a large spike in spot prices in March 2018. Between August and December 2019, its spot prices have a downward trend separate from their future prices.

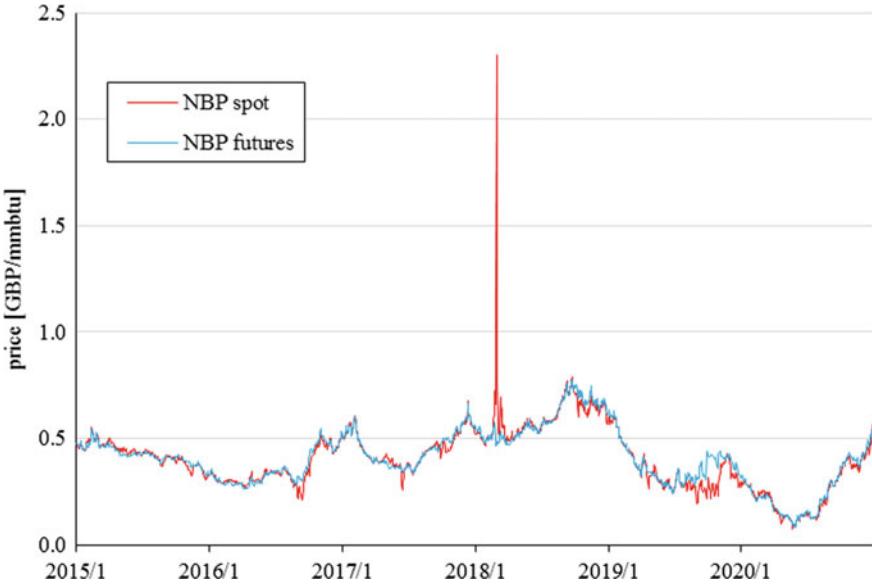


Fig. 4.2 NBP spot and futures prices

4.3 Optimal Hedge Ratio and Hedge Effectiveness

To hedge the spot price fluctuation from $t - 1$ to t , suppose that we hold θ_t units of its future for one unit of its spot during that period. The return series for this hedged portfolio is then

$$R_t = s_t - \theta_t f_t, \quad (4.1)$$

where s_t and f_t are the spot and future return series, respectively.

According to Johnson [12], the conditional variance of the return series is

$$\text{var}(R_t | I_{t-1}) = \text{var}(s_t | I_{t-1}) - 2\theta_t \text{cov}(s_t, f_t | I_{t-1}) + \theta_t^2 \text{var}(f_t | I_{t-1}), \quad (4.2)$$

where I_{t-1} is the information set available at time $t - 1$, $\text{var}(s_t | I_{t-1})$ is the conditional variance of the spot return series, $\text{var}(f_t | I_{t-1})$ is the conditional variance of the future return series, and $\text{cov}(s_t, f_t | I_{t-1})$ are the conditional covariance of the spot and future return series.

By partially differentiating Eq. (4.2) with respect to θ_t , we obtain

$$\frac{\partial \text{var}(R_t | I_{t-1})}{\partial \theta_t} = -2\text{cov}(s_t, f_t | I_{t-1}) + 2\theta_t \text{var}(f_t | I_{t-1}). \quad (4.3)$$

If this partial derivative is equal to 0, then the conditional variance of the portfolio is minimal. Setting Eq. (4.3) equal to zero and solving for θ_t , the OHR_t conditional on I_{t-1} is

$$OHR_t = (\theta_t | I_{t-1}) = \frac{\text{cov}(s_t, f_t | I_{t-1})}{\text{var}(f_t | I_{t-1})}. \quad (4.4)$$

Because both $\text{cov}(s_t, f_t | I_{t-1})$ and $\text{var}(f_t | I_{t-1})$ are conditional, we cannot easily calculate the OHR using ordinary least squares (OLS). We must obtain the OHR_t by estimating a multivariate GARCH model. Then, we calculate the following average OHR:

$$\text{average } OHR = \frac{1}{T} \sum_{t=1}^T OHR_t, \quad (4.5)$$

where T denotes sample size. To compare the performance of the OHRs calculated using different techniques, Ku et al. [16] propose the HE index. The HE_t conditional on I_{t-1} is

$$HE_t = \frac{\text{var}(s_t | I_{t-1}) - \text{var}(R_t | I_{t-1})}{\text{var}(s_t | I_{t-1})}. \quad (4.6)$$

We calculate the following average HE:

$$\text{average } HE = \frac{1}{T} \sum_{t=1}^T HE_t. \quad (4.7)$$

A higher HE indicates higher hedge effectiveness and larger risk reduction. We must thus select a hedging strategy that maximizes HE.

4.4 Multivariate GARCH Model

In general, the multivariate GARCH model is expressed as

$$\begin{aligned} \mathbf{r}_t &= \boldsymbol{\mu}_t + \mathbf{e}_t \\ \mathbf{e}_t &= \mathbf{H}_t^{1/2} \boldsymbol{\zeta}_t, \end{aligned} \quad (4.8)$$

where \mathbf{r}_t is the return vector, $\boldsymbol{\mu}_t$ is the mean vector of the return, \mathbf{H}_t is the covariance matrix, and each element of $\boldsymbol{\zeta}_t$ is generated from an independently and identically distributed random number following the standard normal distribution ($iid N(0, 1)$). Researchers proposed various multivariate GARCH models depending on the formulation of \mathbf{H}_t .

In this chapter, we examine a portfolio consisting of two variables. Therefore, we describe each multivariate GARCH as a bivariate GARCH model. We write the mean equation as follows, assuming that the lag order is 0:

$$\begin{aligned} f_t &= m_f + \varepsilon_{f,t} \\ s_t &= m_s + \varepsilon_{s,t} \\ \varepsilon_{f,t} &= \sqrt{h_{f,t}} z_{f,t} + \sqrt{h_{fs,t}} z_{s,t} \\ \varepsilon_{s,t} &= \sqrt{h_{fs,t}} z_{f,t} + \sqrt{h_{s,t}} z_{s,t} \\ z_{f,t} &\sim iidN(0, 1) \\ z_{s,t} &\sim iidN(0, 1), \end{aligned} \tag{4.9}$$

where $h_{f,t}$ is the conditional variance of f_t , $h_{s,t}$ is the conditional variance of s_t , and $h_{fs,t}$ is the conditional covariance of f_t and s_t .

In the following explanation of each bivariate GARCH model, both the lag order of the ARCH term and that of the GARCH term in the variance equation are set to 1 for simplicity.

4.4.1 Diagonal VECH Model

Bollerslev et al. [3] propose the following VECH model:

$$\begin{bmatrix} h_{f,t} \\ h_{fs,t} \\ h_{s,t} \end{bmatrix} = \begin{bmatrix} f_c \\ fs_c \\ s_c \end{bmatrix} + \begin{bmatrix} fa_f & fa_{fs} & fa_s \\ fs_a_f & fs_a_{fs} & fs_a_s \\ sa_f & sa_{fs} & sa_s \end{bmatrix} \begin{bmatrix} \varepsilon_{f,t-1}^2 \\ \varepsilon_{f,t-1}\varepsilon_{s,t-1} \\ \varepsilon_{s,t-1}^2 \end{bmatrix} + \begin{bmatrix} fg_f & fg_{fs} & gs_s \\ fs_g_f & fs_g_{fs} & fs_g_s \\ sg_f & sg_{fs} & sg_s \end{bmatrix} \begin{bmatrix} h_{f,t-1} \\ h_{fs,t-1} \\ h_{s,t-1} \end{bmatrix}. \tag{4.10}$$

We can understand this model as a simple extension of a univariate GARCH model to a bivariate GARCH model. However, it has a disadvantage in that the number of parameters to estimate increases sharply as the number of variables and lag order increase. Therefore, Bollerslev et al. [3] propose a diagonal VECH model in which both the ARCH and GARCH matrices are diagonal matrices. We can express this as follows:

$$\begin{bmatrix} h_{f,t} \\ h_{fs,t} \\ h_{s,t} \end{bmatrix} = \begin{bmatrix} f_c \\ fs_c \\ s_c \end{bmatrix} + \begin{bmatrix} fa_f & 0 & 0 \\ 0 & fs_a_{fs} & 0 \\ 0 & 0 & sa_s \end{bmatrix} \begin{bmatrix} \varepsilon_{f,t-1}^2 \\ \varepsilon_{f,t-1}\varepsilon_{s,t-1} \\ \varepsilon_{s,t-1}^2 \end{bmatrix}$$

$$+ \begin{bmatrix} {}^f g_f & 0 & 0 \\ 0 & {}^f s g_{fs} & 0 \\ 0 & 0 & {}^s g_s \end{bmatrix} \begin{bmatrix} h_{f,t-1} \\ h_{fs,t-1} \\ h_{s,t-1} \end{bmatrix}. \quad (4.11)$$

The expression as simultaneous equations is

$$\begin{aligned} h_{f,t} &= {}^f c + {}^f a_f \varepsilon_{f,t-1}^2 + h_{f,t-1} \\ h_{s,t} &= {}^s c + {}^s a_s \varepsilon_{s,t-1}^2 + {}^s g_s h_{s,t-1} \\ h_{fs,t} &= {}^{fs} c + {}^{fs} a_{fs} \varepsilon_{f,t-1} \varepsilon_{s,t-1} + {}^{fs} g_{fs} h_{fs,t-1}. \end{aligned} \quad (4.12)$$

Although the bivariate VECH model requires us to estimate 21 parameters, the bivariate diagonal VECH model requires us to estimate only 9 parameters: ${}^f c$, ${}^f a_f$, ${}^f g_f$, ${}^s c$, ${}^s a_s$, ${}^s g_s$, ${}^{fs} c$, ${}^{fs} a_{fs}$, and ${}^{fs} g_{fs}$.

4.4.2 Diagonal BEKK Model

Baba et al. [1] and Engle and Kroner [8] propose the following BEKK model:

$$\begin{bmatrix} h_{f,t} & h_{fs,t} \\ h_{fs,t} & h_{s,t} \end{bmatrix} = \begin{bmatrix} {}^f c & {}^{fs} c \\ {}^{fs} c & {}^s c \end{bmatrix} + \begin{bmatrix} {}^f a_f & {}^{fs} a_s \\ {}^{fs} a_f & {}^s a_s \end{bmatrix} \begin{bmatrix} \varepsilon_{f,t-1}^2 & \varepsilon_{f,t-1} \varepsilon_{s,t-1} \\ \varepsilon_{f,t-1} \varepsilon_{s,t-1} & \varepsilon_{s,t-1}^2 \end{bmatrix} T \begin{bmatrix} {}^f a_f & {}^{fs} a_s \\ {}^{fs} a_f & {}^s a_s \end{bmatrix} + \begin{bmatrix} {}^f g_f & {}^{fs} g_s \\ {}^{fs} g_f & {}^s g_s \end{bmatrix} \begin{bmatrix} h_{f,t-1} & h_{fs,t-1} \\ h_{fs,t-1} & h_{s,t-1} \end{bmatrix} T \begin{bmatrix} {}^f g_f & {}^{fs} g_s \\ {}^{fs} g_f & {}^s g_s \end{bmatrix}. \quad (4.13)$$

The BEKK model has fewer parameters to estimate than the VECH model does. We must estimate as many as 21 parameters in Eq. (4.10). However, Eq. (4.13) requires only 11 parameters. This effect becomes very large as the number of variables increases. Assuming that the number of variables is three and the lag order of both the ARCH and GARCH matrices are 1, the number of parameters to estimate is 78 and 24 in the VECH and BEKK models, respectively. Moreover, to reduce the number of parameters to estimate, Baba et al. [1] and Engle and Kroner [8] propose a diagonal BEKK model in which both the ARCH and GARCH matrices are diagonal matrices. We can express this model as

$$\begin{bmatrix} h_{f,t} & h_{fs,t} \\ h_{fs,t} & h_{s,t} \end{bmatrix} = \begin{bmatrix} {}^f c & {}^{fs} c \\ {}^{fs} c & {}^s c \end{bmatrix} + \begin{bmatrix} {}^f a_f & 0 \\ 0 & {}^s a_s \end{bmatrix} \begin{bmatrix} \varepsilon_{f,t-1}^2 & \varepsilon_{f,t-1} \varepsilon_{s,t-1} \\ \varepsilon_{f,t-1} \varepsilon_{s,t-1} & \varepsilon_{s,t-1}^2 \end{bmatrix} T \begin{bmatrix} {}^f a_f & 0 \\ 0 & {}^s a_s \end{bmatrix}$$

$$\begin{aligned}
& + \begin{bmatrix} {}^f g_f & 0 \\ 0 & {}^s g_s \end{bmatrix} \begin{bmatrix} h_{f,t-1} & h_{fs,t-1} \\ h_{fs,t-1} & h_{s,t-1} \end{bmatrix}_T \begin{bmatrix} {}^f g_f & 0 \\ 0 & {}^s g_s \end{bmatrix} \\
& = \begin{bmatrix} {}^f c & {}^{fs} c \\ {}^{fs} c & {}^s c \end{bmatrix} + \begin{bmatrix} \left({}^f a_f \varepsilon_{f,t-1} \right)^2 & {}^f a_f {}^s a_s \varepsilon_{f,t-1} \varepsilon_{s,t-1} \\ {}^f a_f {}^s a_s \varepsilon_{f,t-1} \varepsilon_{s,t-1} & \left({}^s a_s \varepsilon_{s,t-1} \right)^2 \end{bmatrix} \\
& + \begin{bmatrix} \left({}^f g_f \right)^2 h_{f,t-1} & {}^f g_f {}^s g_s h_{fs,t-1} \\ {}^f g_f {}^s g_s h_{fs,t-1} & \left({}^s g_s \right)^2 h_{s,t-1} \end{bmatrix}. \tag{4.14}
\end{aligned}$$

We calculate Eq. (4.14) and express it using the following simultaneous equations:

$$\begin{aligned}
h_{f,t} &= {}^f c + {}^f a_f^2 \varepsilon_{f,t-1}^2 + {}^f g_f^2 h_{f,t-1} \\
h_{s,t} &= {}^s c + {}^s a_s^2 \varepsilon_{s,t-1}^2 + {}^s g_s^2 h_{s,t-1} \\
h_{fs,t} &= {}^{fs} c + {}^f a_f {}^s a_s \varepsilon_{f,t-1} \varepsilon_{s,t-1} + {}^f a_f {}^s a_s \varepsilon_{f,t-1} \varepsilon_{s,t-1}. \tag{4.15}
\end{aligned}$$

Although the bivariate BEKK model requires us to estimate 11 parameters, the bivariate diagonal VECM model requires only 7 parameters: ${}^f c$, ${}^f a_f$, ${}^f g_f$, ${}^s c$, ${}^s a_s$, ${}^s g_s$, and ${}^{fs} c$.

4.4.3 CCC Model

Bollerslev [2] proposes a CCC model in which the correlation matrix is constant, although each diagonal component of \mathbf{H}_t , that is, the variance of each variable, is conditional. The off-diagonal components of \mathbf{H}_t are 0. When we represent the variance of each variable using a univariate GARCH model with all lag orders of 1, we can write the CCC model as the following equations using the correlation coefficient ${}^{fs} c$ between f_t and s_t :

$$\begin{aligned}
h_{f,t} &= {}^f c + {}^f a_f \varepsilon_{f,t-1}^2 + {}^f g_f h_{f,t-1} \\
h_{s,t} &= {}^s c + {}^s a_s \varepsilon_{s,t-1}^2 + {}^s g_s h_{s,t-1} \\
h_{fs,t} &= {}^{fs} c \sqrt{h_{f,t} h_{s,t}}. \tag{4.16}
\end{aligned}$$

The bivariate CCC model requires the estimation of only 7 parameters: ${}^f c$, ${}^f a_f$, ${}^f g_f$, ${}^s c$, ${}^s a_s$, ${}^s g_s$, and ${}^{fs} c$.

4.5 Analysis Results

We estimate the three types of multivariate GARCH models (the diagonal VECM model, diagonal BEKK, and CCC model) for the HH and NBP, and calculate the average OHR and HE.

4.5.1 HH Market

First, we generate the HH futures and spot return series. Table 4.2 shows the descriptive statistics for each variable. For these two variables, the mean is positive, whereas the median is negative. This result implies that these distributions have a larger number of positive outliers than negative ones. As Fig. 4.1 shows, because the spot price series has large spikes, its return series has an extremely large maximum value and a very small minimum value. These ranges and standard deviations indicate that the spot return series tends to fluctuate more than the future series. These distributions have a positive skewness, implying that these distributions have a long right tail. These distributions have kurtosis values greater than 3, meaning that these distributions have sharp peaks and fat-tails. According to the Jarque–Bera statistics, both series are not significantly normally distributed in terms of skewness and kurtosis.

Figure 4.3 plots the HH spot and futures return series. Referring to Fig. 4.1, we can observe that when the spot price series spikes, the spot return series also spikes and then plummets. Throughout this period, the fluctuations in the futures return series are not as large as in the spot return series.

Table 4.3 shows the estimation results for each multivariate GARCH model for the HH. Each parameter of the variance equations is significant at the 1% level, whereas the parameters of these mean equations are not significant, even at the 10% level.

Table 4.2 Descriptive statistics (HH return series)

	Futures	Spot
Observations	1483	1483
Mean	0.036%	0.124%
Median	– 0.042%	– 0.055%
Maximum	21.9%	101.6%
Minimum	– 16.5%	– 35.5%
Standard deviation	0.032	0.056
Skewness	0.625	5.119
Kurtosis	8.311	90.092
Jarque–Bera	1839 (0.00)	475,171 (0.00)

Note p -values are in parentheses

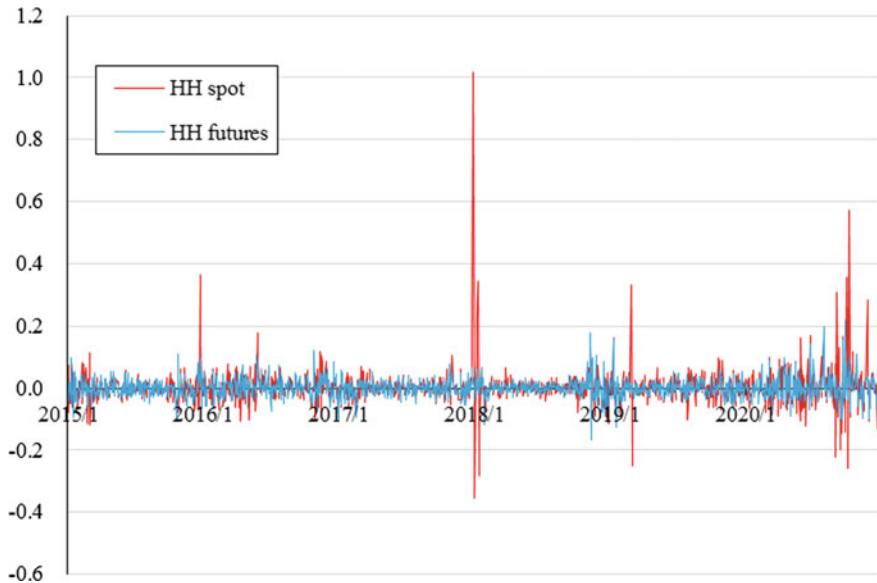


Fig. 4.3 HH spot and futures return series

Table 4.3 Estimated multivariate GARCH model (HH)

		Diagonal VECM	Diagonal BEKK	CCC
Mean equation	m_f	6.11×10^{-4} (0.31)	6.75×10^{-4} (0.28)	4.64×10^{-4} (0.42)
	m_s	3.94×10^{-4} (0.53)	1.73×10^{-4} (0.78)	2.26×10^{-4} (0.71)
Variance equation	f_c	2.62×10^{-5} (0.00)	1.32×10^{-5} (0.00)	1.07×10^{-5} (0.00)
	f_{af}	1.09×10^{-1} (0.00)	3.20×10^{-1} (0.00)	1.01×10^{-1} (0.00)
	f_{gf}	8.71×10^{-1} (0.00)	9.49×10^{-1} (0.00)	8.96×10^{-1} (0.00)
	s_c	8.82×10^{-5} (0.00)	5.76×10^{-5} (0.00)	4.97×10^{-5} (0.00)
	s_{as}	3.21×10^{-1} (0.00)	5.36×10^{-1} (0.00)	3.76×10^{-1} (0.00)
	s_{gs}	7.02×10^{-1} (0.00)	8.68×10^{-1} (0.00)	7.01×10^{-1} (0.00)
	f^s_c	4.05×10^{-5} (0.00)	1.75×10^{-5} (0.00)	4.45×10^{-1} (0.00)
	f^s_{af}	7.65×10^{-2} (0.00)		
	f^s_{gs}	8.32×10^{-1} (0.00)		

Note p-values re in parentheses

Figure 4.4 plots the covariance series calculated using each multivariate GARCH model. Any model can represent the conditional covariance between these return series. Moreover, these three covariance series fluctuate almost in tandem.

Figure 4.5 plots the OHR series calculated using the estimated diagonal VECM model. The OHR series is about 50% throughout this period. Referring to Fig. 4.1, we can see that the OHR series soars in December 2015 when both the futures and

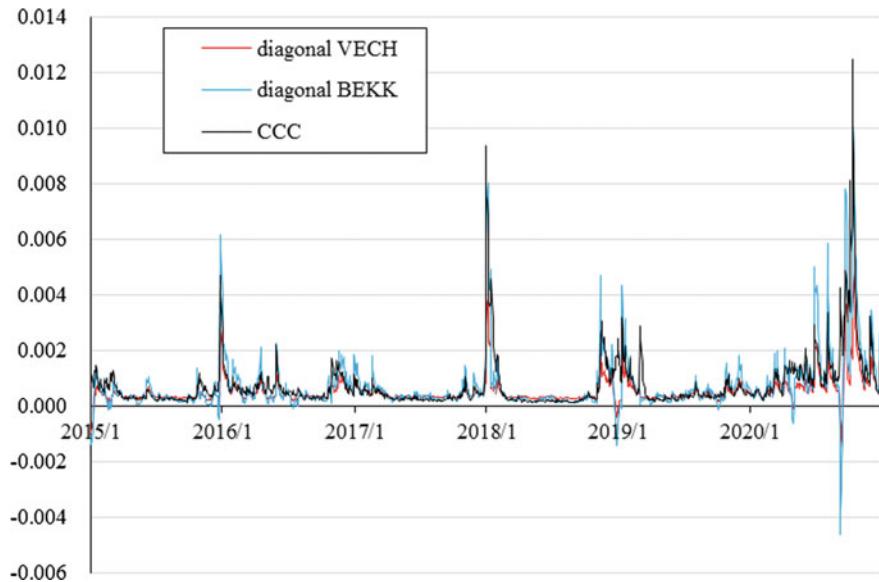


Fig. 4.4 Covariance series between futures and spot prices (HH)

spot price series spike, January 2018 when only the spot price series spikes, and October 2020 when only the futures price series spikes. The OHR series plummets to the negative in September 2020 when the spot price series separates and moves downward relative to the futures price series. Referring to Fig. 4.4, we can see that the OHR series is synchronized with the covariance series.

Figure 4.6 depicts the time plots of the OHR series calculated using the estimated diagonal BEKK model. The OHR series was approximately 50% during this period. Referring to Fig. 4.1, we can observe that the OHR series spikes in December 2015 when both the futures and spot price series spike, January 2018 when only the spot price series spikes, and October 2020 when only the futures price series spikes. The OHR series dips into the negative in September 2020 when the spot price series separates and moves downward relative to the futures price series. This result is almost the same as that of the diagonal VECM model. However, the OHR in the above four periods is approximately twice that of the diagonal VECM. Figures 4.4 and 4.6 indicate that the OHR series is linked to the covariance series.

Figure 4.7 displays the OHR series calculated using the estimated CCC model. The OHR series was approximately 60% throughout this period, although some spikes occurred. These spikes are even larger than in the diagonal BEKK model. There are no negative OHR values, which appear in the OHR series calculated by the diagonal VECM and diagonal BEKK models. Referring to Fig. 4.1, we see that the OHR series soars in December 2015 when both the futures and spot price series spike, January 2018 when only the spot price series spikes, and October 2020 when only the futures price series spikes. This is the same as the OHR series calculated

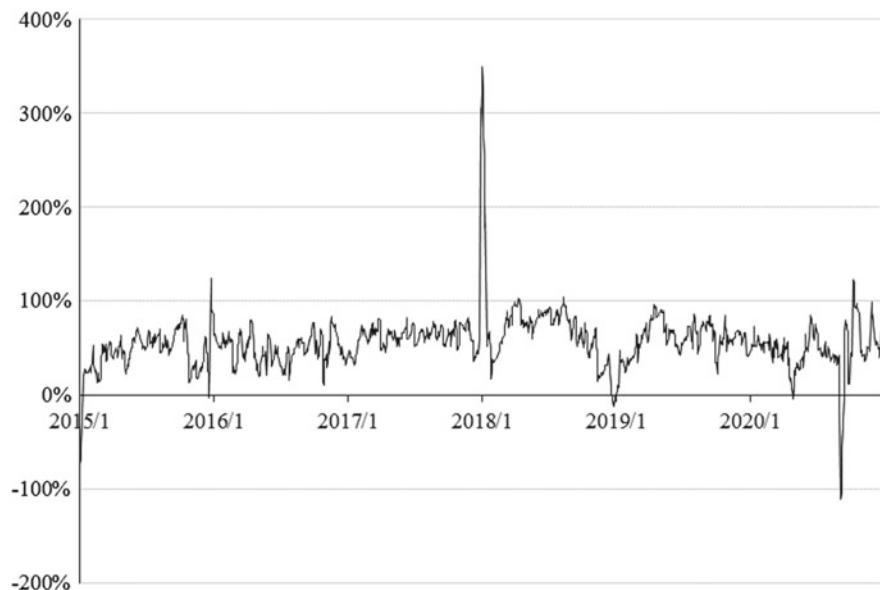


Fig. 4.5 OHR calculated using the estimated diagonal VECM model for the HH

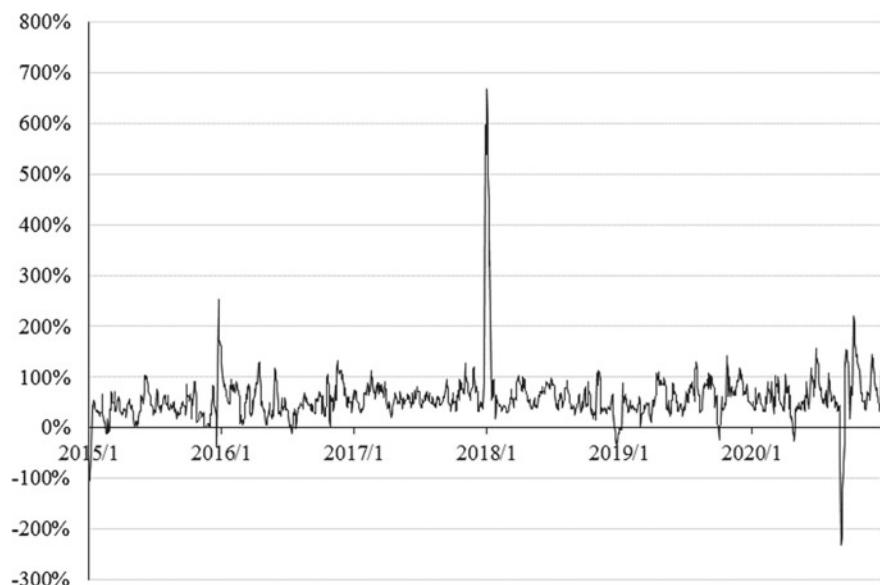


Fig. 4.6 OHR calculated using the estimated diagonal BEKK model for the HH

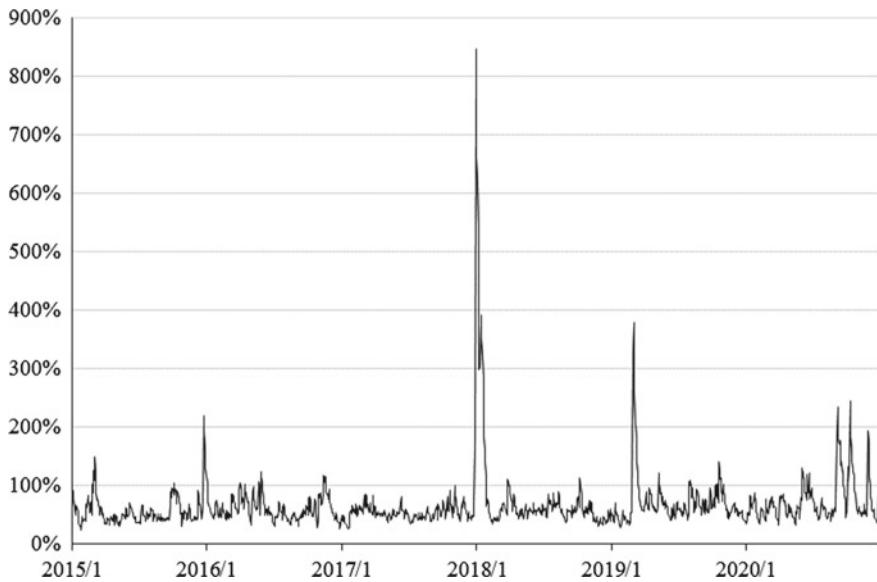


Fig. 4.7 OHR calculated using the estimated CCC model for the HH

Table 4.4 Average OHR and HE (HH)

	Average OHR (%)	Average HE (%)
Diagonal VECM	56.7	21.4
Diagonal BEKK	57.7	22.7
CCC	65.9	19.8

Note Bold indicates the average HE for the optimal hedging strategy

using the diagonal VECM and BEKK models. In addition, the OHR series spikes in March 2015 and March 2019, when only the spot price series spikes were small, and in November 2020, when only the spot price series plummets. From Fig. 4.4, we can see that the OHR series is synchronized with the covariance series.

Table 4.4 lists the average OHR and HE calculated using each multivariate GARCH model. A portfolio built using the OHR obtained by the diagonal BEKK model has the largest hedging effect. Therefore, we can conclude that it is the best hedging strategy among these three. However, the average HE is 22.7%, which is not so large.

4.5.2 NBP Market

Table 4.5 shows the descriptive statistics for the NBP futures and spot return series. For these two return series, the means are larger than the medians, implying that these

Table 4.5 Descriptive statistics (NBP return series)

	Futures	Spot
Observations	1516	1516
Mean	0.065%	0.168%
Median	– 0.027%	0.000%
Maximum	40.9%	91.7%
Minimum	– 14.0%	– 66.1%
Standard deviation	0.034	0.056
Skewness	1.921	2.551
Kurtosis	21.694	68.214
Jarque–Bera	23,007 (0.00)	270,282 (0.00)

Note *p*-values are in parentheses

distributions have a larger number of outliers in the right tail than in the left tail. As Fig. 4.2 illustrates, because the spot price series has a large spike, its return series has extreme minimum and maximum values. The range of the spot return series is wider than that of the futures return series and the standard deviation of the spot return series is larger than that of the futures return series. This result implies that the spot return series tends to be more volatile than the futures return series. Both distributions had positive skewness, implying that each distribution has a long right tail. These distributions had kurtosis values of over 3, indicating that each distribution has a sharp peak and fat-tail. The Jarque–Bera test rejects the normal distribution hypothesis for both series.

Figure 4.8 plots the NBP spot and futures return series. In Fig. 4.2, we can observe that when the spot price series spikes, the spot return series also spikes.

Table 4.6 shows the estimation results for each multivariate GARCH model for the NBP return series. All the parameters of these variance equations are statistically significant at the 1% level, whereas all the parameters of the mean equations are not statistically significant, even at the 10% level.

Figure 4.9 plots the covariance series calculated using each multivariate GARCH model. Each model can represent the conditional covariance between these return series, and these covariance series are synchronized.

Figure 4.10 plots the OHR series calculated using the estimated diagonal VECM model. The OHR series is about 75% throughout this period. Referring to Fig. 4.2, we can see that the OHR series plummets to about –80% in March 2018 when only the spot price series spikes. The OHR series is negative in August 2016 and October 2019 when the spot price series separates and moves downward relative to the futures price series. The OHR series spikes to nearly 200% in June 2017 when only the spot price series swoops down.

Figure 4.11 plots the OHR series calculated using the estimated diagonal BEKK model. The OHR series is about 75% throughout this period. In Fig. 4.2, we can see that the OHR series plummets to about –200% in March 2018 when only the spot price series spikes. The OHR series plummets to the negative in August 2016 and

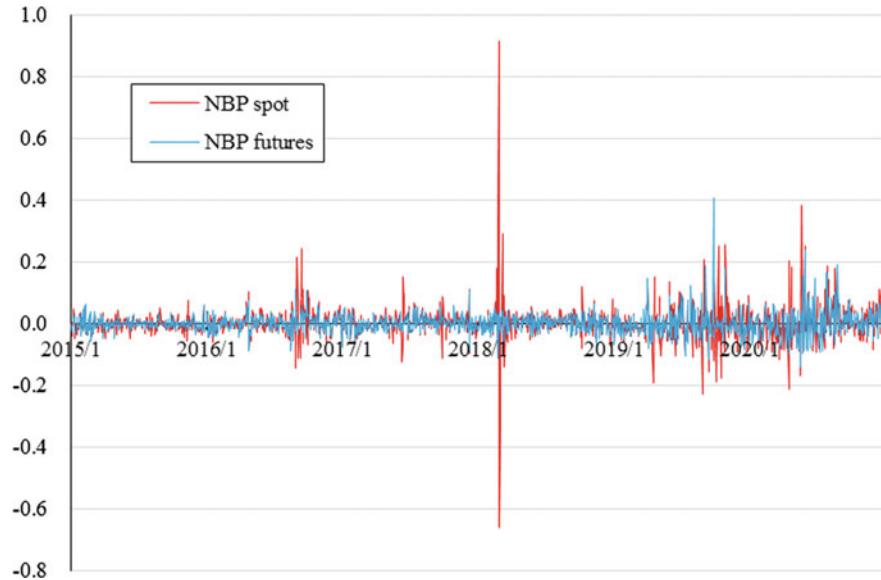


Fig. 4.8 NBP spot and futures return series

Table 4.6 Estimated multivariate GARCH model (NBP)

		Diagonal VECM	Diagonal BEKK	CCC
Mean equation	m_f	4.82×10^{-4} (0.37)	5.25×10^{-4} (0.34)	3.48×10^{-4} (0.48)
	m_s	3.17×10^{-4} (0.59)	5.70×10^{-4} (0.34)	3.79×10^{-4} (0.53)
Variance equation	f_c	8.65×10^{-6} (0.00)	1.31×10^{-5} (0.00)	4.86×10^{-6} (0.00)
	f_{af}	8.09×10^{-2} (0.00)	3.03×10^{-1} (0.00)	9.31×10^{-2} (0.00)
	f_{gf}	9.16×10^{-1} (0.00)	9.52×10^{-1} (0.00)	9.11×10^{-1} (0.00)
	s_c	4.88×10^{-5} (0.00)	7.68×10^{-5} (0.00)	4.51×10^{-5} (0.00)
	s_{as}	2.11×10^{-1} (0.00)	5.71×10^{-1} (0.00)	2.82×10^{-1} (0.00)
	s_{gs}	7.82×10^{-1} (0.00)	8.27×10^{-1} (0.00)	7.33×10^{-1} (0.00)
	$f^s c$	1.50×10^{-5} (0.00)	3.55×10^{-5} (0.00)	6.13×10^{-1} (0.00)
	$f^s a_{fs}$	8.01×10^{-2} (0.00)		
	$f^s g_{fs}$	8.95×10^{-1} (0.00)		

Note p -values are in parentheses

August and October 2019 when the spot price series separates and moves downward relative to the futures price series. The OHR series spikes to over 200% four times. Comparing Fig. 4.10 with Fig. 4.11, we can see that the OHR series calculated using the diagonal BEKK model oscillates more frequently and greatly than that from the diagonal BEKK model.

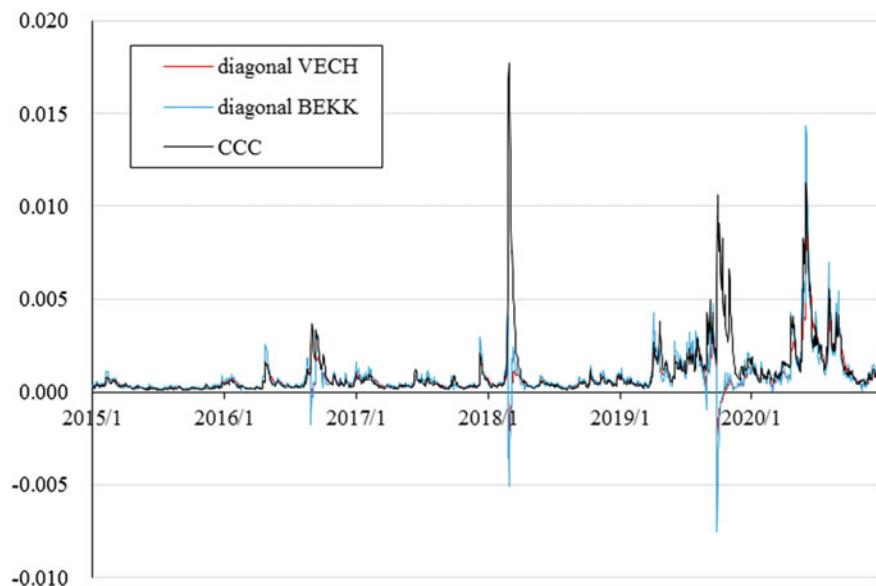


Fig. 4.9 Covariance series between futures and spot prices (NBP)

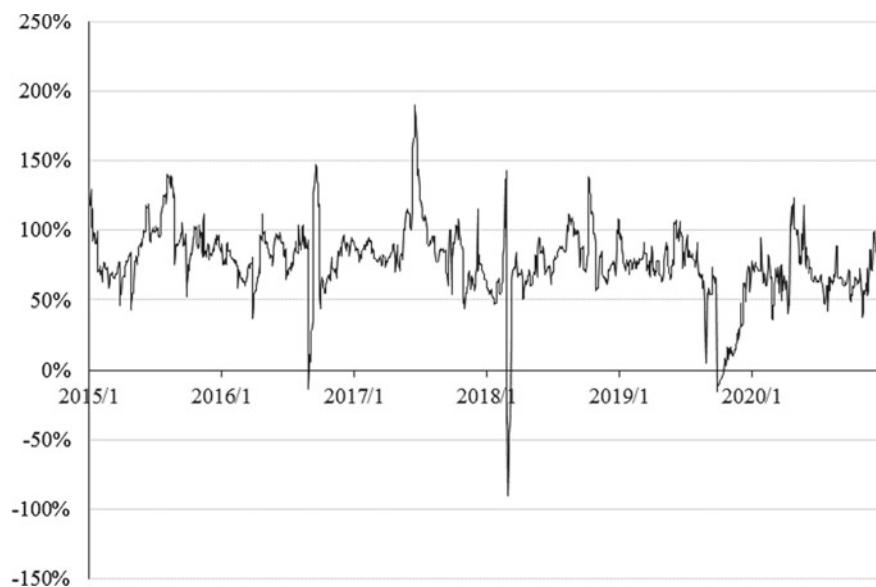


Fig. 4.10 OHR calculated using the estimated diagonal VECM model for the NBP

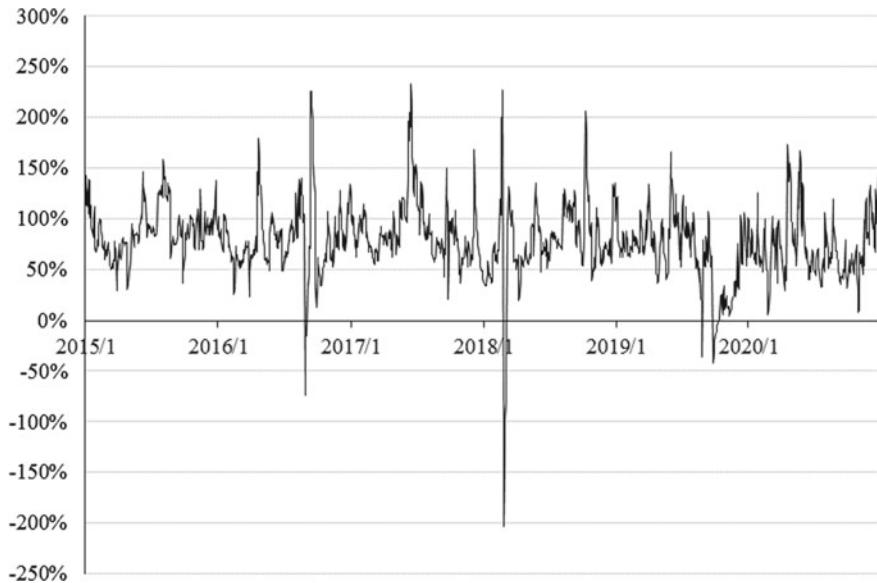


Fig. 4.11 OHR calculated using the estimated diagonal BEKK model for the NBP

Figure 4.12 plots the OHR series calculated using the estimated CCC model. Throughout this period, the OHR series is about 80% and is not negative. From Fig. 4.2, we can see that the OHR series spikes to over 300% in June 2017, when only the spot price series plummets, and in March 2018, when only the spot price series spikes. Besides these events, the OHR series spikes to over 200% four times.

Table 4.7 shows the average OHR and HE calculated using these three multivariate GARCH models. If we construct a portfolio using the OHR obtained from the diagonal BEKK model, we obtain the highest average HE, of 42.4%. Therefore, we can conclude that this is the best hedging strategy among these three.

4.6 Concluding Remarks

When firms procure energy with a high price fluctuation risk, they often trade futures to hedge the risk. Such firms must work to curb the volatility of the portfolio return, which consists of a spot and its futures. By dividing the covariance of the spot and futures return series by the variance of the future return series, we can obtain the ratio of futures positions to spot positions that minimizes volatility, which is defined as the OHR. In other words, by estimating the multivariate GARCH model that formulates the conditional covariance and variance, we can obtain the conditional OHR series, which helps us construct a timely optimal portfolio.

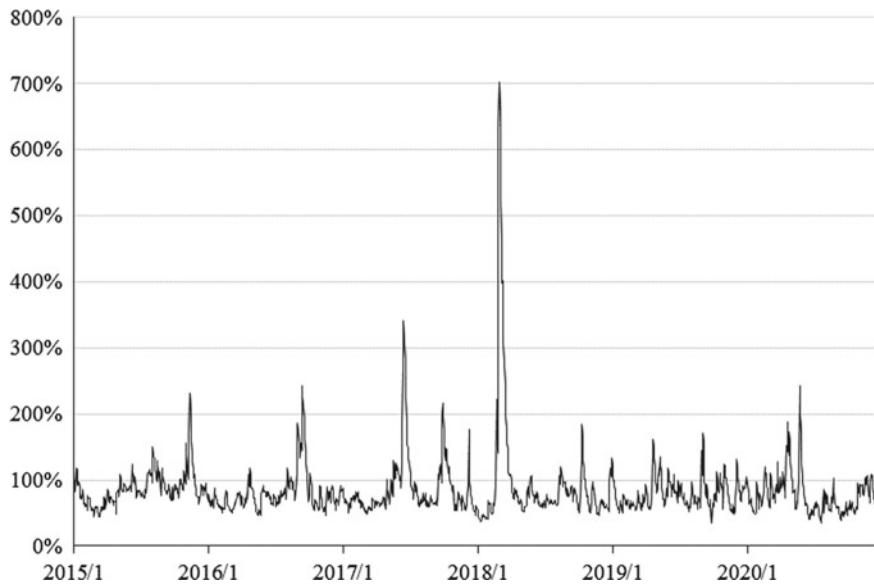


Fig. 4.12 OHR calculated using the estimated CCC model for the NBP

Table 4.7 Average OHR and HE (NBP)

	Average OHR (%)	Average HE (%)
Diagonal VECM	77.2	41.0
Diagonal BEKK	77.9	42.4
CCC	84.5	37.6

Note Bold indicates the average HE for the optimal hedging strategy

This chapter estimates three types of bivariate GARCH models consisting of the spot and futures return series in the US and UK natural gas markets to calculate OHR and HE. We adopt the diagonal VECM, diagonal BEKK, and CCC models as multivariate GARCH models. The OHR series fluctuates drastically depending on the spot and futures market conditions. All the multivariate GARCH models here can capture the time dependence of the covariance and variance in the same way. However, comparing the average HE values reveals that constructing the portfolio using the diagonal BEKK model is the best hedging strategy for both the HH and NBP markets.

While this chapter adopts primitive multivariate models (i.e., the diagonal VECM model, diagonal BEKK model, and CCC model), various multivariate GARCH models exist. Section 4.4 explains the VECM and BEKK models. Advancing the CCC model, Engle [7] proposes a dynamic conditional correlation (DCC) model that assumes that the correlation coefficient is conditional. Moreover, Cappiello et al. [4]

propose the ADCC model, which incorporates the asymmetry in which the correlation tends to be stronger after a negative return than after a positive return. We recommend that readers consider the optimal hedging strategy by applying these various methods to compare the HE values for the securities and commodities of interest.

References

1. Baba, Y., Engle, R. F., Kraft, D., & Kroner, K. F. (1987). *Multivariate simultaneous generalized ARCH*. Unpublished manuscript, Department of Economics, University of California.
2. Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model. *Review of Economics and Statistics*, 72(3), 498–505.
3. Bollerslev, T., Engle, R. F., & Wooldridge, J. M. (1988). A capital asset pricing model with timevarying covariances. *Journal of Political Economy*, 96(1), 116–131.
4. Cappiello, L., Engle, R. F., & Sheppard, K. (2006). A symmetric dynamics in the correlations of global equity and bond returns. *Journal of Financial Econometrics*, 4(4), 537–572.
5. Chang, C. L., McAleer, M., & Tansuchat, R. (2010). Analyzing and forecasting volatility spillovers, asymmetries and hedging in major oil markets. *Energy Economics*, 32(6), 1445–1455.
6. Chang, C. L., McAleer, M., & Tansuchat, R. (2011). Crude oil hedging strategies using dynamic multivariate GARCH. *Energy Economics*, 33(5), 912–923.
7. Engle, R. F. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economics Statistics*, 20(3), 339–350.
8. Engle, R. F., & Kroner, K. F. (1995). Multivariate simultaneous generalized ARCH. *Economic Theory*, 11(1), 122–150.
9. Furió, D., & Torró, H. (2020). Optimal hedging under biased energy futures markets. *Energy Economics*, 88, 104750.
10. Ghoddusi, H., & Emamzadehfard, S. (2017). Optimal hedging in the US natural gas market: The effect of maturity and cointegration. *Energy Economics*, 63, 92–105.
11. Hanly, J. (2017). Managing energy price risk using futures contracts: A comparative analysis. *Energy Journal*, 38(3), 93–112.
12. Johnson, L. L. (1960). The theory of hedging and speculation in commodity futures. *Review of Economic Studies*, 27(3), 139–151.
13. Knill, A. M., Minnick, K., & Nejadmalayeri, A. (2006). Selective hedging, information asymmetry, and futures prices. *Journal of Business*, 79(3), 1475–1501.
14. Li, J., Huang, L., & Li, P. (2021). Are Chinese crude oil futures good hedging tools? *Finance Research Letters*, 38, 101514.
15. Lv, F., Yang, C., & Fang, L. (2020). Do the crude oil futures of the Shanghai International Energy Exchange improve asset allocation of Chinese petrochemical-related stocks? *International Review of Financial Analysis*, 71, 1–1537.
16. Ku, Y. H. H., Chen, H. C., & Chen, K. H. (2007). On the application of the dynamic conditional correlation model in the estimating optimal time-varying hedge ratios. *Applied Economics Letters*, 14(7), 503–509.
17. Ripple, R. D., & Moosa, I. A. (2007). Hedging effectiveness and futures contract maturity: The case of NYMEX crude oil futures. *Applied Financial Economics*, 17(9), 683–689.
18. Toyoshima, Y., Nakajima, T., & Hamori, S. (2013). Hedging strategy: New evidence from the data of the financial crisis. *Applied Financial Economics*, 23(12), 1033–1041.
19. Wang, Y., Geng, Q., & Meng, F. (2019). Futures hedging in crude oil markets: A comparison between minimum-variance and minimum-risk frameworks. *Energy*, 181(15), 815–826.
20. Wang, Y., Wu, C., & Yang, L. (2015). Hedging with futures: Does anything beat the naïve hedging strategy? *Management Science*, 61(12), 2870–2889.

Chapter 5

Market Risk of a Power Generation Business



5.1 Introduction

Power generation businesses have unstable profit environments. While the volatility of fuel and electricity markets is extremely high, companies must make large-scale capital investments that take a long time from decision making to the start of operation and long-term mass fuel procurement under take-or-pay contracts. Therefore, the daily marginal profit (total electricity sales minus total fuel costs) is unstable, and can even be negative. In addition, the amount of electricity sold, that is, the amount of fuel required, is uncertain. It is extremely difficult to maximize profits while minimizing risk.

In the long run, such firms should focus on optimizing power generation equipment, long-term contract fuel procurement, and long-term contract power sales portfolios. Assuming some future cases that consider both energy indicators (e.g., price, demand, and supply) and general economic indicators (e.g., inflation rate, interest rate, and exchange rate), we should aim to design an optimum portfolio by adjusting business parameters (e.g., the ratio of each fuel type, performance of each generator, ratio of long-term contracts, and price determination formula). In the short term, as a prerequisite for existing equipment and long-term contracts, measures should be taken (e.g., strengthening equity capital, increasing accounting allowances, and restructuring portfolios through short-term trading contracts) that consider only normal market price fluctuations as a risk factor. For this purpose, it is necessary to monitor the spillover between markets related to the power generation business and measure market risk in business operations appropriately.

Several existing studies analyze the relationship between the value and/or risk of energy-related business companies and other indexes. Reboredo [8] examines the systemic risk and dependence between oil prices and six clean energy stock price indexes using copulas to express the dependence structure and measure the conditional value-at-risk (VaR). Siburg et al. [9] propose forecasting the VaR of bivariate portfolios using copulas calibrated using nonparametric sample estimates of the

coefficient of lower tail dependence. We demonstrate the superiority of the proposed model over the conventional parametric model by analyzing the benchmarks of electricity, crude oil, natural gas, and coal, and the equities of five global electric utility service providers. Boubaker and Sghaier [2] propose a Markov-switching copula model to investigate the presence of regime changes in the time-varying dependence structure between crude oil benchmark price returns and stock price index returns in six Gulf Cooperation Council countries. Zhang et al. [11] introduce the volatility threshold dynamic conditional correlations approach, which is an extension of the dynamic conditional correlation (DCC) and asymmetric DCC (ADCC), to investigate the spillover of stock market volatility indexes on oil and gas markets. Ji et al. [5] introduce six time-varying copulas to measure four types of delta conditional VaRs and examine the impact of uncertainty on crude oil, natural gas, and clean energy companies. This study considers three proxies for economic policy uncertainty, financial market uncertainty, and energy market uncertainty and reveals the magnitude and asymmetric effects of their influence. Hanif et al. [4] examine the dependence structure of major energy firm equities in the United States and the European Union (EU). This study discusses the effect on the sensitivity of energy equity portfolios to crude oil prices to the selection of bivariate or multivariate copulas. Zhang et al. [12] analyze both the return and volatility spillover between energy and electricity utility stock markets and conclude that investors should monitor current economic events to hedge their risks through proper portfolio diversification. Wu et al. [10] construct a total systemic risk index for global energy companies and investigate whether stock market volatility, energy market risks, and exchange rate risks are factors driving this index. Liu et al. [6] calculate the spillover index between a newspaper-based index that reflects uncertainty in the stock market caused by infectious diseases and three renewable energy stock indices. Mzoughi et al. [7] study the dependence structure and risk spillover between green financial securities and the energy commodity index.

This chapter focuses on the market risk in the power generation business for short-term risk management. As an analysis case, we take an electric power company in the European market. The company procures three types of fuel at their respective prices linked to representative crude oil, natural gas, and coal futures prices and sells electricity at prices linked to typical wholesale electricity futures. Furthermore, the company buys carbon credit futures. First, we provide an overview of the descriptive statistics and time plots of each futures price. Second, we measure the spillover effects of both the return and volatility series between these price indices as well as the total connectedness. Third, we decompose them spectrally. Finally, we measure the market risk of this business using these futures markets. Although this example is extremely simple, it can clarify the market risk of actual power generation companies as an extension of this study.

The remainder of this chapter is organized as follows. Section 5.2 explains the methodologies. Section 5.3 presents the data and the preliminary analyses. Section 5.4 provides the results of the main analyses and measurements. Section 5.5 discusses the overall concluding remarks and considerations.

5.2 Methodology

We examine the spillover effect among five markets, namely the crude oil, natural gas, coal, electricity, and carbon credit markets. Then, we measure the risk of a portfolio consisting of these five indices.

First, we confirm the descriptive statistics and time plots for these price series as an overview. Then, we generate the return (see Sect. 3.4.1) and volatility series using the exponential generalized autoregressive conditional heteroscedasticity (EGARCH) model (see Sect. 3.3.3) and confirm the descriptive statistics and time plots for these series. We test the stationarity of these variables using the augmented Dickey Fuller (ADF) unit root test (see Sect. 2.2.2) to confirm whether we can represent these variables by the vector moving average (VMA) model.

Second, by measuring the connectedness proposed by Diebold and Yilmaz [3] (see Sect. 3.3.1) and spectrally decomposing the connectedness as in Baruník and Křehlík [1] (see Sect. 3.3.2), we grasp the spillover effects of returns and volatility among these markets.

Finally, estimating four types of copulas, (the Gaussian, *t*, Clayton, and Gumbel copulas, see Sect. 2.5.2), we measure the risk of the portfolio. This portfolio consists of long positions in crude oil, natural gas, coal, and carbon credits, and short positions in electricity. The risk of a long position is in the left tail of the return distribution, and the risk of a short position is in the right tail of the return distribution. We generate the distribution of daily returns for the portfolio by simulating each estimated copula 500,000 times. We then calculate the VaR and expected shortfall.

5.3 Data and Preliminary Analyses

This chapter uses an index to represent each European commodities market. Specifically, we adopt Brent futures as the crude oil index, Title Transfer Facility (TTF) futures as the natural gas index, Rotterdam futures as the coal index, French baseload (FrenchBL) futures as the electricity index, and EU allowance (EUA) futures as the carbon credit index. We use daily data from January 1, 2015 to December 31, 2021, though we extract only the days on which all five prices are available. The Brent, TTF, Rotterdam, FrenchBL, and EUA prices are expressed in EUR per barrel, mmbtu, ton, MWh, and tonneCO₂, respectively. We obtain all the data from Bloomberg.

5.3.1 Price Series

We confirm the representative statistics for all the price series. Table 5.1 summarizes the descriptive statistics. For four indexes (TTF, Rotterdam, FrenchBL, and EUA) other than Brent, the mean is larger than the median. These distributions contain

Table 5.1 Descriptive statistics (price series)

	Brent	TTF	Rotterdam	FrenchBL	EUA
Period	From 1/2/2015 to 12/31/2021				
Observations	502	502	502	502	502
Mean	50.53	20.81	65.64	57.24	19.79
Median	52.40	17.52	55.69	44.55	15.98
Maximum	74.94	180.27	236.54	772.11	88.88
Minimum	17.83	3.51	34.69	18.58	3.93
Standard deviation	10.78	17.56	26.60	57.09	16.79
Skewness	-0.23	4.12	2.48	5.91	1.48
Kurtosis	2.64	23.42	11.88	50.29	4.92
Jarque–Bera	25 (0.00)	36,139 (0.00)	7702 (0.00)	177,117 (0.00)	927 (0.00)

Note *p*-values are in parentheses

many outliers in the right tail. Although the range of these four variables was wide, the standard deviation was not large. The outliers in the right tail may be accidental. These four distributions have a positive skewness and a long right tail, consistent with the fact that these four variables have maximum values much larger than their respective means. These four distributions had kurtosis values greater than 3, meaning that their distributions have sharp peaks and long fat-tails. Brent has the opposite characteristics. Its median is larger than its mean and its skewness is negative. Its distribution contained many outliers in the long left tail. This is consistent with the difference between its maximum and mean being smaller than the difference between its mean and minimum. The distribution had a kurtosis of less than 3, indicating that the distribution has a rounded peak and short, thin tails. For all variables, the Jarque–Bera test rejected the normal distribution hypothesis.

Figure 5.1 shows the time plots of the five prices. In the short period after the spring of 2021, the four markets besides Brent rose sharply. These time plots are consistent with the descriptive statistics.

5.3.2 Return Series

Table 5.2 shows the basic statistics for these five return series. Similar to the price series results, only Brent is characteristic. For TTF, Rotterdam, FrenchBL, and EUA, the mean was larger than the median. This result implies that these distributions have more outliers in the right tail than in the left tail. On the other hand, the distribution of the Brent return series had a smaller mean than the median and more outliers in the left tail than in the right tail. The standard deviation of the FrenchBL return series was relatively large and the range was relatively wide. We can conclude that FrenchBL was the riskiest commodity during this period. The distributions of the

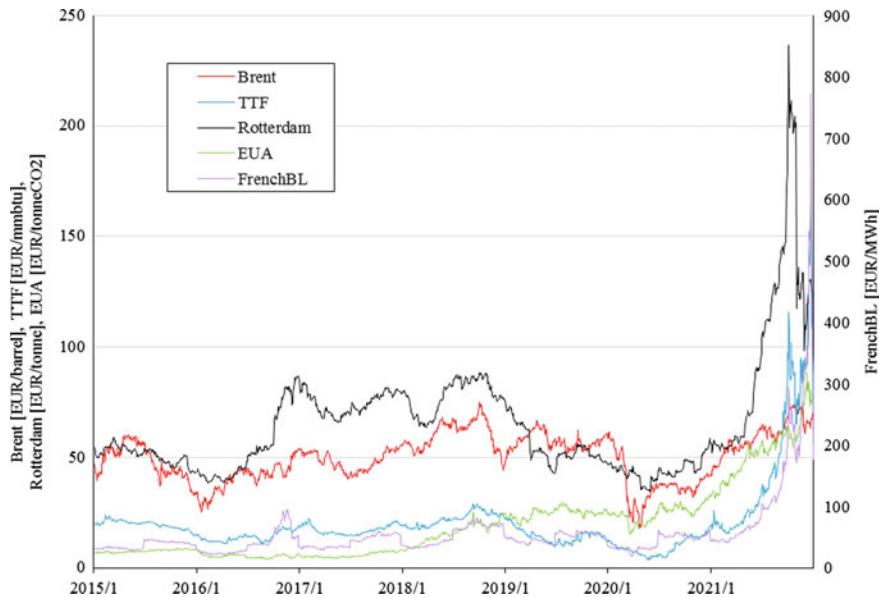


Fig. 5.1 Commodity price series, Europe

Brent, Rotterdam, and EUA return series have a negative skewness, implying a long left tail. In contrast, the skewness for TTF and FrenchBL is positive, indicating a long right tail. Each return series had a kurtosis larger than 3. This result means that each distribution had a sharp peak and a long fat-tail. The Jarque–Bera test rejects the hypothesis that each series has skewness and kurtosis following a normal distribution.

Table 5.2 Descriptive statistics (return series)

	Brent	TTF	Rotterdam	FrenchBL	EUA
Observations	1788	1788	1788	1788	1788
Mean	0.057%	0.136%	0.068%	0.196%	0.178%
Median	0.115%	– 0.037%	0.043%	0.000%	0.134%
Maximum	21.9%	37.3%	19.1%	71.9%	13.6%
Minimum	– 25.1%	– 23.3%	– 41.8%	– 50.9%	– 17.7%
Standard deviation	0.0267	0.0371	0.0210	0.0458	0.0289
Skewness	– 0.24	1.36	– 3.83	3.57	– 0.24
Kurtosis	17.21	17.19	108.29	78.85	6.64
Jarque–Bera	15,066 (0.00)	15,555 (0.00)	830,296 (0.00)	432,403 (0.00)	1001 (0.00)

Note p -values are in parentheses

Figure 5.2 displays the time plots of these return series. As each standard deviation of the return series indicates, we observe a highly volatile FrenchBL return series and less volatile Brent, Rotterdam, and EUA return series. The Brent return series is more volatile during January 2015 to December 2016 and April to June 2020 than during the other periods. For the TTF, the return series fluctuates more from May 2019 to December 2021 than in the other periods. The Rotterdam and FrenchBL return series are stable for some time; however, they often spike. The spikes for FrenchBL were larger and more frequent than those for Rotterdam. The EUA return series has almost the same fluctuation range over the entire period. As this figure illustrates, these five return series do not seem synchronized.

Table 5.3 shows the results of the ADF unit root test for all return series. All tests adopt a model that has neither a constant term nor time-trend time. All unit root hypotheses were rejected at the 1% significance level and all return series are stationary. Therefore, we can represent these vector autoregressive (VAR) model using the VMA model.

5.3.3 Volatility Series

To measure risk spillover effects, we generate a volatility series for these commodities. Table 5.4 shows the estimated autoregressive (AR)-EGARCH model. The lag length of each term was selected based on the Schwarz Bayesian information criterion (SBIC). We determine the lag order of the AR, autoregressive conditional heteroskedasticity (ARCH), generalized ARCH (GARCH), and asymmetric terms in this order. The ARCH term (α_1) and GARCH term (β_1) are statistically significant at the 1% level for all commodities. Each generalized error distribution (GED) parameter estimate is less than 2, meaning that each error term has a fat-tailed distribution.

Table 5.5 shows the descriptive statistics of these volatility series. For all these variables, the mean is larger than the median; that is, these volatility distributions have more outliers in the right tail than in the left tail. For TTF and FrenchBL, the mean is relatively large, the maximum is extremely large, and the minimum is relatively small. These two commodities are riskier than the other three commodities. All volatility series have a positive skewness; that is, these distributions have a long right tail. All volatility series have kurtosis greater than 3; that is, these distributions have a sharp peak and fat-tails. The Jarque–Bera test results reject the hypothesis that each series has skewness and kurtosis following a normal distribution.

Figure 5.3 displays the time plots of these volatility series. The spikes for the TTF and FrenchBL volatility series are very high. Although the volatility for Brent, Rotterdam, and EUA is relatively small, we see some spikes in their series. We cannot deny the synchronization between these spikes.

We confirm whether Diebold and Yilmaz's [3] approach is applicable to these volatility series to measure the spillover effects between these volatility series. The ADF test, which uses a model with a constant term and without a time-trend term,

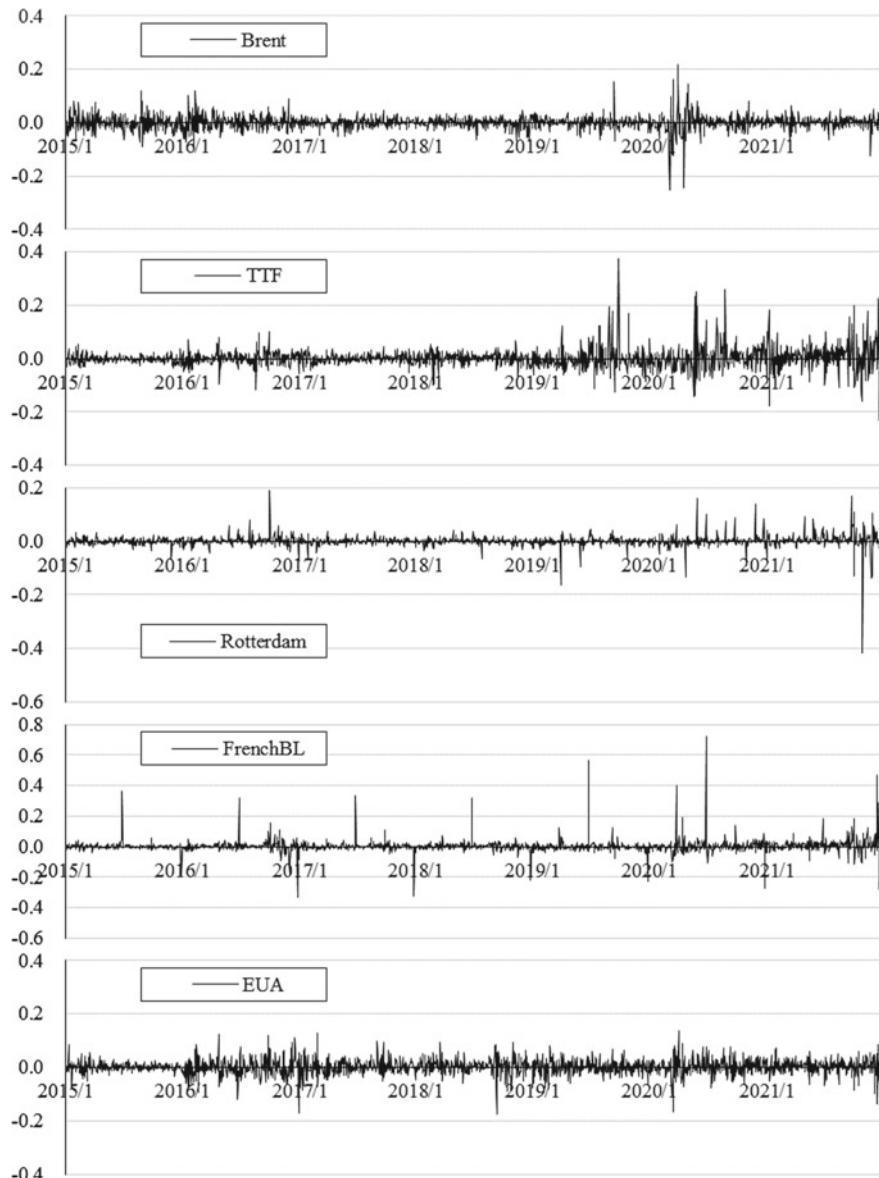


Fig. 5.2 Commodity return series, Europe

Table 5.3 ADF unit root test results (return series)

Return series	ADF-t statics
Brent	- 41.48* (0.00)
TTF	- 38.53* (0.00)
Rotterdam	- 40.15* (0.00)
FrenchBL	- 20.67* (0.00)
EUA	- 45.22* (0.00)

Note * indicates rejection of the unit-root hypothesis at the 1% significance level. p -values are in parentheses

Table 5.4 Estimated AR-EGARCH model

		Brent	TTF	Rotterdam	FrenchBL	EUA
Mean equation	c_0	0.001 (0.01)	0.000 (0.97)	0.000 (0.02)	0.000 (0.75)	0.002 (0.00)
	c_1	- 0.042 (0.05)	0.051 (0.03)	0.074 (0.00)	0.081 (0.00)	- 0.079 (0.00)
Variance equation	θ	- 0.278 (0.00)	- 0.262 (0.00)	- 0.124 (0.02)	- 0.709 (0.00)	- 0.340 (0.00)
	α_1	0.147 (0.00)	0.263 (0.00)	0.062 (0.00)	0.191 (0.00)	0.207 (0.00)
	β_1	0.978 (0.00)	0.991 (0.00)	0.990 (0.00)	0.917 (0.00)	0.975 (0.00)
	γ_1	- 0.095 (0.00)	0.006 (0.71)	0.010 (0.21)	- 0.011 (0.58)	- 0.010 (0.50)
GED	v	1.235 (0.00)	1.353 (0.00)	0.722 (0.00)	0.687 (0.00)	1.340 (0.00)

Note p -values are in parentheses

Table 5.5 Descriptive statistics (volatility series)

		Brent	TTF	Rotterdam	FrenchBL	EUA
Observations		1787	1787	1787	1787	1787
Mean		6.42×10^{-4}	1.33×10^{-3}	2.64×10^{-4}	1.17×10^{-3}	8.39×10^{-4}
Median		4.53×10^{-4}	6.89×10^{-4}	2.00×10^{-4}	7.00×10^{-4}	7.03×10^{-4}
Maximum		8.84×10^{-3}	1.80×10^{-2}	1.75×10^{-3}	7.33×10^{-2}	3.92×10^{-3}
Minimum		1.10×10^{-4}	7.12×10^{-5}	1.11×10^{-4}	8.00×10^{-5}	1.39×10^{-4}
Standard deviation		7.94×10^{-4}	1.92×10^{-3}	2.12×10^{-4}	2.64×10^{-3}	5.57×10^{-4}
Skewness		5.31	3.73	4.26	16.74	1.95
Kurtosis		37.6	21.5	23.7	383.6	8.0
Jarque–Bera		97,639 (0.00)	29,595 (0.00)	37,359 (0.00)	10,869,514 (0.00)	2970 (0.00)

Note p -values are in parentheses

rejects the unit root hypothesis for all volatility series. Table 5.6 provides the ADF unit root test results. The vector of these volatility series is a stationary covariance process. Therefore, we can represent the VAR model as the VMA model.

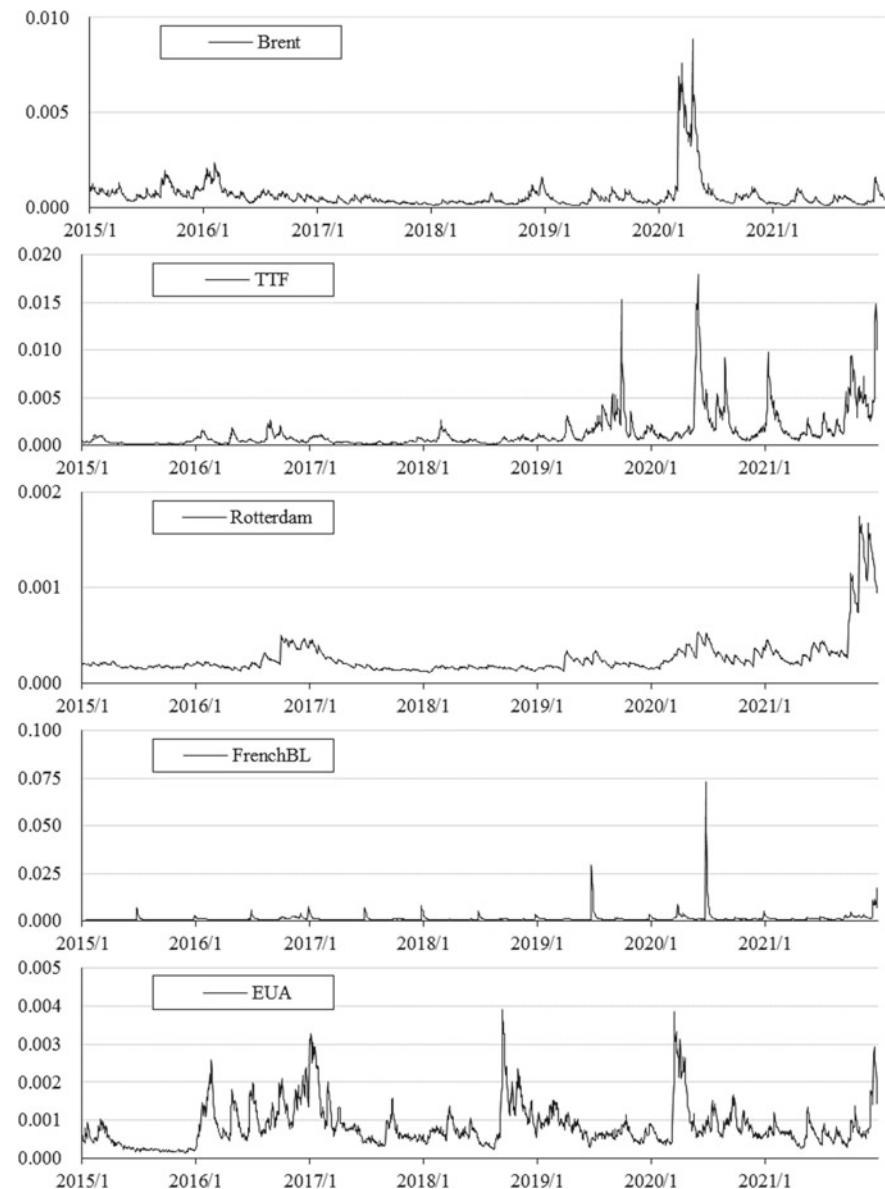


Fig. 5.3 Commodity volatility series, Europe

Table 5.6 ADF unit root test results (volatility series)

Volatility series	ADF-t statics
Brent	- 5.71* (0.00)
TTF	- 5.43* (0.00)
Rotterdam	- 3.55* (0.01)
FrenchBL	- 18.03* (0.00)
EUA	- 5.89* (0.00)

Note * indicates rejection of the unit-root hypothesis at the 1% significance level. *p*-values are in parentheses

5.4 Analysis Results

This section provides the spillover effects of the return series and the volatility series between the Brent, TTF, Rotterdam, FrenchBL, and EUA markets. We then present the results of the spectral decomposition of these effects. Moreover, after estimating four types of quinquevariate copulas, we measure the market risk of the example power generation business.

5.4.1 Return Series

Table 5.7 presents the estimated coefficients of the VAR model for the Brent, TTF, Rotterdam, FrenchBL, and EUA returns series. The lag order is 1, which we selected based on the SBIC.

Table 5.8 lists the variance–covariance matrix of the residuals of this VAR model.

Figure 5.4 illustrates the spillover indexes, which we obtain from Tables 5.7 and 5.8. Only indexes of more than 5% were extracted. We observe spillover effects among the natural gas, electricity, and carbon credit markets. Furthermore, the spillover effects from TTF to Rotterdam and from the EUA to Brent are 5.85 and 5.10%, respectively. Finally, the total connectedness is 18.71%. We can conclude that the level of market integration is not very high.

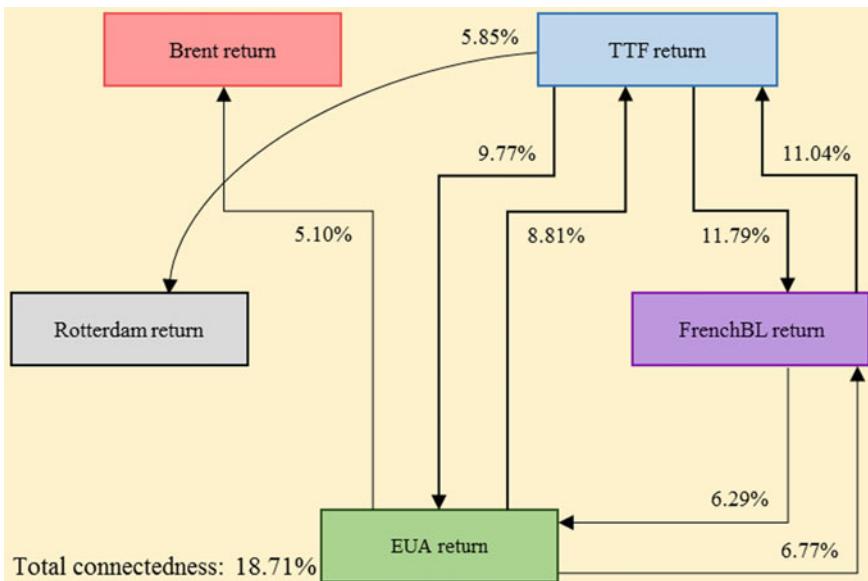
We present the spectral analysis of these connectedness indices in Table 5.9. The total connectedness from 1 to 5 business days, from 6 to 20 business days, and over 21

Table 5.7 VAR model (return series)

	Brent (<i>t</i>)	TTF (<i>t</i>)	Rotterdam (<i>t</i>)	FrenchBL (<i>t</i>)	EUA (<i>t</i>)
Brent (<i>t</i> -1)	0.023	- 0.022	- 0.001	- 0.049	- 0.039
TTF (<i>t</i> -1)	0.014	0.048	0.079	0.125	- 0.053
Rotterdam (<i>t</i> -1)	0.017	0.076	0.024	0.068	- 0.095
FrenchBL (<i>t</i> -1)	- 0.004	0.077	0.008	0.032	- 0.012
EUA (<i>t</i> -1)	- 0.029	- 0.032	- 0.040	0.093	- 0.021

Table 5.8 Residual variance–covariance matrix, VAR model (return series)

	Brent	TTF	Rotterdam	FrenchBL	EUA
Brent	7.12×10^{-4}	1.22×10^{-4}	6.53×10^{-5}	7.47×10^{-5}	1.80×10^{-4}
TTF	1.22×10^{-4}	1.36×10^{-3}	1.71×10^{-4}	6.13×10^{-4}	3.63×10^{-4}
Rotterdam	6.53×10^{-5}	1.71×10^{-4}	4.23×10^{-4}	2.02×10^{-4}	1.05×10^{-4}
FrenchBL	7.47×10^{-5}	6.13×10^{-4}	2.02×10^{-4}	2.05×10^{-3}	3.63×10^{-4}
EUA	1.80×10^{-4}	3.63×10^{-4}	1.05×10^{-4}	3.63×10^{-4}	8.23×10^{-4}

**Fig. 5.4** Spillover effects between return series

business days is 14.06, 3.46, and 1.19%, respectively. Since the total connectedness for the entire period is 18.71%, the short-term factors contribute the most to the return spillover; the spillover effect is largely dependent on events occurring within approximately one week. Each spillover effect between these return series tends to be almost the same as total connectedness. The short-term factor is the largest, the medium-term factor is the second largest, and the long-term factor is the smallest.

5.4.2 Volatility Series

We estimate the VAR model for the volatility series using a lag order of 1 based on the SBIC. Table 5.10 reports the estimated results.

Table 5.9 Spillover index and spectral analysis (return series)

To	From						Bandwidth
	Brent (%)	TTF (%)	Rotterdam (%)	FrenchBL (%)	EUA (%)	Others (%)	
Brent	72.73	1.06	0.96	0.29	4.22	1.31	0 \leqq date \leqq 5
TTF	0.93	56.98	2.73	7.42	6.72	3.56	
Rotterdam	0.92	3.46	66.71	2.83	2.07	1.85	
FrenchBL	0.24	7.84	2.65	58.86	4.58	3.06	
EUA	3.96	8.89	2.93	5.62	63.20	4.28	
Others	1.21	4.25	1.85	3.23	3.52	14.06	
Brent	14.31	0.27	0.25	0.05	0.66	0.25	6 \leqq date \leqq 20
TTF	0.16	13.37	0.99	2.67	1.55	1.08	
Rotterdam	0.21	1.76	14.40	1.07	0.44	0.70	
FrenchBL	0.04	2.93	0.91	13.73	1.63	1.10	
EUA	0.38	0.69	0.10	0.52	9.97	0.34	
Others	0.16	1.13	0.45	0.86	0.86	3.46	
Brent	4.79	0.09	0.09	0.02	0.22	0.08	21 \leqq date
TTF	0.05	4.59	0.35	0.94	0.54	0.38	
Rotterdam	0.07	0.63	4.90	0.39	0.15	0.25	
FrenchBL	0.01	1.02	0.32	4.68	0.56	0.38	
EUA	0.12	0.19	0.02	0.15	3.27	0.10	
Others	0.05	0.39	0.15	0.30	0.29	1.19	
Brent	91.83	1.42	1.29	0.35	5.10	1.63	Total
TTF	1.14	74.94	4.07	11.04	8.81	5.01	
Rotterdam	1.20	5.85	86.00	4.29	2.66	2.80	
FrenchBL	0.29	11.79	3.88	77.27	6.77	4.55	
EUA	4.46	9.77	3.06	6.29	76.43	4.71	
Others	1.42	5.77	2.46	4.39	4.67	18.71	

Table 5.10 VAR model (volatility series)

	Brent	TTF	Rotterdam	FrenchBL	EUA
Brent ($t - 1$)	0.968	- 0.000	0.002	0.046	0.015
TTF ($t - 1$)	- 0.001	0.945	0.001	0.060	- 0.002
Rotterdam ($t - 1$)	0.030	0.285	0.990	0.678	0.069
FrenchBL ($t - 1$)	- 0.000	0.001	- 0.000	0.658	0.001
EUA ($t - 1$)	0.009	0.002	- 0.001	0.125	0.964

The variance–covariance matrix of the residuals of this VAR model, which we require for the Diebold and Yilmaz [3] approach, is shown in Table 5.11.

We depict the spillover indices, which we calculate using Tables 5.10 and 5.11, in Fig. 5.5. Only indexes of more than 5% are extracted. We can observe spillover effects from TTF to the EUA and Rotterdam, and from the EUA to Brent. These results are similar to those for the return series. However, besides these effects, Fig. 5.5 looks completely different from Fig. 5.4. We should note that return management and risk management differ. The spillover from Brent to the EUA and from Rotterdam to TTF is very strong. The connectedness from Rotterdam to the EUA and FrenchBL is 14.91 and 7.17%, respectively. Finally, total connectedness is 20.14%. From a risk perspective, we can conclude that the market integration is at a medium level.

Table 5.12 provides the results of the spectral decomposition of these connectedness indices. The total connectedness of these volatility series is 20.14%; specifically,

Table 5.11 Residual variance–covariance matrix, VAR model (volatility series)

	Brent	TTF	Rotterdam	FrenchBL	EUA
Brent	4.47×10^{-8}	5.90×10^{-11}	1.62×10^{-10}	5.41×10^{-10}	4.47×10^{-9}
TTF	5.90×10^{-11}	3.15×10^{-7}	7.92×10^{-10}	1.26×10^{-7}	2.51×10^{-9}
Rotterdam	1.62×10^{-10}	7.92×10^{-10}	9.03×10^{-10}	6.26×10^{-9}	2.65×10^{-10}
FrenchBL	5.41×10^{-10}	1.26×10^{-7}	6.26×10^{-9}	3.75×10^{-6}	1.62×10^{-8}
EUA	4.47×10^{-9}	2.51×10^{-9}	2.65×10^{-10}	1.62×10^{-8}	2.53×10^{-8}

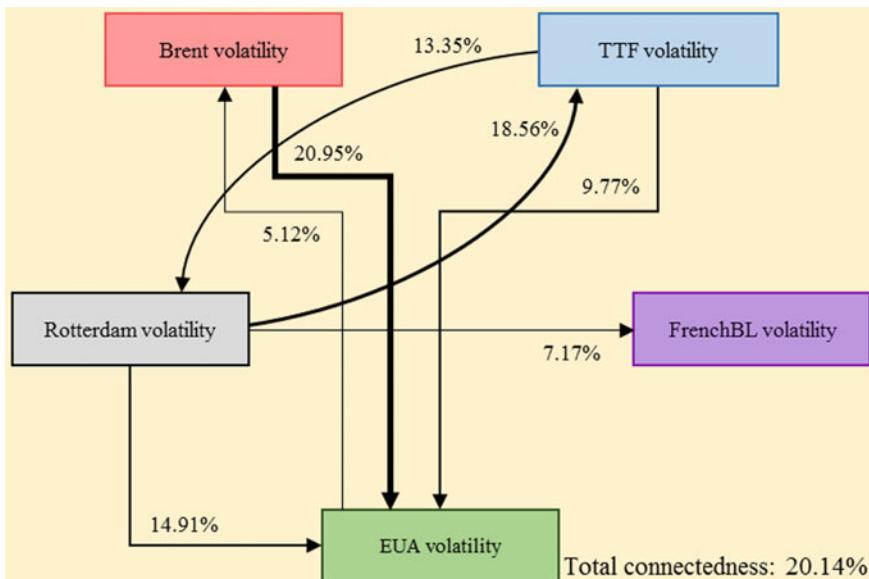


Fig. 5.5 Spillover effects between volatility series

0.20% for short-term, 0.56% for medium-term, and 19.38% for long-term factors. The spillover effects of volatility can be explained primarily by long-term factors. Most spillover is caused by events that occurred more than one month previously. Each spillover effect between these volatility series tends to be almost the same as total connectedness. Events within a week have the least impact, events that occurred between a month and two weeks prior have the second least impact, and events that occurred more than one month prior have the largest impact.

Table 5.12 Spillover index and spectral analysis (volatility series)

To	From						Bandwidth
	Brent (%)	TTF (%)	Rotterdam (%)	FrenchBL (%)	EUA (%)	Others (%)	
Brent	1.70	0.00	0.03	0.00	0.01	0.01	$0 \leq \text{date} \leq 5$
TTF	0.02	2.86	0.13	0.02	0.00	0.03	
Rotterdam	0.05	0.12	0.04	0.00	0.00	0.03	
FrenchBL	0.01	0.24	0.12	30.06	0.06	0.09	
EUA	0.04	0.01	0.13	0.00	1.38	0.03	
Others	0.02	0.07	0.08	0.01	0.01	0.20	
Brent	6.47	0.02	0.11	0.00	0.05	0.04	$6 \leq \text{date} \leq 20$
TTF	0.08	9.98	0.44	0.10	0.00	0.13	
Rotterdam	0.17	0.40	0.13	0.00	0.00	0.12	
FrenchBL	0.02	0.40	0.13	34.01	0.13	0.13	
EUA	0.17	0.09	0.47	0.01	5.28	0.15	
Others	0.09	0.18	0.23	0.02	0.04	0.56	
Brent	83.44	0.05	3.00	0.05	5.06	1.63	$21 \leq \text{date}$
TTF	0.80	65.64	17.99	1.74	0.21	4.15	
Rotterdam	4.06	12.84	80.62	1.44	0.14	3.69	
FrenchBL	1.50	3.60	6.92	21.57	1.23	2.65	
EUA	20.74	0.40	14.31	0.84	56.13	7.26	
Others	5.42	3.38	8.45	0.81	1.33	19.38	
Brent	91.61	0.08	3.14	0.05	5.12	1.68	Total
TTF	0.89	78.48	18.56	1.86	0.21	4.30	
Rotterdam	4.27	13.35	80.79	1.44	0.14	3.84	
FrenchBL	1.53	4.25	7.17	85.64	1.41	2.87	
EUA	20.95	0.50	14.91	0.85	62.79	7.44	
Others	5.53	3.64	8.76	0.84	1.38	20.14	

Table 5.13 Parameters of each marginal distribution function (return series)

	Brent	TTF	Rotterdam	FrenchBL	EUA
Mean	5.69×10^{-4}	1.36×10^{-3}	6.83×10^{-4}	1.96×10^{-3}	1.78×10^{-3}
Standard deviation	2.67×10^{-2}	3.71×10^{-2}	2.10×10^{-2}	4.58×10^{-2}	2.89×10^{-2}

5.4.3 Risk Measurement

We measure the market risk of a power company that procures crude oil, natural gas, coal, and carbon credits and sells electricity. Regarding this company as a portfolio consisting of long positions in fuels and carbon credit futures and short positions in electricity futures, we calculate the VaR and expected shortfall of this company through a simulation using random numbers following the four types of quinquevariate copulas: the Gaussian, t , Clayton, and Gumbel copulas. We selected the VaR and expected shortfall obtained by the simulation using the copula with the smallest Akaike information criterion (AIC).

Table 5.13 presents the means and standard deviations, which are the parameters of the marginal distribution function of each return series.

We estimate each copula using the procedure described in Sect. 2.5.2. Chapters 2, 3, and 5 estimate bivariate, trivariate, and quinquevariate copulas, respectively.

Figure 5.6 plots the relationship between the degree of freedom and the log-likelihood for the t copula estimates. We find the maximum log-likelihood when the degree of freedom is 5. Therefore, we estimate the t copula parameters; that is, $\widehat{\Sigma}$, with degree of freedom 5.

Table 5.14 provides all the estimated copulas. We generate the distribution of daily returns for the portfolio by simulating each estimated copula 500,000 times. Then, we calculate the VaR and expected shortfall. Table 5.14 provides the measured VaRs and expected shortfalls based on these copulas. According to the Gumbel copula with the minimum AIC, the VaR is 3.87% and the expected shortfall is 4.83%. This portfolio has a 1% chance of losing more than 3.87% one day later, with an average loss of 4.83% at a probability of 1%.

5.5 Concluding Remarks

This chapter investigates the market risk of a power generation business by defining a power generation company as an institutional investor who manages an energy portfolio consisting of long positions in fuel futures and short positions in electricity futures. If a company trades fuels and electricity at prices pegged to the prices of highly liquid commodities, we can easily provide an overview and measure the market risk of that business.

Electricity is a major source of economic activity. Power companies broadly accommodate economic activity, making power-generation businesses essential.

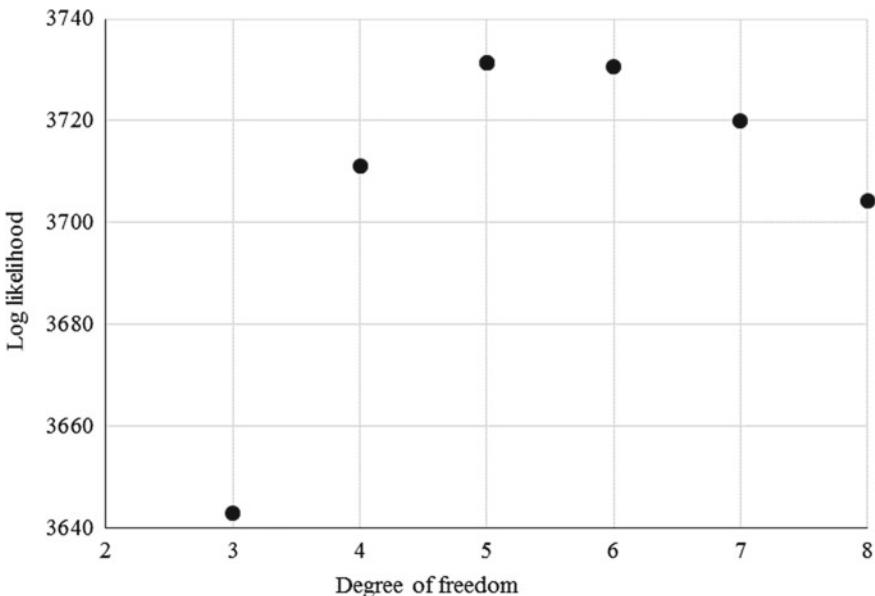


Fig. 5.6 Degree of freedom and log-likelihood in the t copula estimates (European commodity portfolio)

Even if a physically stable supply of electricity is guaranteed, the sudden withdrawal of power-generation companies would have a significant impact on the economy and must be avoided. Because the market risk of the power business is measurable, the government should strictly define the method of measuring the market risk of the power generation business and regulate the capital adequacy and liquidity ratios of power generation companies based on the measured risk. Each power generation company should disclose its market risk and how it measures market risk to shareholders, creditors, and customers as accountability for business continuity.

This chapter analyzes a power generation business in Europe as a case study. This hypothetical business buys fuel oil, natural gas, fuel coal, and carbon credits at prices linked to the futures prices of the Brent, TTF, Rotterdam, and EUA, and sells electricity at a price linked to the futures price of the FrenchBL.

We measure the intermarket spillover effects of the volatility and return series using Diebold and Yilmaz's [3] approach. We then spectrally decompose the results to determine the period during which the spillover factor occurred. Finally, we calculate the VaR and expected shortfall of the generation business by using quinquevariate copulas. The total connectedness of the return series and volatility series are 18.71 and 20.14%, respectively. The level of market integration in terms of the spillover effects of the return series is about the same as that in terms of the spillover effects of the volatility series, however, the breakdown is different. The return spillover effects among the TTF, FrenchBL, and EUA are strong, while the volatility spillover effects from Rotterdam to the EUA, from Brent to the EUA, and bidirectional between TTF

Table 5.14 Risk measurement using estimated copulas (European commodity portfolio)

Gaussian	AIC	34.134				
$\widehat{\Sigma}$	1.000	0.151	0.156	0.098	0.226	
	0.151	1.000	0.311	0.486	0.371	
	0.156	0.311	1.000	0.293	0.182	
	0.098	0.486	0.293	1.000	0.343	
	0.226	0.371	0.182	0.343	1.000	
	Value-at-risk*	3.50%				
<i>t</i>	Expected shortfall*	4.21%				
	AIC	15.551				
	Degree of freedom	5				
	$\widehat{\Sigma}$	1.000	0.165	0.179	0.122	0.204
		0.165	1.000	0.328	0.581	0.382
		0.179	0.328	1.000	0.276	0.163
Clayton	0.122	0.581	0.276	1.000	0.403	
	0.204	0.382	0.163	0.403	1.000	
	Value-at-risk*	3.35%				
	Expected shortfall*	4.01%				
	AIC	– 11.966				
	π_c	1.478				
Gumbel	Value-at-risk*	3.78%				
	Expected shortfall*	4.74%				
	AIC	– 13.095				
	π_g	2.237				
Gumbel	Value-at-risk*	3.87%				
	Expected shortfall*	4.83%				

Note * indicates the value at the 99% confidence level

Bold indicates the minimum AIC and the risk measures at that time

and Rotterdam are strong. The return spillover effects depend mostly on events within five business days, whereas the volatility spillover effects depend mostly on factors before 21 business days. The VaR was 3.87% in the simulation using the estimated Gumbel copula. This business had a 1% chance of losing more than 3.87% one day later. The expected shortfall was 4.83%. If this business loses more than 3.87% one day later, then its expected loss will be 4.83%.

It may be possible to measure the market risk of the oil refining business, which procures crude oil and sells petroleum products (i.e., gasoline, kerosene, and heating oil) if it trades raw materials and products at prices pegged to the prices of highly liquid commodity futures. We might measure the market risk of a business that purchases corn, potatoes, and sugar cane and manufactures and sells ethanol. In addition, if readers find companies that earn their operating revenue by trading goods

in highly liquid markets, they may be able to measure market risk as an extension of this study.

References

1. Baruník, J., & Křehlík, T. (2018). Measuring the frequency dynamics of financial connectedness and systemic risk. *Journal of Financial Econometrics*, 16(2), 271–296.
2. Boubaker, H., & Sghaier, N. (2016). Markov-switching time-varying copula modeling of dependence structure between oil and GCC stock markets. *Open Journal of Statistics*, 6(5), 565–589.
3. Diebold, F. X., & Yilmaz, K. (2012). Better to give than to receive: Predictive directional measurement of volatility spillovers. *International Journal of Forecasting*, 28(1), 57–66.
4. Hanif, W., Arreola-Hernandez, J., Shahzad, S. J. H., Hoang, T. H. V., & Yoon, S. M. (2020). Regional and copula estimation effects on EU and US energy equity portfolios. *Applied Economics*, 52(49), 5311–5342.
5. Ji, Q., Liuc, B. Y., Nehler, H., & Uddin, G. S. (2018). Uncertainties and extreme risk spillover in the energy markets: A timevarying copula-based CoVaR approach. *Energy Economics*, 76, 115–126.
6. Liu, T., Nakajima, T., & Hamori, S. (2022). The impact of economic uncertainty caused by COVID-19 on renewable energy stocks. *Empirical Economics*, 62(4), 1495–1515.
7. Mzoughi, H., Urom, C., & Guesmi, K. (2022). Downside and upside risk spillovers between green finance and energy markets. *Finance Research Letters*, 47, 102612.
8. Reboredo, J. C. (2015). Is there dependence and systemic risk between oil and renewable energy stock prices? *Energy Economics*, 48, 32–45.
9. Siburg, K. F., Stoimenov, P., & Weiβ, G. N. F. (2015). Forecasting portfolio-Value-at-Risk with nonparametric lower tail dependence estimates. *Journal of Banking & Finance*, 54, 129–140.
10. Wu, F., Zhang, D., & Ji, Q. (2021). Systemic risk and financial contagion across top global energy companies. *Energy Economics*, 97, 105221.
11. Zhang, Y. J., Chevallier, J., & Guesmi, K. (2017). “De-financialization” of commodities? Evidence from stock, crude oil and natural gas markets. *Energy Economics*, 68, 228–239.
12. Zhang, W., He, X., Nakajima, T., & Hamori, S. (2020). How does the spillover among natural gas, crude oil, and electricity utility stocks change over time? Evidence from North America and Europe. *Energies*, 13(3), 727.

Chapter 6

Alternative to Postface: Market Risk Transfer in Power Companies



Electric power companies often trade different types of energy at prices linked to futures prices on commodity exchanges. They understand what is traded under the contract and quantify its value to appropriately manage the company. In other words, they often trade both the value of energy and the value of market risk in one contract; therefore, they determine the value of energy quantitatively in terms of profitability, and the value of market risk in terms of profitability and business continuity. Ng et al. [11] adopt an option pricing methodology to compare the cost of a long-term contract with a price cap to that of spot purchases. The following two examples are inspired mainly by Ng et al. [11].

First, we introduce an example fuel procurement contract. Long-term liquefied natural gas (LNG) trading contracts with price formulas, which are a function of benchmark crude oil prices, are not uncommon, especially in the Asia-Pacific region. We discuss the case in which an electric power company continues to buy LNG at the price series $P_{LNG,t}$, which is determined by the following equation:

$$P_{LNG,t} = C_0 + C_1 P_{oil,t}, \quad (6.1)$$

where $P_{oil,t}$ is the price series of crude oil and C_0 and C_1 are the coefficients. We can interpret C_0 and C_1 as the fixed costs and the variable costs to exchange crude oil for LNG, respectively. Additionally, C_0 and C_1 include the seller's profit. We can think of exchanging the currency of crude oil for the currency of LNG. If C_1 properly represents the value of LNG for crude oil, then the risk of crude oil price fluctuations due to the external environment is completely passed through to the buyer; that is, the electric power company. The buyer considers the following measures against this risk:

Measure 1. Passing the risk on to power generation costs and ultimately to selling prices.

Measure 2. Hedging the risk with energy derivatives.

Measure 3. Creating accounting allowances that match the risk amount. In some cases, the capital strategy changes.

It is of utmost importance for electric power companies to secure a stable spread between the selling price and power generation cost. Therefore, we base this study on the measures that depend on selling prices. For simplicity, suppose that the company sells electricity at a price consistent with procurement in Eq. (6.1). In other words, we assume that power sales contracts are limited to those that pass the risk of crude oil price fluctuations to the selling price. Next, suppose that the company procures additional LNG with the following price formula to respond to the increase in electricity sales:

$$\begin{aligned} {}^{ad} P_{LNG,t} &= C_0 + C_1 P_{oil,t}, \text{ if } P_{oil,t} \geq P_s \\ {}^{ad} P_{LNG,t} &= C_0 + C_1 P_s, \text{ if } P_{oil,t} < P_s, \end{aligned} \quad (6.2)$$

where ${}^{ad} P_{LNG,t}$ is the price of the additionally procured LNG. In reality, there is probably no extreme price formula. However, it is not uncommon for the gradient for $P_{oil,t} < P_s$ to be smaller than C_1 to reduce the risk that the seller does not recover its investment in LNG manufacturing equipment. We illustrate Eqs. (6.1) and (6.2) in Fig. 6.1. The power company risks that the spread between the selling price and power generation cost is narrower when $P_{oil,t} < P_s$. Assuming that the business balance in Eq. (6.1) is neutral, we can write the business balance in Eq. (6.2) as in Eq. (6.3) and Fig. 6.2:

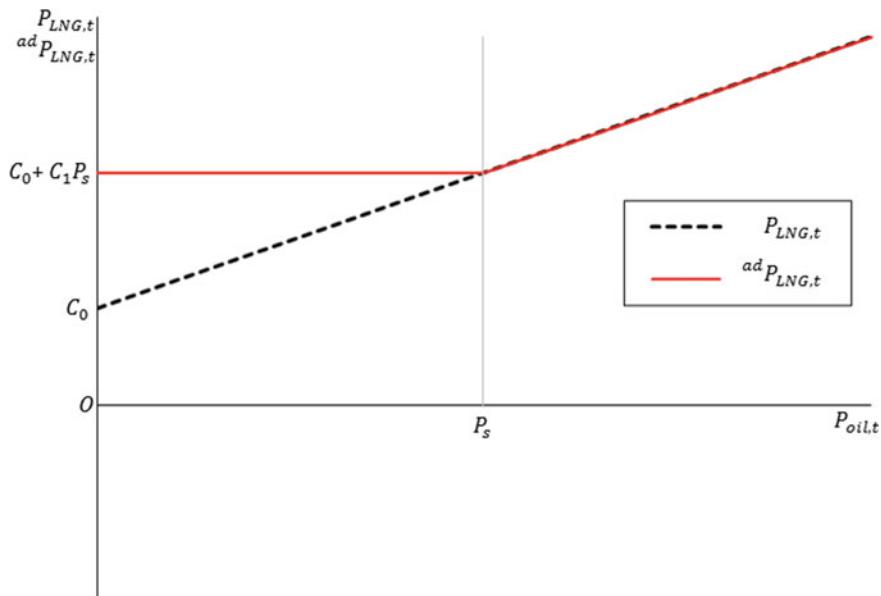
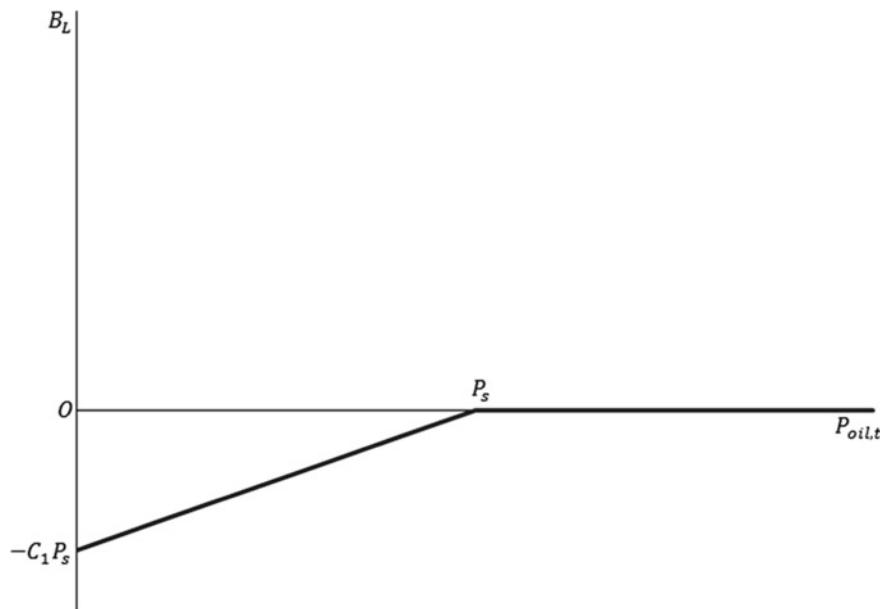
$$\begin{aligned} B_L &= 0, \text{ if } P_{oil,t} > P_s \\ B_L &= C_1(P_{oil,t} - P_s), \text{ if } P_{oil,t} \leq P_s, \end{aligned} \quad (6.3)$$

where B_L is the business balance in Eq. (6.2) assuming that the business balance in Eq. (6.1) is neutral.

We can interpret this as a short position in a crude oil put option according to Kawamoto and Tsuzaki [7]. In practice, taking such risks is unusual for electric power companies. Therefore, the buyer can avoid this risk by taking a long position in a crude oil put option, which we can calculate using the LNG procurement amount in Eq. (6.2) and the value of C_1 . If we do not use derivatives to hedge the risk, it is rational to book an allowance that matches the risk amount. The risk amount is

$$\int_T V_t Put_t dt, \quad (6.4)$$

where T is the contract period, V_t is the position of the put option, and Put_t is the price of the put option. If the put option is listed on an exchange, then we can utilize the price. Otherwise, we must price it theoretically based on market conditions (e.g., the underlying asset price, its volatility, and government bond interest). Equation (6.4) is the swap price of Eq. (6.3) for LNG procured in Eq. (6.2).

**Fig. 6.1** LNG price formulas**Fig. 6.2** Impact of LNG procurement with a lower threshold price on the business balance

Second, we consider an example power sale contract. We discuss the case in which an electric power company continues to sell power in the price series $P_{power,t}$, determined as follows:

$$P_{power,t} = \alpha P_{oil,t} + \beta P_{gas,t} + \gamma P_{coal,t} + \delta, \quad (6.5)$$

where $P_{oil,t}$, $P_{gas,t}$, and $P_{coal,t}$ are the price series for crude oil, natural gas, and coal, respectively. If α , β , γ , and δ appropriately represent fuel power generation efficiency, fuel power generation amount ratio, and fixed cost for each option, then the company can pass on the risk of fuel price fluctuations to the selling price. However, power sales contracts with an upper threshold price are common, as Eq. (6.6) demonstrates.

$$\begin{aligned} {}^N P_{power,t} &= \text{Max}(P_{max}, P_{power,t}) \\ &= \text{Max}(P_{max}, \alpha P_{oil,t} + \beta P_{gas,t} + \gamma P_{coal,t} + \delta), \end{aligned} \quad (6.6)$$

where ${}^N P_{power,t}$ is the power price with an upper threshold price, P_{max} . We illustrate Eqs. (6.5) and (6.6) in Fig. 6.3. The power company takes the risk that the spread between the selling price and power generation cost is narrower when $P_{power,t} > P_{max}$. Assuming that the business balance in Eq. (6.5) is neutral, we can present the business balance in Eq. (6.6) as in Eq. (6.7) and Fig. 6.4.

$$\begin{aligned} B_p &= 0, \text{ if } P_{power,t} \leq P_{max} \\ B_p &= P_{max} - (\alpha P_{oil,t} + \beta P_{gas,t} + \gamma P_{coal,t} + \delta), \text{ if } P_{power,t} > P_{max} \end{aligned} \quad (6.7)$$

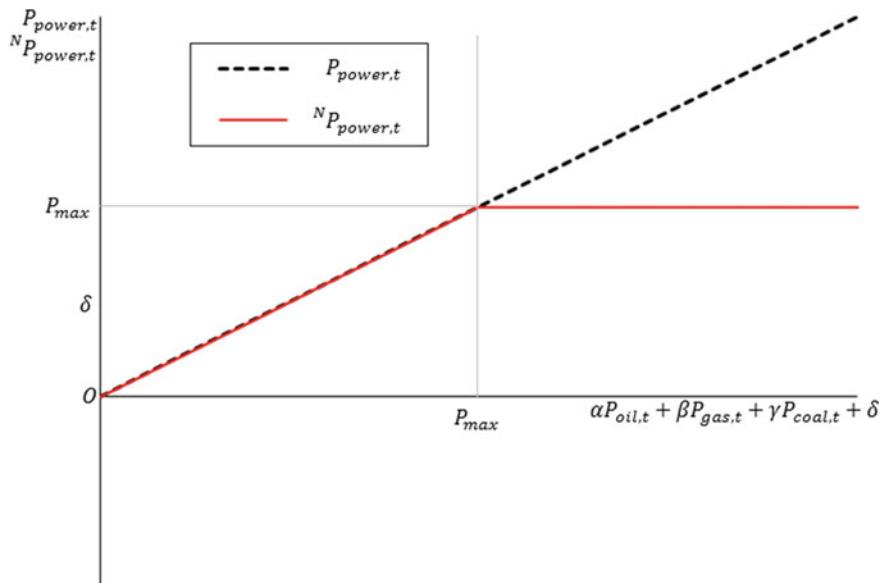
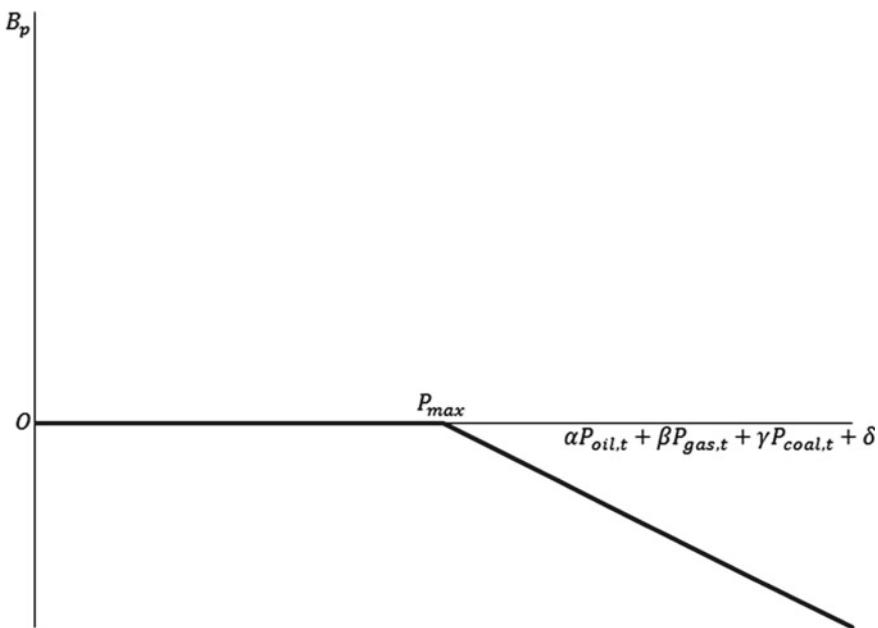
where B_p is the business balance in Eq. (6.6) assuming that the business balance in Eq. (6.5) is neutral.

Just as Asakawa [1] quantifies the gas distribution price with an upper threshold, we regard Eq. (6.7) as a short position in an electricity call option. This call option is not listed on any commodity exchanges. If we design a portfolio equivalent to Eq. (6.7), then we can pass the risk to the energy derivatives market. Although full hedging is impossible, we may gain partial coverage from holding a portfolio of oil, natural gas, and coal call options. Instead, we can clearly see why some rational management theoretically calculates the option value to book an allowance consistent with the risk amount. If each fuel price in Eq. (6.5) that defines the underlying asset price is efficient, then we can reasonably evaluate the option using historical data.

We can now easily estimate the call option price using Black and Scholes's [4] model. The call option price series is

$$C_t = P_{power,t} \Phi(d_1) - P_{max} e^{-r(T_i-t)} \Phi(d_2), \quad (6.8)$$

where Φ is the standard normal distribution function, T_i is the time to sell power, r is the risk-free rate, and

**Fig. 6.3** Power price formulas**Fig. 6.4** Impact of power sales with an upper threshold price on the business balance

$$d_1 = \frac{\ln(P_{power,t}/P_{max}) + (r + \sigma^2/2)(T_i - t)}{\sigma\sqrt{T_i - t}}$$

$$d_2 = d_1 - \sigma\sqrt{T_i - t}, \quad (6.9)$$

where σ is the volatility of $P_{power,t}$. From Eq. (6.5), we can calculate σ as

$$\begin{aligned} \sigma^2 &= \text{Var}(P_{power,t}) \\ &= \text{Var}(\alpha P_{oil,t} + \beta P_{gas,t} + \gamma P_{coal,t} + \delta) \\ &= \alpha^2 \text{Var}(P_{oil,t}) + \beta^2 \text{Var}(P_{gas,t}) + \gamma^2 \text{Var}(P_{coal,t}) \\ &\quad + 2\alpha\beta\text{Cov}(P_{oil,t}, P_{gas,t}) + 2\beta\gamma\text{Cov}(P_{gas,t}, P_{coal,t}) \\ &\quad + 2\gamma\alpha\text{Cov}(P_{coal,t}, P_{oil,t}). \end{aligned} \quad (6.10)$$

Suppose that an M -month power sale contract, whose price is determined by Eq. (6.6), concludes at $t = 0$. In addition, the electric energy of W_i is expected to be sold in month i . In this case, we can calculate the swap value of risk at the time of the contract using the following formula:

$$\sum_{i=1}^M W_i \left(P_{power,0} \Phi(d_1) - P_{max} e^{-r \frac{i}{12}} \Phi(d_2) \right), \quad (6.11)$$

where

$$\begin{aligned} P_{power,0} &= \alpha P_{oil,0} + \beta P_{gas,0} + \gamma P_{coal,0} + \delta \\ d_1 &= \frac{\log(P_{power,0}/P_{max}) + (r + \sigma^2/2)\frac{i}{12}}{\sigma\sqrt{\frac{i}{12}}} \\ d_2 &= d_1 - \sigma\sqrt{\frac{i}{12}}. \end{aligned} \quad (6.12)$$

However, it is difficult to say that this risk value is appropriate because it assumes that the underlying asset price follows geometric Brownian motion and that the variance and covariance are constant. However, in practice, it is rational from the perspective of calculation cost to adjust the numerical value calculated using Eq. (6.11) based on the deep insight of professionals. If higher accuracy is required, one method is as follows. First, we estimate a stochastic process model for the underlying asset price series from historical data. Second, we generate the option cash flow distribution by Monte Carlo simulation using the estimated stochastic process model. Finally, we obtain the option value by calculating the expected value.

A multivariate derivative is a derivative whose future payoff depends on multiple underlying assets. The basket option, whose underlying asset is a portfolio of multiple assets, is a typical multivariate derivative asset. In Eq. (6.7), because the multivariate derivative price depends on the correlation between the underlying asset prices and

their probability distributions, the derivative price is generally estimated using an approximate expression or a Monte Carlo simulation. While several studies examine multivariate derivative assets, none examine power prices. Björk [3] and Dhaene et al. [5] extend the univariate Black and Scholes [4] model to a multivariate case using correlated Brownian motions. However, this model assumes a lognormal distribution for each underlying asset price and a Gaussian dependence structure, which is unrealistic. Luciano and Schoutens [10] discuss a Lévy multivariate model for assets that incorporates jumps, skewness, kurtosis, and stochastic volatility. Their proposed model has the strengths of the univariate variance-gamma process and introduces a non-Gaussian dependence structure. Linders and Schoutens [8] use a one-factor Lévy model to value basket options by replacing the distribution of the portfolios with a reasonable approximation. Linders and Stassen [9] propose a methodology for pricing basket options using a multivariate variance-gamma model. By modeling the underlying assets using time-changed geometric Brownian motions with a common gamma subordinator, they express the basket option price as a linear combination of the Black and Scholes models. Bayer et al. [2] consider basket option pricing in a multivariate Black–Scholes and Variance-Gamma model, and propose using a Monte Carlo simulation designed specifically to solve numerical integration problems with non-smooth integrands. Hanbali and Linders [6] adopt the standard least-squares Monte Carlo approach to generate American basket option prices and discuss the method in terms of calculation time. We expect the readers to advance the study on multivariate derivative assets in energy markets by referring to many past studies cultivated in traditional financial markets. Furthermore, we hope that they contribute to literature for pricing the electricity call option as well as for designing the optimal portfolio consisting of fuel call options, which are discussed in this chapter.

References

1. Asakawa, H. (2010). Real option analysis of price adjustment system for gas distribution. *Journal of Real Options and Strategy*, 3(1), 63–76.
2. Bayer, C., Siebenmorgen, M., & Tempone, R. (2018). Smoothing the payoff for efficient computation of basket option prices. *Quantitative Finance*, 18(3), 491–505.
3. Björk, T. (1998). *Arbitrage theory in continuous time*. Oxford University Press.
4. Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3), 637–654.
5. Dhaene, J., Kukush, A., & Linders, D. (2013). The multivariate Black & Scholes market: Conditions for completeness and no-arbitrage. *Theory of Probability and Mathematical Statistics*, 88, 1–14.
6. Hanbali, H., & Linders, D. (2019). American-type basket option pricing: A simple two-dimensional partial differential equation. *Quantitative Finance*, 19(10), 1689–1704.
7. Kawamoto, K., & Tsuzaki, K. (2008). Market valuation of LNG price formulas. *Journal of Japan Society of Energy and Resources*, 29(2), 1–7.
8. Linders, D., & Schoutens, W. (2016). Basket option pricing and implied correlation in a one-factor Lévy model. In *Innovations in derivatives markets* (pp. 335–367). Springer-Open
9. Linders, D., & Stassen, B. (2016). The multivariate variance gamma model: Basket option pricing and calibration. *Quantitative Finance*, 16(4), 555–572.

10. Luciano, E., & Schoutens, W. (2006). A multivariate jump-driven financial asset model. *Quantitative Finance*, 6(5), 385–402.
11. Ng, F., Björnsson, H. C., & Chiu, S. S. (2004). Valuing a price cap contract for material procurement as a real option. *Construction Management and Economics*, 22(2), 141–150.

Index

A

- Akaike Information Criterion (AIC), 16, 39, 43, 44, 71, 72, 79–81, 119, 121
Arbitrage, 2, 6, 21–23, 26, 28, 47, 48, 53, 85
Archimedean copula, 33
Asymmetric term, 62, 66, 76, 110
Augmented Dickey-Fuller (ADF), 2, 14, 16, 17, 48, 64, 67, 68, 73, 74, 76, 77, 107, 110, 112, 114
Autoregression term (AR term), 62, 66, 76
Autoregressive Conditional Heteroskedasticity term (ARCH term), 62, 66, 76, 91, 110

B

- Baruník and Křehlík, 2, 54, 58, 60, 82, 107
Basket option, 128, 129
BEKK, 92, 93, 98, 103
Black and Scholes, 126, 129
Brent, 54, 55, 63–72, 82, 107–110, 112, 114–120

C

- Call option, 126, 129
Carbon credit, 3, 106, 107, 114, 119, 120
Clayton copula, 35, 40, 43, 46
Coal, 1, 3, 5, 106, 107, 119, 120, 126
Cointegrating vector, 17–20, 24, 48, 49
Cointegration (Cointegrating, Cointegrated), 6, 16–19, 21–23, 48, 53, 54
Cointegration test, 2, 16, 18, 19
Commodity future, 85, 121

- © The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022
T. Nakajima and S. Hamori, *Energy Trading and Risk Management*,
Kobe University Monograph Series in Social Science Research,
<https://doi.org/10.1007/978-981-19-5603-4> 131

Conditional covariance, 86, 89, 91, 95, 99, 102

Conditional variance, 62, 86, 89–91

Connectedness, 2, 53, 54, 57–62, 65, 66, 69, 73, 76, 78, 79, 81, 82, 106, 107, 114, 115, 117, 118, 120

Constant Conditional Correlation (CCC), 3, 87, 93–96, 98, 100, 102, 103

Contango, 8, 87

Convenience yield, 9

Copula, 2, 6, 31–48, 53, 54, 58, 70–72, 79–81, 86, 105–107, 114, 119–121

Crude oil, 1–3, 41, 53–56, 58, 62–72, 80, 82, 85, 86, 106, 107, 119, 121, 123, 124, 126

Crude oil portfolio, 71, 72, 79, 80, 82, 86

D

Descriptive statistics, 2, 7, 47, 54, 55, 57, 63, 66, 67, 72, 73, 77, 87, 88, 94, 98, 99, 106–110, 112

Diagonal BEKK, 3, 86, 87, 92, 94–100, 102, 103

Diagonal VECM, 3, 86, 91–101, 103

Diebold and Yilmaz, 2, 54, 57–60, 63, 67, 73, 76, 82, 107, 110, 117, 120

Discrete Fourier Transform (DFT), 60

Dubai-Oman, 54, 55, 63–72, 82

Dynamic correlation, 2

Dynamic ordinary least squares, 2

E

- Electricity, 1–3, 5–7, 12, 13, 21, 24, 41, 44, 47, 82, 85, 105–107, 114, 119, 120, 124, 126, 129
 Electric power, 5, 11, 24, 44
 Electric power business, 49
 Electric power company, 106, 123, 126
 Elliptical copula, 33
 EU allowance (EUA), 107, 108, 110, 114, 117, 120
 Expected shortfall, 2, 30, 31, 40, 42, 43, 48, 58, 62, 70, 71, 79, 80, 82, 107, 119–121
 Explosive process, 13, 16
 Exponential generalized autoregressive conditional heteroscedasticity, Exponential, GARCH, EGARCH, 2, 54, 107

F

- Frank copula, 33, 38, 41, 43, 47, 48
 French baseload (FrenchBL), 107
 Fuel portfolio, 2

G

- Gaussian copula, 33, 40, 45
 Generalized Autoregressive Conditional Heteroscedasticity term (GARCH term), 3, 62, 86
 Generalized Error Distribution (GED), 62, 66, 76, 110
 Gumbel copula, 36, 37, 40, 41, 43, 46, 70, 71, 79, 80, 107, 119, 121

H

- Hedge Effectiveness (HE), 89, 90
 Hedge, Hedging strategy, 90, 98, 102–104
 Henry Hub (HH), 7–9, 11, 12, 17–19, 21–25, 41, 43, 47, 48, 54, 56, 72, 73, 76, 78, 82, 86–88, 94, 95, 103

I

- Independently and identically distributed random number following the standard normal distribution, *iid* $N(0, 1)$, 90

J

- Japan Korea Marker (JKM), 54, 56, 72, 73, 76, 78–80, 82

- Jarque-Bera, 11, 55, 56, 63, 67, 73, 76, 87, 94, 99, 108–110
 Johansen, 2, 17, 19, 48
 Joint distribution, 6, 31, 39, 41

K

- Kurtosis, 10, 11, 47, 48, 55, 56, 63, 67, 73, 76, 87, 94, 99, 108–110, 112, 129

L

- Liquefied Natural Gas (LNG), 54, 79, 123–125
 Long-term equilibrium, 2, 16, 17, 20–26, 48

M

- Marginal distribution, 31, 32, 40, 42, 43, 48, 70, 79, 119
 Market integration, 53, 54, 65, 114, 117, 120
 Market risk, 3, 58, 105, 106, 114, 119–121, 123
 Mean, 8–11, 25, 30–32, 41, 47, 55, 56, 62, 63, 66, 70, 72, 76, 79, 85, 87, 90, 94, 98, 99, 107–110, 119
 Mean absolute deviation, 9
 Median, 8, 9, 47, 55, 56, 63, 66, 72, 76, 87, 94, 98, 107, 108, 110
 Multivariate derivative, 128, 129
 Multivariate generalized autoregressive conditional heteroscedasticity (Multivariate GARCH), 3, 86

N

- National Balancing Point (NBP), 86–89, 94, 98–103
 Natural gas, 1, 2, 5–7, 9, 12, 13, 21, 22, 24, 41, 47, 53–57, 71, 73–78, 82, 86, 87, 106, 107, 114, 119, 120, 126
 Natural gas portfolio, 58, 62, 71, 79–82
 Nonstationary process, 13, 18
 Normal distribution, 11, 14, 30, 31, 33, 40, 48, 55, 56, 87, 99, 108–110, 126

O

- Optimal Hedge Ratio (OHR), 2, 86, 87, 89, 90, 94–96, 98–100, 102, 103
 Optimal portfolio, 86, 102, 129

P

- Percentile, 9, 31
 PJM, 7–12, 16–19, 21–25, 41, 43, 47, 48
 Portfolio, 2, 3, 6, 29–31, 40–43, 48, 53–55,
 57, 59, 65, 69–71, 73, 81, 82, 85, 86,
 89–91, 98, 102, 103, 105–107, 119,
 126, 128
 Power, 1, 2, 5–7, 9, 11, 16, 21, 22, 24, 25,
 44, 48, 49, 53, 70, 82, 105, 106, 119,
 123, 124, 126–129
 Power generation business, 3, 5, 6, 24, 105,
 106, 114, 119, 120
 Put option, 124

R

- Range, 2, 9, 17, 42, 47, 55, 56, 63, 87, 94,
 99, 108, 110
 Return, 2, 6, 20, 30, 41–43, 48, 53, 54, 58,
 61–66, 68, 70–76, 78, 79, 82, 85, 86,
 89, 90, 94, 95, 98, 99, 102, 104,
 106–110, 114, 115, 117, 119, 120
 Risk management, 43, 106, 117
 Risk measurement, 9, 41, 43, 44, 47, 72, 81,
 82, 119, 121
 Risk, Portfolio risk, 29, 30, 48, 53, 82
 Risk transfer, 123
 Rotterdam, 107, 108, 110, 112, 114, 117,
 120, 121

S

- Schwarz Bayesian Information Criterion
 (SBIC), 16, 20, 59, 62, 66, 68, 73,
 76, 110, 114, 115
 Skewness, 9, 10, 47, 55–57, 63, 67, 73, 76,
 87, 94, 99, 108–110, 112, 129
 Spectral analysis, 54, 58, 65, 73, 78, 114,
 116, 118
 Spectral decomposition, 2, 60, 62, 69, 73,
 79, 82, 114, 117
 Spillover, 2, 53, 54, 57, 59–66, 68, 70, 71,
 73, 75, 78, 79, 81, 82, 105–107, 110,
 114, 115, 117, 118, 120, 121
 Spot-future arbitrage, 7, 26–28, 48
 Spread, 6, 9, 21–24, 41, 48, 76, 124, 126

Spurious regression, 13–16

- Standard deviation, 9, 10, 32, 41, 47, 55,
 56, 59, 63, 66, 67, 70, 72, 77, 79, 87,
 94, 99, 108–110, 112, 119
 Stationarity, 12, 73, 107
 Stationary process, 12–14, 16–18, 21, 58,
 64, 67
 Statistical arbitrage, 2, 6, 7, 21, 23–29, 41,
 43, 47–49

T

- T* copula, 33–35, 40, 42, 43, 45, 70, 71, 80,
 119, 120
 Title Transfer Facility (TTF), 54, 56,
 72–74, 76, 78–80, 82, 107, 109, 110,
 114, 117, 120
 Trading strategy, 6, 21, 24, 26

U

- Unit root, 14, 16–18, 48, 67, 112
 Unit root process, 13, 14, 16, 17, 22
 Unit root test, 2, 12, 14, 16, 17, 20, 107,
 110, 112, 114

V

- Value-at-Risk (VaR), 2, 29, 44, 58, 81, 105,
 121
 Variance-covariance matrix, 64, 65, 68, 69,
 73, 74, 76, 114, 117
 VECH, 91, 92
 Vector Autoregressive (VAR), 2, 110
 Vector Moving Average (VMA), 2, 54, 59,
 63, 64, 76, 107, 110, 112
 Volatility, 2, 3, 48, 53, 54, 58, 61, 62,
 66–69, 71, 73, 76–78, 82, 102,
 105–107, 110, 112–115, 117, 118,
 120, 124, 128, 129

W

- West Texas Intermediate (WTI), 54, 55,
 63–68, 70, 71, 82