STRATEGY, VALUE AND RISK

INDUSTRY DYNAMICS AND ADVANCED FINANCIAL MANAGEMENT

JAMIE ROGERS

FOURTH EDITION

GLOBAL FINANCIAL MARKETS

Global Financial Markets

Global Financial Markets is a series of practical guides to the latest financial market tools, techniques and strategies. Written for practitioners across a range of disciplines it provides comprehensive but practical coverage of key topics in finance covering strategy, markets, financial products, tools and techniques and their implementation. This series will appeal to a broad readership, from new entrants to experienced practitioners across the financial services industry, including areas such as institutional investment; financial derivatives; investment strategy; private banking; risk management; corporate finance and M&A, financial accounting and governance, and many more.

More information about this series at
http://www.palgrave.com/gp/series/15011

Jamie Rogers

Strategy, Value and Risk

Industry Dynamics and Advanced Financial Management

Fourth edition

palgrave
macmillan

Jamie Rogers
New York, NY, USA

Global Financial Markets
ISBN 978-3-030-21977-2 ISBN 978-3-030-21978-9 (eBook)
https://doi.org/10.1007/978-3-030-21978-9

This Palgrave Macmillan imprint is published by the registered company Springer Nature Switzerland AG.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The objective of the fourth edition of Strategy, Value and Risk (SVR4e) is to provide a framework that integrates advanced financial management with industry dynamics. Those factors in the business environment that have a significant impact on firm strategy, business models, investments and value are identified, and the accounting, finance, economic and quantitative principles that provide a foundation for the analysis of these issues are addressed.

Part I (Strategy, Value and Risk) provides the strategy, economic, accounting and finance framework. Strategy as value is the discovery and development of sources of profitability that maximize value, while delivering on operations in the short term and investing to maintain long-term continuity. A firm's external environment presents the value opportunities, while its resources and capabilities leverage those opportunities.

The external environment is defined within a framework that includes technology and innovation, industry dynamics, globalization and climate change. Technologies transform labour, capital, materials and information into products and services of greater value. Innovation refers to a change in one of these technologies. Joseph Schumpeter developed theories on technology waves within the context of economic growth, while neo-Schumpeterians moved the emphasis to the technological revolutions themselves and the resulting transformative effects on the economy.

'Industry dynamics' discusses how successive technology waves have influenced industries and the factors that drove change. Globalization is framed within the context of market concentration that is the result of a succession of mergers globally and in the US, followed by the enormous scale reached by today's dominant technology firms. Climate is discussed in relation to industrialization and its residual effects, and impact on industry in the future.

'Transformation and the firm' covers innovation and the firm, industry boundaries, strategy and the environment and future value. The firm has played a consistent role in innovation, with invention and research developing into processes, research and development becoming institutionalized, and in today's open innovation systems. The late nineteenth century saw the rise across many industries of large vertically integrated corporations, along with a series of innovations in systematic management. A century later, new technologies led to a transition from vertical to horizontal boundaries within the computer industry, with dominant vertically integrated firms virtually disappearing, and in their place the ascent of open platforms and the emergence of ecosystems.

The development of strategy over the last 60 years has largely been driven more by business pragmatism than theory. Themes have included diversification, vertical integration, scale, portfolio planning matrices, competitive advantage, shareholder value and resources and capabilities. Today, information technology is increasingly having an influence on strategy, with factors such as network effects, intellectual property, platforms, and revenue and cost structures that are unique to technology industries.

'Future value' discusses themes that are central to industries, strategy, new business models, management systems and organization forms. Open platforms create business ecosystems that transform industries, and include technologies such as cloud computing, data, analytics and software services. Digitization will drive a fundamental realignment of industry boundaries, with firms that have a horizontal dominance in technologies leveraging those resources across other vertical sectors. The energy industry is in the process of a fundamental transformation, with global oil demand forecast to potentially peak as early as 2025. Accelerating the peak is an array of competitive alternatives to fossil fuels that include solar, wind power, batteries and electric vehicles. Biopharmaceuticals have the potential to become the foundation of the pharmaceutical industry. Replicating large molecules on an industrial scale itself requires new capabilities in manufacturing, while DNA sequencing platforms, or biofoundries, have the potential to become a new industry.

The value section frames a firm's short-term and long-term financial objectives. In the short term, the focus is assessing current performance. A firm's financial statements reflect its business model and where and how firm value is derived. Free cash flows are defined within the context of stocks and flows, and ROE (return on equity) and ROIC (return on invested capital) measures are discussed. There is a discussion on intangible assets and their increasing relevance in the digital economy, and a pro forma analysis to illustrate the concepts. In the long term, capital investments and divestments ensure the continued generation of value and firm continuity. Discounted cash flow (DCF) techniques and the assumptions are discussed, and the advantages of real options techniques are then outlined. Corporate finance covers the theory and builds a framework that extends on the initial model with the inclusion of growth options, an abandonment option, modularity and financial options. Finally, a firm's capital structure is discussed within this framework.

The risk section covers theories of investment risk and discount rates, corporate risk management, the risk drivers, and value and risk. Investors' risk appetite is reflected in investment analysis through the discount rate, and there is a discussion on the Capital Asset Pricing Model and Risk-Adjusted Discount approach. Diversifying a firm over a portfolio of independent businesses can decrease the variance of the combined cash flows, thereby reducing risk. Shareholders can, however, reduce risk by maintaining diversified portfolios themselves. Corporate financial risk management encompasses the use of derivatives to manage risk exposures, which can include financial, commodity, legal, operational, strategy, technology, product and political risk. Finally, value is created through the management of a firm's strategic portfolio and its real options portfolio. While a firm's real assets are a significant component of its risk profile, its real options also contribute to value. A firm can therefore enhance its capabilities by integrating real options analysis into its strategic management, corporate finance and risk management processes.

Part II (Quantitative Analytics) discusses financial statistics, derivatives and derivative applications. 'Financial Statistics' covers time series analysis, volatility, empirical distributions, the lognormal distribution and a discussion on volatility. 'Derivatives' discusses futures, forwards and options, the replicating portfolio and risk-neutral valuation, Brownian motion, Black–Scholes, and numerical techniques. 'Derivative applications' examines advanced energy derivatives, stochastic volatility, convertible bonds and compound options. The overall goal is to provide a framework that outlines the underlying financial economics for the analysis of physical and intangible assets, financial assets and contracts.

Part III (The Analysis of Investments, Transformation and Value) has five chapters that cover platforms, data and analytics, the transformation of the energy sector, pharmaceutical and biotech, growth firms and business model migration. These chapters discuss industry and technology trends, and illustrate valuation methods such as discounted cash flow, advanced accounting analysis, advanced corporate finance methods that include binomial trees and real options, and advanced derivatives that include spread options and compound options to illustrate the various valuation techniques.

Who Should Read This Book

SVR4e lays out scenarios that will likely influence and shape firms and industries in the coming decades. A central theme currently influencing economies, industries and firms across all sectors is the fundamental shift from an industrial to a digital economy, which has had a major influence on strategy, business models, and corporate investment and divestment.

The business environment in many ways today is similar to that at the turn of the twentieth century, with the dominance of large vertical firms and their associated new technologies and innovations, industry concentration and impact on industry boundaries. As management accounting and other technologies laid the foundation for the management of the large vertical firms at the turn of the twentieth century, data and analytics today is the new technology for systematic management. Climate is also a fundamental driver of industry transformation, and will shape strategy and investment in the coming decades. Within this context, valuation is one of the most relevant topics in finance, with the emergence of the digital economy, new industries, business models and technologies.

SVR4e has relevance to CFOs and private equity, corporate finance and investment professionals in general. The role of the CFO is expanding from a focus on external reporting and fiduciary duties to collaborating with other key business partners in business development, operations, marketing and executive management. This requires an analysis of the business drivers of success and operational strategies that integrates finance with business and strategy. Concepts such as business model disruption, data and analytics, intangible assets and dynamic analysis are key issues that need to be addressed within the CFO role. CFOs can enhance corporate value by promoting a well-designed set of financial analytics to secure assets and maintain value.

Investment professionals are required to see the larger economic setting and environment in which firms compete, assess a firm's industry and its position within that industry, understand which investments best serve its broad strategic goals and identify a firm's capabilities and options. In addition, analysts must translate these broad insights and communicate them to management and investors. SVR4e lays out accounting, finance, economic and quantitative concepts that are relevant to the current transformation in the economy, the driver of these themes, and where valuation is essential.

The content covers disciplines as diverse as accounting, finance, economics, econometrics, quantitative finance, statistics, management science and data science. These subjects have their own bodies of knowledge, and there are limitations on how much of the material can be covered in one text. In all these disciplines, however, much of the value of an analysis is in the process, as opposed to just producing an end number.

New York, NY, USA Jamie Rogers

Acknowledgements

The fourth edition of this book represents an accumulation of experience and feedback on the first, second and third editions from friends, colleagues, practitioners and academics. I would like to thank Carliss Baldwin, Ian Clark, Ben Garrard, David Kellogg, Elayne Ko, Karl Mattingly, Nick O'Brien and Stephen Penman. I wish to thank and acknowledge Maria Barbera and Bill Birkett, who are no longer with us for contributions on the first edition. I also would like to thank Les Clewlow and Chris Strickland for content in Chaps. 5, 6 and 8. I thank Tula Weis at Palgrave Macmillan for her commitment to the fourth edition and Jacqueline Young for her assistance. I would also like to thank my wife Meg for her continued patience, guidance and support.

Every effort has been made to trace all the copyright holders, but if any have been inadvertently overlooked, the publishers will be pleased to make the necessary arrangements at the first opportunity.

Jamie Rogers

Contents

List of Figures

List of Tables

Part I

Strategy, Value and Risk

1

Strategy

1.1 Strategy and Value

The objective of all firms is the creation of value. A firm's strategies describe how it intends to create value over an immediate time frame, and the value opportunities it is searching for over the long term. Value, however, has different meanings to different stakeholders. Firms manage resources to create value through their capabilities to deliver products or services that provide customer value, maintain relationships with resource providers and customers, and organize activities through governance, management systems and processes. To do this, a firm has to create an equitable balance between stakeholders such as management, customers, employees, financiers, unions, suppliers, shareholders, government and society in general.

Although value has to be established and maintained by firms in order to offer incentives to various stakeholders, a premium is associated with value creation for customers and shareholders. The creation of value for other stakeholders is dependent on the success of creating value for customers and the incentives offered to shareholders—the residual stakeholders. Offering unsuitable incentives to stakeholders, however, is likely to lead to the destruction of shareholder value.

Strategy as a search for value is the discovery and development of sources of profitability to maximize firm value. To achieve sustainable shareholder value, firms have to simultaneously manage and deliver on operations in the short term, while investing and divesting to maintain long-term continuity. The issue then becomes the durability of a firm's competitive advantage to maintain a rate of return on the firm's assets greater than its cost of capital. A firm's

J. Rogers, *Strategy, Value and Risk*, Global Financial Markets,
https://doi.org/10.1007/978-3-030-21978-9_1

external environment presents the value opportunities, while a firm's resources and capabilities determine how to leverage these opportunities.

Analysing how innovation and technology have influenced industries, firms, strategy, business models and investments provides a foundation for identifying what lies ahead. The themes that dominate the external environment today are new technologies and their diffusion, globalization, industry dynamics and climate. The question then becomes how these factors will influence industry transformation, innovation and the firm, strategy and future value.

1.2 The External Environment

1.2.1 Technology and Innovation

Technology is defined as the processes by which an organization transforms labour, capital, materials and information into products and services of greater value. All firms have technologies. Innovation refers to a change in one of these technologies. So spectacular was the wave of innovation in the late nineteenth century that the Commissioner of the United States Office of Patents recommended in 1899 that the office be abolished with the words 'Everything that can be invented has been invented'.

Since the First Industrial Revolution, economic growth has been driven by science and research funded by financial speculation, with financial bubbles being a persistent feature of this process. Periodically, the focus of the financial speculation is an innovation that fundamentally transforms the economy. This relationship has repeatedly created transformative infrastructures such as canals, railroads, electrification, cars, airplanes, computers and the internet, with the fundamental value to the economy realized decades later.

Business cycles are the recurring levels of economic activity over time, and were once described as having long predictable durations. Kondratiev waves are long macroeconomic cycles lasting 50–60 years that the Russian economist Nikolai Kondratiev theorized existed in capitalist economies. The economist Joseph Schumpeter extended Kondratiev's concept with his own theories on long technology waves. In Schumpeter's theory of business cycles and economic development, the circular flow of income, an economic model depicting the circulation of income between consumers and producers, is stationary. Entrepreneurs disturb this equilibrium through innovations, and in doing so, create the economic development that drives the economic cycle.

Schumpeter formed two theories in regards to entrepreneurship. The first (1909) was that individuals and small firms were more innovative, which he expanded in a second theory (1942) in which large firms drive innovation by investing in research and development (R&D) through their access to capital and resources. Schumpeter's 'gale of creative destruction' is the fundamental driver of new industries or industry combinations, the result of entrepreneurs producing innovative new products, processes or business models across markets and industries that either partially or entirely displace previous innovations. Entrepreneurship can, therefore, be framed as recognizing and exploiting value opportunities, and applies to individuals, small firms or large institutions.

Schumpeter also recognized and analysed the dynamic interaction between competition and industry structures. Schumpeter focused on innovation as the central component of competition and the driving force behind industry evolution. In Schumpeter's view, each long wave of economic activity is unique, driven by entirely different clusters of industry. Each upswing stimulates investment and an expansion of the economy, resulting in an economic boom. Each long boom eventually declines as the technologies mature and investors' returns decline, only to be followed by a new wave of innovations that replace the old methods and create the conditions for a new upswing.

The long wave theories of Kondratiev and Schumpeter both focused on economic growth. While Kondratiev does not identify a specific causal factor and Schumpeter tied these waves to technological revolutions, both were attempting to describe long-term deviations in GDP and the economy in general.

Neo-Schumpeterians moved the emphasis to the technological revolutions themselves, the diffusion processes that result with each wave and the resulting transformative effects on the economy. Technological revolutions are viewed as creating clusters, following Schumpeter's long wave theory, where interconnected innovative new products, processes and infrastructure initially lead to new fundamental industries, which are then followed by their diffusion to incumbent industries.

The economic historian Carlota Perez identified a regular pattern of technological revolutions over the past 250 years that materialized every 50–60 years. These cycles have discrete phases, where the emergence of general-purpose technologies signals massive changes in the economy. These general-purpose technologies lay the foundations for generating clusters of products, processes and innovations, initially with the rise of new fundamental industries, followed by the diffusion of the technologies to more mature industries.

Table 1.1 illustrates the waves of general-purpose technologies that laid the foundations for successive technology revolutions, starting with the Industrial

Table 1.1 The general-purpose technology revolutions

1st wave: 1770s	2nd wave: 1830s	3rd wave: 1870s	4th wave: 1900s	5th wave: 1970s
Water power	Steam power	Electricity	Oil	Computers
Iron	Railways	Steel	Automobiles	Telecommunications

Revolution in the late eighteenth century, the Second Industrial or Technological Revolution, the Third and Fourth Technological Revolutions, and finally, the Information Revolution. The waves refer to a starting period instead of a specific year.

The Industrial Revolution, with its origins in the UK during the 1770s, began with factory automation transforming the English economy, followed by infrastructure such as roads, ports and, in the 1790s, the emergence of canals, all of which stimulated the flow of trade. The Second Industrial Revolution that began in the 1830s in the UK saw the emergence of steam, iron and railway technologies. The Third Industrial Revolution, with its foundation in the mid-1870s, facilitated the first globalization. Steel, heavy engineering and electrification technologies emerged, which led to transcontinental railways, fast steamships and the intercontinental telegraph, which facilitated the flow of information and trade.

In the early 1900s, mass production, oil and the automobile emerged in the US with the Fourth Industrial Revolution, leading to enormous investment until the 1929 financial crash and the 1930s Depression. The 1970s saw the emergence of information and telecommunications technologies that are driving the current technological revolution, which—following the technology cycle theory—is midway through its evolution.

During the course of the Information Revolution over the last quarter of the twentieth century, a number of fundamental changes emerged in the global economy. The burst of innovation in information technology reduced the cost of communications, which in turn facilitated the globalization of production and capital markets. This laid the groundwork for the use of innovative new business models. A new knowledge economy also emerged, driving the transformation from an industrial to a post-industrial economy centred around intangible assets and services. This led to the growing significance of intellectual assets relative to physical assets. Finally, there was a shift from closed to open innovation systems, an important trend that identifies and acquires new technologies from outside the firm. As a result of these factors, stock market indexes are now dominated by service industries such as information technology, financial services, pharmaceuticals, telecommunications and retail industries that leverage technology and science.

1.2.2 Industry Dynamics

Industry dynamics focus on how industries are currently organized, how they differ from earlier periods, what factors brought about the reorganization of the industry and how these factors changed over time.

The successive technology waves over the nineteenth and twentieth centuries are reflected in the transformations of the listed stock market industry sectors. Figures 1.1 and 1.2 illustrate the industry shifts in the US and UK indexes for the years 1899, 1950 and 2000. The industry sectors are weighted in the figures and based on the industry classifications in use in 1900, with a few additional sectors that, although minor in 1900, grew to significance through to 1950.

Railroad companies were the first true industrial giants at the beginning of the twentieth century. As they gradually became regulated and nationalized, however, the industry was marginalized by new industries such as steel, chemicals, rubber, mechanical engineering, machinery and consumer durables. Some sectors that were insignificant in 1900 grew to dominance through to 1950, and had declined by 2000. One example is the chemical industry, with US growth increasing from 0.5% to 13.9% in weighting between 1900 and 1950, and declining to 1.2% by 2000. The UK

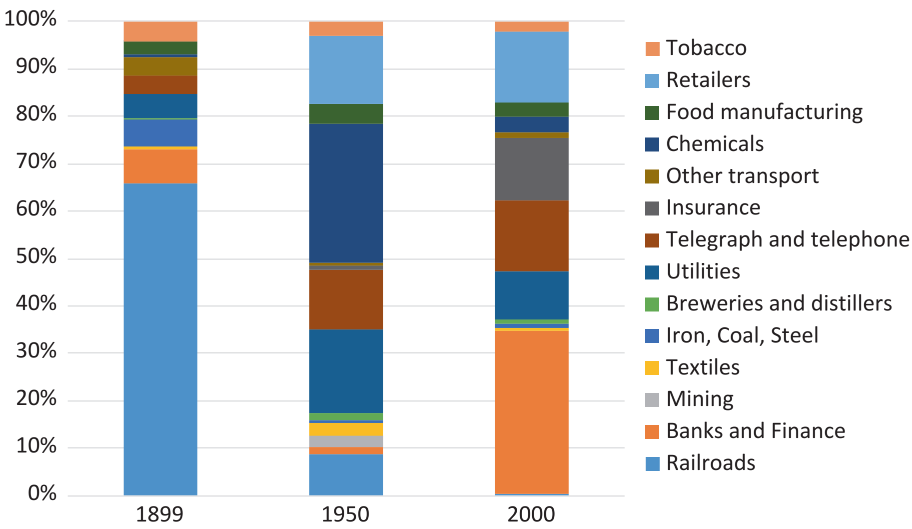

Fig. 1.1 US industry sector weights based on the 1899 classification system. (Source: Dimson et al. 2002). UK 1899 and 1950 sector weights are based on the largest 100 stocks. 2000 sector weights indicate the entire market. US 1899 weights include all NYSE and New York City banks. 1950 and 2000 weights reflect the entire US market

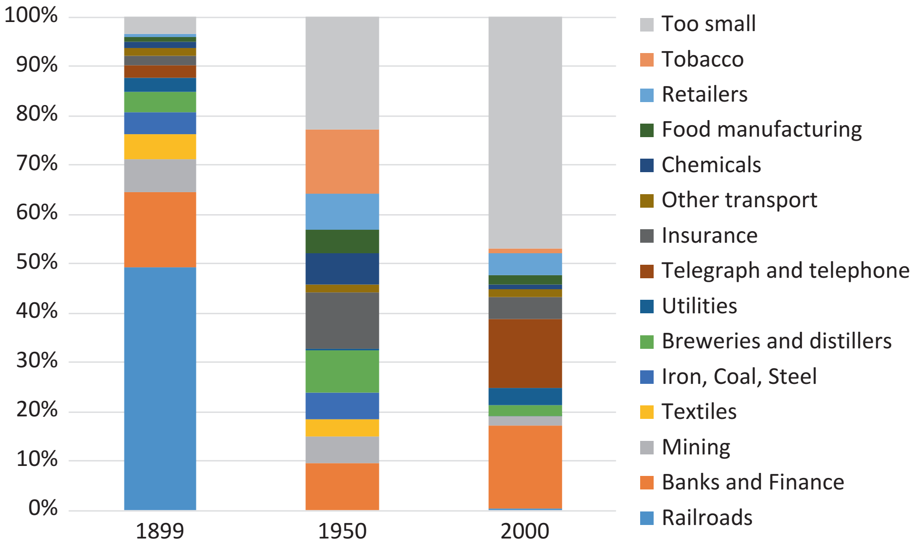

Fig. 1.2 UK industry sector weights based on the 1899 classification system. (Source: Dimson et al. 2002)

chemical sector followed a comparable pattern, with a huge weighting increase from 1900 to 1950, followed by a dramatic decline by 2000. The banks and finance sector in the US and the UK, however, declined and then grew in weightings from 1900 to 1950, and then to 2000—from 6.7% to 0.7% to 12.9%, and 15.4% to 9.7% to 16.8%, respectively. The telegraph and telephone sector saw moderate growth in the US from 3.9% to 6.0% to 5.6%, and significant growth in the UK from 2.5% to 0.0% to 14.0% in weightings over the twentieth century.

Figures 1.3 and 1.4 illustrate firms classified under the industry sector definitions as at 2000 and listed according to their US and UK significance. In 2000, the three largest US sectors were information technology, banks and finance, and pharmaceuticals, which combined, accounted for over a third of US firms. Sectors such as oil and gas and pharmaceuticals in the US and UK were nearly non-existent in 1900, and information technology had a zero weight in the years 1900 and 1950. In 1950, pharmaceuticals were still relatively insignificant, while oil companies in the US had reached dominance, followed by a decline in relative weighting. The telecommunications sector, while relatively small in 1900, grew to approximately 6% of the US market, where it has since remained. The UK telecommunications sector was nationalized up to the 1980s, when it was then privatized, ultimately reaching 14% of the UK 2000 weightings.

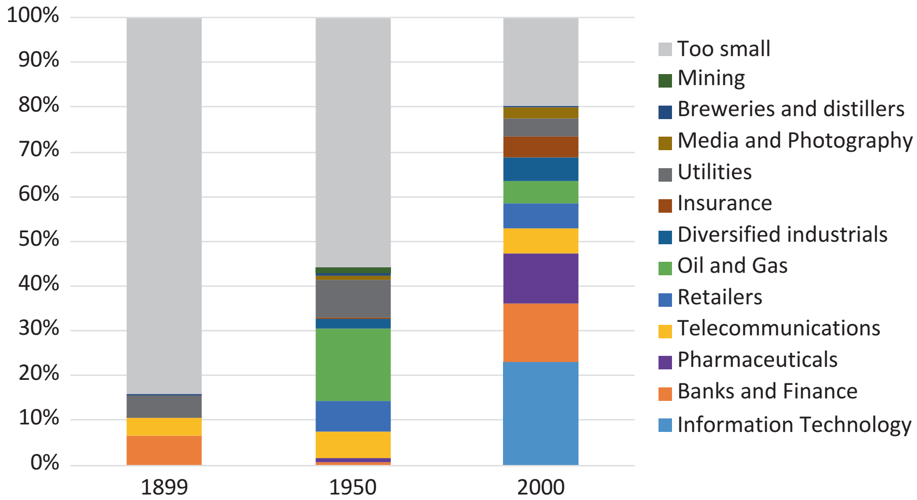

Fig. 1.3 US industry sector weights based on the 2000 classification system. (Source: Dimson et al. 2002)

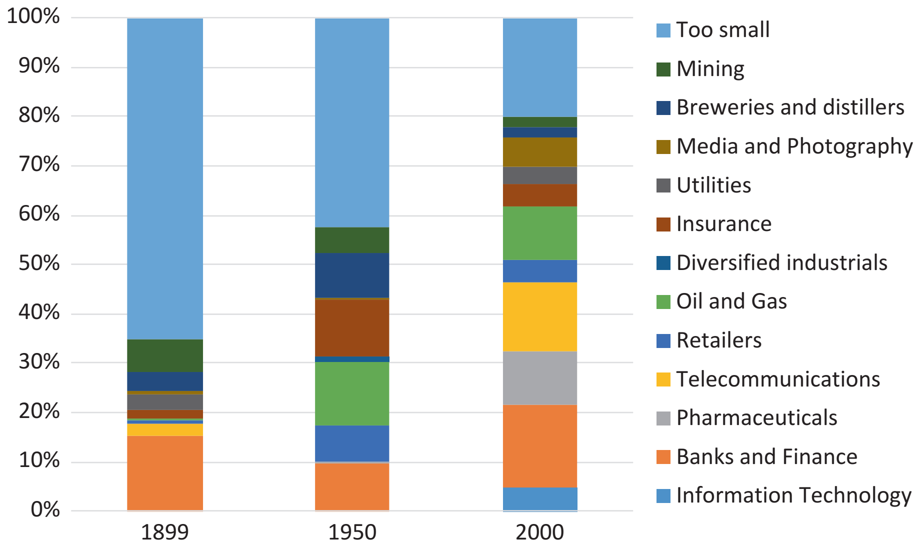

Fig. 1.4 UK industry sector weights based on the 2000 classification system. (Source: Dimson et al. 2002)

The composition of the stock market indexes has always been shifting. Over the first 75 years of the twentieth century, these shifts were gradual, with new industrials replacing older ones and manufacturing dominating. US industrials, however, began a relative decline after the mid-1950s, with the decline accelerating in the 1970s due to soaring energy costs and increased competition domestically and from overseas. On the supply side, US industrials were facing low-cost foreign competitors that were producing products that were increasingly improving in quality. On the demand side, growth in the US domestic market had ceased, with demand for industrial goods diminishing by the end of the 1960s.

The final quarter of the twentieth century saw a significant change in the US economy with the decline in manufacturing and growth in service industries, and a shift to information technology, a revolution that also significantly expanded the extent of services. Figure 1.5 illustrates the transformation of the 100 largest US firms as measured by revenues for the years 1955, 1975 and 2000. The 1955 and 1975 sectors show a relatively small decline in manufacturing sales and market value and the rise of financial services, information technology, and pharmaceuticals and healthcare. Both lists are fundamentally the same, with manufacturing firms dominating. By 2000, however, manufacturing had declined significantly, with financial services, information technology, and telecoms and media dramatically increasing, as measured by revenues.

There were a number of fundamental changes in the US economy over the last quarter of the twentieth century. The first was the appearance and growth in the components of a new knowledge economy. These included computer

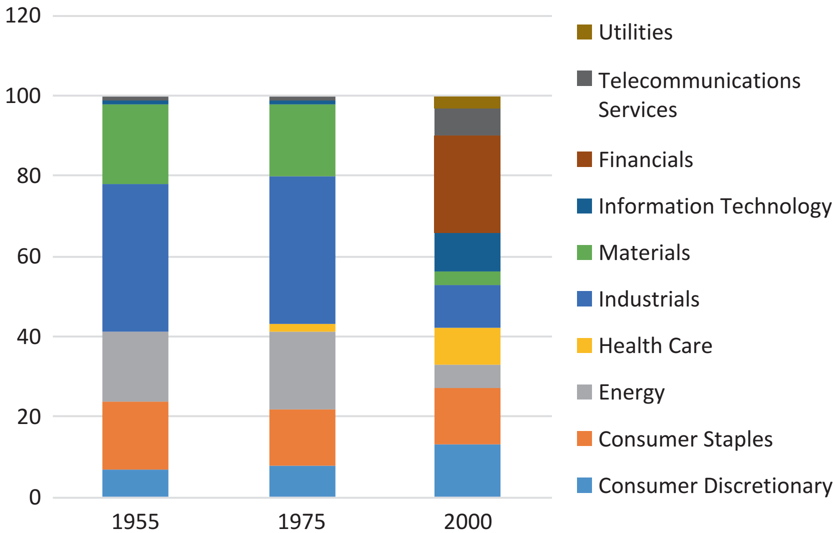

Fig. 1.5 Industry sector weights for the largest 100 firms by revenues—1955, 1975, 2000. (Source: Fortune 100–1955, 1975 and 2000)

hardware, software and services, providers of internet services and content, and the telecoms that develop and manage the infrastructure over which information flows. The combination of these components created an information revolution equivalent to the Industrial Revolution, which generated the expansion of the industrial economy. Analogies can be drawn between firms in the information economy and those from the industrial economy, with chips the equivalent of basic materials such as metals and plastics, software equal to process control and electrical mechanical systems, and telecommunications infrastructure and services the equivalent to transportation and related services. These information economy firms have grown significantly and now have considerable influence, as did their industrial equivalents, as information technology diffused and integrated into the economy.

The second is the growing significance of intellectual assets compared to physical assets. The foundations for value generation in industrial firms are physical assets such as plant and factories. In the knowledge economy, science and technology are increasingly driving business, as firms in all industries use these factors to create a competitive advantage. This is reflected in how firms manage R&D and commercialize innovation, with the shift to the formation of networks, partnerships, and other structures to exploit discoveries and mitigate R&D risks.

The third attribute was the manner in which new entrants were financed compared to their industrial counterparts. Risk capital was provided by venture capital, which financed entrepreneurs and laid the foundation for the rise of new information technology and biotechnology firms with a relatively small amount of initial capital. The fourth was the use of innovative new business models. The fifth was the emergence of a new economic focus. While economies of scale and scope, the primary drivers of industrial growth, continued to be significant, network economies rose to become of equal importance and included the drive for standards and alliances.

The result of all these factors was the shift of large industrial firms from the centre to the fringe of the modern global economy. The forces that created the new knowledge economy will continue to drive the transformation from an industrial economy, with its foundations in physical assets and production, to a post-industrial economy centred around intangible assets and services.

The shift in the economy has also influenced the turnover in the composition of the major stock market indexes. In 1917, Forbes published their first list of the 100 largest American firms ranked by assets, with steel, oil and gas, mining, food and telecoms the dominant sectors. In 1997, only 15 survived, and by 2017, only 2 firms—AT&T and GE—from the original 1917 Forbes list endured under their original trade names.

The same result was found for Standard & Poor's S&P 500. Changes in the S&P 90 index were relatively slow for the first 20 years after its initial start in the 1920s, with an average turnover in companies of 1.5% per year. A firm that was included in the index during this period would remain there for an average of more than 65 years. By 1998, the rate of change in the S&P index had shifted dramatically, with an increase to nearly 10%. The time a firm spent in the index over the last 70 plus years had decreased from an average of 65 years to 10 years. Over the last 40 years, only 74 of the original 500 firms that comprised the S&P 500 in 1957 still remained on the index in 1997.

Turnover in the stock market indexes is often cited in economic analysis as a metric for productivity and innovation. Annual turnover, the number of firms entering and exiting the S&P 500, on average was relatively normal in the late 1950s and decreased in the 1960s and 1970s. Turnover increased on average however to historically high levels from the early 1980s, and in the late 1990s, reached new peaks, with high M&A (mergers and acquisitions) volume reflected in the higher volatility. Turnover since 2000 has however declined to the levels seen towards the end of the 1950s.

While turnover can be represented as the volume of firms entering and exiting the S&P 500, there are a number of underlying factors as to why this single metric cannot be used as an indicator for innovation. The omission of service firms before the mid-1990s, the cycle of many firms re-entering and exiting, and the concentration of turnover at the low end of the S&P 500 all convolute turnover as a single aggregate number.

Turnover is also a reflection of economic dynamics, such as transformations in financial markets, mergers and acquisitions and initial public offerings that are a function of temporal change, new technologies, sector transformations and geopolitical considerations. Other factors include the unbundling of firms as a result of the Information Revolution, the growth in the information and telecommunications sectors in which innovation is especially important, service industries having less barriers to entry than manufacturing, and the persistence of firms that can be attributed to the advantage of scale and innovation.

1.2.3 Industry and Globalization

Information and communications technologies have been a fundamental driver of the integration of the global economy over the last 40 years. The 1980s saw conglomerates broken up, a wave of mergers and acquisitions, widespread industry restructuring and deregulation that resulted in industries such as telecoms being subjected to competition. In the 1990s, US firms focused on overseas expansion through a boom in mergers and acquisitions (M&A). Firms from high-income economies massively expanded their international operations, with industry concentration rising significantly across diverse industries as firms integrated global business systems through M&A.

Almost every industry sector saw the emergence of firms with leading technologies and brands that dominated their global market sector. This process cascaded into global supply chains, with a broad restructuring of value chains around these core firms, which further intensified the concentration of industries. These dynamics stimulated intense competition and unparalleled advances in technology as a result, and created firms that became the foundation of the global economy, with technologies and brands concentrated among a small cluster of firms.

Since 2008, the US has seen a further wave of mergers with a total value of $10 trillion, one of the largest in history. While the previous M&A wave focused on building global dominance, the post-2008 wave saw consolidation and increased market share across US industries. The result of these waves is that 10% of firms now produce 80% of global profits according to a McKinsey estimate, while 60% of revenues and nearly 65% of the total market capitalization are sourced to firms with greater than $1 billion in revenues.

In the US, profits as a ratio of GDP are larger than at any other time since 1929. From 1994 to 2013, the percentage of nominal GDP produced by the Fortune 500's largest US firms increased from 33% to 46%, while the Fortune 100 percentage of revenues increased from 57% to 63%. The number of US listed firms almost halved from 6797 to 3485 from 1997 to 2013, with a fifth of total US corporate profits now sourced abroad. Dominant firms have also built up massive cash reserves equal to 10% of US GDP, with domestic free cash flows of approximately $800 billion per annum.

Today's technology firms have reached enormous scale within a few decades, and dominate in terms of market share, revenues and the information economy's infrastructure. In the industrial economy, firms leveraged economies of scale to achieve size. Today's giant technology firms have adapted this economic law to the digital economy by focusing on demand and network effects over supply side and production efficiencies. Innovations in technologies have moved the industrial economy's cost focus to the leveraging of network effects in the digital economy.

These dominant technology firms also have built huge cash reserves, which are used to further consolidate market dominance through the acquisition of start-ups that have patent portfolios and new technologies while also removing potential competitors. Today's technology firms are also platforms that provide the capabilities to quickly reach scale. These platforms also provide data and software cloud computing services that offer cost advantages and flexibility, and reduce the lead time for start-ups to become cash flow positive. Technology firms can also leverage their platforms by moving into other industries and further scale across other sectors. Therefore, firms that have the balance sheets capable of making considerable investments in the physical and digital assets that are necessary in the digital economy will continue to dominate.

1.2.4 Industry and Climate Change

Industrialization has had a significant influence on economic growth over the last 250 years. A residual effect of the technologies and innovations that drove this growth has been the impact on climate. The same drivers of innovation also have the potential, however, to address the factors that contributed to climate change over the last 200 years.

The factory was an innovation that introduced scale, efficiency and standardization in production, and increased the output of many products. This increase in output intensified pressure on transportation infrastructure. Innovations in the steam engine produced the locomotive and steam-driven ship, which along with the expansion of railways, transformed the transportation of people and freight. The demand generated by railways resulted in the expansion of iron followed by steel production. This created a surge in coal mining to supply the fuel for iron smelting and the developments in steam engines. Nineteenth-century railways also dramatically increased immigration, integrated markets and spurred other industries. This further created an enormous increase in the use of coal, iron, steel and petroleum-based products.

Car manufacturing combined numerous innovations, including the internal combustion engine and new steel manufacturing methods. The mass production that followed in the early twentieth century led to lower prices and product access for the masses. Industry and agriculture also began using gasoline and diesel-powered transportation and machinery.

The advent of electricity saw the development of long-distance power transmission and the expansion of electricity utilities and grids. Electrical power became available for street lighting, residential use, public transport and industrial uses such as heat for the refining and manufacturing of copper and aluminium. Coal became the fuel source for electrical power with the development of coal-fired electrical generating plants towards the end of the nineteenth century.

While industrialization has produced tremendous economic benefits, it has also generated significant burdens. Accompanying the Industrial Revolution was a massive growth in energy consumption, largely through the burning of coal, a fossil fuel. Since the onset of the Industrial Revolution, societies have increased their use and dependence on fossil fuels such as oil, coal and gas, primarily to generate electricity, and power transportation and industrial processes. One of the great challenges of climate change today is that greenhouse gas emissions result from almost every major function in society, including electricity production, transportation, agriculture and industry.

Climate change generates systematic risks throughout the economy, and will have an influence on agriculture, energy, health, national income, regulation and reputations at the industry and firm levels. Sectors such as agriculture, fisheries, forestry, health care, insurance, real estate, tourism and the energy infrastructure will disproportionally feel the effects of climate change. The consequence is that climate change is changing the competitive environment, with particular sectors, industries and organizations more at risk than others.

As there are uncertainties in regards to how climate change will impact future states of the world, any number of risk factors will have an impact on a firm. These include exposures to financial, commodity, legal, operational, strategic, technology, product, political and reputational risks. Climate change could include some or all of these risk factors, depending on the nature of a firm's activities. Two specific risk categories can, however, be defined. Sector-specific risk, the risk exposure to firms within an industry sector, includes regulatory and physical risks—for example, severe weather directly affecting economic sectors such as insurance, agriculture, health care, real estate, water and tourism. Firm-specific risk includes competitive, litigation and reputational risks, where a firm's operations could result in repercussions from consumers, shareholders and stakeholders.

Climate-related events represent risks to all firms at some level. These can occur as events at regular frequencies, and regionally, such as disruptions to agricultural or energy production, supply chains or infrastructure. Most firms have strategies and processes to manage the regular changes in climate. Firms in the future, however, cannot depend on climate conditions being consistent with those over the last century. Climate trends are anticipated to undermine the notion of continuity, with deviations both in general conditions and the number and severity of extreme weather events.

1.3 Transformation and the Firm

1.3.1 Innovation and the Firm

Transformations in technology have been driven by momentum from needs and end users in some cases, and developers and system builders in others. Firms, however, have played a consistent role as participants, and while not always the initiator, have been leading players in innovation as invention and research developed into processes in the nineteenth century. In the early twentieth century, many firms had internalized innovation and focused on efficiency and rationalization as a means to secure their technologies. Other firms had leveraged innovation to pursue new products or processes, which became an important development that, while riskier, was potentially more rewarding.

Over the course of the twentieth century, innovation and research and development (R&D) were institutionalized, which influenced both the trends and speed of change in technology throughout the industrial and industrializing worlds. A principal driver of new technologies over the twentieth century was the exploitation of science by US industry, which is reflected in the shifts in the research environment. The industrialization of research began with the establishment of centralized research laboratories at the turn of the nineteenth century at large US industrial and telecom firms. These new science-based firms were confronting hostile business environments that included new competing technologies for expiring patents and antitrust activities, and these laboratories were established as a defence.

During the period between the First and Second World Wars, US corporate laboratories pushed the limits of innovation strategies. In the 1920s, the focus was on optimizing and rationalizing production, which reflected the final phase of scientific management. The 1930s saw a shift from engineering departments to corporate laboratories as the principal focus for innovation, with new products being given the utmost priority. This trend laid the foundations for the post-World War II recovery.

Corporate interest in technology as a driver of business development appeared after World War II, with the tying of technology investments to strategy. Numerous science-related products emerged from corporate laboratories that were also driven to some extent by huge increases in military spending. This created the linear model of innovation that existed for a number of decades. Motivated by large World War II projects, such as the atom bomb and radar, many US firms in the 1950s and 1960s embraced the concept that R&D investment was all that was required for commercial innovation. The linear innovation model reinforced this perception, with the innovation process starting with a scientifically developed concept, followed by methodical development stages. The perception was that by basing innovation on science, large payoffs could be expected through the opening of new markets.

Cynicism with this approach began in US industry during in the 1970s and followed soon after in Europe. Up to the 1960s, demand from the reconstruction of the industrialized economies and the lack of any major competition resulted in a focus on the optimization and enhancement of system operations as opposed to productivity and innovation. A large component of US R&D was also derived from government funding in the high tech sector, especially the military.

By the 1980s, it became obvious in many sectors that an innovation system based on research had problems executing the later phases of innovation. Another issue was that the expectations of significant new products based on science had been overstated, with final success often elusive. Invention on demand did not fit the process model, with a number of product failures challenging corporate research, and management calling into question R&D expenditure levels.

Open innovation systems then gained traction, with global firms successfully coordinating design and manufacturing communities to deliver their requirements with speed, efficiency and flexibility. Networks led to successes in innovations and have typically included both small and large firms that swapped expertise and information. Innovation prospers on the diversity and flow of information, and having access to knowledge networks proved to be far more valuable than the centralized corporate laboratory with its long project cycles and large overheads.

The business environment today is similar to that at the turn of the twentieth century when large vertical corporations dominated, with today's huge technology firms seen as sources of innovation. Approximately 90% of successful start-ups today are acquired in private markets by incumbent firms. This represents a fundamental change to the venture capital technology model—from the development of successful new firms to the funding of private research and development start-ups that will be acquired by established incumbents, an approach used by biotech venture capital. Dominant technology firms today are valued by investors for their future market share, continuing network effects and amassing of data, and ultimately, monopoly profits.

1.3.2 Industry Boundaries

Industry dynamics focuses on how industries are currently organized, how they differ from earlier periods, what factors brought about the reorganization of the industry, and how these factors changed over time. The factors that drive innovation, entry and exit, growth and decline, and ultimately, an industry's evolution can be framed within its business infrastructure. An industry's business infrastructure includes assets that are used in the production and distribution of goods and services that the firm is unable to provide. These include technologies of production, transportation, communications and financing, while government influences both the firm's regulatory environment and provides infrastructure. Government infrastructure investments are important public goods, as firms are unable to capture the benefits, and therefore, reluctant to bear the costs of these investments.

A firm's boundaries describe its business model, and include scale, scope and in what businesses to conduct operations. A firm's horizontal boundaries are defined by the size of its product markets, while its vertical boundaries are those activities that the firm conducts internally versus those bought from external markets. A firm's corporate boundaries are the portfolio of discrete businesses in which the firm competes.

Horizontal boundaries are defined by scale and scope. Scale defines the range of output for a production process, in which the average cost declines over that range as output increases, and the marginal cost of the last unit of production is less than the average cost. Scope relates to the cost savings a firm can realize as it increases the diversity of goods and services produced.

Vertical boundaries include a firm's make-versus-buy decisions, whether an activity is performed in-house or procured on external markets. Make-versus-buy decisions can include long-term contracts, joint ventures and alliances in which firms can pool resources for strategic purposes. Goods and services in a production process generally flow from upstream to downstream along a vertical chain, initially with materials, then components, manufacturing, and finally, distribution and retailing. A firm's position along this vertical chain defines its vertical boundaries, and therefore, the costs and benefits of make-versus-buy decisions.

The Third Industrial Revolution that began in the 1870s saw huge investments in mechanization initially in the UK then in the US, and, by the late nineteenth century, the rise across many industries of large vertically integrated corporations. The emergence of these organizational structures was facilitated by flow rationalization, a process that addressed internalized bottlenecks within the boundaries of the firm, used organizational hierarchies and provided owners and management direct authority over the supervision of labour and work design. A series of innovations in systematic management that included cost accounting, schedule planning, and production and inventory control also emerged as solutions to management control. These innovations, as opposed to pure fundamental technologies, were an essential condition in the design of industrial step processes and the ultimate source of value.

Big business during the Fourth Industrial Revolution was defined by a number of concepts by the 1920s. The high volume mass production and distribution of goods relied on complex step processes. Systematic management, combined with industrial engineering, increased production volume and lowered costs. Budget planning, financial control and the vertical integration of industries—in itself a comparatively new phenomenon—also provided management with the critical components for control over the entire process and the organizational hierarchies.

While innovations in systematic management enabled the ascent of dominant vertically integrated firms leading into the twentieth century, new technologies led to a transition from vertical to horizontal boundaries in the computer industry a century later. The final two decades of the twentieth century saw dominant vertically integrated firms in the computer industry virtually disappearing, and in their place, the ascent of open platforms, and from 1985, the emergence of ecosystems. The effects of this transformation were enormous in regards to firm turnover and value created and lost.

What separated computers from the industrial era is that they are both complex systems with separate functional components, and platforms that offer multiple options, which facilitated both their exponential growth in functional development and decline in costs. Platform systems also differ from step processes in regards to organizational design. While step processes led to integrated ownership, vertical integration, hierarchical information flows, central planning and direct authority, the component optionality within platform systems removed the need for vertical controls and supply chain logistics, promoted open innovation, and enhanced overall value within the platform system.

Platforms can be closed or open systems. Open platforms can be built intentionally by a sponsor to generate fee revenues, or created as a closed system by a sponsor. In closed or product platforms, components are supplied by one firm, and by definition, are vertically integrated. Open platforms are ecosystems by definition, where components are supplied by multiple firms and communities. Industry structures will ultimately shift from closed to open platforms, as platforms facilitate open innovation, offer network effects, an increase in value through optionality and the ability to generate fee revenues.

1.3.3 Strategy and the Environment

Strategy is defined as the process by which a firm deploys its resources and capabilities within its business environment in order to achieve its goals. Corporate strategy is concerned with where a firm competes, while business strategy is concerned with how a firm competes.

The development of business strategy over the last 60 years was driven more by business pragmatism than theory. The 1950s and 1960s were a period of relative stability, and firms focused on growth through diversification, vertical integration, mass marketing, efficiencies through scale and long-term investments. Corporate planning grew in popularity as a result of the increasing size and complexity of these firms and the problems associated with management and control. Although financial budgeting offered some means for addressing these issues, the main strategic objective was the long-term planning of investments, which required a longer time horizon than that provided by annual budgets.

In the 1970s, analytical concepts such as portfolio planning matrices became popular strategy and resource allocation frameworks. The matrix approach was designed to assess business unit performance and the corporate portfolio's performance and strategies in general. Boston Consulting Group's growth-share matrix was an innovation in corporate strategy, and became a principal framework for resource allocation in diversified firms. A number of economic events during the 1970s, however, ended the post-war period of relative stability. The oil shocks, high interest rates and the increased international competition from Asia and Europe created an unstable business environment in which diversification and planning no longer provided the expected synergies.

Firms moved towards more flexible strategic management methods that focused on competitiveness as a result, with competitive advantage becoming the main objective. This had a significant impact on strategic concepts at the beginning of the 1980s. One development pioneered by Michael Porter was the use of industrial organization economics in the analysis of profitability, which emphasized a firm's competition, market environment and industry structure. Capital market developments and the profit incentives in reviving non-performing corporations also created a fertile environment for the emergence of corporate raiders and leveraged buyout firms. The activities of these players exposed the vulnerability of many large diversified corporations, which led to several takeovers. Management became focused on the stock market valuations of their firms as a result.

In the 1990s, shareholders and the financial markets continued to pressure management to maximize shareholder returns, and as a result, the shareholder value concept was included in all aspects of strategy. Growth strategies such as diversification, vertical integration and corporate planning developed over the previous 30 years were replaced with a focus on profitability. The focus also moved from the external environment to the analysis of a firm's resources and capabilities as the basis for competitive advantage, where those resources and capabilities that are unique to the firm are identified as delivering value.

Since the start of the twenty-first century, information technology has had an increasing influence on strategy analysis. Technology intensive businesses have unique investment requirements. These include markets with strong network effects, the creation of value from intellectual property assets, and leveraging technology to build platforms. The cost structures in technology industries also fundamentally differ from those found in manufacturing and service industries.

Profit is derived as revenue from demand minus costs that are a function of the technical aspects of production. There are two business drivers have unique attributes in technology businesses. The first are network effects, which have a significant impact on product demand. The second are cost structures. Almost all costs are fixed and sunk, which influences the cost per unit associated with a firm's level of output, while marginal costs are almost zero, which implies increasing returns with scale.

A third characteristic of technology businesses is the value found in intellectual property assets. This includes both the creation of intellectual property and business models that can capture the value. The patent litigation seen today is driven, in part, by smartphone developments that are providing a foundation for the next open platform.

Digital platforms share these characteristics with previous transformational innovations, such as railroads and utilities. Marginal costs will move towards zero and below average costs. Price will also be driven to the marginal cost, generating operating losses for competitors and providing a potential source of competitive advantage. The initial cash flow deficits funded by internal cash flow and investors, and the amortization required for the huge investments made in assets, are a function of the expected growth and massive profits that can result when a monopoly emerges. The payoff is that platforms ultimately offer scale and dominance, and the ability to create large ecosystems of consumers and suppliers.

Value is created when a firm earns a return greater than the cost of capital employed to generate that return through the efficient management of resources. The continuing rise of the information economy will therefore require the realignment of a firm's resources and capabilities as technology continues to redefine business models, industry boundaries, strategic alliances and networks, strategies based on IP and platforms, and ultimately, the creation of value.

1.3.4 Future Value

The digital economy is defined in terms of the Internet and related information, communications and technology (ICT). The digital sector are those core activities that include digitalization, ICT, platforms and platform activities, while digitalization includes an array of new applications of information technologies within business models and products that are transforming the economy and industry sectors. While innovation is often defined in terms of new products or processes, it also includes new business models, management systems, organization forms, value chains, processes, contractual relationships and investment approaches.

Platforms have the potential to reach the scale and scope seen with the huge vertical corporations that dominated at the turn of the twentieth century. Since the 1980s, the PC (personal computer) has evolved into a huge open platform system, while in the 1990s, the Internet generated a still greater open platform, which led to open platform exchanges that facilitated the exchange of information, products and other transactions. Open platforms create ecosystems, which transform industries dominated by vertically integrated firms into networks of specialized modular firms.

How value is created in platforms is a function of the economic relationships within the Internet, data, analytics, software, computer capacity, intellectual property and the ecosystem generated by the platform and the terms under which users participate. Platforms in many situations are also disrupting the structure of economic activity through regulatory arbitrage, the rearrangement of the barriers to entry, and ultimately, how value is created.

Platform technologies include cloud computing, data, analytics and software services. Data and analytics are the new innovations in systematic management in the digital economy, and are equivalent to the innovations in systematic management seen in the huge vertical firms at the turn of the twentieth century. These technologies also form a foundation for other technologies such as the Internet of Things, the Industrial Internet, autonomous vehicles and mobile technologies.

Externalities and public goods are economic concepts that have a much larger influence within the digital economy relative to the industrial economy. Two concepts used in digital business models are open source products and patients.

While the open source model is economically efficient—with a zero marginal cost of providing a good and therefore a theoretical zero price—there is, however, an initial cost. Free digital goods in the open source model are therefore cross-subsidized through income sources that typically include advertising.

Patents, by definition, are legal monopolies that provide an exclusive right to leverage an idea, and are central to technology standards. Patents that create industry standards also have leverage, as they provide the ability to disrupt other businesses. Patient litigation is the result of the inevitable disruption that emerges with new expansive markets. The current legal wars over smartphone patents were also seen when the telegraph and radio technologies emerged, as firms attempted to position themselves for these new markets. Also indicative of a fundamental change in the economy is the increased patient litigation seen today in industries not related to the technology industries, such as the automotive, food and mining industries.

Digitization will also drive a fundamental realignment of industry boundaries. While these dynamics may not affect every firm and sector, many industry boundaries will however be redefined. Vertical to horizontal transitions are seen today in industries such as biotech and pharmaceuticals, telecommunications, and media and entertainment. Firms that have a horizontal dominance in technologies will also leverage those resources across other vertical sectors. Those firms with operations within established industry sectors will compete with firms from other sectors as a result, and require the resources and capabilities to manage these cross-sector dynamics. Industries exposed to these dynamics include health care, financial services and energy.

The energy industry is in the process of a fundamental transformation. Global oil demand is forecast to potentially peak as early as 2025. Accelerating the peak is an array of competitive alternatives to fossil fuels that include solar, wind power, batteries and electric vehicles. Government restrictions on greenhouse gas emissions are also having an impact. The intensity of climate trends and events is likely to increase in the future, and also have an impact on global supply chains and other industries to various degrees. Although investments in renewable energy are currently double that of coal, natural gas and nuclear combined for power generation, as of 2017, 85% of the global energy system is still originated from fossil fuels.

To address the impact of global warming and move towards zero carbon clean electricity, the power supply system will need to expand by an estimated factor of four to 2050, with electricity demand expected to reach approximately 100,000 terawatt-hours. The global demand for electricity, currently 20% in the energy mix, will potentially rise to 60% by the 2050s. To meet future electricity demand, the energy mix will likely consist of intermittent renewables and firm low-carbon energy resources such as natural gas, carbon capture technologies, nuclear power and bioenergy.

The storage of energy can be achieved through a number of methods that include thermal, mechanical, which includes hydroelectric, electrical and electrochemical technologies. Electrochemical energy storage, including lithium-ion batteries, has seen the largest growth in scale capacity in energy storage since the start of the twenty-first century. The transformation of electricity grids to an intermittent renewable and firm generation mix requires the ability to smooth out demand spikes, and large-scale battery deployment offers a solution for grid management. Hydrogen fuel-cell vehicles offer a solution to the decarbonization of road transport. There are, however, significant challenges that need to be addressed, which include its manufacture, the price of fuel cells in which hydrogen is used, and its transportation.

Biopharmaceuticals have the potential to become the foundation of the pharmaceutical industry. Replicating large molecules on an industrial scale, however, requires new capabilities in manufacturing. DNA sequencing platforms, or biofoundries, have the potential to become a new industry. Biofoundries offer a solution for the centralization of process work in genetic engineering research, can provide scale by centralizing the cost of biotechnology firms operating their own laboratories, and therefore offer a shift from the biotech vertical business model. This will facilitate a new synthetic lifeform design process that can be scaled up from the current boutique business model to a global industry.

While the digital revolution has its seeds in the information technology sector, it is consistent in scale and scope to that seen in the railroad and electrification technology revolutions. The diffusion of these technologies took approximately 50 years to be realized in the economic transformations that followed. It has taken an equivalent period to deploy the fundamental technologies of the digital economy. As with these prior technologies, it is therefore likely to take another 50 years to see the full economic impact of digitization, and as such, the digital revolution and its diffusion is only half way there.

Appendix: Classifying Industry Sectors Today

Industry classification systems categorize firms into groups using a number of different factors, which can include similar production processes, products or financial market behaviour. In general, industries are identified with relatively broad markets, while markets themselves refer to specific products. The rise of the information and service economies, however, has blurred the boundaries between manufacturing and services, and created an issue in how to define an industry's boundaries.

One distinction is the difference between high technology and mature industries. High technology or technology intensive industries are science-based manufacturing industries that have above average R&D levels. Measures for high technology industries include the level of R&D intensity, derived by dividing industry R&D expenditures with industry sales, and levels of patent activity. Examples of high tech industries include the information technology sector, aerospace, pharmaceuticals and communications.

Mature industries are those that have moved through the emerging and growth phases, and have reached a stage in their life cycle where they grow in line with the economy. Examples include financial services, insurance, food, energy, construction, automotive, tobacco, steel and textiles. R&D expenditure is typically less than 1% of sales, which contrasts to high technology industries, where R&D spending can be up to 65% of sales.

In 1999, Standard & Poor's and Morgan Stanley Capital International (MSCI) together launched the Global Industry Classification Standard (GICS) to establish consistent industry definitions. The GICS system was designed to classify firms into groups that have similar stock market behaviour, and today, consists of 11 sectors aggregated from 24 industry groups, 69 industries and 158 sub-industries. The 11 sectors are:

- Consumer discretionary
- Consumers staples
- Energy
- Financials
- Health care
- Industrials
- Information technology
- Materials
- Communication services
- Utilities
- Real estate

Bibliography

Baldwin, C. Explaining the Vertical-to-Horizontal Transition in the Computer Industry. Working Paper 17-084. 2017.

Baldwin, C. Keynote Speech, Real Options Conference. Boston. 2017.

Barbera, M. and Coyte, R. Shareholder Value Demystified. UNSW Press. 1999.
Besanko, D. Dranove, D. Schaefer, S. Shanley, M. Economics of Strategy. 6th edition, Wiley, 2012.
Birkett, W. Value Creation and Risk Management in Financial Services, Towards a Taxonomy of Risk, UNSW White Paper, 2001.
Bodrozic, Z., Adler, P. The Evolution of Management Models: A Neo-Schumpeterian Theory, Administrative Science Quarterly, Vol. 63(1) 85–129, 2018.
Bureau of Economic Analysis. Defining and measuring the digital economy. Working paper March 15, 2018.
Cassis, Y. Big Business, The Oxford Handbook of Business History, Oxford University Press, 2007.
Collis, D. and Montgomery, C. Corporate Strategy: A Resource Based Approach, Irwin McGraw-Hill, 1997.
CVC, Standard & Poor's, Global cost of capital report, 3rd Qtr, 2004.
Davies, H., Lam, P.L. Managerial Economics, An Analysis of Business Issues: Third edition. Financial Times/Prentice Hall Books. 2001.
Díez, F. J., Leigh, D. Tambunlertchai, S. Global Market Power and its Macroeconomic Implications, IMF working paper, WP/18/137, 2018.
Dimson, E.. Marsh, P., and Staunton, M. Triumph of The Optimists: 101 Years of Global Investment Returns, Princeton University Press, 2002.
Financial Times, Wolf M, Why China will not buy the world, July 9, 2013.
Financial Times. Diane Coyle. Digital economy is disrupting our old models. April 25, 2018.
Financial Times. Richard Waters. Technology: Mired in a legal morass. May 19, 2012.
Foster R., and Kaplan S., Creative Destruction, Doubleday, 2001.
GICS Global Industry Classification Standard, S&P Global Market Intelligence, MSCI, 2018.
Graham, M. Technology and Innovation, The Oxford Handbook of Business History, Oxford University Press, 2007.
Grant, R.M. Contemporary Strategy Analysis, Blackwell, 2005.
IMF. Measuring The Digital Economy. February. 2018.
Janeway, W. Doing Capitalism in the Innovation Economy: Reconfiguring the Three-Player Game between Markets, Speculators and the State, 2nd Edition, Cambridge University Press, 2018.
McCraw, T K. Prophet of Innovation: Joseph Schumpeter and Creative Destruction, Belknap Press, 2010.
Nohria, N., Dyer D, and Dalzell, F. Changing Fortunes, Wiley, 2002.
Nolan, P. Is China buying the world? Polity, 2013
Perez, C. Technological Revolutions and Financial Capital, The Dynamics of Bubbles and Golden Ages, Edward Elgar, 2003.
Stearns, P. The Industrial Revolution in World History. Westview Press. 2007.
The Economist. Global heroes, a special report on entrepreneurship, March 14, 2009.

The Economist. Schumpeter. America's antitrust apparatus prepares to act against big tech. April 26, 2018.

The Economist. The battle for digital supremacy. March 15, 2018.

The Economist. Too much of a good thing, March 26, 2016.

The Economist. Why giants thrive. September 17, 2016.

The Economist. Why large firms are often more inventive than small ones, December 17, 2011.

2

Value

2.1 Overview

To deliver sustainable shareholder value, management has to simultaneously manage operations in the short term while delivering on plans for the long term. Commitments in the short term include delivering on earnings and maintaining liquidity, while in the long term, they involve developing and executing on strategy and investments. The following framework discusses the metrics that encompass current and future performance.

The accounting foundations describe how the financial statements translate a business model and provide a framework for identifying value. The articulation of the financial statements identifies their stocks and flows, which provide the means to identify how a business generates value. Financial ratios provide a lens into the current performance of a firm to identify the factors that drive value. Return on shareholders' common equity (ROCE) and return on invested capital (ROIC) ratios are examined within this context.

Residual earnings (RE) account explicitly for the cost of equity in equity valuation, where returns greater than the cost of equity are required to create value. Free cash flow is a corporate finance concept that identifies a firm's cash flows that are available for distribution to various parties, which include equity and debt holders, while also continuing operations.

Intangible assets have increasingly become far more relevant with the rise of the information economy, along with their recognition and valuation on balance sheets. A pro forma analysis projects the financial statements of a firm, and has applications in valuation, strategy, credit analysis, M&A (mergers and acquisitions) and budgeting.

J. Rogers, *Strategy, Value and Risk*, Global Financial Markets,
https://doi.org/10.1007/978-3-030-21978-9_2

Corporate investment methods are then reviewed. Discounted cash flow (DCF) techniques and its shortcomings in the assumptions and decision rules are discussed. The advantages of real options techniques that address these issues are then outlined. The valuation of real options covers the issues associated with this technique, followed by the various types and definitions.

Corporate finance covers theory and builds a framework that extends on the initial model with the inclusion of growth options, an abandonment option, modularity and financial options. Finally, a firm's capital structure is discussed within this framework.

2.2 The Accounting Foundations

A business model provides a framework for identifying and creating value. Business models describe how the components of a business combine as a system. The phrase is widely used to describe the diverse features of a business, and its scope can include strategy, purpose, offerings, processes, operations, organizational aspects and trading practices. A good business model identifies the customer, what customers' value and how value is created in a business. The ability to identify how a business model functions and creates value provides a foundation for valuation.

Financial statements provide a framework for identifying how firm value is generated for shareholders and other stakeholders. The attributes of a business model are translated into accounting metrics that provide a lens into how and where value is created. Accounting principles define how financial statements are organized, and therefore, how value is measured. Firms generally issue three primary financial statements—the balance sheet, the income statement and the cash flow statement. One additional report usually required is the statement of shareholder's equity.

The balance sheet itemizes a firm's assets, liabilities and shareholder equity. Assets are a firm's investments that are anticipated to generate future payoffs. Liabilities are claims to payoffs on the firm by non-owner claimants, while shareholders' equity is a claim on the firm by its owners. The balance sheet is therefore a statement of the firm's investments and the payoff claims on those investments. Assets and liabilities are also identified as being either current or long-term, where current defines those assets that produce cash or how cash is used to pay liabilities within one year. The balance sheet's three components are linked through the following accounting relationship:

$$\text{Shareholders' Equity} = \text{Assets} - \text{Liabilities}$$

This accounting equation states that shareholders' equity is equivalent to a firm's net assets, or the net difference between the firm's assets and liabilities. Shareholders' equity is therefore the residual claim on a firm's assets after liabilities have been deducted.

The income statement provides an account of the increases or decreases in shareholder's equity that result from a firm's operations and activities. The value that is added to shareholder value is described by convention as the bottom line, net income, net profit or earnings. The income statement also itemizes the firm's revenues and expenses that are the foundation of net income. This is established though the following accounting relationship:

$$\text{Net Income} = \text{Revenues} - \text{Expenses}$$

The cash flow statement shows the cash generated and used by a firm over an accounting period. The various cash flows in the statement are identified as cash flows from operating activities, cash flows from investing activities and cash flows from financing activities. The total cash flows from the three definitions identify a firm's increase or decrease in cash activities:

$$\begin{aligned}\text{Change in cash} = {} & \text{Cash from Operations} + \text{Cash from Investment} \\ & + \text{Cash from financing}\end{aligned}$$

The statement of shareholders equity explains how a firm's equity has changed over an accounting period:

$$\begin{aligned}\text{Ending equity} = {} & \text{Beginning equity} + \text{Total}\left(\text{Comprehensive}\right) \\ & \text{Income} - \text{Shareholders' net payout}\end{aligned}$$

A firm's equity increases when value is added through operations as net income in the income statement, along with other comprehensive income and shareholders' investments, and decreases with payouts to shareholders.

2.2.1 Stocks and Flows

The articulation of the financial statements describes their relationships or the manner in which they fit together. The balance sheet provides the stock of owner's equity and cash at a point in time, while the cash flow statement accounts for how the stock of cash has changed over time. The statement of shareholder equity, which identifies the change in owner's equity or flows over two balance sheet dates, describes the relationship between the income statement and balance sheet. The income statement, adjusted for other comprehensive income in the equity statement, describes the change in owners' equity derived from the value added from operations.

Identifying the articulation of the financial statements reveals their stocks and flows, which provide a foundation for the analysis of how a business generates value. The balance sheet describes the stock of value in a firm at a point in time, while the income statement and cash flow statement account for the flows, or change in stocks, between two points in time in the balance sheet. The statement of shareholders equity equation given earlier is the stocks and flows equation for equity, as it describes how the stocks of equity have changed with the equity flows. The cash flow relations described in the cash flow statement equation is the stocks and flows equation for cash.

The stocks and flows concept can be extended to define value, with the balance sheet providing the shareholders' net worth as a stock, while flows are the value added through a firm's operations in the income statement, and in the cash flow statement as changes in cash. Therefore, the value that flows to a firm's owners is the change in equity over an accounting year.

The value of a firm should always equal the value of the claims on the firm:

$$\text{Firm value} = \text{Value of debt} + \text{Value of equity}$$

This relationship illustrates that total firm value is distributed to the various claimants on that value. Firm valuation can therefore be defined as either valuing the firm itself, or valuing and summing the claims on the firm. The firm also has a portfolio of projects, with the value of the firm represented by the present value of the expected cash flows from operations, or free cash flows, from these projects. Firms seek continuity by investing in new projects while letting existing projects terminate. The cash generated from a firm's assets and operations flows to the claimants on the firm. Therefore, the analysis of a firm's operations, financing activities and investments provides the foundation on how firm value is generated and the sustainability of that value creation.

Figure 2.1 illustrates all the stocks and flows of a firm. The firm's debt and equity financing activities are transacted with claimants in the capital markets. At inception, a firm begins with funding sourced as cash from shareholders. This cash is initially invested in liquid financial assets such as short-term

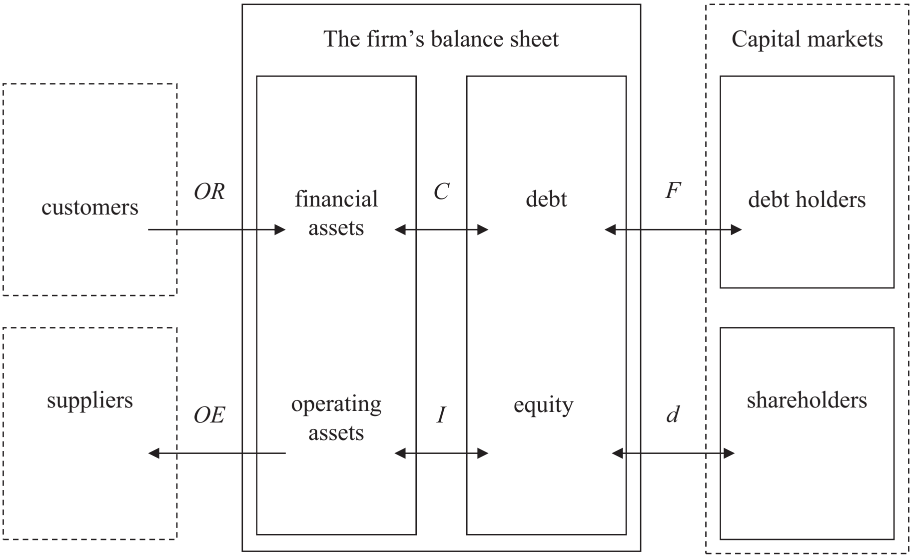

Fig. 2.1 The stocks and flows of a firm

money market securities before being invested in operating assets. Additional funds are also raised as debt to fund balance sheet assets. Cash that moves between the debt holders and the firm is defined as the net debt financing flow, F. This net cash flow consists of cash paid as interest and principal repayments to debt holders minus cash borrowed by the firm from its creditors. Similarly, the net dividend to shareholders d represents cash paid as dividends and stock repurchases minus capital cash injections from shareholders.

Firms divest the cash in financial assets to invest in operating assets, described by the firm's investing activities, I. These cash flows can move in either direction, with investments in financial assets also flowing from the proceeds of liquidating operational assets such as discontinued operations. Net cash flows are then generated from the operating assets, defined as cash from operations, C, through the operating income generated by operating revenues, OR, minus the operating expenses, OE. This cash from operations is then invested in liquid financial assets, and so, the cycle continues.

An important identity is the cash conservation equation, or sources and uses of cash equation. The four cash flows—cash flow from operations, C; cash investment, I; net cash flow to debt holders and issuers, F; and net cash flow to shareholders, d—always observe the relationship:

$$C - I = F + d$$

Or, free cash flow equals the net payments to debt holders and debt issuers plus the net dividends to shareholders. Cash flow from operations minus the cash investment in operations, therefore, is always equivalent to the net cash flows to debt holders and shareholders.

The left hand side of the relationship, $C - I$ represents a firm's free cash flow. Free cash flow is positive if operations produce more cash than required for new investment, and negative if operations create less cash than required for new investment. A positive free cash flow is either invested in financial assets, F, or distributed as dividends, d. A negative free cash flow necessitates either the issuing of bonds, a negative F, or the issuing of stock, a negative d, to meet the cash shortfall.

The following identities also hold for corporate cash management. If

$$C - I - i > d,$$

where i is the net interest cash flow, or the interest paid minus the interest received, then either lend or buy down the firm's debt. If,

$$C - I - i < d,$$

then either borrow or reduce lending.

2.2.2 Ratio Analysis

A multiple is the ratio of the market price of a firm's stock to some accounting measure per share that is used as an estimate of relative value. A price multiple summarizes the relationship between a firm's stock price and a measure such as earnings, book value or sales per share.

The price-earnings (P/E) ratio compares the current stock price with earnings, and anchors a valuation to an income statement. The ratio is interpreted as:

- the price or numerator reflecting future earnings, or the market's expectations of value added from future sales, and
- earnings, the denominator, reflecting current earnings, or the value added from current sales

The P/E ratio therefore evaluates the forecast of future earnings in relation to current earnings. Higher future earnings expectations relative to current earnings should result in a higher P/E ratio, while lower future earnings expectations relative to current earnings should result in a lower P/E. The P/E ratio is, therefore, an indication of anticipated earnings growth.

The price-to-book ratio, or P/B ratio, compares a firm's book value to its current market price. The P/B is derived as:

- the ratio of the firm's market capitalization over the firm's total book value, or
- a per-share value, the ratio of the firm's current share price over its book value per share, or the ratio of book value over the number of shares issued

By convention, book value does not include intangible assets.

A firm's book value represents the shareholders' investment in the firm, with the value derived on the expectations of how much the net assets will earn in the future. Book value can either increase or decrease, depending on the firm's future earnings expectations. While book value does not accurately determine value, the missing component is ultimately realized in the future earnings created by book value.

The stock price in the P/B ratio's numerator is based on expected future earnings. Therefore, the higher expected earnings are in relation to book value, the higher the P/B ratio. The book value rate of return, or profitability, is a measure that principally determines P/B ratios. The market price-to-book value ratio is the price-to-book ratio or the market-to-book ratio, while the intrinsic value-to-book value ratio is the intrinsic price-to-book ratio.

Return on shareholders' equity (ROE), or more specifically, return on shareholders common equity (ROCE) measures the rate of return on common stock:

$$\text{ROCE} = \frac{\text{Comprehensive income}}{\text{Average CSE}}$$

The measure assesses a firm's profitability efficiency per unit of shareholders' equity or book value. ROCE can be decomposed into three drivers:

$$\text{Net profit margin} \times \text{Asset turnover} \times \text{Financial leverage.}$$

or,

$$\text{Net income / Sales} \times \text{Sales / Total assets} \times \text{Total assets /}$$
$$\text{Average shareholder equity}$$

Net profit margin is a relative measure of the rate at which profitability is generated from operating assets, or the dollar of net profit generated from a dollar of sales revenue. Asset turnover measures the efficiency of operating assets, and describes the relationship between the use of assets and profitability. The ratio focuses on the volume of sales generated from an investment in operating assets, or the dollar of sales revenue generated for each dollar invested in operating assets. Asset turnover centres on two operating asset groups—the working capital assets, such as cash, inventory and receivables; and the fixed assets that include plant, property and equipment. The leverage ratio describes the degree to which a firm relies on debt to create profitability. A firm can increase its asset base through financial leverage or borrowing, which can enhance the returns to shareholders.

The first two ROCE drivers—net profit margin x asset turnover—define the return on assets (ROA) ratio. ROA establishes a firm's efficiency in the use of assets and is also a profitability measure. The Du Pont Formula integrates the analysis of a firm's profitability and investments in assets, and provides a window into the sources of a firm's profitability. High net margins indicate that customers are prepared to pay more for a firm's products, while a high asset turnover indicates that a firm uses its assets relatively more efficiently in generating sales, and therefore, invests less capital. The return to shareholders can therefore be increased by either increasing the profit per dollar of sales, or increasing the sales dollars generated from the operating assets.

Return on invested capital (ROIC) measures a firm's success in generating cash flow relative to its invested capital. The measure is derived as net operating profit after taxes (NOPAT) divided by invested capital, which includes working capital, debt, and common and preferred stock:

$$\text{ROIC} = \frac{\text{NOPAT}}{\text{Invested capital}}$$

Firm value is created when the ROIC is greater than the cost of capital, and value lost if the spread is negative. ROIC provides a better metric for the analysis of a firm's performance than ROCE and ROA as it centres on a firm's actual operations. ROCE combines operations with leverage, while ROA understates a firm's profitability as it does not include the leverage from operating liabilities or the profitability from financial assets. ROIC driver patterns reveal a fade rate or persistence where the ROIC reverts to a long run level. Economic factors typically influence firms in a comparable manner within industry sectors, and drivers tend to fade to levels that are representative for an industry.

2.2.3 Residual Earnings and Free Cash Flows

Residual earnings is net income minus a deduction that represents the common shareholder's opportunity cost in generating net income. For each earnings period, residual earnings (RE) is derived as:

$$\text{RE} = (\text{ROCE} - \text{Required return on equity}) \times \text{Book value of common equity}$$

where ROCE equals $\text{earnings}_t/B_{t-1}$ and is the rate of return on common equity. Two value drivers therefore determine residual earnings, the ROCE and the book value B_{t-1} for each period.

Firm value is generated over book value by increasing the ROCE spread over its cost of capital. Value is additionally increased by the growth in book value, or net assets, which earn at the firm's ROCE. A value strategy can therefore be framed as increasing firm value through investments and strategies that increase ROCE over the required return, and grow book value or net assets.

A firm and its equity can also be valued by discounting the free cash flow to the firm and the free cash flow to equity. Free cash flow to the firm (FCFF) measures the net cash generated by a firm, while the free cash flow to equity (FCFE) measures the cash distributed to the firm's equity shareholders after all reinvestments, debt repayments and expenses. Both measures can be derived from the financial statements, from either the net income statement or the statement of cash flows.

In the case of the income statement:

$$\text{FCFF} = \text{NI} + \text{NCC} + \text{IE}\,(1 - \text{Tax rate}) - \text{IWC} - \text{IFA}$$

where:

NI = net income, or profit after tax
NCC = net non-cash charges
IE = interest expense
IWC = investments in working capital
IFA = investments in fixed assets

In the case of the statement of cash flows:

$$\text{FCFF} = C + \text{IE}\left(1 - \text{Tax rate}\right) - \text{IFA}$$

where:

C = cash flow from operations
IE = interest expense
IFA = investments in fixed assets

Free cash flow to equity is derived from free cash flow to the firm as:

$$\text{FCFE} = \text{FCFF} - \text{IE}\left(1 - \text{Tax rate}\right) + \text{NB}$$

where:

IE = interest expense
NB = net borrowing, or change in debt

Free cash flows also vary over the life cycles of products and firms. Product and firm life cycles are related in that the product life cycle is the demand side counterpart to the industry life cycle. Firms and products progress through the stages of emerging, growth, mature and decline, or in some cases, can stay at a mature stage indefinitely. Some product cycles have a long lifespan, such as steel, paper and cement manufacturing, while products such as electronics and pharmaceuticals can have relatively short lifespans.

2.2.4 Intangible Assets and Intellectual Property

The value of a firm as a going concern lies in its income stream, with its assets the resource that generates value. The rise of the information economy has increased attention on the recognition and valuation of intangible assets on corporate balance sheets. Balance sheets explicitly exclude assets such as brands, distribution and supply chains, and knowledge, organization and human capital. This issue is especially relevant when firm value is derived more from intangible assets rather than tangible assets. The majority of intangible assets on balance sheets, however, cannot be identified and independently valued from other assets, as their value is derived from the cash flow streams generated with other assets. Knowledge capital is used in processes, marketing and management, and does not exist without tangible assets, while value from organizational capital is derived from its combination with other assets. As such, the firm itself is the asset, or an organization of assets that underlies a business plan to create value.

Asset values can however be determined from the income statement in addition to a balance sheet. The articulation of the income statement and balance sheet can reveal firm value, with each statement correcting for the shortcomings of the other. In the case of intangible assets, an income statement or flow valuation is available when a balance sheet or stock valuation cannot be determined. Although intangible assets are absent from the balance sheet, the earnings from intangible assets still flow through the income statement. As such, value can be determined by either measuring the asset value directly, or through the capitalization of the earnings from that asset itself. Although a balance sheet does not provide a summary amount for the value of assets used in combination, the income statement in principle does so. Earnings measure the value added from tangible assets in conjunction with entrepreneurship, knowledge, organizational capital and brands. Identifying the intangible assets on a balance sheet therefore is not required, as the earnings generated by the business plan provide a summary measure of value.

Intangible assets, when compared to tangible assets, generally not only have no physical identity, they also cannot be identified for the purpose of writing contacts against them for delivery. Legal rights such as patents and copyrights, and in some cases, brands are exceptions, however. As firms move to more open models of innovation and external sources of knowledge, the management of intellectual property rights (IPR) has become a significant issue. Examples of IPRs that can be licensed include patents, copyrights and trademarks; however, patent licenses are the most frequent in technology exchanges. This increasing reliance on external sources of innovation means that it is essential to have consistent measurements of what is being traded.

2.2.5 Pro Forma Analysis

Pro forma financial statements are projected forecasts that have a variety of applications, which include:

- strategic planning, such as merger and acquisition transactions and new capital investments
- financial planning, including revenue and expenditure planning; working capital modelling, capital structure analysis and short- and long-term borrowings

- credit analysis for debt covenants such as debt-to-equity ratios and debt service reserve coverage, and
- the design and valuation of securities

Pro forma financial statements provide an integrated projection of a firm's future operating prospects and financial condition, based on the current financial statements. A pro forma income statement provides an earnings estimate, and a pro forma balance sheet, the book value of equity as bottom line numbers. Forecasted free cash flows are derived from the pro forma cash flow statement, where net income and depreciation are used to construct the statement, which then provides data for the projected pro forma balance sheet. The pro forma cash flow statement is then available for forecasting free cash flows for use in financial planning, DCF analysis and liquidity analysis. Finally, the pro forma statements are modelled for multi-year projections, the data translated into values, and the projected firm value divided among the shareholders, debt holders and any hybrid security holders.

The consequences of changes in business conditions and the available choices in managing these changes can also be analysed in a pro forma. Scenarios that represent transformations in industries and the external environment can be integrated into an analysis of future performance. Building a pro forma therefore requires identifying factors that are a function of business conditions and those that are management choices. These factors include changes in products, markets, technologies, industries and regulations, while management decisions include identifying those key drivers to which the firm has an exposure, which can add value and which ensure firm continuity.

Sensitivity analysis is the modelling of the set of possible future balance sheets and income statements and identifying what is at risk. Value is framed based on whether a firm can grow book value and where it will be positioned in future years. Risk can be analysed by using different scenarios in the pro forma statements, including the best and worst scenarios, and the base case to find the margin of safety. Macro and micro economic effects and event risks can also be included in the scenarios. As financial reporting moves stock prices through earnings releases, the set of alternative accounting outcomes that will influence a firm's stock price can also be modelled.

A pro forma analysis can include an industry's driver patterns, industry and economic forecasts, how a firm's key drivers will diverge from conventional patterns, management's options versus the external environment, and the firm's projected book value. A percentage-of-sales pro forma framework example is used as an illustration, with quarterly intervals up to 12 months, followed by yearly intervals up to three years.

The following assumptions were made:

- the rate of sales growth is constant at 5%
- cost of sales are 80% of sales
- the current assets are 30% of sales
- the fixed assets are also a ratio of sales at 100%
- the parameters are assumed to be constant over the three-year forecast
- a total of 1000 shares are issued

Table 2.1 summarizes the data for the pro forma analysis example:

Table 2.2 shows the pro forma income statement and balance sheet. The pro forma statements illustrate the articulation of the accounting statements. Retained earnings, the last line of the income statement, represents the change in the retained earnings line item in the balance sheet, while changes in the income statement and balance sheet form the foundation for the statement of cash flows.

Table 2.3 shows the pro forma free cash flows. The calculations start with profit after taxes, and reverse the accruals to arrive at free cash flow.

The GAAP Statement of Cash Flows mingles free cash flows with the flows from financing activities, where cash flow from operations minus the cash used for investing activities plus the cash from financing activities equals the change in cash and cash equivalents. Realigning the statement of cash flows draws a distinction that follows the four cash flows that were linked together in the cash conservation equation $C - I = F + d$ in Sect. 2.2.1 (Stocks and Flows), as illustrated in Table 2.4.

Table 2.5 shows the calculation of the valuation of equity from the free cash flows to equity, which is derived by discounting the equity cash flows for each year up to three years, with a terminal value added to the equity cash flow at Year Three.

The terminal value at Year Three is derived as a perpetuity, with the Year Three equity free cash flow projected through multiplying by 1 plus the FCFE growth rate, and dividing by the equity discount rate minus the growth rate.

Table 2.1 The pro forma example data

Operations		Assets		Financing	
Initial sales	1000	Current assets (% sales)	30%	Interest expense	3%
Sales growth (annual)	5%	Current liabilities (% sales)	10%	Dividend payout ratio	65%
Cost of sales	80%	Fixed assets (net, % sales)	100%	Cost of equity	8%
Taxes	40%	Depreciation	10%	Debt/equity ratio	60%

Table 2.2 The pro forma income statement and balance sheet

Income statement	Y0	Q1	Q2	Q3	Q4	Y1	Y2	Y3
Sales	1000.0	262.3	262.4	262.6	262.7	1050.0	1102.5	1157.6
Cost of sales		(209.8)	(209.9)	(210.1)	(210.2)	(840.0)	(882.0)	(926.1)
Interest		(3.0)	(3.0)	(3.1)	(3.1)	(12.2)	(12.8)	(13.5)
Depreciation		(1.4)	(1.4)	(1.4)	(1.4)	(5.6)	(5.8)	(6.1)
Profit before taxes		48.1	48.1	48.0	48.0	192.2	201.9	211.9
Taxes		(19.2)	(19.2)	(19.2)	(19.2)	(76.9)	(80.7)	(84.8)
Net income		28.9	28.8	28.8	28.8	115.3	121.1	127.2
Dividend		(19.0)	(19.0)	(19.0)	(19.0)	(76.1)	(79.9)	(83.9)
Retained earnings		9.8	9.8	9.8	9.8	39.2	41.2	43.2
Balance sheet	**Y0**	**Q1**	**Q2**	**Q3**	**Q4**	**Y1**	**Y2**	**Y3**
Assets								
Current assets	300.0	303.7	307.4	311.2	315.0	315.0	330.8	347.3
Fixed assets	1111.1	1124.7	1138.6	1152.5	1166.7	1166.7	1225.0	1286.3
Depreciation	(111.1)	(112.5)	(113.9)	(115.3)	(116.7)	(116.7)	(122.5)	(128.6)
Net fixed assets	1000.0	1012.3	1024.7	1037.3	1050.0	1050.0	1102.5	1157.6
Total assets	1300.0	1316.0	1332.1	1348.5	1365.0	1365.0	1433.3	1504.9
Liabilities								
Current liabilities	100.0	101.2	102.5	103.7	105.0	105.0	110.3	115.8
Debt	387.5	392.3	397.1	401.9	406.9	406.9	427.2	448.6
Equity								
Stock	802.5	802.7	803.0	803.4	803.9	803.9	805.4	806.9
Retained earnings	10.0	19.8	29.6	39.4	49.2	49.2	90.4	133.6
Total liabilities	1300.0	1316.0	1332.1	1348.5	1365.0	1365.0	1433.3	1504.9

Table 2.3 The pro forma free cash flow forecast

	Y0	Q1	Q2	Q3	Q4	Y1	Y2	Y3
Profit after taxes		28.9	28.8	28.8	28.8	115.3	121.1	127.2
+ net non-cash charges (depreciation)		1.4	1.4	1.4	1.4	5.6	5.8	6.1
+ interest expense after taxes		1.8	1.8	1.8	1.9	7.3	7.7	8.1
– change in net working capital		(2.5)	(2.5)	(2.5)	(2.5)	(10.0)	(10.5)	(11.0)
– change in fixed assets		(13.6)	(13.8)	(14.0)	(14.1)	(55.6)	(58.3)	(61.3)
Free cash flow		15.9	15.8	15.6	15.4	62.7	65.8	69.1

The growth rate is derived as $g = (1 - \text{payout ratio}) * \text{ROCE}$, and is assumed to be constant for illustration. The present value of the equity cash flows is derived as $\text{FCFEt}/(r_{equity} - g)$, where r_{equity} is the cost of equity, which is divided by the number of shares to arrive at the value per share.

Table 2.6 shows the ROCE, ROIC and projected book value as the sum of the balance sheet stock and retained earnings line items.

Table 2.4 Free cash flows and financing flows

		Q1	Q2	Q3	Q4	Y1	Y2	Y3
Cash flow from operations	*C*	32.0	32.0	32.1	32.1	128.2	134.6	141.4
Investments	*I*	(16.1)	16.3)	(16.5)	(16.7)	(65.6)	(68.8)	(72.3)
Free cash flow	*C – I*	15.9	15.8	15.6	15.4	62.7	65.8	69.1
Debt financing flows:								
Change in financial assets (net)		–	–	–	–	–	–	–
Financial asset interest		–	–	–	–	–	–	–
Debt issuance (net)		(4.8)	(4.8)	(4.9)	(4.9)	(19.4)	(20.3)	(21.4)
Debt interest expense		1.8	1.8	1.8	1.9	7.3	7.7	8.1
	F	(2.96)	(2.99)	(3.03)	(3.07)	(12.05)	(12.65)	(13.29)
Equity financing flows:								
Dividends and stock repurchases		19.0	19.0	19.0	19.0	76.1	79.9	83.9
Stock issuance		(0.2)	(0.3)	(0.4)	(0.5)	(1.4)	(1.5)	(1.6)
	d	18.9	18.7	18.6	18.5	74.7	78.5	82.4
Total financing flows	*F + d*	15.9	15.8	15.6	15.4	62.7	65.8	69.1

Table 2.5 Valuation of equity from free cash flow to equity

	Y0	Y1	Y2	Y3
Free cash flow		62.7	65.8	69.1
– interest expense after taxes		(7.3)	(7.7)	(8.1)
+ change in debt		19.4	20.3	21.4
Free cash flow to equity		74.7	78.5	82.4
Terminal value				2198.2
Total cash flows		74.7	78.5	2280.6
PV total cash flows		73.5	74.7	2103.2
Cost of equity	8.0%			
Growth rate	4.7%			
Value of equity	2251.5			
Value per share	2.25			

Table 2.6 ROCE, ROIC and projected book value

	Y0	Q1	Q2	Q3	Q4	Y1	Y2	Y3
ROCE		3.5%	3.5%	3.4%	3.4%	13.8%	13.8%	13.8%
ROIC		2.2%	2.2%	2.2%	2.1%	8.7%	8.7%	8.7%
Book value	812.5	822.5	832.6	842.8	853.1	853.1	895.8	940.6

2.3 Corporate Investments

2.3.1 Investment Methods

Investment can be defined as the sacrifice of current dollars for future dollars (Sharpe, Alexander & Bailey). Understanding how investments are valued is important for investors, finance executives or management generally. Assets are only worth what someone is willing to pay for them, and as that person could be an analyst, a trader, a fund manager or a competitor, a background in the available valuation methods is essential. Investment valuation is used for a wide range of real and financial assets, including companies, bonds, stocks, real estate and derivatives. Although most investment valuation models are generalized rather than specific to particular markets, it is also probably one of the most difficult tasks in finance.

Many factors can have an impact on the value of investments. Errors in forecasts can result from unforeseen changes in factors such as financial variables, markets, competitors and technology. Unexpected changes in asset values can also result from factors that are completely unrelated to a firm, an industry or the economy generally. Another major influence today on the value of investments is the corporate emphasis on short-term results. A firm will either be rewarded or penalized every quarter through its share price, depending on whether earnings satisfy investor expectations. A consequence often seen is the corporate behaviour of managing quarterly financial results to reduce share price volatility.

The basis for an investment will depend on the investment philosophy. Generally, the value of an asset should be a function of the cash flows it is expected to produce. A wide variety of models are used for investment valuation with various levels of complexity; however, there are some common features. Two common approaches are DCF valuation and relative valuation, which is similar to DCF in the sense that the value of an asset is derived from the cash flows of comparable assets. DCF analysis can be performed either from the viewpoint of equity holders, in which case, the expected cash flows to equity are discounted, or by considering the firm from all perspectives and discounting the firm's expected cash flows.

2.3.2 DCF Valuation

There are two basic approaches to discounted cash flow methods—the net present value (NPV) and the internal rate of return (IRR). The NPV is the difference between the present value of the net cash inflows generated by the asset and the initial cash outlay. The IRR is the rate of return that equates the present value of the net cash inflows generated by the asset with its initial cash outlay. The IRR is the equivalent to interest rates quoted in financial markets. The NPV approach is the most popular, and has the following features:

- NPV recognizes the time value of money.
- NPV is a function of the future cash flows from an investment and reflects the opportunity cost of capital.
- Net present values can be aggregated as they are measured in today's values.

The NPV of an investment is the sum of the present values of the expected benefits, generally in the form of cash flows, from which the present values of all expected cash outlays are deducted. The variable k is defined as the rate of return that can be earned on an alternative investment. If I_0 is defined as the initial outlay, and CF_t the cash flow at the end of period t, the net present value is:

$$NPV = \frac{CF_1}{(1+k)^1} + \frac{CF_2}{(1+k)^2} + \ldots\ldots\ldots\ldots + \frac{CF_n}{(1+k)^n} - I_0 \tag{2.1}$$

2.3.3 The Net Present Value Rule

The net present rule is to accept investments that have positive net present values, that is, when the present value of the investment's cash inflows are at least as large as the present value of the cost outlays. There are some implicit assumptions underlying the NPV method (Dixit and Pindyck 1994):

- the investment can be reversed or unwound and any outlays recovered if circumstances prove to be less than expected, or,
- the choice is either a now or never decision, even if the investment cannot be reversed—if the investment is not made immediately, then it cannot be made in the future

Most investments actually do not meet these requirements. Instead, the majority of investment decisions have three central features that interact at various levels:

- The investment cannot be reversed, if not totally, then as least to some degree, in which case, the initial investment is, to some extent, a sunk cost, or a cost that cannot be recovered.
- The investment's future payoffs have an associated uncertainty.
- There is some flexibility in the investment timing. There is an opportunity to delay a decision until further information becomes available.

The opportunity to delay and the inability to reverse an investment decision are significant features found in most investments. NPV valuations, however, compare investing immediately with never investing, and reflect a static value derived from assumptions that only consider a single scenario. An NPV analysis does not recognize any flexibility management has to assess and react to future circumstances that were not initially anticipated. The NPV approach assumes a static commitment to a static strategy. Any business case based on NPV analysis will therefore ignore management's potential to modify a decision alternative in the future.

Although the NPV approach assumes a predetermined path no matter how events materialize, the business environment today is anything but static. NPV techniques are not designed to capitalize on future opportunities when they arise, or to manage any potential downside risks. NPV valuation methods do not include the value of the opportunity to wait and act in the future as more information becomes available. NPV valuations also imply that risk is a single dimension that reduces value. All uncertainties and decisions are reduced to one single scenario that is adjusted for risk through the level of a discount rate.

2.3.4 Real Options

Real options analysis is a valuation and strategic decision paradigm that applies financial option theory to real assets. Stewart Myers (Myers 1987) first referred to the term in a discussion about the gap between strategic planning and finance theory. DCF analysis, developed from finance theory, made sense when applied to businesses such as 'Cash Cows'. However, the dynamics associated with today's business environment are putting limitations on the DCF techniques used to analyse them. Risk can also be leveraged to create, rather than suppress value. Hedges can protect investments from downside risks while an exposure is maintained to any upside potential. Real options offer a framework and the metrics for managing strategy, value and risk in today's business environment.

Firms will typically invest in projects that generate a return greater than a hurdle rate. Hurdle rates can, however, be often observed at three to four times the cost of capital (Dixit and Pindyck 1994). One explanation is the implied option value, or opportunity cost associated with a capital investment. Instead of the investment decision being that discounted cash inflows must equal or exceed discounted cash outflows as per the NPV rule, the investment's cash inflows must exceed the cash outflows by the value of keeping open any optionality in the investment. If a decision is made to proceed with an irreversible investment, the opportunity to delay the investment is forfeited, and so, the rights to any option implied in the investment are exercised. This opportunity cost should therefore be included in the valuation of an investment.

The NPV rule should therefore be revised by subtracting the opportunity cost of exercising any options, and then, investing if the modified NPV is positive. The alternative is to keep the conventional NPV and the option value distinct. The investment framework can therefore include two identified value components—the NPV and the real option value. A strategic NPV can be defined as (Trigeorgis 1996):

$$\text{Strategic NPV} = \text{Standard NPV} + \text{Option Premium}$$

The modified NPV rule is now to invest if the Strategic NPV is greater than zero.

2.3.5 Valuing Real Options

Financial options are asymmetric relationships, where the option holder has a right but not the obligation to transact at a contracted price (the exercise price) on or before a predetermined date (the exercise or maturity date). A call option is the right to buy, and a put option is the right to sell the underlying instrument at the exercise price. A European Option can only be exercised at the end of its life, while American Options can be exercised at any time during its life. In the case of a real option, it is the right but not the obligation to act, such as deferring, expanding, contracting or abandoning a project or investment at a predetermined cost (the exercise price) for a predetermined period of time (Trigeorgis 1996 and Copeland and Antikarov 2001). Value is created in a financial option from the volatility in an underlying financial asset, and the same concept is applied to real options, where value is derived from the uncertainty or the volatility associated with a real asset.

A relatively simple argument has been developed in financial economics to price an option under the assumption that no arbitrage opportunities exist. An economy exists that has an abundant set of traded assets from which a portfolio can be created. This portfolio consists of buying a specific number of shares of a stock, against which a certain amount is borrowed at a risk-free rate such that the portfolio replicates an option's returns in any state of nature. In the absence of any arbitrage opportunities, or risk-free profits, the option and the corresponding portfolio must sell for the same price as they provide the same future return. Therefore, the price of the option is the equivalent to the cost of setting up the replicating portfolio.

The no-arbitrage replicating portfolio concept used to price options can be applied to real options by employing the assumptions used in deriving the NPV of an asset or project. The discount rate used in DCF analysis typically estimated using the Capital Asset Pricing Model (CAPM) is based on the price of traded twin security with the same risk characteristics as the investment or project being analysed. Therefore, the same traded twin security can be used to replicate the real option's returns. This leads to an important assumption in valuing real options—that existing assets in the economy span the risks in the asset or project being valued. Capital markets must be adequately complete so that an asset exists such that its price is perfectly correlated with the asset underlying the real option. Real options can, however, have risks that are not priced or spanned in the financial markets. These risks that cannot be represented by the price of a traded security are known as private risks.

Incomplete markets can be found in all real asset markets, and even in financial markets. Incomplete markets are likely to remain in regards to a specific risk if the costs exceed the benefits of creating the securities required to span a specific risk, or if there are problems associated with making such securities legitimate. Other market imperfections include intermittent trading, sporadic price discovery and a lack of liquidity. Robert Merton (1998) presented a framework in his 1997 Nobel Prize lecture for determining the value and risk of a non-traded asset by using a portfolio of traded securities. There are two aspects that can be drawn from Merton's address. The first is that it is probable that some kind of tracking of the risks in a corporate investment can be established through a portfolio of traded securities, in spite of market imperfections. The second is the rigorous definition Merton offers of private risk. Merton defines and measures private risk as the size of the tracking error between the portfolio of traded securities and the value of the underlying asset. Private risk can therefore be identified through the data, rather than through subjective breakdowns of market and private risks.

Other techniques that can be used when spanning does not hold are decision analysis and dynamic programming. Decision analysis is a structured quantitative approach for the evaluation of decisions that have complex alternatives, competing objectives and major sources of uncertainty. The origins of decision analysis began at Harvard Business School in the early 1960s as a continuation of the quantitative advances in operations research and management science. Decision analysis combines systems analysis, which considers the interactions and dynamic behaviour of complex situations, and statistical decision theory, which focuses on logic in simple uncertain situations. Merging these two concepts into decision analysis provided a focus on logic in complex dynamic and uncertain situations.

Real options and decision analysis both have the common goal of modelling the decisions and uncertainties associated with investments. Where there is a distinction between the real options and decision analysis method is in the definition of valuing risky cash flows. Valuation in decision analysis is derived from the values and preferences of an individual or organization, whereas valuation in real options is derived from prices in traded markets. As value in real options is based on markets, risk-neutral probabilities and risk-free discount rates, the utility functions and risk adjustments to discount rates as used in decision analysis are unnecessary.

Dynamic programming was developed as an approach to the optimal control problem found in an area of economics called dynamic optimization. Optimal decisions problems, where current decisions influence future payoffs, can be solved using dynamic programming, and it is particularly useful when dealing with uncertainty. The method derives possible values of the underlying asset by extrapolating out over the duration of the option, and then, folding back the value of the optimal future value to the present. Dynamic programming can deal with complex decision structures that include constraints and complex relationships between the option value and the underlying asset. The binomial option pricing method is a form of dynamic programming.

Dynamic programming and contingent claims analysis are based on similar partial differential equations. There are also similarities in the way the Bellman equation used in dynamic programming is interpreted in terms of an asset value and to what degree investors are prepared to retain that asset. In contingent claims analysis, boundary conditions define where investors decide on the optimal exercise date that maximizes asset value. The main difference lies in the definition of the rate of return. Dynamic programming specifies the discount rate exogenously, and is therefore considered a subjective valuation of risk. In contingent claims analysis, the rate of return on an asset is derived from assets traded in financial markets.

The holder of a financial option has an exclusive right over exercising that option. The same, however, is not always the case in real options. Some real options will be exclusive or proprietary, and therefore, the holder of the real option will have sole exercise rights without the threat of competitors. Other investment opportunities however will have shared real options and may also be available to competitors or other potential participants. Other possible situations are where shared real options have no value as they collectively belong to a whole industry, or where they are a public good.

In option markets, the best strategy for the holder of a non-dividend paying American call option on a stock is typically to delay the exercise until the option maturity. There is no opportunity cost associated with waiting to exercise the option, and therefore, the holder of the call option would rationally wait as long as possible before exercising that option. If a stock does pay a dividend, however, its value will typically fall after the dividend payout, and so, reduce the payoff for a dividend-paying American call option if it is exercised immediately after the dividend payout.

There is, therefore, an associated opportunity cost in waiting to exercise if a stock option does pays a dividend, in which case, early exercise would be a better strategy. In a similar sense, if there were no opportunity costs associated with delaying an investment, the holder of a real option would wait until its maturity before exercising. In circumstances where competitors can enter a market however, the real option holder would forgo any potential value from waiting to exercise so as to pre-empt competitors. Competitors entering a market can reduce the value of the cash flows from an investment made in that market, and therefore, the value of any investment opportunities.

While there are many issues associated with identifying and valuing real options, in the final analysis, the critical issue is to be able to think in terms of real options. Projects and investments can be conceptualized as portfolios of assets that have opportunities, option portfolios that can be managed dynamically as the future unfolds, uncertainty is resolved and new information becomes available. Real options analysis draws on a range of techniques that include market values, quantitative methods, and also, qualitative assessment. Even if objective market based valuations are not always obtainable, a qualitative interpretation of real options is essential, as a real options framework provides management with a structure for decisions that have to be made in any case.

2.3.6 Types of Real Options

Real options can exist in almost every business decision, although they are not always easily identified. Many types of real options have been recognized and analysed (Dixit and Pindyck (1994) and Trigeorgis (1996)), and the following is a summary of common categories:

Option to Defer The opportunity to invest can be more valuable than investing immediately, as it provides management with the flexibility to defer the investment until conditions become more favourable, or to cancel completely if they become unfavourable. The opportunity to defer is the equivalent to a call option on the value of a project. These investment opportunities can still be beneficial even though the investment may have a negative NPV.

Option to Expand or Contract Options can exist in projects and operations to expand, to contract, and to shut down and restart. Management can expand production or increase resource deployment if the market environment develops more favourably than expected. This is the equivalent to a call option. On the other hand, operational scale can be reduced if market developments are less than initial expectations, which is the equivalent to a put option. The option to expand is specified as:

$$\text{Payoff} = \max\left[S_t - K, (1+x)S_t - K^*\right] \tag{2.2}$$

The option to contract is specified as:

$$\text{Payoff} = \max\left[S_t - K, (1+y)S_t - K^*\right] \tag{2.3}$$

Finally, the option to temporarily shut down is specified as:

(a) temporarily shut down operations:

$$\text{Payoff} = \max\left(S_t - VC, C\right) \tag{2.4}$$

(b) restart temporarily closed operations:

$$\text{Payoff} = \max\left(0, S_t - VC\right) \tag{2.5}$$

where,

S_t = initial underlying value
K = investment cost at t
K^* = K plus the increase (expand) or decrease (contract) in the investment cost at t

x = is the percentage increase in firm value
y = reduction in firm value
S = the project value
VC = variable costs
C = cash payout

Option to Abandon Management can abandon an operation if market conditions deteriorate, and liquidate any capital and other assets. The option to abandon is the equivalent to a put option. If the value of the asset or project falls below its liquidation value, the owners or holder of the option can exercise the put.

Option to Switch Management can change a project or an operation by restarting an operation that has been shut down, the equivalent to a call option; or shut the operation down, the equivalent to a put option. The cost of starting up or shutting down is the equivalent to the strike of the call or put. The option to switch is specified as:

$$\text{Payoff} = \left(S_1 - S_2 - CS, 0\right) \tag{2.6}$$

where,

S_1 = the NPV of the current operating mode one,
S_2 = the NPV of the current operating mode two,
CS = the cost of switching from the first to the second mode.

Growth Options Investments such as research and development, undeveloped land, oil and gas reserves, acquisitions and information networks can connect a chain of interrelated projects and create future growth opportunities, such as new products or processes and new markets.

Compound Options Projects frequently involve a collection of options, with combinations of upside value and downside protection present. The combined value of interacting options can differ from the sum of the separate parts due to their interaction. Some real options are relatively simple as their value, if exercised, is limited to the value of the underlying project. Other real options, however, can lead to further investment opportunities when exercised. These are options on options, or compound options, where the option payoff is another option.

Rainbow Options These are options that have multiple sources of uncertainty. Options that have payoffs that depend on two or more assets are called rainbow options. In the financial world, rainbow options can refer to the maximum or minimum of two or more assets, or other options, for example, where the payoff depends on the spread between two assets, the better of two assets and cash, portfolio options and dual strike options. In the case of real options, numerous sources of uncertainty can exist in the form of prices, quantities, technologies, regulation and interest rates.

2.4 Corporate Finance

2.4.1 Overview

Corporate finance has the objective of optimizing firm value while minimizing the associated risks. This encompasses the management of real assets that create firm value, minimizing the costs associated with the financing of these investments, and maintaining the firm's working capital.

Corporate finance is also framed within short- and long-term domains. The short-term domain focuses on a firm's working capital, defined as the net of current assets and current liabilities, and includes cash management, inventory and short-term lending and borrowing. The goal of working capital management is to optimize a firm's liquid assets. The long-term domain focuses on the capital investment decisions that involve a firm's fixed assets and capital structure. These decisions involve capital expenditure, the balance sheet debt and equity financing choices and dividend decisions. Capital investment decisions consist of an investment, a financing and a dividend decision, and are usually framed with the goal of maximizing firm value by investing in projects with a positive NPV.

Firm value is equivalent to the firm's total capitalization, which is equivalent to the market value aggregate of the firm's equity, bonds and any other claims, or the present value of all the claims on the firm. The value of the firm is therefore the present value of all free cash flows created from the firm's business model that are available to claimants on the firm. The concepts behind the analysis of real asset investments are equivalent for either the value for specific projects, or the firm itself, as the firm represents a collection of projects.

Capital structure is defined as the way in which a firm finances its balance sheet through the weighting of equity, debt and other security types. A firm's leverage is the ratio of firm debt to total financing. The goal of defining a firm's capital structure is to finance the assets so as to maximize firm value.

2.4.2 Theories of Firm Value

Initial theories of firm value were proposed by Miller and Modigliani, who examined the associations between a firm's operations in the real economy and its financing decisions in the financial economy. Miller and Modigliani showed that under an assumption of no taxes, firm value is the same, regardless of whether it is financed through equity or debt. The only impact the type of financing has is on the distribution of a firm's value between its investor types.

Miller and Modigliani also suggested that establishing firm value enabled the valuation of the firm's stock, bonds and other claims on the firm. A firm is represented by the present value of the firm's free cash flows discounted at a risk-adjusted interest rate, with the assumption that financing, the ratio of equity and debt, had no influence on the firm's operating cash flows and therefore on firm value. Once the value of the firm is established, the market value of debt is deducted to arrive at the firm's capitalization.

A firm's capital structure defines the manner in which it finances its assets and structures its liabilities, which include equity, debt and other claims. The Miller and Modigliani theory provides a foundation for the analysis of capital structure. Using the assumptions of perfect markets, no taxes, a universal borrowing interest rate, no bankruptcy or transaction costs and financing decisions not affecting investments, Modigliani and Miller drew two conclusions on capital structure. The first, defined as their first proposition, was that a firm's value is not influenced by capital structure. Their second proposition was that a leveraged firm's cost of equity is the same as that of a firm with no leverage. Miller and Modigliani later revised some of the assumptions—in particular, in regards to taxes.

The Modigliani–Miller theory provides a framework to examine how a firm's value is influenced by capital structure decisions and determining optimal capital structures. The Modigliani–Miller representation is defined as the primitive firm, for which its value is represented by the sum of the expected free cash flows discounted by the weighted average cost of capital. The primitive firm represents the DCF model of the firm, and provides a foundation for the analysis of a firm's financial structure through the financial options on the primitive firm.

Black, Scholes and Merton were the first to formalize the association between a firm's equity and debt. The insight was that equity can be defined as an option on a firm's assets, with the value of debt being equivalent to the residual of the value of assets over the value of equity. The Black and Scholes theory of the firm considered equity as a call option, with a strike equal to the notional value of zero coupon debt, on the value of the primitive firm. Merton also considered equity as an option on a firm's assets to define the firm's debt value and credit risk. The model developed by Merton uses the value and volatility of the firm's assets and the notional value of debt.

Geske extended the Black and Scholes model by specifying a call option on the stock, which itself is an option of the firm's assets, or the equivalent to a compound option. Whereas the Black and Scholes model assumes that the volatility of a stock price is constant, the Geske model recognizes that volatility is not constant. The compound option model identifies volatility as a function of the level of the firm's stock price—or more fundamentally, on firm value. To achieve this, the Geske model adds an additional variable—the firm's notional debt—to the Black and Scholes model, as it is financial leverage that influences the volatility or risk of a firm's equity.

The next development in defining firm value focused on the actual firm as the underlying asset. Myers proposed that a firm's investments can be represented as options. Firm value had been defined as the primitive model, or a pool of projects that represents the present value of free cash flows. Investing in product markets can, however, produce cash flows from an initial investment, and value from growth options if a product market continues to expand. A firm's initial investments therefore provide a base for a sequence of potential follow-on investment decisions.

The identification of this time series of investments is an extension of the primitive firm, and these discretionary future investments were defined by Myers as real options, or options on real assets. Real options identify the investment decisions within a firm as a right without an obligation, or as state contingent decisions on real assets. A firm has the choice in the future whether or not to exercise the option on follow-on investments.

2.4.3 Developments in the Theories of Firm Value

2.4.3.1 Overview

The theories of the primitive firm, financial options and real options can be combined in a value framework that provides the flexibility required by a firm to adapt to its external environment. Copeland (2007) defines the combination as a 'three layer cake', with the primitive firm as the foundation, real options consisting of a portfolio of growth options and a firm abandonment option, and a portfolio of financial options. The three-layer framework identifies the relationships between a firm's real and financial options. A firm has both an optimal real options investment structure and an optimal capital structure, with a trade-off between the two. The firm's operating and financial decisions, therefore, are not unrelated as per the Modigliani and Miller theory.

The following illustrates the components of the three layers which, when combined, offers a framework to manage firm value in a dynamic environment where both operating and financial flexibility are essential.

2.4.3.2 Primitive Firm Valuation

The primitive firm is defined as an underlying security that represents the firm's business risks. The firm's value, market capitalization, debt and other claims are defined as contingent claims on this underlying security—the primitive firm itself. The valuation of the underlying security is represented by the expected free cash flows to the firm, $E(\mathrm{FCF}_t)$ discounted at the cost of capital, w, with the assumption that systematic risk is the only risk factor. It is also assumed that the firm has no debt or other claims, is financed only with equity, and pays no taxes to segregate the tax issues from business risks. The value of the primitive firm, V is equal to V_0, the expected present value at $t = 0$:

$$E\left(V_t\right) = \sum_{t=1}^{N} \frac{E\left(\mathrm{FCF}_{i+1}\right)}{\left(1+w\right)^{i-t}} t = 0,1,\ldots.N. \tag{2.7}$$

The assumptions underlying the primitive firm are naïve, as firms will delay investments until uncertainty is reduced and also divest. These investment alternatives can be reduced to growth and expansion, defined as a European call option, or abandonment, or the equivalent to an American put option.

2.4.3.3 Growth Options

Given that it is possible to define the notion of a capital structure, it is also possible to define a firm's investment structure that includes its real options portfolio. This investment structure can be optimized to provide a firm the flexibility to adapt to its environment, and includes growth options, an abandonment option and a trade-off between scale and modularity.

The discretion a firm has on exercising its future investment opportunities is identified as call options on real assets. These options are growth options, and can be defined as a sequence of growth opportunities embedded in a firm's investments that have an impact on a firm's value.

Refer to the case study in Chap. 10 for an illustration of growth options.

2.4.3.4 Modularity

A firm's operational capacity is also a fundamental component of its investment structure. The firm has the option to expand capacity to meet increased demand, or not to expand if it has excess capacity. A firm's operating leverage is defined as the ratio of its fixed to variable costs. A high operating leverage is associated with less flexibility to adapt to change. Flexibility is therefore a function of a firm's operating leverage, and the capability to invest in modules has an impact on the firm's operating leverage.

The term 'modularity' is defined as a specific design structure where, within each unit or module, the parameters and tasks are mutually dependent, whereas across each module, they are independent. Modularity is a concept that can define a firm's operating leverage, or the degree to which the firm lacks operational flexibility due to its fixed costs. A firm's excess capacity is the variation between the firm's output capacity and expected output. The trade-off between modularity and economies of scale has an impact on a firm's excess capacity, and therefore, its investment and capital structure. Firm value can therefore be optimized through its investments in growth, abandonment and modularity.

Refer to the Appendix for an overview on modularity.

2.4.3.5 The Abandonment Option

A firm has value in the decision to discontinue operations and liquidate the firm's assets. This liquidation value is the equivalent to an option to abandon the firm to the firm's investors, with the exit value equal to the firm's total collateral or the total cash proceeds from liquidation. This option is an American put option, with the value of the option increasing as the value of the firm's total collateral or exit value increases.

Refer to the case study in Chap. 11 for an illustration of the abandonment option.

2.4.3.6 Financial Options

A firm's capital structure is defined by the financial options on the primitive firm as in the theories of Merton, Black and Scholes and Geske. The only decision variable in this case is the firm's debt policy, a choice that does not have an effect on primitive firm value. Under these assumptions, the same value will be generated by a marginal investment, no matter what debt policy is selected.

2.5 Optimizing the Firm Structure

An optimal firm structure that provides the flexibility and capabilities to adapt to its external environment consists of the primitive firm, an optimal investment structure, modularity and an optimal capital structure. The optimal investment structure includes the real options in the European call option growth portfolio and the American put option to abandon the firm. A firm therefore has three layers that consist of the primitive firm, the real options portfolio and a capital structure consisting of debt and equity that is represented by financial options.

The variables of interest in an investment policy are:

- Capital structure, which is influenced by the trade-off between a firm's real options portfolio, debt ratio, and the tax benefits of leverage. Copland presents a model for an optimal capital structure using the variables described here.
- Modularity, which has a function in establishing a firm's investment structure. Flexibility in a firm's capacity provides the ability to adapt to changing markets and industries.
- Volatility, which creates value in real options along with a firm's ability to exercise these options. Volatility and flexibility also influence a firm's level of cash.

The variables of interest in an optimal financial structure are:

- The abandonment or collateral value, which has a positive relationship with firm value.
- A firm's debt policy, or leverage, which has an influence on growth and abandonment.
- Taxes, which will raise the value of the firm's equity if the balance sheet has debt that has tax benefits. This upside will however start to roll off or fall beyond a leverage threshold, and ultimately goes to zero if the firm is abandoned.
- Cash and cash management, which are related to flexibility. Firms with large cash reserves are able to react quickly to market and industry conditions, and exercise real options when compared to firms with higher debt

ratios and external funding requirements. Growth through expansion and abandonment are both influenced by the level of a firm's debt, or its leverage.
- Volatility, which also has an influence on the firm's financial option values. Increases in volatility will influence the firm's debt costs, and therefore the firm's capital structure.

A firm's investment and capital structure is ultimately a function of the industry in which it operates, its external environment and the firm's strategy. The interactions between a firm's investment structure, modularity versus economies of scale and capital structure create trade-offs in framing the firm's structure. Leverage will have an impact on the firm's equity when viewed as a call option on the firm's assets—and therefore, its growth and abandonment options. The use of tax to optimize a firm's capital structure will influence the flexibility of its operations. A firm's overall structure should therefore consider a range of variables and trade-offs when defining its capabilities to adapt to its environment.

Appendix: Modularity

The waves of innovation since the start of the industrial evolution have created an economic system that is increasingly sophisticated and complex. This economic system consists of objects that result from human intelligence and endeavours. These objects, or artefacts, include physical activities such as technologies and products, and intangible objects such as systems of law, organizations, strategies, science and designs.

Artefacts develop and evolve over time, as do the firms and markets that create and support these objects. These markets, technologies, products and firms evolve interactively to produce adaptive complex systems that ultimately become industries. An artefact is described by its design, and this designing of artefacts is a continuing process that accumulates at all levels to transform industries and economies.

Modularity is a theory of complex adaptive systems, design and industrial evolution that describes the creation of complex products and processes from smaller subsystems or modules that, although are independent in their design, nonetheless function together as a complete system. The modularity concept facilitates the management of complex systems by dividing them into smaller, more manageable components. This is achieved by creating a particular design structure with a set of design principles that separate the knowledge and tasks required for complex designs and artefacts.

The modularity concept has numerous applications, which include production scale and scope, mass customization and organization theory. An example can be found in computers, an artefact that has grown in complexity over the twentieth century. In the 1970s, the computer business evolved from a highly concentrated industry to modular clusters that manufactured components of larger computers systems. New designs created the opportunity for the emergence of new firms, which focused on manufacturing specialized components, or modules, that were linked by design rules for the creation of computer systems.

At the start of a modular design process, mandatory design rules are established for all stages of design and production. These design rules allow pieces of a modular system to be changed without the need to change the system as a whole. This capability creates the flexibility for the design to evolve at the module level, and therefore creates options for designers. These options provide opportunities for innovation and capabilities for firms to complete in today's environment.

In the twenty-first century the dynamics of global commerce will continue. The new technologies, markets, products and competitors that emerge from this process present both risk and rewards. In this context, modularity can address three issues—it increases the ability to manage complexity, it facilitates various components of a design to be worked on simultaneously, and it accommodates uncertainty. Modularity therefore offers the capabilities to manage the complexities and uncertainties in this environment, and provides a framework for creating value, growth and innovation.

Bibiography

Amram, M, and Kulatilaka, N. Strategy and Shareholder Value Creation, the Real Option Frontier, Harvard Business School Press, 1999.

Baldwin, C. and Clark, K. Design Rules, MIT, 2000.

Benninga, S. and Sarig, O. Corporate Finance, A Valuation Approach, McGraw-Hill, 1997.

Black, F. and Scholes, M. The pricing of options and corporate liabilities, Journal of Political Economy, 1973.

Brealey, R and Myers, S. Principles of Corporate Finance, McGraw Hill, 1996.

Briys, E, Bellalah, M, Minh Ma, i H, and De Varenne, F. Options, Futures and Exotic Derivatives, Wiley, 1998.

Clayman, M., Fridson, M. and Troughton, G. Corporate Finance – A Practical Approach, J. Wiley & Sons, 2008.

Copeland T and Antikarov V, Real Options, Texere, 2001.

Copeland, T. The Firm as a Three-Layer Cake – Optimal Investment and Financial Structure, MIT Sloan School of Business, 2007.
Cuthbertson K, Nitzsche D, Financial Engineering: Derivatives and Risk Management, Wiley 2001.
Damodaran, A. Applied Corporate Finance, Wiley, 1999.
Damodaran, A. Investment Valuation, Wiley, 1996.
Dixit, A, and Pindyck, R. Investment under uncertainty, Princeton University Press, 1994.
Ferris, K. and Pecherot Petitt, B. Valuation: Avoiding the Winner's Curse, Prentice Hall, 2002.
Geske, R. and Zhou, Y. Capital Structure Effects on Prices of Firm Stock Options: Tests Using Implied Market Values of Corporate Debt, November 2007, revised January 2009.
Geske, R. The Valuation of Compound Options, Journal of Financial Economics, 1978.
Higgins, R. Analysis for Financial Management, McGraw-Hill Irwin, 2007.
Ho, T. and Lee, S. The Oxford Guide to Financial Modeling, Oxford, 2004.
Hooke J C, Security Analysis on Wall Street, Wiley, 1998.
Howard R A & Matheson J E, The Principles and Applications of Decision Analysis, SDG, 1989.
Jiang, B, and Koller, T. Data focus: a long-term look at ROIC, McKinsey Quarterly 2006 Number 1.
Luehrman, T. Strategy as a portfolio of Real Options, Harvard Business Review, September–October 1998a.
Luehrman, T. Investment Opportunities as Real Options: Getting Started on the Numbers, Harvard Business Review, July–August 1998b.
M. I. Leone and R. Oriani, "Explaining the Remuneration Structure of Patent Licenses", 14th Annual International Real Options Conference, LUISS Business School, Rome, Italy, June 16–19, 2010.
Magretta, J. What Management Is, The Free Press, 2002.
Merton, R. Applications of Option-Pricing Theory: Twenty Five Years Later, The American Economic Review, June 1998.
Merton, R. On the Pricing of Corporate Debt: The Risk Structure of Interest Rates, Journal of Finance, 1973.
Miller, M, Modigliani, F. Dividend Policy, Growth and the Valuation of Shares, Journal of Business, 1961.
Morin, R, and Jarrell, S. Driving Shareholder Value, McGraw Hill, 2000.
Myers, S. Financial Theory and Financial Strategy, Midland Corporate Finance Journal, 1987.
Myers, S. The Determinants of Corporate Borrowing, Journal of Financial Economics, 1977.
Penman, S. Accounting for Intangible Assets: There is Also an Income Statement, Abacus, Vol. 45, No. 3, 2009a.
Penman, S. Accounting For Value, Columbia University Press, 2010.

Penman, S. Financial Statement Analysis and Security Valuation, McGraw Hill/ Irwin, 2004.
Penman, S. Accounting for Intangible Assets: There is also an Income Statement. CEASA Center for Excellence in Accounting & Security Analysis, June 2009b.
Seitz, N and Ellison, M. Capital Budgeting and Long-Term Financing Decisions, The Dryden Press, 1995.
Sharpe, W, Alexander, G, and Bailey, J. Investments, Prentice Hall, 1999.
Smith, J. Much Ado About Options? Fuqua School of Business, Duke University, July 1999.
Stowe J, Robinson T, Pinto, McLeavey W. Equity Asset Valuation, Wiley 2002/2007.
The Economist, Corporate finance, 27 January 2001.
Trigeorgis, L. Real Options: Managerial Flexibility And Strategy In Resource Allocation, MIT Press 1996.

3

Risk

3.1 Investment Risk

Corporate management will typically develop strategies and allocate resources to increase shareholder value. Shareholders, on the other hand, will focus on the cash growth of their investments. As to whether there is value in any potential cash flow growth will depend on the risks associated with these investments. Investors will generally demand a higher rate of return from investments that are perceived to be relatively riskier. The risks associated with corporate investments are found in variables such as prices, quantities, costs, competition, market share and project lifecycles. These variables can be unpredictable and result in cash flow volatility, which will therefore have an impact on any net present value (NPV) calculations.

The risk appetite of investors should therefore be included in investment analysis. As a result, the cash flows used in NPV estimates are modified by a discount rate that reflects both the time value of money and any related cash flow risks. Risk is represented in the discount rate determined from the risk appetites of investors and the financial markets. The hurdle rate is the discount rate that is the minimum acceptable rate of return that a firm will accept for a project. A number of methods have been developed to determine the discount rate used in NPV calculations. These include the Capital Asset Pricing Model (CAPM) and Risk-Adjusted Discount approach.

The foundations of quantifying risk can be found in Markowitz (1959), regarded as the origin of modern portfolio theory. Markowitz's solution was to assume that a portfolio could be structured as a function of the mean and

J. Rogers, *Strategy, Value and Risk*, Global Financial Markets,
https://doi.org/10.1007/978-3-030-21978-9_3

the standard deviation of the portfolio return. His conclusions from this construct were that the risk of the portfolio will generally be less than the weighted average of the separate asset risks, and the lower the correlations between the component asset returns, the lower the portfolio risk (the diversification principle). Each asset's risk consists of two components—the diversifiable risk, which will disappear through the right combination of assets, and the non-diversifiable risk, which will always be carried by investors. The portfolio selection problem can therefore be defined as consisting of maximizing the return while minimizing the risk.

Sharpe, Lintner and Mossin extended Markowitz's portfolio theory with the assumption of homogeneous expectations, where all investors agree on expected returns, standard deviations and correlations, and therefore, choose the same portfolio. This concept led to the Capital Asset Pricing Model. The CAPM is a general model for formulating an asset's risk and return. The variance of the return is defined as the risk measure, with only the non-diversifiable or systematic component of the variance being rewarded. The relevant risk in pricing an asset is that part of an asset's risk, or variance of the assets return, that is correlated with the overall risk of a market, and not the overall risk of the asset. An asset's beta coefficient measures its systematic risk.

The CAPM model can be used to illustrate those businesses in which a firm should have operations. Diversifying a firm over a portfolio of independent businesses decreases the variance of the combined cash flows if the cash flows of the various businesses are not completely correlated. Reducing risk would therefore be consistent from the perspective of shareholders as they are typically risk adverse.

Shareholders can, however, reduce risk by maintaining diversified portfolios themselves. Individual shareholders can accomplish a broader diversification, typically at lower transaction costs, than that offered by the majority of firms diversifying through mergers and acquisitions. Creating shareholder value through diversification requires the existence of market imperfections that firms can exploit more efficiently than investors. As a firm can be described as a portfolio of assets and projects, the value of a new project would be conditional on the total risk of the firm, in which case, NPVs would not be additive. If, however, a project's risks are not correlated to existing assets or projects, NPVs will generally then be additive.

As the value of a firm can be considered as the value of all the firm's assets, the firm can therefore also be viewed as the value of all sources of financing. The weighted average cost of capital is a discount rate that represents the costs of the various sources of finance, which can consist of a firm's equity, debt and any hybrid securities. The cost of equity can be derived using a CAPM model, the cost of debt from interest rates and bond ratings, and the cost of hybrid securities through the characteristics of each of their components. The weighted cost of capital can be used in project analysis decisions to determine which project NPVs do not change a firm's business risk, and also provide a hurdle rate for projects.

The objective of the Risk-Adjusted Discount Rate (RADR) method is to maximize a firm's market value by using discount rates from investments that have the similar risk characteristics as the investment projects under analysis. The NPV derived by discounting future cash flows at the RADR reflects the opportunity cost of capital, or the rate of return required by the firm or investors for similar investments. The RADR includes the time value of money, the risk-free rate and the discount risk premium. Conventional projects that have similar risk characteristics as existing businesses should not influence the aggregate risk of a firm, and would therefore be discounted at the opportunity cost of capital. The discount risk premium is adjusted upwards for projects with above average risk, and down for projects with below average risk. The return of an investment project can also be compared to a hurdle rate to determine whether the project should proceed.

3.2 Corporate Risk Management

Business risks are those that a firm assumes to create a competitive advantage and add value. The motivation for firms to better understand and measure risk is being driven by:

- the increasing awareness that earnings volatility can significantly affect stock valuations and shareholder value
- the increasing size and types of interrelated risk exposures firms are facing that include globalization, changing markets and industry dynamics, and
- organizational requirements for improved exposure and risk-related information to define management's risk appetite and improve decision making

Business risk management is a process where risk exposures are identified, measured and managed, where possible, within the context of strategy and corporate finance, and is essentially a core competency of all business activities. The focus is moving from individual price exposures to a firm's exposure as a portfolio of interrelated risks.

An effective risk management framework can address issues such as:

- using more transparent risk management methods to manage the external factors that can influence a firm's performance
- translating risk management practices to analysts, investors and rating agencies
- evaluating the potential impact of adverse market movements on a firm's capital
- defining risk and return targets for businesses and projects
- the use of risk adjusted measures that can influence management decisions, and,
- whether the rewards are adequate for a given level of performance

Financial theory defines risk as a dispersion of unexpected outcomes due to movements in market or risk factors, where positive and negative deviations are viewed as sources of risk. Changes can be expressed as either absolute or relative returns, and probabilities can be derived for the distributions of these returns. Risk can therefore be evaluated and measured in a probability context, where risk is conceptualized as the probability that an event will occur. Measures of risk can now be defined as the volatility of unexpected outcomes, such as the variance or volatility of an asset's returns.

Two risk measures—Value At Risk (VAR) and Cashflow At Risk (CFAR)—are based on volatility. These measures of market risk use probabilities to interpret risk exposures as a potential loss. VAR summarizes the expected maximum loss over a target horizon within some confidence interval. VAR however is not always a suitable risk measure for many firms, as it focuses on the potential loss in the market value of assets and liabilities over a short horizon. Many firms have physical assets, brand names and intangible assets such as capitalized research and development, for which market or liquidation values are only relevant for a small portion of the balance sheet. An alternative risk measure in these situations is CFAR, where an aggregate risk exposure is derived from the variability of projected cash inflows and outflows over a multi-year planning horizon.

Financial risk management encompasses the use of derivatives to manage foreign exchange, interest rate, credit and commodities risk. Derivative techniques can identify and measure corporate risk exposures, and provide a basis to build a framework that integrates financial risk management with strategy. Applications include identifying natural hedges, mitigating risk exposures, and risk reduction in mergers, acquisitions, privatizations and financing.

3.3 The Risk Drivers

Risk factors are any market price, value or index that can have an influence on a firm's cash flows. Event definitions of risk differentiate risk types by the nature of the event that can cause a loss. A number of factors, or risk drivers, that account for the volatility in value can generally be identified and analysed within a framework that includes:

- Financial risk: Financial risks are generally identified and classified as:

  - Market risk: the changes in prices of financial assets and liabilities
  - Basis risk: the price difference between the forward and spot price, or the potential for loss resulting from a hedge and the instrument being hedged not being perfectly matched (correlation risk); and includes:
    - Spread risk: the risk relative to a particular group of securities
    - Curve (or shape) risk: changes in the shape of yield and forward curves
  - Volatility risk: sensitivity with respect to the volatility of securities (typical of options)
  - Currency risk: relates to adverse changes in currency rates
  - Credit risk: possible changes in credit ratings, outright defaults or counterparties unwilling or unable to fulfil their contractual obligations
  - Liquidity risk: has two forms; when a transaction cannot be conducted at prevailing market prices due to insufficient market activity; and the inability to meet cash obligations

- Commodity risk: occurs when an organization is affected by fluctuations in the price of some commodity. A wide range of physical assets are considered as commodities:

  - Metals, such as gold and copper
  - Agricultural products, such as wheat, timber, and wool
  - Energy products, such as oil and gas

Commodities and energies are increasingly traded like financial instruments. The underlying price drivers, however, are fundamentally different from those found in financial assets. The dynamics of production and use, transport and storage, buying and selling, and advances in technology all add to the complexity of energy and commodity markets.

- Legal risk: is the loss from an organization's activities judged to be outside the relevant legal and/or regulatory framework governing such activities, and includes but is not limited to the enforceability of contracts. Legal risk includes disclosers, disclaimers, compliance and regulatory risks, and can take the form of shareholder lawsuits against firms that suffer large losses.
- Operational risk: the Group of Thirty defines operational risk as 'the risk of losses occurring as a result of inadequate systems and control, human error, or management failure.' This definition includes fraud and regulatory risk. Disaster recovery, a contingency plan to cope with a disaster, is also included in operational risk.
- Strategic risk: corporate strategy is a firm's pattern of decisions that determines its objectives, purposes or goals. Strategic planning processes include the identification of areas susceptible to changes within the firm's environment that can affect its future. Strategy includes identifying opportunities and threats in the firm's environment, and attaching some estimate of risk to alternatives.
- Technology risk: technology can lower operating costs, increase value and capture new markets. Technology risk occurs when the investments do not produce the anticipated cost savings in economies of scale (lower average costs of operations by expanding output), or scope (generate cost synergies through producing more than one output with the same inputs). Technology risk can result in major losses in competitive efficiency and can ultimately result in long-term failure.
- Product risk: can take a number of forms. Products go through a life cycle, growing in sales, declining and eventually being replaced. The length of the product cycle and actual product failure are examples.
- Political risk: arises from actions taken by policymakers that can significantly influence a firm's business operations.

3.4 Value and Risk

Firms manage sustainable competitive advantage by selecting markets that match its capabilities and abandoning markets in which it has a competitive disadvantage. Value is created through the management of a firm's strategic portfolio and its real options portfolio, where real options that have value are identified and exercised and those that do not are abandoned. A firm's capabilities will therefore include managing both its strategic portfolio and its real options portfolio.

Firm value is a stock at a point in time, representing the present value of the firm's future cash flows and its real options. Any one or combination of risk factors can impact on value, and therefore, what is at risk is the likelihood that a firm will be unable to maintain the creation of value. A firm's real options can therefore be viewed as either sources of risk or sources of opportunity. A key distinguishing characteristic will be found in those firms with a risk management process that aims at value enhancement, where risk exposures are identified and managed in the context of strategy, investments and revenue optimization as opposed to risk control. While a firm's real assets are a significant component of its risk profile, its real options also contribute to value. A firm can therefore enhance its capabilities by integrating real options analysis into its strategic management, corporate finance and risk management processes.

Bibliography

Barbera, M. and Coyte, R. Shareholder Value Demystified, UNSW Press, 1999.
Beckers S, A Survey or Risk Measurement Theory and Practice, The Handbook of Risk Management and Analysis, Carol Alexander, Editor, Wiley 1996.
Birkett W, Value Creation and Risk Management In Financial Services, Towards a Taxonomy of Risk, UNSW White Paper, 2001.
Brealey, R and Myers, S. Principles of Corporate Finance, McGraw Hill, 1996.
Culp C, The Risk Management Process, Wiley, 2001.
Damodaran A, Investment Valuation, Wiley, 1996.
Dembo R, Seeing Tomorrow: Rewriting the Rules of Risk, Wiley, 1998.
Grant R M, Contemporary Strategy Analysis, Blackwell, 1998.
Jorion P, Value At Risk, Irwin, 2000.
Lintner, John (1965). "The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets". Review of Economics and Statistics. 47 (1): 13–37.
Markowitz H, Portfolio Selection: Efficient Diversification of Investment. John Wiley and Sons, New York, 1959.
Mossin, Jan (1966). "Equilibrium in a Capital Asset Market". Econometrica. 34 (4): 768–783.
Sharpe, William F. (1964). "Capital asset prices: A theory of market equilibrium under conditions of risk". Journal of Finance. 19 (3): 425–442.
Shimko D, VAR for Corporates, Risk, September 1995
Treynor, Jack L. (8 August 1961). Market Value, Time, and Risk. no. 95–209. Unpublished manuscript.
Treynor, Jack L. (1962). Toward a Theory of Market Value of Risky Assets. Unpublished manuscript. A final version was published in 1999, in Asset Pricing and Portfolio Performance: Models, Strategy and Performance Metrics. Robert A. Korajczyk (editor) London: Risk Books, pp. 15–22.
Trigeorgis L, Real Options: Managerial Flexibility And Strategy In Resource Allocation, MIT Press 1996.

Part II

Quantitative Analytics

4

Financial Statistics

4.1 Time Series Analysis

The econometric analysis of economic, financial and business time series has become an integral part in the research and application of quantitative descriptions of the real world. A time series typically consists of a set of observations of some observational unit or variable, y, which is taken at equally spaced intervals over time (Harvey 1993). A time series can be considered from two aspects—analysis and modelling. The objective of a time series analysis is to identify and summarize its properties and describe its prominent characteristics. The analysis can be framed in either the time domain or the frequency domain. In the time domain, the focus is on the relationship between observations at various points in time, whereas in the frequency domain, the analysis focuses on the cyclical movements of a series.

Economic, business and financial time series will have at least one of the following key features:

- *Trends:* are one of the main features of many time series. Trends can have any number of attributes, such as upward or downward, with relatively different slopes, and linear, exponential or other functional forms.
- *Seasonality:* time series can often display a seasonal pattern. Seasonality is a cyclical pattern that occurs on a regular calendar basis.
- *Irregular observations:* there can be periods or samples within a time series that are inconsistent with other periods, and therefore the series is subject to regime changes.

J. Rogers, *Strategy, Value and Risk*, Global Financial Markets,
https://doi.org/10.1007/978-3-030-21978-9_4

- *Conditional heteroskedasticity:* is a time series condition where there is variation (as opposed to constancy) in the variance or volatility and patterns emerge in clusters, that is, high volatility is followed by high volatility, and low volatility is followed by low volatility.
- *Non-linearity:* generally, a time series can be described as non-linear when the impact of a shock to the series depends if it is positive or negative and is not proportional to its size.

A stochastic time series is generated by a stochastic process, that is, each value of y in a series is a random draw from a probability distribution. Inferences can be made about the probabilities of possible future values of the series by describing the characteristics of the series randomness. Much of the research in time series has focused on investigating the hypothesis as to whether a series is a random walk or reverts back to a trend after a shock. The simplest random walk process assumes that each successive change in y_t is drawn from a probability distribution with zero mean:

$$y_t = y_{t-1} + \varepsilon_t \tag{4.1}$$

where ε_t is an error term which has a zero mean and whose values are independent of each other. The price change $\Delta y_t = y_t - y_{t-1}$ is therefore the error ε_{t} and is independent of price changes.

The question of whether economic variables follow random walks or tend to revert back to a long-run trend after a shock is an important issue for modelling. Most financial models of futures, options and other instruments tied to an underlying asset are based on the assumption that the spot price follows a random walk. In some markets, however, the prices of such assets as energies and commodities are tied in the long run to their marginal production cost. Although the price of an energy or commodity may be subject to sharp short-run fluctuations, it typically tends to return to a mean level based on cost.

A number of methods exist to test hypotheses about the properties of a time series. One technique is to examine its *autocorrelation* properties. Time series can be characterized by a set of autocorrelations, which can provide insights into possible models to describe the time series. A *correlogram* displays the autocorrelation and partial autocorrelation functions up to the specified order of lags. These functions characterize the pattern of temporal dependence in time series data. Another method for testing the hypothesis that the process is a random walk against the alternative that it is stationary, that is, the stochastic process in fixed time, is the unit root test introduced by Dickey and Fuller. Formally stated, the simplest model tested is:

$$y_t = \phi y_{t-1} + \varepsilon_t \quad t = 2, \ldots T \tag{4.2}$$

where the null hypothesis is $\phi = 1$ and the alternative hypothesis is $\phi < 1$. The generalization of the test for a unit root is known as the augmented Dickey–Fuller (ADF) test (1979, 1981).

Most statistical tools are designed to model the conditional mean of a random variable. Autoregressive conditional heteroskedasticity (ARCH) models are specifically designed to model and forecast *conditional variances*. ARCH models were introduced by Engle (1982) and generalized as GARCH (generalized ARCH) by Bollerslev (1986). These models are widely used in econometrics, especially in financial time series analysis. The modelling of variance or volatility can be used, for example, in the analysis of the risk of holding an asset or in the valuation of an option. In a GARCH model, there are two separate specifications—one for the conditional mean and one for the conditional variance. The standard GARCH(1,1) specification is:

$$y_t = \alpha_0 + \sigma_t \varepsilon_t \tag{4.3}$$

$$\sigma_t^2 = \varpi + \alpha_1 \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \tag{4.4}$$

where y_t is the log return of a series, and the mean equation in (4.3) is written as a function of exogenous variables with an error term. As σt^2 is the one-period ahead forecast variance based on past information, it is called the conditional variance. The conditional variance equation specified in (4.4) is a function of three terms—the mean, news about volatility from the previous period, measured as the lag of the squared residual from the mean equation (the ARCH term), and last period's forecast variance (the GARCH term).

4.2 Regression Models

Regression analysis is a statistical tool that can identify the correlation between two or more variables as a causal relationship by formulating a hypothesis that a dependent variable is a function of one of more independent variables. Applications include the Capital Asset Pricing Model (CAPM), which represents the relationship between a financial asset's risk and return, and Factor Models, which use multiple explanatory variables for asset returns to decompose risk and return into observable and unobservable components.

4.3 Volatility

Volatility, defined as the annualized standard deviation of price returns, is one of the critical concepts in option pricing and risk management. A percentage is derived as:

$$r_t = \frac{S_t}{S_{t-1}} - 1 \tag{4.5}$$

where S_t is the spot price at time t. Price returns are typically calculated by taking the natural logarithms of the price ratios:

$$r_t = \ln\left(\frac{S_t}{S_{t-1}}\right) \tag{4.6}$$

which is an approximation of the percentage change. Log returns are usually used to calculate volatility, as the natural log of S_t/S_{t-1} is equivalent to the natural log of $1 + r$, which is approximately equal to r. Another advantage is the log of a product is equal to the sum of the logs, and therefore, a log return over a time period can be calculated as the sum of log returns for the sub-periods. Figure 4.1 illustrates the inflation adjusted S&P 500 index from January 4, 1950 to December 31, 2018, and Fig. 4.2, the S&P returns from February 4, 2015 to August 17, 2015.

Volatility, rather than standard deviations or variances, is used as a measure of uncertainty so that any comparisons of distributions are equivalent. Normalizing a price return's standard deviation into a volatility measure creates a consistent measure of magnitude of random behaviour, and therefore, facilitates the comparison of various markets and models. The volatility of a price process also measures the annualized distribution of price returns, whereas standard deviations can measure the width of any distribution. The probability of exceeding an option's exercise price increases as a result of the volatility of the underlying asset, which is why volatility increases the value of options. Typically, the greater the volatility associated with an underlying asset, the greater the value of an option on that asset.

Volatility can be estimated from historical data or implied from option market prices. If there is a reasonably liquid market for traded options, then the implied volatility can be derived through an iteration process using an analytical pricing formula, such as the Black–Scholes model, the option price and the

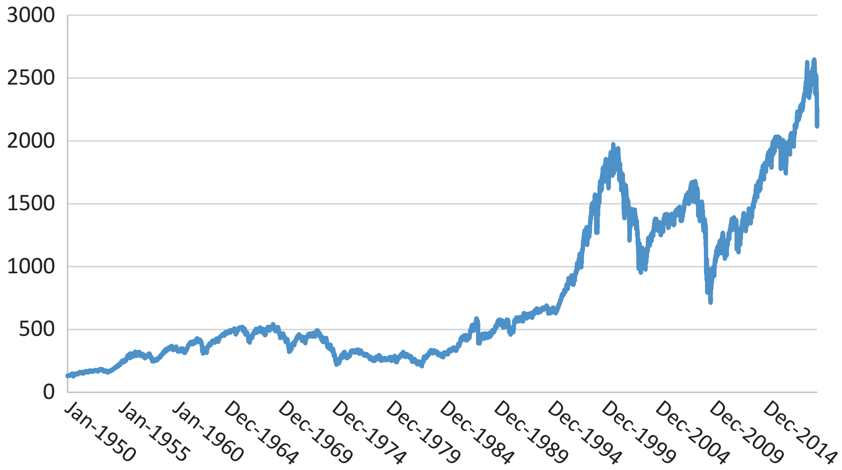

Fig. 4.1 S&P 500—January 4, 1950 to December 31, 2018, GDP inflation adjusted, 2012 = 100. US Bureau of Economic Analysis, Gross Domestic Product: Implicit Price Deflator [GDPDEF], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/GDPDEF, March 5, 2019. Note—the index contained 90 stocks up to 1957, and then, expanded to the current 500

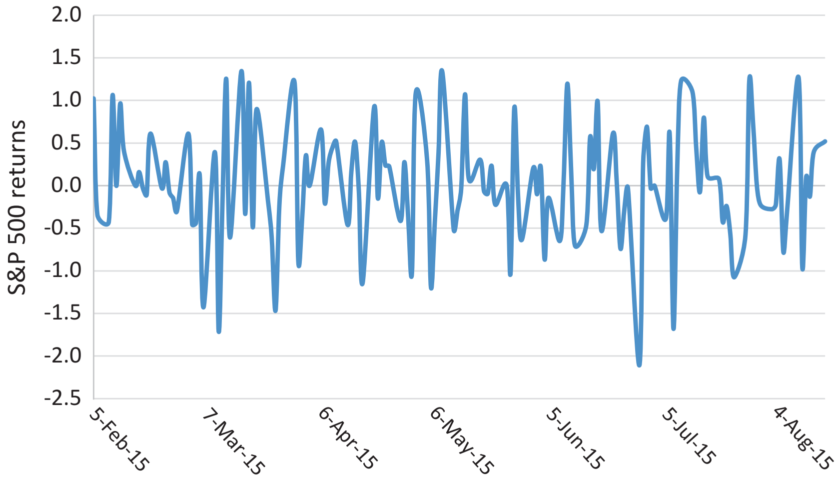

Fig. 4.2 S&P 500 returns (×100)—February 4, 2015 to August 17, 2015

known variables such as the interest rate, time to maturity and exercise price. The result is a forecast of the volatility implied in the quoted price of the option, with the forecast horizon being the maturity or expiry of the option. Volatility can also be derived from historical data by annualizing the standard deviation of the log returns through a scaling factor defined as the square root of time. The annualization factor depends on the price data frequency. If the data is monthly, the factor is √12, for weekly data, √52, and for the daily data for each calendar day, it is √365. If the data is available for trading days only, the relevant number may vary from √250 to √260, according to public holidays.

While volatility provides a comparative risk parameter, other test statistics can provide insights as to how well the assumptions capture the behaviour of a time series. The properties of a time series can be depicted by its descriptive statistics. The mean and standard deviation are descriptive measures of the properties of a time series. Other descriptive measures can be illustrated using a histogram, which displays the frequency distribution of a series. A histogram divides the range between the maximum and minimum values of a series into a number of equal length intervals or bins, and exhibits the number of observations within each bin. Figure 4.3 illustrates the histogram of the S&P 500 index log returns from February 2, 2015 to August 17, 2015, chosen as there was no trend within the sample.

The descriptive statistics of the S&P returns sample are:

- The *mean:* the average value of the series sample, derived by adding up the series sample and dividing by the number of observations.
- The *median:* a measure of central tendency, or the middle value (or average of the two middle values) of a series sample sequenced from the smallest to the largest. The median is a more robust measure of the centre of the distribution than the mean, as it is less sensitive to outliers.
- The *maximum* and *minimum* values of the series sample.
- The *standard deviation:* a measure of dispersion or spread in the series.

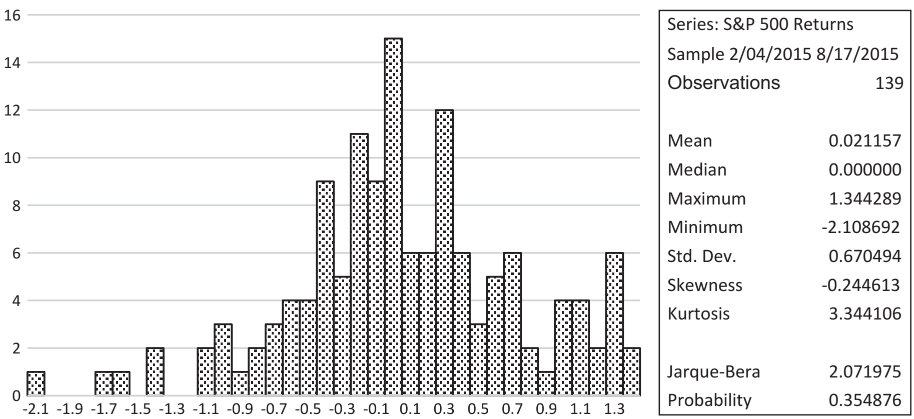

Fig. 4.3 Histogram of the S&P 500 returns (×100)—February 4, 2015 to August 17, 2015

- *Skewness:* a measure of the asymmetry of a series distribution around its mean. The skewness of the normal distribution, which is symmetric, is zero. Positive skewness implies that a distribution has a long right tail, while negative skewness indicates a long left tail.
- *Kurtosis:* measures the peakness or flatness in the distribution of a series. A normal distribution has a kurtosis of three. If the kurtosis exceeds three, the distribution is leptokurtic or relatively peaked compared to the normal distribution, while if the kurtosis is less than three, the distribution is platykurtic or relatively flat to the normal distribution.
- *Jarque-Bera:* a test statistic for testing whether the series approximates the normal distribution. The test statistic measures the differences in the skewness and kurtosis of the series with those from the normal distribution. The null hypothesis is that a series has a normal distribution.

The annualized volatility for the S&P 500 index sample period is 10.6%, which is 0.670494, the standard deviation multiplied by $\sqrt{250}$. The histogram in Fig. 4.3 also illustrates the presence of fat tails in the distribution of the S&P 500 index returns. Fat tails refers to the probability of extreme outcomes in an observed series exceeding the assumed theoretical probability distribution. Distributions displaying fat tails are described as leptokurtic and are measured by kurtosis, which in this case is 3.344106, and therefore greater than three. The skewness, which is zero in a normal distribution, is negative in this case, and is typical of many financial assets, such as stock prices.

4.4 The Lognormal Distribution

A variable has a *lognormal distribution* if the natural logarithm of the variable is normally distributed. Figures 4.4 and 4.5 illustrate the distributions of a simulated series and its natural log equivalent, respectively. A lognormal variable can have any value between zero and infinity. As a result, the lognormal distribution has a positive skew, and therefore, is unlike the normal distribution, as indicated by the skewness and kurtosis statistics. The log series, however, has a skewness close to zero and a kurtosis that is approximately three, and therefore can be described as being normally distributed.

The use of the log of financial variables is popular in derivative modelling as the price can never become negative, and the return is the relative change in the level of the log price. Figure 4.6 illustrates the distribution of the log returns of the simulated series. The returns can also be described as being normally distributed. The lognormal property of asset prices also can be used

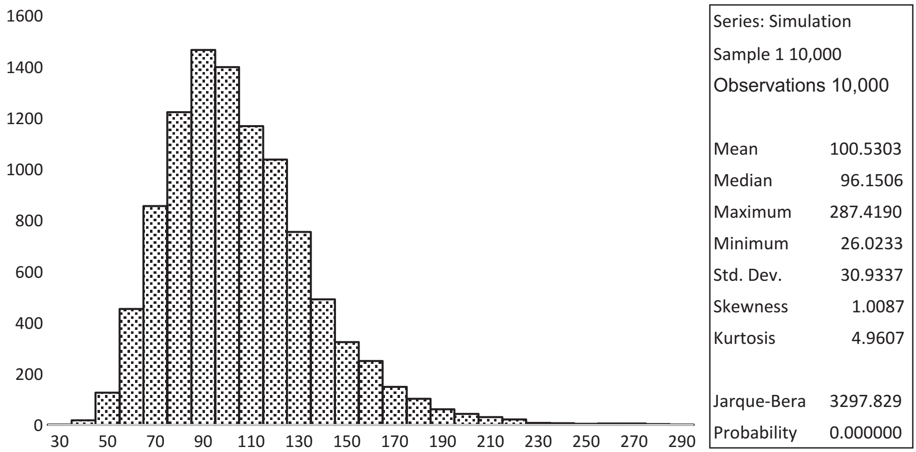

Fig. 4.4 The simulated series

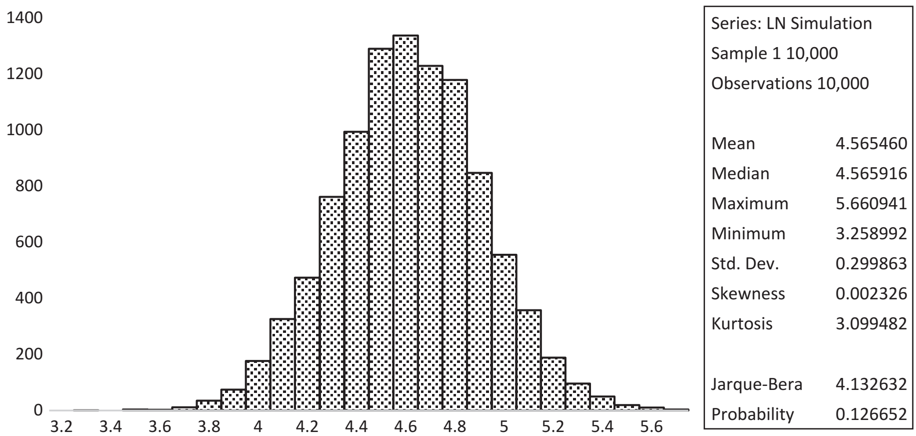

Fig. 4.5 The natural log of the simulated series

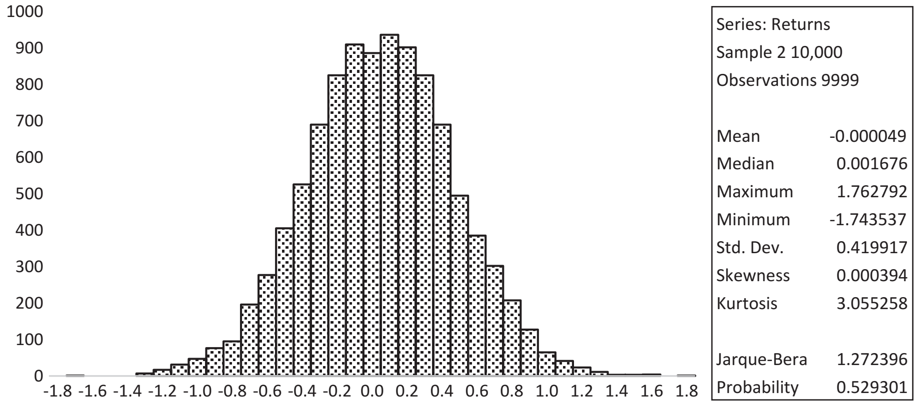

Fig. 4.6 The simulated log returns

to describe a price process and its probability distribution. If an asset price follows a geometric Brownian motion, then the natural log of an asset price follows a process called a generalized Weiner process. This implies that, given an asset's price today, the price at T has a lognormal distributed. The standard deviation of the logarithm of an asset is $\sigma\sqrt{T}$, that is, it is proportional to the square root of the length of time into the future. This stochastic process is the basis for the Black–Scholes option pricing model.

4.5 Volatility and the Firm

The volatility of a project, asset or firm is not necessarily the same as the volatility of one of its components. One example is the difference between the volatility of a firm's market value and the volatility of its equity. A firm's capital structure is the mixture of debt, equity and other liabilities that the firm uses to finance its assets. Merton (1974) defined the value of a firm's equity as a call option on the assets of the firm, where the strike is the book value of the firm's liabilities, and the underlying asset is the total value of the firm's assets. Merton's approach illustrated the link between the market value of the firm's assets and the market value of its equity, and provided a framework for determining the value of a firm's equity by reference to the underlying market value of the firm.

The analysis can be reversed to estimate a firm's value and volatility from the market value of its equity, the volatility of its equity and the book value of its liabilities. KMV (now a division of Moody's Analytics) extended Merton's approach to estimate probabilities of default for credit analysis. If the market price of equity is available, the market value and volatility of assets can be determined directly using an options pricing-based approach, which recognizes equity as a call option on the underlying assets of the firm. The limited liability of equity provides equity holders with the right but not the obligation to pay off the debt holders and acquire a firm's remaining assets. A call option on the underlying assets has the same properties. The holder of a call option on a firm's assets has a claim on those assets after fulfilling the option's strike value, which in this case, is equal to the book value of the firm's liabilities. If the value of the assets is not sufficient to meet the firm's liabilities, the shareholders, the holders of the call option, will not exercise the option and will abandon the firm to its creditors. KMV utilizes the optional nature of equity to derive the market value and volatility of a firm's underlying assets implied by its equity market value by solving backwards for the implied asset value and asset volatility.

Bibliography

Bollerslev, T. Generalized autoregressive conditional heteroskedasticity, Journal of Econometrics, 31: 307–27, 1986.

Clewlow, L. and Strickland, C. Implementing Derivative Models, Wiley, 1998.

Clewlow, L. and Strickland, C. Energy Derivatives, Pricing and Risk Management, Lacima, 2000.

Culp C.L. The Risk Management Process, Wiley, 2001.

Dickey, D.A. and Fuller, W.A. Distribution of the estimators for autoregressive time series with a unit root, Journal of the American Statistical Association, 74, 1979.

Dickey, D.A. and Fuller, W.A. Likelihood ratio statistics for autoregressive time series with a unit root, Econometrica, 49: 1057–72, 1981.

Engle, R.F. Autoregressive conditional heteroskedasticity with estimates of the variance of UK inflation, Econometrica, 50: 987–1008, 1982.

Franses, P.H. Time Series for Business and Economic Forecasting, Cambridge University Press, 1998.

Harvey, A.C. Time Series Models, Harvester Wheatsheaf, 1993.

Hull, J.C. Options, Futures and Other Derivatives, Prentice Hall, ninth edition, 2014.

Merton, R. On the pricing of corporate debt: the risk structure of interest rates, Journal of Finance, 29: 449–70, 1974.

Mills, T.C. The Econometric Modelling of Financial Time Series, Cambridge University Press, third edition, 2008.

Pindyck, R. and Rubenfield, D. Econometric Models and Economic Forecasts, McGraw Hill, 1998.

Taylor, S.J. Modelling Financial Time Series, Wiley, 1984.

Vasicek, O. An equilibrium characterisation of the term structure, Journal of Financial Economics, (5): 177–88, 1977.

5

Derivatives

5.1 Futures, Forwards and Options

A derivative is a financial instrument whose payoff depends on the values of other more basic variables. The variables underlying derivatives are often the prices of traded securities. Derivatives separate market and credit risks from the underlying assets and liabilities, and offer the ability to reduce a risk exposure through its transfer to a party that is prepared to take on and manage those risks. Derivative securities are also known as contingent claims, and can be contingent on almost any variable—from the price of a commodity to weather outcomes. There are two basic types of derivatives—futures/forwards and options.

Forward and futures contracts are agreements to buy or sell an underlying asset at a predetermined time in the future for a specified price. Futures are exchange standardized contracts, whereas forward contracts are direct agreements between two parties. The cash flows of the two contracts also occur at different times. Futures are marked-to-market daily, with cash flows passing between the long and the short position to reflect the daily futures price change, whereas forwards are settled once at maturity. If future interest rates are known with certainty, then futures and forwards can be treated as the same for pricing purposes.

There are two sides to every forward contract. The party who agrees to buy the asset holds a long forward position, while the seller holds a short forward position. At the maturity of the contract (the 'forward date'), the short position delivers the asset to the long position in return for the cash amount

J. Rogers, *Strategy, Value and Risk*, Global Financial Markets,
https://doi.org/10.1007/978-3-030-21978-9_5

agreed in the contract, often called the delivery price. Figure 5.1 shows the profit and loss profile to the long forward position at the maturity of the contract. If T represents the contract maturity date, a long forward payoff is expressed as $S_T - K$, where S_T represents the asset price at time T, and K represents the agreed delivery price. The payoff can be positive or negative, depending on the relative values of S_T and K. The short position has the opposite payoff to the long position, that is, $K - S_T$, as every time the long position makes a profit, the short incurs a loss and vice versa. As the holder of a long forward contract is guaranteed to pay a known fixed price for the spot asset, futures and forwards can be seen as insurance contracts providing protection against the price uncertainty in the spot markets.

For an arbitrage relationship to exist, the forward price has to equal the cost of financing the purchase of the spot asset today and holding it until the forward maturity date. Let F represent a forward contract price on a spot asset that is currently trading at S, T the maturity date of the contract, c the cost of holding the spot asset (which includes the borrowing costs for the initial purchase and any storage costs) and d the continuous dividend yield paid out by the underlying asset. The price of a forward contract at time t and the spot instrument on which it is written are related via the 'cost of carry' formula:

$$F = S\mathrm{e}^{(c-d)(T-t)} \tag{5.1}$$

where $T - t$ represents time in years. The continuous dividend yield, for example, can be interpreted as the yield on index futures, the foreign interest

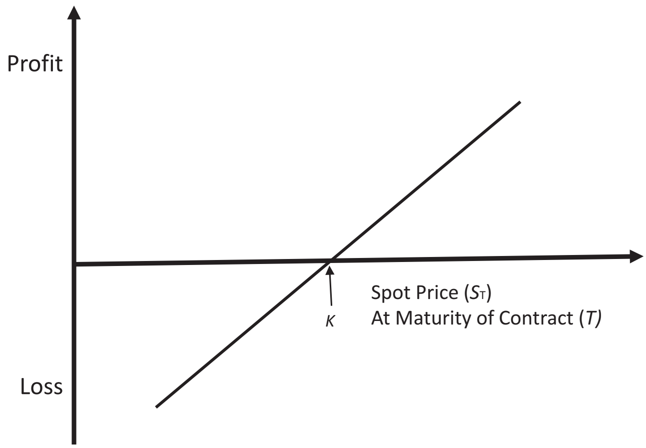

Fig. 5.1 Payoff to a long forward position

rate in foreign exchange futures contracts and the convenience yield for various energy contracts.

Options contracts are the second foundation to the derivatives markets. Options are asymmetrical relationships where the option holder has a right, but not an obligation, to transact at a contracted price called the exercise price. There are two basic types of options. A call option gives the holder the right but not the obligation to buy the spot asset on or before a predetermined date (the maturity date) at a certain price (the strike price), which is agreed today. A put option is the right to sell at the exercise price. Option sellers, or writers, are obliged to commit to the purchaser's decision. Figure 5.2 shows the payoff to the holder of a call option.

Options differ from forward and futures contracts in that a payment, or the option price or premium, must be made by the buyer, usually at the time when entering the contract. If the spot asset price is below the agreed strike or exercise price K at the maturity or expiration date, the holder lets the option expire worthless, forfeits the premium and buys the asset in the spot market. For asset prices greater than K, the holder exercises the option, buying the asset at K and has the ability to immediately make a profit equal to the difference between the two prices less the initial premium. The holder of the call option therefore essentially has the same positive payoff as the long forward contract without the downside risk.

The payoff to a call option is defined as:

$$\max\left(S-K,0\right) \tag{5.2}$$

The second basic type of option, a put option, gives the holder the right, but not the obligation, to sell the asset on or before the maturity date at the strike price.

The payoff for a put option is defined as:

$$\max\left(K-S,0\right) \tag{5.3}$$

Figure 5.3 shows the payoff to the holder of a put option.

As with forwards, there are two sides to every option contract. One party buys the option and has the long position, while the other party writes or sells the option and takes a short position. Figure 5.4 shows the four possible combinations of payoffs for long and short positions in European call and put options at the maturity date T. Options are also classified with respect to their exercise conventions. European options can only be exercised on the maturity

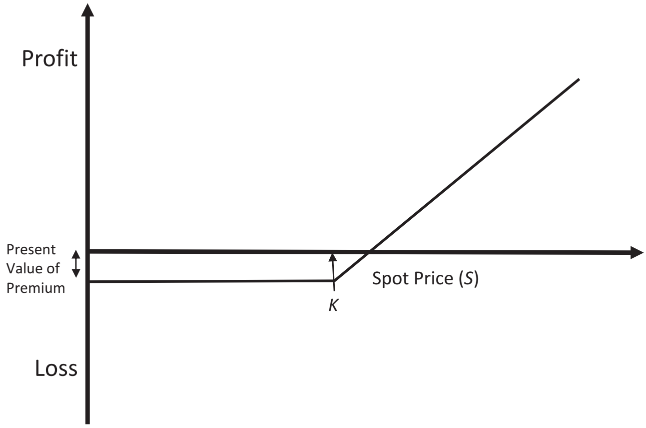

Fig. 5.2 Payoff for a call option

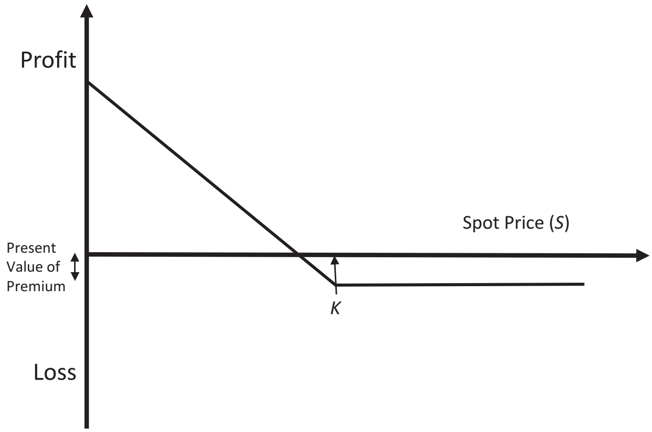

Fig. 5.3 Payoff for a put option

date itself, whereas American-style options can be exercised at any time, up to and including the expiration date. While early exercise of an American option is generally not optimal, there are exceptions to the rule. One example is where the underlying asset pays dividends, reducing the value of the asset and any call options on that asset, in which case, the call option may be exercised before maturity.

Forwards and options are also the key building blocks of more complex derivatives, and these building blocks are themselves interdependent. The decomposition of derivatives into their components assists in identifying a derivative's risk characteristics, which promotes more accurate pricing and

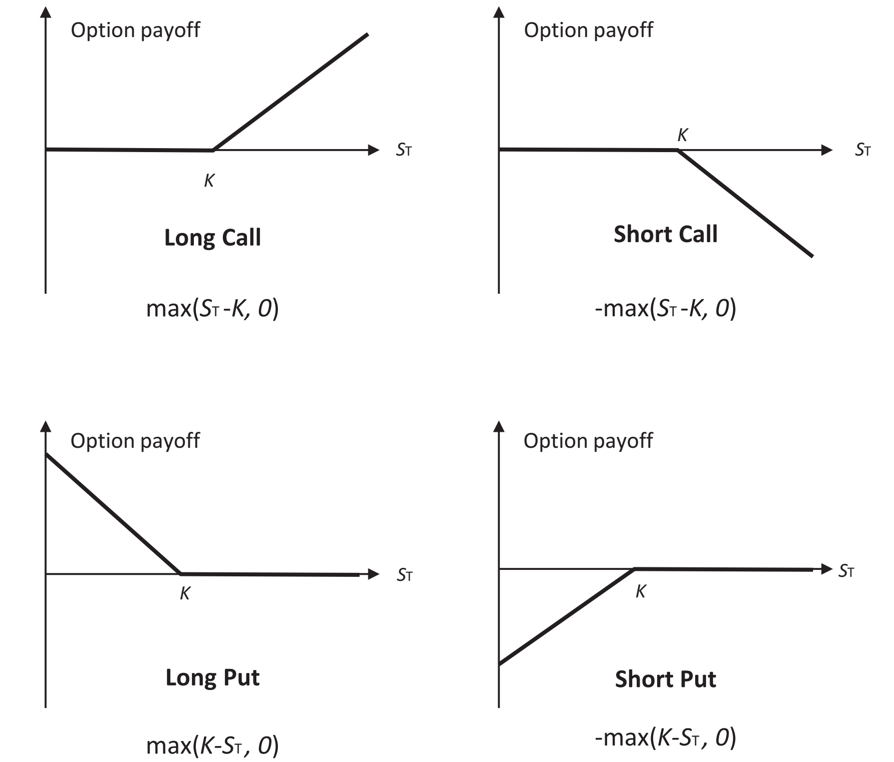

Fig. 5.4 Payoffs for European options

better risk management strategies. The basic futures and options described are the building blocks of all derivative securities, and the principles are consistent across all underlying markets. In some markets, however, derivative structures exhibit a number of important differences from other underlying markets. These differences arise due to the complex contract types that exist in these industries, as well as the complex characteristics of the relevant underlying prices. Both the type of derivative and the associated modelling need to capture the evolution of the underlying prices to reflect these differences.

5.2 The Replicating Portfolio and Risk-Neutral Valuation

The modern theory of option pricing is possibly one of the most important contributions to financial economics. The breakthrough came in the early 1970s, with work by Fisher Black, Myron Scholes and Robert Merton (Black and Scholes 1973; Merton 1973). The Black–Scholes–Merton (BSM) modelling approach not only proved to be important in providing a computationally efficient and relatively easy way of pricing an option, but also demonstrated the principal of no-arbitrage risk-neutral valuation. Their analysis showed that the payoff to an option could be perfectly replicated with a continuously adjusted holding in an underlying asset and a risk-free bond. As the risk of writing an option can be completely eliminated, the risk preferences of market participants are irrelevant to the valuation problem, and it can be assumed that they are risk-neutral. In this construct, all assets earn the risk-free rate of interest, and therefore, the actual expected return on the asset does not appear in the Black–Scholes formula.

Options can be valued by deriving the cost of creating the replicating portfolio such that both the option and the portfolio provide the same future returns, and therefore must sell at the same price to avoid arbitrage opportunities. The portfolio consists of Δ units of an underlying asset S and an amount B borrowed against Δ units at the risk-free rate r. This combination of the borrowing and the underlying asset creates the same cash flows or returns as an option. A binomial model can be used to illustrate the replicating portfolio. The binomial model assumes that the underlying asset price follows a binomial process, where at any time, the asset price S at t_0 can only change to one of two possible values over the time period Δt, either up to uS or down to dS at time t_1, where $u = e^{\sigma\sqrt{\Delta t}}$ and $d = 1/u$. Figure 5.5 is a binomial model for a one-period process, in which a risk-free portfolio consisting of the underlying asset and the call option is illustrated.

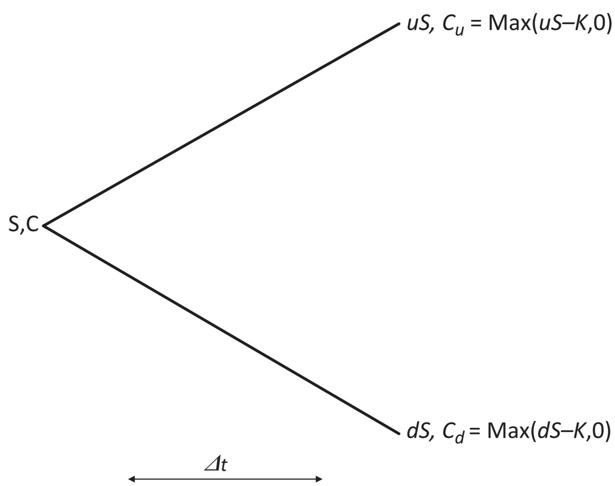

Fig. 5.5 Binomial model of an asset price and call option

The call option is defined as:

$$C \approx (\Delta S - B)$$

The value of the portfolio is the same, regardless of whether the asset price moves up or down over the period Δt:

$$C_u = \Delta uS - (1 + r)B$$

and

$$C_d = \Delta dS - (1 + r)B$$

which after rearranging becomes:

$$-C_u + \Delta uS = -C_d + \Delta dS \tag{5.4}$$

This is the equivalent to:

$$\Delta = \frac{C_u - C_d}{(u - d)S} \tag{5.5}$$

The portfolio must earn the continuously compounded risk-free rate of interest as it is risk-free:

$$(-C + \Delta uS) = \mathrm{e}^{r\Delta t}(-C + \Delta S) \tag{5.6}$$

Substituting into Eq. (5.6) for ΔS, using Eq. (5.5) and rearranging for the call price at t_0 obtains:

$$C = \mathrm{e}^{-r\Delta t}\left(\frac{\mathrm{e}^{r\Delta t} - d}{u - d}C_u + \frac{u - \mathrm{e}^{r\Delta t}}{u - d}C_d\right) \tag{5.7}$$

The actual probabilities of the asset moving up or down have not been used in deriving the option price, and therefore, the option price is independent of the risk preferences of investors. Equation (5.7) can be interpreted as taking discounted expectations of future payoffs under the risk-neutral probabilities.

This provides a means to derive the risk-neutral probabilities directly from the asset price:

$$uSp + dS(1-p) = Se^{r\Delta t} \tag{5.8}$$

for which the return can now be assumed as being the risk-free rate. Rearranging gives:

$$p = \frac{e^{r\Delta t} - d}{u - d} \tag{5.9}$$

for the risk neutral probability for uS, and $1 - p$ for dS. Equation (5.7) can now be written as:

$$C = e^{-r\Delta t}\left(pC_u + (1-p)C_d\right)$$

This is the price of the call option with one period to maturity.

5.3 A Model for Asset Prices

The evolution of uncertainty over time can be conceptualized and modelled as a mathematical expression, known as a stochastic process, which describes the evolution of a random variable over time. Models of asset price behaviour for pricing derivatives are formulated in a continuous time framework by assuming a stochastic differential equation (SDE) describes the stochastic process followed by the asset price. The most well-known assumption made about asset price behaviour, which was made by Black and Scholes (1973), is geometric Brownian motion (GBM).

The GBM assumption in the Black–Scholes model is the mathematical description of how asset prices evolve through time. In the GBM assumption, proportional changes in the asset price, denoted by S, are assumed to have constant instantaneous drift, μ, and volatility, σ. A non-dividend paying asset S following a GBM is represented by the following stochastic differential equation (SDE):

$$dS = \mu S dt + \sigma S dz \tag{5.10}$$

where dS represents the increment in the asset price process during a (infinitesimally) small interval of time dt, and dz is the underlying uncertainty driving the model, representing an increment in a Weiner process during dt. The risk-neutral assumption implies that the drift can be replaced by the risk-free rate of interest (i.e., $\mu = r$). Any process describing the stochastic behaviour of the asset price will lead to a characterization of the distribution of future asset values. An assumption in Eq. (5.10) is that future asset prices are lognormally distributed, or that the returns to the asset are normally distributed. Dividing through by S gives:

$$\frac{dS}{S} = \mu dt + \sigma dz \tag{5.11}$$

In Eq. (5.11), the percentage change or return in the asset price dS/S has two components. The first is that during the small interval of time dt, the average return on the asset is μdt, which is deterministic. The parameter μ is known as the drift. Added to this drift is the random component made up of the change dz, in a random variable z, and a parameter σ, which is generally referred to as the volatility of the asset. The random variable z, or equivalently, the change dz is called a Weiner process. A Weiner process is defined by two key properties. The first is that dz is normally distributed with mean zero and variance dt or the standard deviation of the square root of dt. The second is that the values of dz over two different non-overlapping increments of time are independent. Equations (5.10) and (5.11) are examples of an Itô process, as the drift and volatility only depend on the current value of the variable (the asset price) and time. In general, the stochastic differential equation for a variable S following an Itô process is:

$$dS = \mu\left(S,t\right)dt + \sigma\left(S,t\right)dz \tag{5.12}$$

where the functions $\mu(S,t)$ and $\sigma(S,t)$ are general functions for the drift and volatility. Many models for the behaviour of asset prices assume that the future evolution of the asset price depends only on its present level and not on the path taken to reach that level. A stochastic process possessing this property is known as Markovian.

The stochastic process followed by any derivative can be inferred from the assumption of the behaviour of the asset price on which the derivative's payoff is dependent. It follows that, using the Black–Scholes concept of constructing a riskless portfolio, a partial differential equation can be derived that governs the price of the derivative security.

5.4 The Black–Scholes Formula

The stochastic differential equation for the asset price S is the starting point for any derivative model. As the process for the asset and the process for the derivative have the same source of uncertainty, it is possible to combine the two securities in a portfolio in such a way as to eliminate that uncertainty. A portfolio consisting of a short position in an option and a long position in an underlying asset can be constructed such that the change in its value over an infinitesimal increment of time is independent of the source of randomness, and is therefore risk-free. This relationship leads to the Black–Scholes partial differential equation. The Black–Scholes formulae for standard European call and put options are the result of solving this partial differential equation.

As the expected return on the underlying asset does not appear in the Black–Scholes partial differential equation, the value of the derivative is independent of the risk preferences of investors. The implication of this risk-neutral pricing is that the present value of any future random cash flow—for example, the payoff for an option—is given by the expected value of the random future value discounted at the riskless rate. Replacing the expectation with the integral and solving obtains the Black–Scholes equation.

The Black–Scholes formula for a European call option on a non-dividend paying stock is:

$$c = S_0 N(d_1) - K\mathrm{e}^{-r(T-t)} N(d_2) \tag{5.13}$$

where,

$$d_1 = \frac{\ln\left(\frac{S_0}{K}\right) + \left(r + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}$$

$$d_2 = d_1 - \sigma\sqrt{T-t}$$

while the corresponding equation for the European put is:

$$p = K\mathrm{e}^{-r(T-t)} N(-d_2) - S_0 N(-d_1) \tag{5.14}$$

and the parameters are:

S_0 = the value of S at time zero,
K = the strike price of the option,
r = the risk-free interest rate,
t = a point in time,
T = Time at maturity of a derivative.

One of the qualities that has led to the enduring success of the Black–Scholes model is its simplicity. The inputs of the model are defined by the contract being priced or are directly observable from the market. The only exception to this is the volatility parameter, and there is now a vast amount of published material in the finance literature for deriving estimates of this figure, either from historical data or as implied by the market prices of options.

One widely used relaxation of the original formula takes into account assets that pay a constant proportional dividend. Assets of this kind are handled by reducing the expected growth rate of the asset by the amount of the dividend yield. If the asset pays a constant proportional dividend at a rate d over the life of the option, then the original Black–Scholes call formula (5.13) can be used with the adjustment where the parameter S is replaced by the term $Se^{-d(T-t)}$. This adjustment has been applied to value options on broad-based equity indices, options on foreign exchange rates, and real options that allow for competition, where the fall in value due to competition is equivalent to the dividend yield.

The intuition of the replicating portfolio concept can be illustrated with the Black–Scholes formula. The Black–Scholes formula can be defined as a combination of two binary options—a cash-or-nothing call and an asset-or-nothing call:

Asset-or-nothing call:

$$S\mathrm{e}^{-\delta(T-t)}N\left(d_1\right) \tag{5.15}$$

Cash-or-nothing call:

$$K\mathrm{e}^{-r(T-t)}N\left(d_2\right) \tag{5.16}$$

A European call option represents a long position in an asset-or-nothing call and a short position in a cash-or-nothing call, where the cash payoff on the cash-or-nothing call is equivalent to the strike price. A European put is a long position in a cash-or-nothing put and a short position in an asset-or-nothing put, where the strike price represents the cash payoff on the cash-or-nothing put. $N(d_1)$, the option delta, is the number of units of the underlying asset required to form the portfolio, and the cash-or-nothing term is the number of bonds, each paying \$1 at expiration.

Although it is possible to obtain closed-form solutions such as Eq. (5.13) for certain derivative pricing problems, there are many situations when analytical solutions are not obtainable, and therefore, numerical techniques need to be applied. Examples include American options and other options where there are early exercise opportunities, 'path-dependent' options with discrete observation frequencies, models that incorporate jumps and models dependent on multiple random factors. The description of two of these techniques is the subject of the next section.

5.5 Numerical Techniques

Two numerical techniques that are most commonly used by practitioners to value derivatives in the absence of closed-form solutions are binomial and trinomial trees and Monte Carlo simulation. Practitioners also use other techniques such as finite difference schemes, numerical integration, finite element methods and others. It is possible to price not only derivatives with complicated payoff functions dependent on the final price using trees and Monte Carlo simulation techniques, but also derivatives whose payoff is determined also by the path the underlying price follows during its life.

5.5.1 Monte Carlo Simulation

Monte Carlo simulation provides a simple and flexible method for valuing complex derivatives for which analytical formulae are not possible. The method can easily deal with multiple random factors, can be used to value complex path-dependent options, and allows the inclusion of price processes such as price jumps. In general, the present value of an option is the expectation of its discounted payoff. Monte Carlo simulation derives an estimate of this expectation by simulating a large number of possible paths for the asset price from time zero to the option maturity, and computing the average of the discounted payoffs.

GBM for non-dividend spot prices with constant expected return μ and volatility σ is represented by the SDE in Eq. (5.10). The Black–Scholes perfect replication argument leads to the risk-neutral process in which the actual drift of the spot price μ is replaced by the interest rate r:

$$dS = rSdt + \sigma Sdz \tag{5.17}$$

If the asset pays a constant continuous dividend yield δ, then the risk-neutral process becomes:

$$dS = (r - \delta)Sdt + \sigma Sdz \tag{5.18}$$

Transforming the spot price to the natural log of the spot price $x = \ln(S)$ gives the following process for x:

$$dx = vdt + \sigma dz \tag{5.19}$$

where $v = r - \delta - \frac{1}{2}\sigma^2$. The transformed GBM process represented in Eq. (5.10) can be discretized as:

$$x_{t+\Delta t} = x_t + \left(v\Delta t + \sigma\left(z_{t+\Delta t} - z_t\right)\right) \tag{5.20}$$

In terms of the original asset price, the discrete form is:

$$S_{t+\Delta t} = S_t \exp\left(v\Delta t + \sigma\left(z_{t+\Delta t} - zt\right)\right) \tag{5.21}$$

Equations (5.20) or (5.21) can be used to simulate the evolution of the spot price through time. The change in the random Brownian motion, $z_{t+\Delta t} - z_t$, has a mean of zero and a variance of Δt. It can therefore be simulated using random samples from a standard normal distribution multiplied by $\sqrt{\Delta t}$, that is, $\sqrt{\Delta t}\varepsilon$ where $\varepsilon \sim N(0,1)$. In order to simulate the spot price, the time period $[0, T]$ is divided into N intervals such that $\Delta t = T/N$, $t_i = i\Delta t$, $i = 1, \ldots, N$. Using, for example, Eq. (5.21) gives:

$$S_{t_i} = S_{t_{i-1}} \exp\left(v\Delta t + \sigma\sqrt{\Delta t}\varepsilon_i\right) \tag{5.22}$$

As the drift and volatility terms do not depend on the variables S and t, the discretization is correct for any chosen time step. Therefore, the option can be simulated to the maturity date in a single time step if the payoff is a function of the terminal asset value and does not depend on the asset's path during the life of the option. Repeating this process N times, and choosing ε_i randomly each time, leads to one possible path for the spot price for each simulation.

At the end of each simulated path, the terminal value of the option C_T is evaluated. Let $C_{T,j}$ represent the payoff to the contingent claim under the jth simulation. For example, a standard European call option terminal value is given by:

$$C_{T,j} = \max\left(S_{T,j} - K, 0\right) \tag{5.23}$$

Each payoff is discounted using the simulated short-term interest rate sequence:

$$C_{0,j} = \exp\left(-\int_0^T r_u du\right) C_{T,j} \tag{5.24}$$

In the case of constant or deterministic interest rates, Eq. (5.24) simplifies to:

$$C_{0,j} = P\left(0,T\right) C_{T,j} \tag{5.25}$$

This value represents the value of the option along one possible asset price path. The simulations are repeated M times and the average of all the outcomes is taken to compute the expectation, and hence, the option price:

$$\hat{C}_0 = \frac{1}{M} \sum_{j=1}^{M} C_{0,j} \tag{5.26}$$

Therefore, $\hat{C}_0$ is an estimate of the true value of the option, C_0, but with an error due to the fact that it is an average of randomly generated samples, and so, is itself random. In order to obtain a measure of the error, the standard error SE(.) is estimated as the sample standard deviation, SD(.), of $C_{t,j}$ divided by the square root of the number of samples:

$$SE\left(\hat{C}_0\right) = \frac{SD\left(C_{0,j}\right)}{\sqrt{M}} \tag{5.27}$$

where $SD(C_{0,j})$ is the standard deviation of C_0:

$$SD\left(C_{0,j}\right) = \sqrt{\frac{1}{M-1} \sum_{j=1}^{M} \left(C_{0,j} - C_0\right)^2} \tag{5.28}$$

For many American-style options, early exercise can be optimal, depending on the level of the underlying price. It is rare to find closed-form solutions for prices and risk parameters of these options, so numerical procedures must be applied. Using Monte Carlo simulation for pricing American-style options, however, can be difficult. The problem arises because simulation methods generate trajectories of state variables forward in time, whereas a backward dynamic programming approach is required to efficiently determine optimal exercise decisions for pricing American options. Therefore, practitioners usually use binomial and trinomial trees for the pricing of American options.

5.5.2 The Binomial and Trinomial Method

The binomial model of Cox et al. (1979) is a well-known alternative discrete time representation of the behaviour of asset prices following GBM. This model is important in several ways. First, the continuous time limit of the proportional binomial process is exactly the GBM process. Second, and perhaps most importantly, the binomial model is the basis of the dynamic programming solution to the valuation of American options. Section 5.2 discussed a one-step binomial tree as part of the overview of the replicating portfolio. To price options with more than one period to maturity, the binomial tree is extended outwards for the required number of periods to the maturity date of the option. Figure 5.6 illustrates a binomial tree for an option that expires in four periods of time.

A state in the tree is referred to as a node, and is labelled as node (i,j), where i indicates the number of time steps from time zero and j indicates the number of upward movements the asset price has made since time zero. Therefore, the level of the asset price at node (i,j) is $S_{i,j} = Su^j d^{i-j}$ and the option price will be $C_{i,j}$. At the lowest node at every time step $j = 0$, and j will remain the same when moving from one node to another via a downward branch, as the number of upward moves that have occurred would not have changed. It is generally assumed that there are N time steps in total, where the Nth time step corresponds to the maturity date of the option. As is the case with the one period example, the value of a call option at the maturity date is the payoff:

$$C_{N,j} = \max\left(S_{N,j} - K, 0\right) \tag{5.29}$$

As the value of the option at any node in the tree is its discounted expected value, at any node in the tree before maturity:

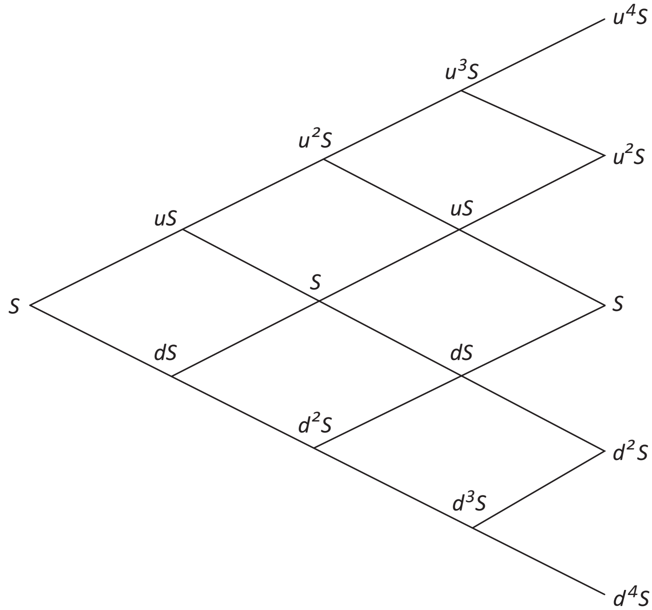

Fig. 5.6 A four-step binomial tree for an asset

$$C_{i,j} = \mathrm{e}^{-r\Delta t}\left(pC_{i+1,j+1} + (1-p)C_{i+1,j}\right) \tag{5.30}$$

where the binomial risk neutral probabilities p and $(1-p)$ are derived as:

$$p = \frac{\mathrm{e}^{r\Delta t} - d}{u - d}$$

and r is the risk free rate.

Using Eqs. (5.29) and (5.30), the value of the option can be computed at every node for time step $N-1$. Equation (5.30) can then be reapplied at every node for every time step, working backwards through the tree to compute the value of the option at every node in the tree. The value of a European option can be derived using this procedure. To derive the value of an American option, the value of the option, if it is exercised, is compared at every node to the option value if it is not exercised, and the value at that node set to the greater of the two.

Although binomial trees are used by many practitioners for pricing American-style options, trinomial trees offer a number of advantages over the binomial tree. As there are three possible future movements over each time period, rather than the two of the binomial approach, the trinomial tree provides a better approximation to a continuous price process than the binomial tree for the same number of time steps. The trinomial tree is also easier to work with because of its more regular grid and is more flexible, allowing it to be fitted more easily to market prices of forwards and standard options, an important practical consideration. A discussion of trinomial trees follows.

In the following, it is more convenient to work in terms of the natural logarithm of the spot price as defined in Eq. (5.19). Consider a trinomial model of the asset price in which, over a small time interval Δt, the asset price can increase by Δx (the space step), stay the same or decrease by Δx, with probabilities p_u, p_m and p_d, respectively. This is depicted in terms of x in Fig. 5.7.

The drift and volatility parameters of the asset price are now captured in this simplified discrete process by Δx, p_u, p_m and p_d. It can be shown that the space step cannot be chosen independently of the time step, and that a good choice is $\Delta x = \sigma\sqrt{3\Delta t}$. The relationship between the parameters of the continuous time process and the trinomial process is obtained by equating the mean and variance over the time interval Δt and requiring that the probabilities sum to one, that is:

$$E\left[\Delta x\right] = p_u\left(\Delta x\right) + p_m\left(0\right) + p_d\left(-\Delta x\right) = \nu\Delta t \tag{5.31}$$

$$E\left[\Delta x^2\right] = p_u\left(\Delta x^2\right) + p_m\left(0\right) + p_d\left(\Delta x^2\right) = \sigma^2\Delta t + \nu^2\Delta t^2 \tag{5.32}$$

$$p_u + p_m + p_d = 1 \tag{5.33}$$

Solving Eqs. (5.31–5.33) yields the following explicit expressions for the transitional probabilities:

$$p_u = \frac{1}{2}\left(\frac{\sigma^2\Delta t + \nu^2\Delta t^2}{\Delta x^2} + \frac{\nu\Delta t}{\Delta x}\right) \tag{5.34}$$

$$p_m = 1 - \frac{\sigma^2\Delta t + \nu^2\Delta t^2}{\Delta x^2} \tag{5.35}$$

$$p_d = \frac{1}{2}\left(\frac{\sigma^2\Delta t + \nu^2\Delta t^2}{\Delta x^2} - \frac{\nu\Delta t}{\Delta x}\right) \tag{5.36}$$

The single period trinomial process in Fig. 5.7 can be extended to form a trinomial tree. Figure 5.8 depicts such a tree.

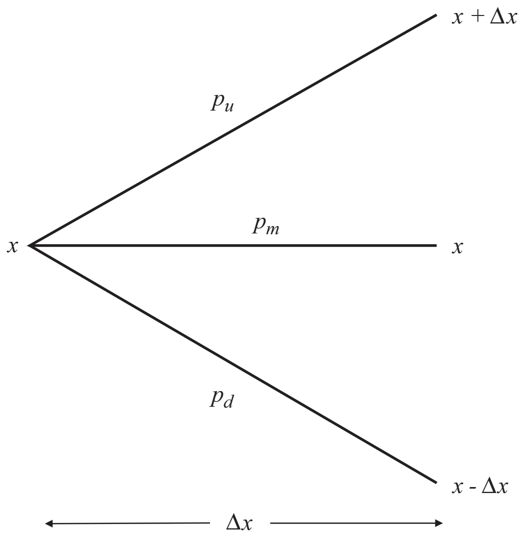

Fig. 5.7 The trinomial model of an asset price

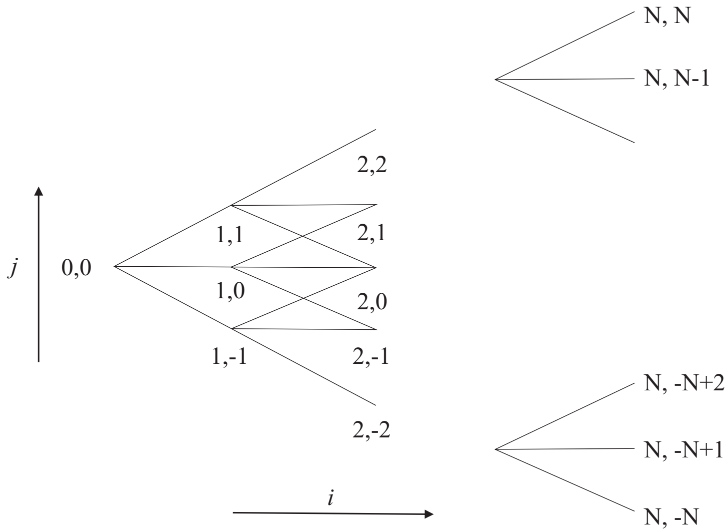

Fig. 5.8 A trinomial tree model of an asset price

Let i denote the number of the time step and j, the level of the asset price relative to the initial asset price in the tree. If $S_{i,j}$ denotes the level of the asset price at node (i,j), then $t = t_i = i\Delta t$, and an asset price level of $S\exp(j\Delta x)$. Once the tree has been constructed, the spot price is known at every time and every state of the world consistent with the original assumptions about its behaviour process, and the tree can be used to derive prices for a wide range of derivatives.

The procedure is illustrated with reference to pricing a European and American call option with a strike price K on the spot price. The value of an option is represented at node (i,j) by $C_{i,j}$. In order to value an option, the tree is constructed as representing the evolution of the spot price from the current date out to the maturity date of the option. Let time step N correspond to the maturity date in terms of the number of time steps in the tree, that is, $T = N\Delta t$. The values of the option at maturity are determined by the values of the spot price in the tree at time step N and the strike price of the option:

$$C_{N,j} = \max\left(S_{N,j} - K, 0\right); j = -N, \ldots, N \tag{5.37}$$

It can be shown that option values can be computed as discounted expectations in a risk-neutral world, and therefore, the values of the option at earlier nodes can be computed as discounted expectations of the values at the following three nodes to which the asset price could jump:

$$C_{i,j} = \mathrm{e}^{-r\Delta t}\left(p_u C_{i+1,j+1} + p_m C_{i+1,j} + p_d C_{i+1,j-1}\right) \tag{5.38}$$

where $\mathrm{e}^{-r\Delta t}$ is the single period discount factor. This procedure is often referred to as 'backwards induction' as it links the option value at time i to known values at time $i + 1$. The attraction of this method is the ease with which American option values can be evaluated. During the inductive stage, the immediate exercise value of the option is compared with the value if not exercised as computed from Eq. (5.38). If the immediate exercise value is greater, then this value is stored at the node, that is:

$$C_{i,j} = \max\left\{\mathrm{e}^{-r\Delta t}\left(p_u C_{i+1,j+1} + p_m C_{i+1,j} + p_d C_{i+1,j-1}\right), S_{i,j} - K\right\} \tag{5.39}$$

This method also provides the optimal exercise strategy for the American option, since for every possible future state of the world, that is, every node in the tree, it can be determined whether to exercise the option or not. The value of the option today is given by the value in the tree at node $(0,0)$, $C_{0,0}$.

Bibliography

Black, F. and Scholes, M. The pricing options and corporate liabilities, Journal of Political Economy, 81: 637–59, 1973.

Clewlow, L. and Strickland, C. Implementing Derivative Models, Wiley, 1998.

Clewlow, L. and Strickland, C. Energy Derivatives, Pricing and Risk Management, Lacima, 2000.

Cox, J.S., Ross, S. and Rubinstein, M. Option pricing: a simplified approach, Journal of Financial Economics, 7 (September): 229–63, 1979.

Hull, J.C. Options, Futures and Other Derivatives, Prentice Hall, 2000.

Merton, R. Theory of rational option pricing, Bell Journal of Economics and Management Science, 4: 41–183, 1973.

Trigeorgis, L. Real Options: Managerial Flexibility and Strategy in Resource Allocation, MIT Press, 1996.

6

Derivative Model Applications

6.1 Spot Price Models

The Black–Scholes GBM (geometric Brownian motion) model can be generalized to other models that are more realistic for particular markets. The various simple extensions to the Black–Scholes model assume constant parameters for ease of calculation. In reality, the properties of time series such as volatility, mean reversion, long-term levels and jump behaviour will at the very least vary through time with reasonably predictable patterns. These characteristics can be included in spot models.

6.1.1 Geometric Brownian Motion

The GBM assumption defined in Eq. (5.10) as a process that describes the dynamics of the prices of financial instruments is an approximation of the behaviour observed in real markets. GBM models are frequently used for security prices, interest rates, commodities and other economic and financial variables, and follow what has been defined as a random walk. The Weiner process is the continuous limit of a discrete time random walk. A generalized Weiner process introduces the concept of an expected drift rate. The drift rate is the average increase in a stochastic variable for each unit of time. In models for financial variables, the expected drift rate is replaced with a constant drift rate. Another issue in GBM models is that the uncertainty associated with the price path is greater the longer the time horizon. As the variance of the Weiner process increases linearly as the time horizon increases, the standard deviation

J. Rogers, *Strategy, Value and Risk*, Global Financial Markets,
https://doi.org/10.1007/978-3-030-21978-9_6

grows as the square root of the time horizon. This is the equivalent to the definition of volatility, where scaling the standard deviation by the square root of T annualizes the volatility σ.

The GBM process represented in Eq. (5.10) was discretized in Eqs. (5.20) and (5.21) for the simulation of a spot price. Figure 6.1 illustrates a GBM process simulated 100 times with the parameters S = 100, $r - \delta$ = 0.05, σ = 0.30, and Δt = 1/250. In this example, $r - \delta$ is the drift, and $\sigma\varepsilon\sqrt{\Delta t}$ is the stochastic component. One observation is that the sample paths in Fig. 6.1 tend to wander from the initial starting point of σ = 100. While this may be realistic for some variables, and can be verified in tests for random walks, it may, however, not be suitable for other financial and economic time series.

6.1.2 Mean Reversion

The usual assumption made for asset price evolution in many markets is the GBM model assumption. This model, however, allows prices to wander off to unrealistic levels when applied to markets such as energy and commodities. Mean reversion was first described by Vasicek (1977) for modelling interest rate dynamics and has subsequently been widely adapted. Mean reversion can

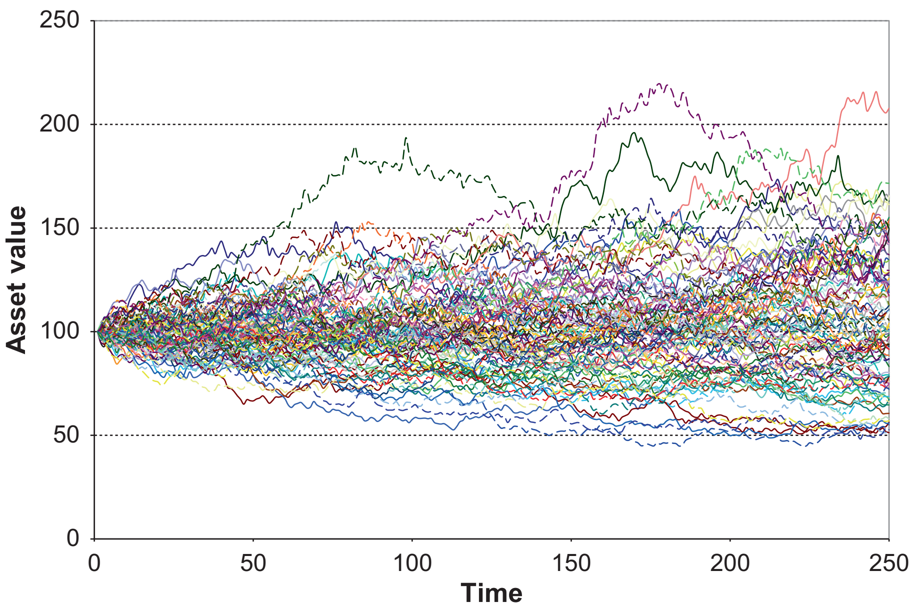

Fig. 6.1 Illustration of 100 simulated GBM paths

be understood by looking at a simple model of a mean reverting spot price (Schwartz 1997), represented by the following equation:

$$dS = \alpha(\mu - \ln S)Sdt + \sigma S dz \tag{6.1}$$

Figure 6.2 illustrates the log form of a mean reverting process simulated 100 times with the parameters S = 100, α = 3, $\overline{S} = 100$, σ = 0.30, and Δt = 1/250. In this model, the spot price mean reverts to the long-term level $\overline{S} = e^{\mu}$ at a speed given by the mean reversion rate, α, that is taken to be strictly positive. If the spot price is above the long-term level $\overline{S}$, then the drift of the spot price will be negative and the price will tend to revert back towards the long-term level. Similarly, if the spot price is below the long-term level, then the drift will be positive and the price will tend to move back towards $\overline{S}$. Note that, at any point in time, the spot price will not necessarily move back towards the long-term level as the random change in the spot price may be of the opposite sign and greater in magnitude than the drift component. This formulation of the mean reversion process represents one of a number of possible equations that capture the same type of market evolution of prices over time. In reality, the spot price does not mean revert to a constant long-term

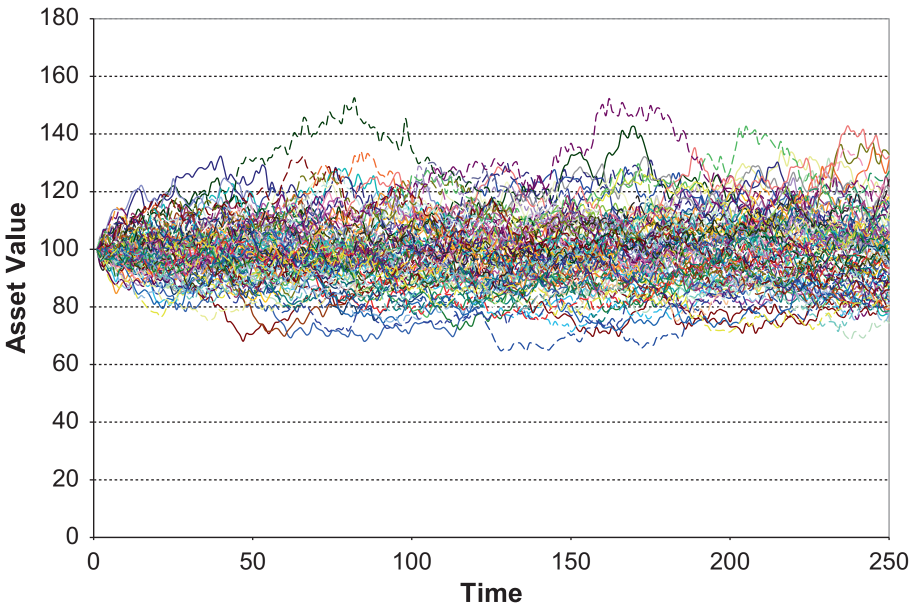

Fig. 6.2 Illustration of 100 simulated mean reversion paths

level. Information on the level to which the spot price mean reverts is contained in the forward curve prices and volatilities.

6.1.3 Jumps and Seasonal Patterns

Jumps can be a significant component of the behaviour of spot prices. This type of behaviour, where the price exhibits sudden, large changes, can be modelled by using jump processes. A simple and realistic model for a spot price, which is identical to the Black–Scholes model except for the addition of a jump process, is the jump-diffusion model introduced by Merton (1976). This model is described by the following stochastic differential equation (SDE):

$$dS = \mu S dt + \sigma S dz + \kappa S dq \tag{6.2}$$

where the lognormal jumps are driven by a Poisson process, and the annualized frequency of jumps is given by ϕ, the average number of jumps per year (ϕ) is defined by prob(dq = 1) = ϕdt). The proportional jump size is κ, which is random and determined by the natural logarithm of the proportional jumps being normally distributed:

$$\ln(1+\kappa) \sim N\left(\ln(1+\bar{\kappa}) - \frac{1}{2}\gamma^2, \gamma^2\right) \tag{6.3}$$

where $\bar{\kappa}$ is the mean jump size and γ is the standard deviation of the proportional jump size. The jump process (dq) is a discrete time process, that is, jumps do not occur continuously, but at specific instants of time. Therefore, for typical jump frequencies, most of the time $dq = 0$ and only takes the value 1 when a randomly timed jump occurs. When no jump occurs, the spot price behaviour is identical to GBM and only differs when a jump occurs. The proportional jumps (or equivalently jump returns) in Eq. (6.2) are normally distributed and therefore symmetrical, that is, the number of positive and negative jumps and the range of sizes of the proportional jumps will be equal on average.

Season patterns can be taken into account by including seasonality as a deterministic process in the stochastic process for the underlying price path. Discrete methods such as Fourier Transforms that include Sine, Cosine and Fast Fourier Transforms can be specified as continuous processes and included in the specification of the underlying price path.

6.2 Stochastic Volatility

The assumption in the Black–Scholes model that volatility is constant does not always hold. The GARCH (generalized autoregressive conditional hetero-skedasticity) process is one representation of a stochastic volatility model. Many other models have been proposed for the behaviour of volatility. The Heston (1993) form of the stochastic volatility model is described by the following processes for the spot price and the spot price return variance $V = \sigma^2$;

$$\frac{dS_t}{S_t} = \mu dt + \sqrt{\upsilon_t} dW_t \tag{6.4}$$

$$dv_t = \kappa(\theta - \upsilon_t)dt + \sigma_v \sqrt{v_t}\, dZ_t \tag{6.5}$$

Equation (6.4) is the GBM model with volatility ν_t, which is not constant and changes randomly. The behaviour of the volatility is determined by Eq. (6.5), which specifies the process followed by the variance, the square of the volatility. The variance mean reverts to a long-term level θ at a rate given by κ. The absolute volatility of the variance is $\sigma_v\sqrt{\nu}$, which is proportional to the square root of the variance, that is, the volatility of the spot price. The source of randomness in the variance, dZ_t, is different from the dW_t driving the spot price, although it may be correlated with correlation coefficient ρ.

The following illustrates the estimation of the parameters for the Heston stochastic volatility model.

Two sources of uncertainty reflected in FX options are the stochastic FX rate and stochastic volatility. The Black–Scholes model addresses the first, while stochastic volatility models address both the first and second. The Heston model (1993) is a common method applied to capture stochastic diffusion volatility:

$$\frac{dS_t}{S_t} = \mu dt + \sqrt{\upsilon_t} dW_t \tag{6.6}$$

$$dv_t = \kappa(\theta - \upsilon_t)dt + \sigma_v \sqrt{v_t}\, dZ_t \tag{6.7}$$

$$dW_t dZ_t = \rho dt \tag{6.8}$$

Stochastic volatility induces smiles and skews that decrease as the option maturity increases. The positive volatility of volatility (σ_v) generates a smile, or fatter tails in the distribution, while a non-zero correlation (ρ) generates skew of the same sign, that is, shifts the probability weight to either one of the tails of the distribution.

Calibrating the Heston model ensures the model matches the market and avoids arbitrage. The Heston model requires the estimation of five parameters:

- kappa (κ), the rate of mean reversion in the volatility
- theta (θ), the long run mean
- the asset volatility (v_t)
- volatility of volatility (σ_v), which influences the kurtosis of the distribution, and
- the correlation (ρ)

This is achieved by finding the set of parameters that produce Heston model prices that match vanilla Black–Scholes option market prices.

The Black–Scholes model has a number of applications in the OTC FX options market:

- Market prices are quoted as Black–Scholes implied volatilities instead of option prices, and are provided at a Black–Scholes delta (δ) instead of the strike.
- Liquidity is typically at five delta levels, 10 δ put, 25 δ put, 0 δ straddle, 25 δ call and 10 δ call.

The EUR-USD is used as an illustration. Calibrating the Heston model consists of:

1. converting the EUR-USD option delta quotes into strikes
2. deriving the Black–Scholes option prices using the implied volatilities, the derived strikes, EUR-USD forwards and interest rates for the two currencies, and
3. calibrating the Heston model parameters to the Black–Scholes prices across the volatility surface

The Black–Scholes option model is equivalent to the Black model when the risk-free interest rate is zero, which reflects the forward price, and then discounted to derive the present value.

Figure 6.3 shows the EUR-USD implied volatility surface.

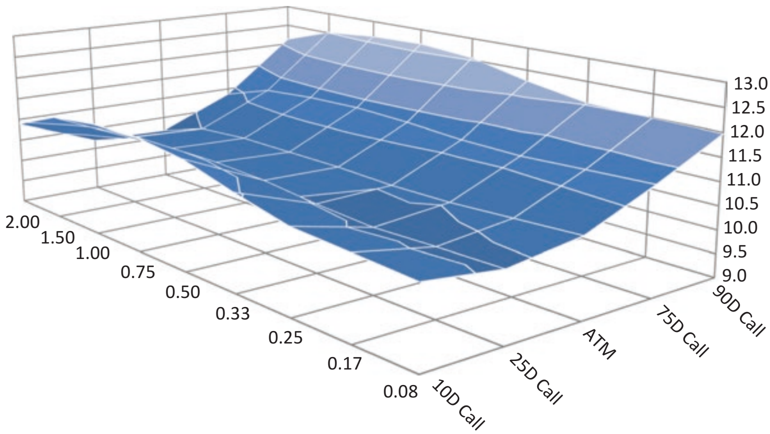

Fig. 6.3 EUR-USD implied volatility surface. *As at 5/15/08*

The five FX option quotes for each maturity are:

- *A delta-neutral straddle (ATMV)* implied volatility. A straddle equals the sum of a call and a put with the same strike. Delta neutral implies that δ(c) plus δ(p) is equal to zero, with $N(d_+)$ equal to 0.5, and d_+ equal to zero. Therefore, ATMV ≡ IV(50 δ c) (= IV(−50 δ p) by put call parity). IV is the implied volatility, c equals call, p equals put, d is the delta and N() is a standard normal function.
- *25-delta Risk Reversal (RR25)*. The RR25 describes the slope of a smile, which represents the skew in the risk neutral distribution of the return. RR25 ≡ IV (25δc) − IV (25δp).
- *25-delta Strangle Margin (SM25)*, or a butterfly spread. A strangle equals the sum of a call and a put with two different strikes, and captures the smile curvature, or the distribution's kurtosis. SM25 ≡ (IV (25δc) + IV (25δp)) /2 − ATMV.
- *10-delta Risk Reversal (10RR)*, and
- *10-delta Strangle Margin (10SM)*

The implied volatility quotes for the five deltas have the following relationships:

- IV (0δs) = ATMV
- IV (25δc) = ATMV + RR25/2 + SM25
- IV (25δp) = ATMV − RR25/2 + SM25
- IV (10δc) = ATMV + RR10/2 + SM10, and
- IV (10δp) = ATMV − RR10/2 + SM10

Table 6.1 The Heston parameters

Kappa (κ)	2.044
Theta (θ)	1.16%
Asset volatility (V)	1.57%
Vol of Vol (σ)	35.62%
Correlation (ρ)	0.173

The EUR-USD option deltas are converted into the strike surface as:

$$K = F\exp\left[\pm IV(\delta,\tau)\sqrt{\tau}N^{-1}\left(\pm e^{r_f\tau}\delta\right)+\frac{1}{2}IV(\delta,\tau)^2\tau\right] \tag{6.9}$$

where for each maturity:

K = the option strike,
F = FX Forward,
IV = implied volatility,
δ = strike delta,
τ = time to expiry,
$N^{-1}()$ = the inverse of the standard normal cumulative distribution, and
r_f = the USD interest rate.

Out-of-the-money Black prices are then calculated across the surface using the implied volatility quotes and the derived strikes. Finally, the stochastic volatility parameters are derived by minimizing the squared error, scaled by a weight derived as the inverse of the delta implied volatility bid/offer spread quote, between the Heston and Black prices across the whole surface.

Table 6.1 illustrates the Heston model parameter estimates for the EUR-USD FX option.

6.3 Forward Curve Models

Forwards and futures markets are often used by risk managers to hedge risk, with liquid forward prices providing a price discovery mechanism to determine the fair value for future delivery. Forward curves contain information about the prices an investor can lock into today to buy or sell at a certain time in the future. Forward curves are well known and understood in the debt markets. Forward rate agreements and exchange traded futures contracts are heavily traded and allow users to lock in borrowing and lending rates for future time periods.

In contrast to futures and forwards, price forecasts are predictions on the likely spot price for periods in the future, and can differ widely between market participants. Forward prices, however, depend on the relationship between traded instruments. Tradable prices today for future spot transactions can be locked in using forward prices, and as such, capture the market reality. Therefore, prices from forwards and futures markets are key inputs to many derivative pricing models, and are as essential in the pricing of derivatives as spot prices.

In the past, the majority of work on modelling prices has focused on stochastic processes for the spot price and other key variables, such as the dividend yields, convenience yields and interest rates. This approach, however, can have some fundamental disadvantages. The first is that key state variables, such as the convenience yield, are unobservable, and second, the forward price curve is an endogenous function of the model parameters, and therefore, will not necessarily be consistent with the market observable forward prices. As a result, many industry practitioners require the forward curve to be an input into the derivative pricing model, rather than an *output* from it.

Term structure consistent models model the dynamics of the entire term structure in a manner that is consistent with the initial observed market data. These models can be further classified into those that fit the term structure of prices such as interest rates, and those that fit the term structure of prices and price volatilities. There are models in the interest rate world and developments in the energy and commodity markets that use term structure approaches. An approach based on modelling the entire forward price curve with multiple sources of uncertainty uses all the information contained in the term structure of futures prices in addition to the historical volatilities of futures returns for different maturities.

6.3.1 A Single Factor Model for the Forward Curve

Forward curve models are defined as models that explicitly model all the forward prices simultaneously instead of just the spot price. A simple single factor model of the forward curve can be represented by the following stochastic differential equation:

$$\frac{dF(t,T)}{F(t,T)} = \sigma e^{-\alpha(T-t)} dz(t) \tag{6.10}$$

The inputs to the model are the observed forward curve $F(t, T)$, which denotes the forward price at time t for maturity date T, and $\sigma e^{-\alpha(T-t)}$, which is the single 'factor' or volatility function associated with the source of risk $dz(t)$. Equation (6.10) also has no drift term. As futures and forward contracts have zero initial investment, their expected return in a risk-neutral world must be zero, implying that the process describing their evolution has zero drift. The volatility function of Eq. (6.10) has a very simple negative exponential form illustrated in Fig. 6.4.

For this volatility function, short-dated forward returns are more volatile than long-dated forwards. Information occurring in the market today has little effect on, say, the 5-year forward price, but can have a significant effect on the 1-month forward price. The parameter values used for Fig. 6.4 are $\alpha = 1.0$ and $\sigma = 0.40$. Here, σ represents the 'overall' volatility of the forward curve, while α explains how fast the forward volatility curve attenuates with increasing maturity. With an α of 100%, the 1-month forward has a volatility of about 37%, decreasing to approximately 2% for the 3-year forward.

The volatility function is not restricted to have the parameterized form of Eq. (6.10). The function can be generalized as:

$$\frac{dF(t,T)}{F(t,T)} = \sigma(t,T)dz(t) \tag{6.11}$$

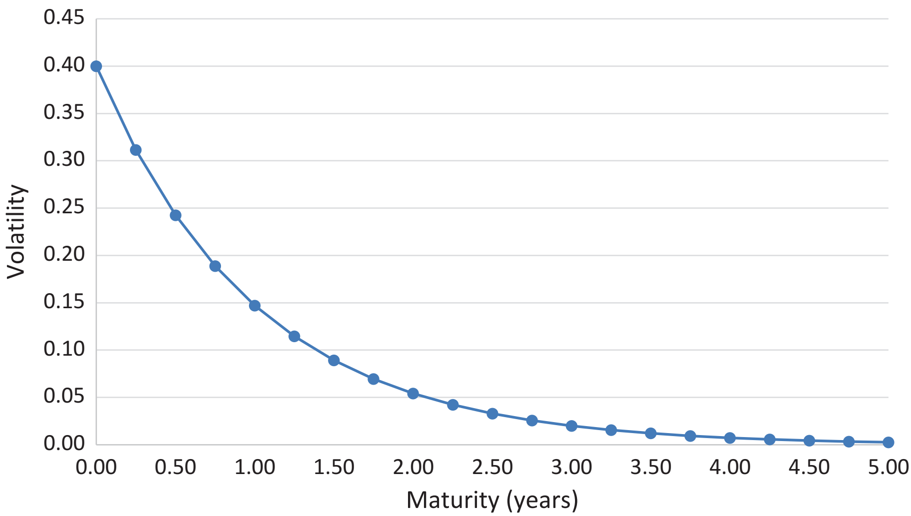

Fig. 6.4 A negative exponential volatility function for forward prices

where $\sigma(t,T)$ is the time t volatility of the T maturity forward price return. The form of $\sigma(t, T)$ can be determined from market data.

6.3.2 The Dynamics of the Forward Curve

An important observation is that forward prices of different maturities are not perfectly correlated. The curves generally move up and down together, but they also change shape in quite complex ways. One method that can be used to determine the set of common factors that drive the dynamics of the forward curve is principal components analysis (PCA), or eigenvector decomposition of the covariance matrix. This procedure can be utilized to simultaneously identify the number of significant factors and estimate the volatility functions. The technique involves calculating the covariances between every pair of forward price returns in a historical time series to form a covariance matrix. The eigenvectors of the covariance matrix yield estimates of the factors driving the evolution of the forward curve.

The implication is that to effectively describe the evolution of the energy forward curve, more than a single factor is required. The model described by Eq. (6.11) can be modified through the addition of sources of risk and volatility functions. For a general multifactor model, the behaviour of the forward curve can be represented by the following equation:

$$\frac{dF(t,T)}{F(t,T)} = \sum_{i=1}^{n} \sigma_i(t,T)\, dz_i(t) \tag{6.12}$$

In this formulation, there are n independent sources of uncertainty, which drive the evolution of the forward curve. Each source of uncertainty has associated with it a volatility function, which determines by how much, and in which direction, that random shock moves each point of the forward curve. Therefore, $\sigma_i(t, T)$ are the n volatility functions associated with the independent sources of risk $dz_i(t)$. In practice, n is usually set to n = 1, 2, or 3.

6.3.3 The Relationships Between Forward Curve and Spot Price Models

Intuitively, a model that describes the evolution of the whole forward curve is implicitly describing the front end of the curve, which is simply the spot energy price, and so, the forward curve models must be related to spot price models. The stochastic differential Eq. (6.12) can be integrated to obtain the following solution:

$$F(t,T) = F(0,T)\exp\left[\sum_{i=1}^{n}\left\{-\frac{1}{2}\int_0^t \sigma_i(u,T)^2\,du + \int_0^t \sigma_i(u,T)\,dz_i(u)\right\}\right] \tag{6.13}$$

This equation expresses the forward curve at time t in terms of its initially observed state (time 0) and integrals of the volatility functions. The spot price is just the forward contract for immediate delivery, and so, the process for the spot price can be obtained by setting $T = t$, that is:

$$S(t) = F(t,t) = F(0,t)\exp\left[\sum_{i=1}^{n}\left\{-\frac{1}{2}\int_0^t \sigma_i(u,t)^2\,du + \int_0^t \sigma_i(u,t)\,dz_i(u)\right\}\right] \tag{6.14}$$

Equation (6.14) can then be differentiated to yield the stochastic differential equation for the spot price:

$$\frac{dS(t)}{S(t)} = \left[\frac{\partial \ln F(0,t)}{\partial t} - \sum_{i=1}^{n}\left\{\int_0^t \sigma_i(u,t)\frac{\partial \sigma_i(u,t)}{\partial t}\,du + \int_0^t \frac{\partial \sigma_i(u,t)}{\partial t}\,dz_i(u)\right\}\right]dt + \sum_{i=1}^{n}\sigma_i(t,t)\,dz_i(t) \tag{6.15}$$

The term in square parentheses in the drift can be interpreted as being equivalent to the sum of the deterministic riskless rate of interest $r(t)$ and a convenience yield $d(t)$, which, in general, will be stochastic. Since the last component of the drift term involves the integration over the Brownian motion, the spot price process will, in general, be non-Markovian—that is, the evolution of the spot price will depend upon its past evolution.

One special case of the general model is the simple single factor model described by Eq. (6.10). For this model, $n = 1$ and $\sigma_1(t,T) = \sigma e^{-\alpha(T-t)}$. Clewlow and Strickland (2000) evaluate Eq. (6.15) with this volatility function and show that the resulting spot price process is given by:

$$\frac{dS(t)}{S(t)} = \left[\frac{\partial \ln F(0,t)}{\partial t} + \alpha\left(\ln F(0,t) - \ln S(t)\right) + \frac{\sigma^2}{4}\left(1 - e^{-2\alpha t}\right)\right]dt + \sigma\,dz(t) \tag{6.16}$$

This implies:

$$\frac{dS(t)}{S(t)} = \left[\mu(t) - \alpha \ln S(t)\right]dt + \sigma dz(t) \tag{6.17}$$

where,

$$\mu(t) = \frac{\partial \ln F(0,t)}{\partial t} + \alpha \ln F(0,t) + \frac{\sigma^2}{4}\left(1 - e^{-2\alpha t}\right)$$

This single factor forward curve model is therefore just the single factor Schwartz (1997) model with a time-dependent drift term. It is this term in the drift that allows the model to now fit the observed forward prices. Note also that this particular form of the forward curve volatility function results in a 'Markovian' spot price process, as the dependence in the drift on the path of the Brownian motion disappears.

The relationship between the forward curve model and the spot return model also shows that the mean reverting behaviour of the spot price is directly related to the attenuation of volatility of the forward curve. By setting $\alpha = 0$, the Black (1976) model is obtained. This is, therefore, a special case of the general model in Eq. (6.12) with $\sigma(t,T) = \sigma$ and $n = 1$. The main advantage of the forward curve modelling approach is the flexibility that the user has in choosing both the number and form of the volatility functions. These can be chosen in one of two general ways—historically from time series analysis or implied from the market prices of options.

6.4 Convertible Bonds

Convertibles are hybrid derivative securities that combine the characteristics of both bonds and stocks, and include options on the issuer's common stock and debt. A convertible gives the bond holder the right to exchange or convert the bond's par amount for the issuer's common shares at a fixed rate during a specified time period. Convertible bonds can be callable by the issuer on specified dates over the life of the convertible, with the call option decreasing the value of the convertible.

Convertibles can also contain puts where the buyer can put the bond to the issuer, with the put option increasing the value of the convertible. When the stock price is relatively low when compared to the conversion price, the convertible is unlikely to be converted into stock, and is therefore effectively a straight bond. When the stock price is relatively high when compared to the conversion price, the convertible is more likely to be converted into stock, and the convertible price is therefore the conversion parity, or the stock price multiplied by the conversion ratio.

The motives for financing through the issuance of convertibles include delaying equity financing until growth has been realized, and financing when the debt markets is not accessible, while for investors, convertibles can offer a higher yield than common stock dividend yields and the potential upside in the firm's growth and stock conversion.

Figure 6.5 illustrates a convertible bond price as a function of the stock price. The stock price is on the horizontal axis, the conversion ratio on the vertical axis, and the horizontal line represents the bond floor. The bond floor is the equivalent of the market value of a fixed income bond, where the coupons and redemption value are discounted at an interest rate that reflects the credit quality of the issuer. The conversion ratio is the number of ordinary shares at which the bond's notional value is converted, is established at issue and typically stays constant over the life of the convertible bond. The conversion price is the bond's nominal value divided by the conversion ratio. The diagonal line represents parity, derived as the conversion ratio multiplied by the share price, and represents the market value of the shares received at the conversion of the bond.

The option to convert into stock or retain the bond implies that the convertible's value should be the minimum of that of the stock or the bond. The stock and bond's minimum values therefore function as lower bounds on the

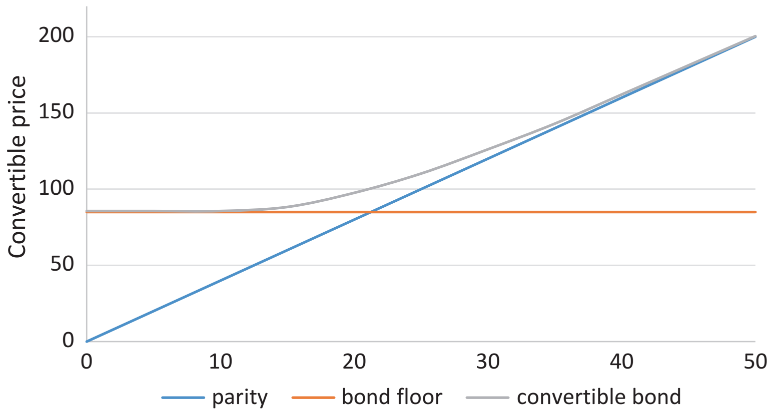

Fig. 6.5 A convertible bond price as a function of the stock price

convertible bond price. The conversion value is the equivalent to a call option on the stock, with the market value of the option to convert reflected in the difference between the bond floor and the convertible value. The conversion becomes more valuable, and the price of the convertible bond will increase as the stock volatility increases, with the conversion line in Fig. 6.5 becoming less convex.

Convertible bonds are typically priced with binomial lattice trees that include the bond's embedded option at each tree node. The model assumes that the convertible bond's value is a function of the underlying stock price volatility. The factors that can influence a convertible bond's value include the market parameters, the terms in the prospectus and their behaviour. A binomial one factor model was used for a simple illustration of a convertible bond's theoretical value. The Cox, Ross and Rubinstein (1979) approach was used for the binomial tree. See Sect. 5.5.2 for background on the binomial model.

The following variables are used to illustrate the convertible bond:

- the bond notional value is 100,
- the bond coupon is 10%
- the convertible bond maturity (T) is 5 years
- the conversion ratio (m) is 4
- calls of \$107.5 in year two, which decline by \$2.5 every year to maturity
- the current stock price is \$25
- the volatility is 20%
- the risk-free rate is 5%

A yield to maturity or flat term structure is assumed for the risk-free interest rate. The bond coupon and interest rate are compounded annually, while the volatility can be estimated using historical volatility.

Figure 6.6 illustrates the convertible binomial lattice tree. The convertible bond payoff at each node at maturity is derived as the maximum of the bond redemption value and coupon, and the binomial tree's stock price $S_{T,j}$ multiplied by the conversion ratio.

$$P_{T,j} = \max\left(mS_{T,j}, 100 + coupon\right) \tag{6.18}$$

The convertible price at each node is then derived recursively to the valuation date (t_0) as:

$$P_{i,j} = \max\left[mS_{i,j}, \min\left(\mathrm{e}^{-r\Delta t}\left(pP_{i+1,j+1} + (1-p)P_{i+1.j-1}\right), C\right)\right] \tag{6.19}$$

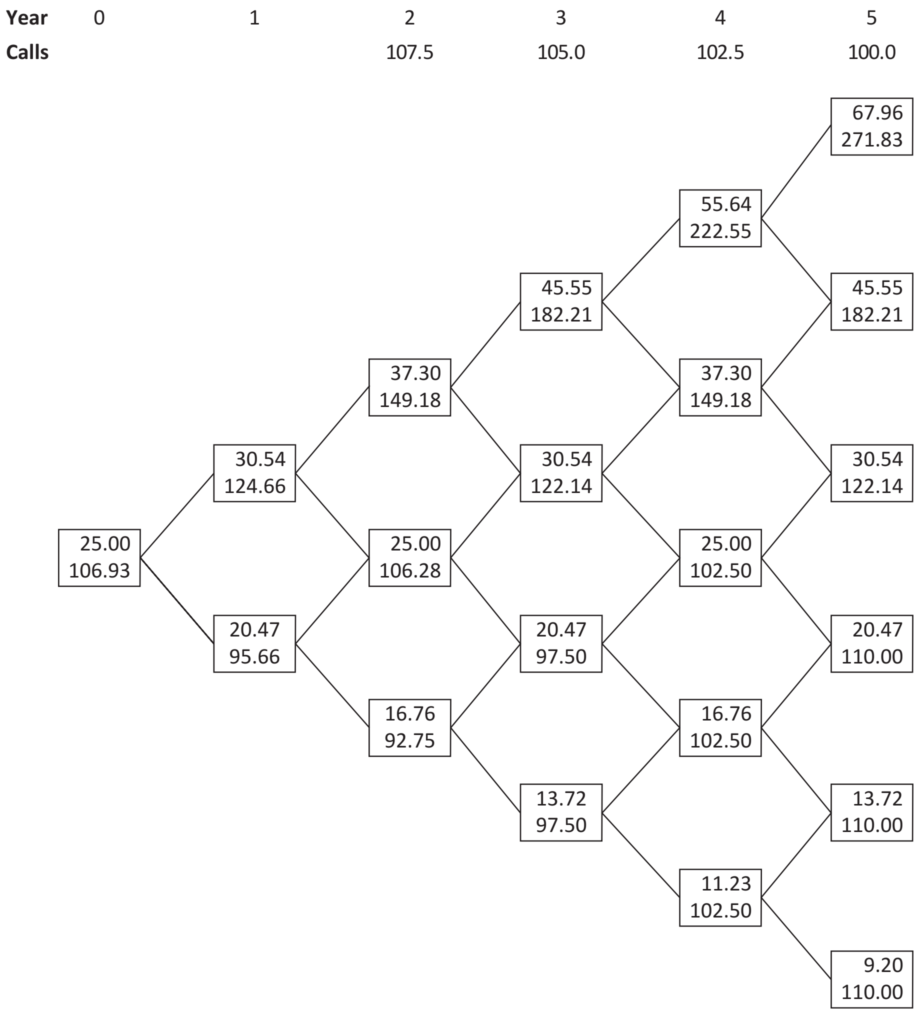

Fig. 6.6 The convertible binomial lattice tree

with the process continuing up the valuation date. The value of the convertible in the example is calculated as $106.93.

This is a relatively simple example of the valuation of a convertible bond. The valuation can be extended to include stochastic interest rates and the probability of default.

Market factors influence the behaviour of the theoretical value of convertibles in a number of ways. A convertible's value rises along with the parity of the underlying stock, as conversion is more probable. The value of a convertible also increases along with the volatility of the underlying stock, as the option value to convert the bond to stock is larger and near-the-money.

A convertible's value also increases with the put option held by the owner of the convertible. A higher put level has a larger value due to the protection provided by the put in declining markets, and is greater at lower parity levels. The increase in value as a result of puts is also more evident with higher interest rates, as the convertible's fixed-income value is lower.

A convertible's value will decrease as interest rates rise, with the rate of decrease larger as the maturity of the convertible increases. The rate of decrease, however, is less than an equivalent conventional bond due to the offsetting influence of the conversion option, which rises in value as interest rates increase. A convertible's value also decreases as the credit spread increases, with the sensitivity to the credit spread increasing with longer maturities.

Call options also decrease the convertible value, with lower call levels providing the issuer a larger probability for early conversion and reducing the conversion option's time value. The decrease in value with calls is higher at higher parity levels. The decrease in a convertible's value with calls is also more evident with lower interest rates, as the convertible's fixed-income value is larger.

6.5 Compound Options

Compound options, or options on options, where the payoff is another option, allow the holder to buy or sell another option for a fixed price. There are four compound option types—a call on a call, a put on a put, a call on a put, and a put on a call. Projects and investments that are staged as a sequence are compound options, where the initial investment cost is the exercise price for the subsequent option on the next stage of the investment. Plant development, product development and research and development are examples of sequential compound options. Compound options are useful for analysing the impact of an investment on a firm. Many projects and investments are not independent, as assumed in a discounted cash flow (DCF) analysis, but are a series of interrelated cash flows where the initial investment is a prerequisite for the following outlays.

Geske (1979) developed the original closed form solution for a compound option as a call option on a firm's equity, which itself is a European call option on the total value of the firm. The compounding in this specification occurs simultaneously, as the firm's equity, a call option on the leveraged value of the firm, and the call option on the equity appear at the same time. Both simultaneous and sequential compound options can be solved in trees, although the valuation progressively more complicated as more options are added.

The Geske compound option model is specified as a call option on a stock which itself is an option on the firm's assets. The functional representation of this relationship is:

$$C = f\left(S,t\right) = f\left(g\left(V,t\right),t\right) \tag{6.20}$$

where C is the value of a call option, S is the firm's stock and V is the value of the firm. Transformations in the call option value are therefore defined as a function of transformations in firm value and time. The Geske model transforms the option's underlying state variable from the firm's stock to the firm's market value (V), or the total market value of the firm's equity and debt. The Geske specification therefore provides a measure of firm value when applied to listed options.

The Geske model is specified as:

$$\begin{aligned} C = {} & VN_2\left(h_1 + \sigma_v\sqrt{T_1 - t}, h_2 + \sigma_v\sqrt{T_2 - t}; p\right) \\ & - M\mathrm{e}^{-r_{F_2}(T_2 - t)}N_2\left(h_1, h_2; p\right) - K\mathrm{e}^{-r_{F_1}(T_1 - t)}N_1\left(h_1\right) \end{aligned} \tag{6.21}$$

where,

$$h_1 = \frac{\ln\left(V / V*\right) + \left(r_{F_1} - \frac{1}{2}\sigma_v^2\right)\left(T_1 - t\right)}{\sigma_v\sqrt{T_1 - t}}$$

$$h_2 = \frac{\ln\left(V / M\right) + \left(r_{F_2} - \frac{1}{2}\sigma_v^2\right)\left(T_2 - t\right)}{\sigma_v\sqrt{T_2 - t}}$$

$$p = \sqrt{\frac{T_1 - t}{T_2 - t}}$$

$V*$ represents the firm's critical total market value, where the firm's stock level S_{T1} is equal to the option strike K. S_{T1} is derived using Merton's definition of the Black–Scholes model, where a firm's stock is the equivalent to an option:

$$S = VN_1\left(h_2 + \sigma_V\sqrt{T_2 - t}\right) - M\mathrm{e}^{-rF_2(T_2 - t)}N_1\left(h_2\right) \tag{6.22}$$

Therefore, at $t = T_1$ when $S_{T1} = K$

$$S_{T1} = V_{T1}^{*} N_1\left(h_2 + \sigma_V \sqrt{T_2 - T_1}\right) - M\mathrm{e}^{-rF_2(T_2 - t)} N_1\left(h_2\right) = K \tag{6.23}$$

where h_2 is defined as given earlier. The variable M is the face value of a firm's debt, while T_2 represents the debt's duration. The addition of the M term to the Black–Scholes model reflects the effects of leverage, where leverage changes the firm's equity volatility. The Black–Scholes model assumes that a firm's equity volatility is not a function of the level of equity. The Geske model, however, considers that a firm's equity volatility has an inverse relationship with a firm's stock level. As a firm's stock level increases, the firm's leverage and stock volatility will fall, and the inverse of this relationship also holds. The Geske model also implies the volatility of a firm's total market value, conforms to Miller and Modigliani, and is the equivalent to the Black–Scholes model when the firm has no debt.

A summary of the Geske model variables follows:

C = current market value of a firm's stock call option
S = current market value of the firm's stock
V = current market value of the firm's securities (debt + equity)
V^* = the critical total firm market value where $V \geq V^*$ which implies $S \geq K$
M = face value of market debt (debt outstanding for the firm)
K = strike price of the option
r_{Ft} = the risk-free rate of interest to date t
σ_V = the instantaneous volatility of the firm market value return
σ_s = the instantaneous volatility of the equity return
t = current time
T_1 = expiration date of the option
T_2 = duration of the market debt
N_1 (.) = univariate cumulative normal distribution function
N_2 (…) = bivariate cumulative normal distribution function
ρ = the correlation between the two option exercise opportunities at T_1 and T_2

Refer to Chap. 10 for an example of an application of the Geske compound option model.

6.6 Model Risk

Over the last 50 years, there has been a huge growth in the use of theoretical models for valuation and pricing in financial markets. A large body of the theory relates to derivatives, financial instruments where value is derived from underlying assets. These theories have been extended into real options, where models have been developed for options on real assets. Relying on models to analyse and quantify value and risk, however, carries its own risks. The term *model risk* has many connotations and is used in many different contexts. The following is based on Rebonato (2001) definition. Model risk is the risk, at some point in time, of a significant difference between the modelled value of a complex and/or illiquid asset and the realized value of that same asset.

In the physical sciences, where quantitative modelling originated, predictions can be made reasonably accurately. Variables in physical science models such as time, position and mass exist, regardless of the existence of humans. The fundamental unknown in financial markets, however, is certainty. Many financial and real assets only trade at certain discrete times, while financial variables also only symbolize human expectations. Risk and return refers to expected risk and return, variables that are unobserved and not realized. In most circumstances, however, models based on financial concepts and theory assume causation and stability between the values of these unobserved variables and asset values.

There a number of ways in which the development of a financial model can go wrong:

- The most fundamental risk is that modelling is just not appropriate. Modelling requires knowledge and context within a discipline. Mathematics is a representation or an abstraction of a discipline, and is a means to an end and is not the end itself.
- All the factors that affect valuation may not have been included in the model.
- Although a model may be theoretically correct, the model variables such as forward prices, interest rates, volatilities, correlations and spreads may be poorly estimated. A model's variables, for example, may be based on historical data from a past regime, and therefore, not provide a good estimate of future value.
- Incorrect assumptions can be made about the properties of the asset values being modelled and the relationships between the variables in a model.

- A model may be inappropriate in the existing market environment, or some of the assumptions such as the distributions of variables may not be valid. Even if a model itself is satisfactory, the world it is predicting may be unstable.

Financial modelling draws on a multitude of disciplines—from business management and financial theory to mathematics and computer science—and is as much art as it is theory and quantitative techniques. An intimate knowledge of markets and how market participants think about valuation and risk are also part of the model practitioner's skill set. Derman (1996) provides some procedures for constructing financial models:

- Identify and isolate the most important variables used by market participants to analyse value and risk, and decide which variables can be used in mathematical modelling.
- Separate the dependent variables and the independent variables.
- Determine which variables are directly measurable and those that are more in the nature of human expectations, and so, are only indirectly measurable.
- Specify which variables can be treated as deterministic and those that must be considered as stochastic. Uncertainty will have little effect on the future values for some variables and these therefore can be approximated. For other variables, however, uncertainty will be a critical issue.
- Build a quantitative picture that characterizes how the dependent variables are influenced by the independent ones.
- Determine how to obtain the market values of independent observable variables, and how to derive the implied values of indirectly measurable ones.
- Create a mathematical picture of the problem, and determine which stochastic process best describes the evolution of the independent stochastic variables. Determine whether an analytical or numerical solution is appropriate.
- Deliberate the issues and difficulties in solving the model, and simplify it if necessary to make the solution as easy as possible. Only give up substance, however, for a relatively easy or elegant analytical solution when it is absolutely necessary.
- Finally, programme the model, test it and apply to the valuation problem.

The application of financial modelling draws from a palette that includes knowledge of the markets, the applicability of the financial model, the relevance of the mathematics used to solve the problem, the systems and software used to implement and present it, and the accurate communication and dissemination of the information and knowledge gained from the analysis. Drawing from these various disciplines can address the issues and reduce the risks associated with the application of financial modelling.

Bibliography

Black, F. The pricing of commodity contracts, The Journal of Financial Economics, 3, (Jan–March): 167–79, 1976.

Carr, P. Wu, L. Stochastic Skew Models for FX Options, 2004.

Clewlow, L. and Strickland, C. Energy Derivatives, Pricing and Risk Management, Lacima, 2000.

Clewlow, L. and Strickland, C. Implementing Derivative Models, Wiley, 1998.

Copeland, T. and Antikarov, V. Real Options, Texere, 2001.

Cox, J.S., Ross, S. and Rubinstein, M. Option pricing: a simplified approach, Journal of Financial Economics, 7 (September): 229–63, 1979.

Derman, E. Quantitative Strategies, Research Notes: Model Risk, Goldman Sachs, 1996.

Dixit, A. and Pindyck, R. Investment Under Uncertainty, Princeton University Press, 1994.

Geske, R. The valuation of compound options, Journal of Financial Economics, 7: 63–81, 1979.

Geske, R. and Zhou, Y. Predicting Risk and Return of the S&P 500: Evidence from Index Options, UCLA Working Paper, 2007.

Geske, R. and Zhou, Y. Capital Structure Effects on Prices of Firm Stock Options: Tests Using Implied Market Values of Corporate Debt, UCLA Working Paper, 2007a.

Goldman Sachs. Valuing convertible bonds as derivatives, Quantitative Strategies Research Notes, November 1994.

Haug, E.G. The Complete Guide to Option Pricing Formulas, McGraw Hill, second edition, 2007.

Heston, S. L. 'A closed-form solution for options with stochastic volatility with applications to bonds and currency options', The Review of Financial Studies 6(2), 327–343, 1993.

Hull, J. and White, A. An analysis of the bias in option pricing caused by stochastic volatility, Advances in Futures and Options Research, 3: 29–61, 1988.

Merton, R. Option pricing when underlying stock returns are discontinuous, Journal of Financial Economics, 3: 125–44, 1976.

Moodley, N. The Heston Model: A Practical Approach with Matlab Code, 2005.

Natixis Asset Management. Investing in global convertible bonds – stylized facts, pricing and strategies, research paper number eight, January 2017.

Rebonato, R. Model risk: new challenges, new solutions, Risk, March 2001.
Schwartz, E. The stochastic behavior of commodity prices: implications for valuation and hedging, Journal of Finance, LII(3): 923–73, 1997.
Vasicek, O. An equilibrium characterization of the term structure, Journal of Financial Economics, (5): 177–88, 1977.
Zhou, Y. Pricing Individual Stock Options on Firms with Leverage, Anderson School of Management at UCLA, 2007.

Part III

The Analysis of Investments, Transformation and Value

Overview

The following five chapters discuss industry dynamics, technology trends, quantitative and statistical applications, the fundamental technologies that are having an impact across industries, and various valuation methodologies. Valuation is the process of determining the worth of an asset, liability or a firm, or an economic relationship derived from an underlying asset or liability. Valuation concepts have an influence on investments, capital structures and dividend policies, and are also used in the evaluation of business models and strategies, capital budgeting, financial reporting, merger and acquisitions, intellectual property and licensing transactions.

Firm value can have a number of interpretations:

- *Book value,* or net asset value, is an asset's value on a firm's balance sheet, recorded at the original cost minus any depreciation, amortization or impairments. By convention, a firm's book value is derived as total assets less intangible assets and liabilities.
- *Break-up value* is the total of the business segment values of a firm.
- *Economic value* represents the ongoing value of a firm, with value defined as the expected future cash flows generated by a firm that flow to its owners over its economic life.
- *Liquidation value* is the value of a firm's assets when it goes out of business, or the current liquidation value for a firm or its parts.
- *Market value* is a firm's market capitalization.

- *Enterprise value* is the sum of a firm's market capitalization, market value of debt, minority interests and preferred shares, less total cash and cash equivalents. As enterprise value is capital structure neutral, it provides a measure for the comparison of firms with diverse capital structures.

These concepts are considered in the following five chapters. Platforms, data and analytics provides a theoretical finance background on platforms, a background on artificial intelligence, quantum computing, data and analytical management, and IT investment and patent options value concepts. Energy provides an overview of the energy sector trends, energy statistics and a power generator valuation example. Pharmaceuticals and biotechnology provides a sector overview and the valuation of a drug development using net present value (NPV), expected net present value (ENPV) and real options methods. A growth firm illustrates the residual earnings and growth options concepts and methods. Finally, firm transformation examines the exit from a firm's existing operations and the abandonment option so as to align its business model with the digital economy.

7

Platforms, Data and Analytics

7.1 Platforms

Platforms now play a significant role in the digital economy, with platform firms reaching the scale and dominance last seen with the huge vertical corporations of the early 1900s. Platform business models have broad implications for firms across all industries seeking to leverage information technology (IT). Digital technologies are increasingly being integrated into strategy frameworks, with platforms and platform complements providing the potential for future growth.

A platform generally is a cluster of technologies that serve as a base for the development of products, applications, processes or various technologies within the same structural framework. Platforms are also business frameworks that facilitate multiple business models that can be developed into operations and systematically leverage technologies across various domains. Although platform terminologies exist across product development, technology strategies and economics, all have common attributes, the repeated use of a core component to both achieve economies of scale and the lowering of costs in the generation of a broad range of complementary products.

A platform is defined as a group of components that remain stable, while facilitating diversity and evolution within the system through the constraint of the relationships between the other components. The architecture of platforms can be generalized as the modulization of complex systems, where the platform components remain constant while the complements are free to change either as a cross-section or sequentially.

J. Rogers, *Strategy, Value and Risk*, Global Financial Markets,
https://doi.org/10.1007/978-3-030-21978-9_7

Platforms also facilitate firm activities within ecosystems, which can take the form of product lines within a firm's boundaries, multi-product systems across co-dependent firms, and as multiple-sided markets. Business ecosystems are networks of independent firms whose products have more value when used collectively than individually, and are defined by the complements between products within the system. Individual firms can partake in multiple ecosystems, while firms from different industries can participate in an ecosystem. The coordination within an ecosystem is maintained through a combination of standards, contracts and prices.

Platform and platform complements, where demand for a product increases when the price of an alternative decreases, as a strategy contrasts to that of a conventional product strategy. An ecosystem is required for the creation of complementary products and services, and to generate positive network externalities between the platform and the complements. These interactions are significant for firms in which platforms play a major role, as they can generate growth at a much greater rate than that of an individual firm. These complements within a business ecosystem can also be supermodular, where an investment by one firm enhances the value of investments made by other firms.

Technology platform ecosystems are especially relevant in software, where they provide the leverage and optionality for large-scale specialization, either within a value chain or across a wider ecosystem. Short lifecycles require specialization, and the product architectures a high level of modularity, the level to which the components within a system can be divided and recombined and provide flexibility and diversity.

A digital platform is a computing environment for software execution that can include hardware, an operating system, web browsers and associated application programming interfaces and any other software. Digital technologies have two attributes that are distinct from the physical flow and step process technologies that were dominant at the turn of twentieth century. The first is that a computer is a composite system that consists of software and distinct functional components and a platform that supports an array of discretionary options and complements that generate value. The second is that the integrated circuits (IC) that provide the foundation for computer hardware have scaling properties, which are enhanced by the IC size, cost and chip speed improvements that occur with each new generation. Advances within the various component technologies of a computer system also occur independently and at different rates of development. All these attributes combined provided the flexibility to customize computer systems and their component upgrades.

Value within a digital platform system therefore differs significantly from that within a step process. All steps within a step process are critical, with the risk of production bottlenecks hindering the flow and throughput of the overall system. Each step is also interdependent, with any variation across the step processes affecting value. Steps processes therefore required a hierarchical management structure that bridged the steps to reduce the risk of bottlenecks. The value of a step process is therefore proportional to its output.

In contrast, value within a platform system is a function of the sum of its optional complements. The options provide the platform user the right and not the obligation to undertake an activity that creates value for the user. Platform users can determine whether value is greater than the cost for each component, and can add the complement to the system if this is the case. The value of the platform system is therefore proportional to the sum of the individual option values. The greater the number of options within the platform system, the greater the value to individual users. The users and options together are therefore supermodular, where a user's decision influences the incentives of other users, and also their complements.

A platform has no value without the combination of one or more complements. Complementors have value when there is optionality, or the right but not the obligation to choose one complement over an alternative within the platform complementary modules, units within a system that, while are structurally independent, can function in combination. The value of the option is low when consumer preferences are homogeneous and predictable, and high when preferences are heterogeneous or unpredictable, and when there is uncertainty around future technology paths.

Complementors can also provide a source of investment capital for the platform sponsor and facilitate platform growth, especially when there are significant network externalities. Platform sponsors also have to initially build scale on one side to attract participants and third parties on the other. The network externalities within these participants create the chicken or the egg problem, whether to initially build the platform's supply side or the demand side. Solutions for platform sponsors to address this issue include cross-subsidization and providing free products and services.

Platforms are diverse in terms of their economies of scale, network effects, investment approaches, openness, growth subsidies and means of monetization. Industrial platforms initially were closed manufacturing systems, where firms designed and engineered their platform and complements to support the product variety and flexibility required to respond to demand dynamics. The rise of open platforms led to the vertical to horizontal transformation of the computer industry towards the end of the twentieth century. Open platforms not only influenced firms within the computer industry, they also transformed the structure of the industry itself.

Two types of digital technology open platforms and related ecosystems emerged during the last two decades of the twentieth century. Open product platforms distribute the design and production of the various complex modular system components over multiple independent firms. Open exchange platforms, which originated with the rise of the Internet and World Wide Web in the 1990s, emerged as websites structured to facilitate transactions in goods and services, information and opinion. The sponsors of open platforms in both cases can allocate essential tasks to third parties while maintaining control over critical and unique platform components.

Product platforms can be further defined as standards-based product platforms and logistical product platforms. Standards-based product platforms establish standards that facilitate the ability for the system components to operate within the platform, and therefore, the design of complex systems for the production of goods and services. Logistical product platforms coordinate the movement of products and services via a system integrator that manages a complex network of step processes. Standards-based and logistical open product platforms are both contingent on ecosystems of various suppliers of components and tasks to produce a final product.

Exchange platforms differ from product platforms as they enable exchanges between agents, and are defined as having two- to *n*-sided markets. Value within an exchange platform, as with product platforms, is a function of the options it facilitates. Exchange platforms enable the efficient connection and transfer of goods and services and information through the medium of the platform.

Exchange platforms can be further split into two sub categories. Transaction exchange platforms support transactions in markets between buyers and sellers to exchange property rights for payments. Digital transaction platforms are one example, where there are two sides—the buy side and the sell side. Communication exchange platforms support the exchange of information and opinion, and can be structured as point-to-point operations such as email, or broadcast through television, radio, newspapers and social media. The senders and receivers of messages comprise the two principle sides within these platforms, while also facilitating the ability of agents to sell advertising and collect data.

Open and closed platforms have a number of similarities. Both have as a foundation the modularization of its core platform and its optional complements. Modular design, or modularity, subdivides a system into smaller parts or modules that can be created independently for use in various systems.

Platform systems also have the upside of numerous options, which include network effects, risk, modularity, the complementary between modularity and risk, and also facilitate the decentralization of tasks and decisions. The fundamental property that underlies all platforms is that they generate options, or the right but not the obligation to transform the product in response to new consumer demand, prices and technologies.

Platforms and their related ecosystems have the potential to redefine industries, business models and the generation of value. Platform technologies include cloud computing, which encompasses data, software, analytics and artificial intelligence (AI), and new technologies such as DNA foundries, fintech, robotics and the industrial Internet. Platforms will also converge technologies such as energy storage, electric vehicles, AI and robots, and ultimately, drive the transformation of industries and the realignment of industry boundaries.

Platforms also provide a foundation for non-proprietary call options on its functionality, which are linked to platform value and growth. The platform options have no downside risk, with more options and upside risk increasing value and growth for the platform sponsor. The division of the platform components into modules also increases value, as modularity provides more options to the platform. Finally, platform sponsors can use patents as options to control standards, to scale and to achieve network effects.

7.2 Cloud Computing

Analytics, data and cloud computing together will have a significant impact on business and the economy in general. Cloud computing systems (the Cloud) are technology platforms that provide the modules, services and the functionality that can be used repeatedly across numerous information technology (IT) applications. The Cloud also provides access to a portfolio of scalable IT services on demand and as required by firms. Third-party developers can embed Cloud modules and services into their own products, which facilitates further adoption of the Cloud platform. Cloud computing platforms, by design, do not connect the third parties and are therefore not two-sided markets, and have no exponential network effects, with growth a function of adoption by third parties.

Cloud computing service offerings can be generalized as three standard models—Software as a Service (SaaS), Platform as a Service (PaaS) and Infrastructure as a Service (IaaS). SaaS is a software distribution model, in contrast to on-site software delivery, where software is centrally hosted and accessed over the Internet through devices that connect to Cloud applications, and is typically a pay-as-you-go service. PaaS provides a development and distribution environment on the Cloud, and the ability to develop scalable product and service applications as a pay-as-you-go service. IaaS provides real-time virtualized computing, storage and networks resources over the Internet that allows clients to build information technology infrastructure that scales with demand with a pay-as-you-go structure.

Cloud computing, storage and software offer technology infrastructure to clients that has cost flexibility and an operating expense as opposed to a capital investment. Cloud computing users include start-ups and SMEs (small and medium-sized enterprises) that are scaling, large corporations migrating legacy systems to the cloud, and private equity firms with investment portfolios that have technology and cost optimization business strategies. All business types can also benefit from an asset-light business model, and a significant decrease in information technology costs generally. It will take time for the migration of corporate information technology to the Cloud to fully materialize as firms write down the sunk cost of their existing legacy systems. Given the size of spending on information technology in general, however, the potential for the Cloud to transform business models and operations is enormous.

7.3 Artificial Intelligence

Artificial intelligence (AI), a subfield of computer science, is the ability of a system to interpret and act on data so as to achieve particular objectives. These objectives can include speech recognition, visual perception, language translation and decision-making. AI has the potential to be as transformative as the impact of computer technologies, with the exponential growth in data and the continuing developments in computing and algorithms driving the world's dominant technology firms to invest substantial sums in their AI capabilities.

The transformative impact of AI is the next phase to follow that of databases, which have continually lowered the cost of storing data since the 1980s. While developments in the first generation of software were driven by developments in databases, so will AI drive the next generation of analytics and prediction software. As with the personal computer and mobile phone technology waves, AI has the potential to redefine the technology industry's business models in both operations and new business ventures.

Cloud computing will be structured into two AI businesses—one that focuses on consumer services and the other on business and corporates.

Machine learning will be offered to firms across industry sectors that do not have the capabilities to build and scale AI through internal technology platforms. AI offers new capabilities to firms across industries that are facing transformative technologies in transportation, energy, manufacturing, biotech and media. Global personal transportation, for example, is worth approximately $10 trillion, and those that succeed in this market will also have spill-over synergies in other AI-related technologies, such as robots and drones. Autonomous vehicle AI investments are also an example of how technology firms are moving beyond software into hardware.

The larger issue, however, is whether AI will further concentrate the market power of the world's dominant digital technology firms. These firms—through their resources and capabilities in data, computing, algorithms, human assets and investments—are likely to capture much of the value in AI market share. As with the rise of databases, software and personal computers in previous technology waves, the likelihood of the transformation of industries, business models, industry boundaries, the decline of incumbents and the concentration of dominant firms is significant.

7.4 Quantum Computing

While full-scale quantum computing is unlikely in the foreseeable future, it will have specific applications as the result of significant advances in the technology, which are likely to be ready much earlier than previously forecast.

The foundation of quantum computing is the behaviour of sub-atomic particles. One property of this behaviour is superposition, where a particle can be in two states simultaneously. Quantum bits, or qubits, can have a value of one and zero at the same time, in contrast to the binary elements, or bits, in current computers, which can only have a value of either zero or one at the same time. Through the threading together of numerous qubits, the number of states represented grows exponentially, and provides the ability to calculate potential outcomes in the millions instantaneously.

A second property of quantum computing is entanglement, in which qubits are able to be in two states simultaneously in a manner that can lead the qubits to act in unison. This second property makes it possible to code algorithms that bypass the current computing sequential logic process, and arrive at results by excluding those that are incorrect at a much faster rate than current computers.

As large-scale qubits appear to be unattainable for at least a decade, the question is whether there are practical applications for the current state of the technology. Researchers are focusing on three areas in which quantum computing is expected to be relevant in the short term.

The first concerns the analysis of the natural environment and the modelling of the behaviour of molecules, as nature itself is quantum mechanical, and therefore aligns with quantum computing principles. Quantum computing can model sub-atomic particle behaviour exactly, while current computing can only arrive at approximations of this behaviour. One sector that can benefit from this technology is the chemicals industry, with computational chemistry facilitating the discovery of new materials.

The second is machine learning, where quantum computing is uniquely applicable to specific problems that can be an issue in current computing. Quantum computing is particularly relevant to specific types of probability-based algorithms, computations that do not follow a logical deterministic step sequence as derived in classical computing.

The third area is complex optimization problems that have variables that are too numerous for current computer processing. While these quantum computing applications appear to be narrow in scope, in combination, they have the potential to address a wide range of problems, making the technology generally applicable to specific areas in the short term.

Other applications of quantum computing include developing cryptography systems to replace the current encryption technology, improved financial risk management systems for the calculation of financial exposures, potential losses and adjusting portfolio risk, and in renewable energy by improving system efficiency and developments in solar cell materials. Quantum computing will also become a component of the dominant tech firms' cloud computing offerings, where scale and capabilities are especially relevant within specific applications.

7.5 Data

Big data is defined as a broad range of new and massive data sets from which new forms of value and insights can be extracted. The data is classified as unstructured, structured and semi-structured data. Unstructured data is information that does not have a defined data model, and includes text, social media, natural language, digital images, communications, science data, health-related data and search data. Structured data is information that is separated into standardized components within a defined data model, which can range from individual data points, dates and text to data that includes multiple data components. Semi structured data is a type of structured data that—while it does not conform to formal data models—does contain separate semantics, and record and field hierarchies, within the data.

Cloud computing is a technology that provides universal access to shared pools of resources and services, typically over the Internet. This common sharing of resources and services can achieve economies of scale and a variable cost for a firm's technology infrastructure. Data as a service is one of these services, where Cloud computing vendors can import large volumes of data, analyse the data and publish the results back to clients.

The economics of big data and machine learning algorithms has largely been a function of the centralizing of data-intensive processing on the Cloud, which has significantly driven down computing costs. AI is also moving towards another computing paradigm, where data processing is performed on a network 'Edge'. Edge computing is defined as devices that interconnect with the physical world, where applications, data and services are located away from central nodes to the edge of the Internet. Examples can include smartwatches, autonomous cars and devices connected through the Internet of Things. This architecture can reduce communication bandwidth to a central data centre by locating data and analytics in proximity to the data source, providing speed and the optimization of computer resources. This technology architecture has added another dimension to the economics of data and analytics, with local processing versus processing data on the Cloud further influencing data and analytics performance and the associated costs.

7.6 Analytical Management

Information has become one of the most powerful commodities in the world today. From the 1980s, the value derived in information technology moved from hardware to software as computers became commoditized. Today, value is being transformed again—in this instance, from software to data. Technology innovations will continue to drive the growth in the volume and types of data firms can access for analysis. Information systems management is increasingly being integrated into business strategy, with data availability improving as a result. This process is providing a rich source of business and financial data, much of which is proprietary to an organization, and which managers and investors can utilize for analysis.

The transformation of data into information and into value by firms that dominate the digital economy is having effects across all industries. Firms will require investments in data and analytics infrastructure, and in many cases, the transformation of the business model to remain competitive. Data and analytics in the digital economy are the new technologies of systematic management, and increasingly, a core capability and a source of competitive advantage and value. Analytics is about finding value in data, and deriving insights that can solve specific business objectives that align with strategy. As innovations in technology continue to add to the volume of data available as a resource, a range of advanced statistical methods are available for the analysis of a range of complex business problems.

Economics, finance, statistics and machine learning combined can provide an analytical framework for management to provide structure to complex business problems and to gain insights into a firm's internal and external environment. The analytical framework can be applied to a broad set of business domains that include product innovation, supply chain optimization, identifying financial drivers, sources of risk, consumer behaviour and advertising, customer profitability, optimal pricing and resource allocation.

The application of explanatory statistics can offer insights into a firm's enterprise and consumer behaviour, revenue, cost, profit and production functions, factor demand and other business relationships. These statistical techniques include structural model equations, discrete event simulation, dynamic simulation models and time series applied to continuous, discrete and categorical data. Supply chain optimization methods can be applied to inventory management, optimizing third-party supply chains and scalable Cloud solutions.

Predictive statistical learning techniques are classified as supervised and unsupervised. Supervised learning involves the building of a statistical model for the prediction or estimation of an output on one or more inputs. These problems can be found in business, medicine and public policy. Unsupervised learning has inputs that, while there is no supervising output, can however learn from the relationships within the data sets. Supervised methods include linear and logistic regression, additive models, LASSO, support vector machines and K-Nearest Neighbors, while unsupervised methods include Principal Component Analysis and Clustering Methods.

7.7 IT Investments

Investment in information technology (IT) has become a major component of the capital budget in many service and manufacturing organizations. Although IT investments can be highly risky, the corporate rewards can be enormous. Managing this capital expenditure in today's business environment raises a number of important issues for decision makers, including how IT investments should be integrated with strategy, and the risk management implications of these investments.

IT projects can be viewed as an activity where resources are allocated with the goal of maximizing shareholder value. Any IT investment such as software, platforms or data analytics is therefore managed with the goal of maximizing value, where value is defined in terms of the market value added to the firm. IT investments are not only costly, they are also often risky due to the uncertainty of the value of future payoffs.

The return on investment (ROI) concept applies principles from finance to maximize the value of IT investments and expenditures. In a conventional discounted cash flow (DCF) valuation analysis, a forecast of the future cash flows is discounted at the risk-adjusted opportunity cost of capital to obtain the present value. Present values are derived for both costs and benefits to obtain the net present value (NPV), and if the NPV is positive, the project is considered viable.

The NPV of a feasible project corresponds to the change in value of a firm if it proceeds with the investment. If the NPV is positive, the decision would be to proceed with the investment, as this would increase firm value. Likewise, if the NPV is negative, the investment should not be made. Valuing IT investments using cost benefit or ROI analysis, however, has always been a problem in computation. It is typically easier to calculate the costs of the investment than it is to calculate the benefits or returns. There are concepts such as partitioning benefit or returns analysis into tangible and intangible benefits that can be used. One approach is to ignore all intangible benefits and focus on tangible benefits, as most tangible benefits can be converted into measurable returns.

The uncertainties associated with a firm's IT investments can also be defined as project- and market-related risks. Project-related risks are associated with the planning, implementation and management of a firm's IT project, such as the technology not delivering, cost overruns and project setbacks. Market-related risks are the factors that can influence the demand for a firm's products and services, such as customer approval and the behaviour of competitors. Even if a project meets management expectations, any capabilities created by the IT investment may not be suitable for the existing market environment at the time of completion.

Other uncertainties that can be associated with software investments include development costs, coding issues and subsequent operational failures, developments in future technologies and standards, user acceptance, and the potential costs associated with changes in processes. Managers cannot always predict how systems will need to be adapted to changing user requirements, market developments and developments in technology.

A one-step binomial tree is used to illustrate the uncertainty associated with the IT investment decision. The payoff is represented as three cash flows, with the costs at t_0 of \$100,000, and two possible payoffs at t_1 (one step into the future) of \$198,000 at one branch if the outcome is favourable, and \$66,000 at the other branch if the outcome is unfavourable.

Arbitrary probabilities of 0.5 are assigned to each branch, and a discount rate of 10% is used. These are simplified parameters for the purpose of illustration. If the decision is made to invest immediately, the benefits are the expected value of the profit stream at time t_0. Over the time period Δt from t_0 to t_1, the benefits can either go up to uS (favourable) or down to dS (unfavourable), with values at t_1 defined as:

$$uS = 198{,}000$$
$$dS = 66{,}000$$

Table 7.1 illustrates the cash flows for the decision to invest immediately.

The return on the outlay of \$100,000 to invest immediately is \$20,000, and therefore, the IT investment would be accepted immediately based on this analysis.

The IT investment is now analysed using the concept that there is value in the form of a real option. This option offers the flexibility to defer any IT investments until one-time step into the future, at which point, new information will become available.

A binomial model is again used to illustrate the option to defer, or a call option on the IT investment at t_1, with a strike of $K = 110{,}000$ (the costs at t_1 equivalent to \$100,000 × 1.1^1).

At t_0, the payoff for the call option on the IT investment is:

$$c = Max\left(S_T - K, 0\right) \tag{7.1}$$

where c is the value of the call option, and T is equal to t_1.

The payoffs for the option at the up and down nodes at t_1 are:

$$c_u = Max\left(uS - K, 0\right) \tag{7.2}$$

$$c_d = Max\left(dS - K, 0\right) \tag{7.3}$$

Table 7.1 IT investment cash flows: invest immediately

Costs at t_0	= (100,000)
Favourable savings at t_1	= 198,000
Unfavourable savings at t_1	= 66,000
NPV at t_0	$-100{,}000 + 0.5 * \left(\frac{198{,}000}{1.10}\right) + 0.5 * \left(\frac{66{,}000}{1.10}\right)$
	= 20,000

The expected value of the call option at t_0 is:

$$c = P(t,T)\left[0.5 * Max(uS - K,0) + (1-0.5) * Max(dS - K,0)\right] \quad (7.4)$$

where $P(t, T)$ is the discount factor. The expected cash flow payoff at t_0 is therefore:

$$c = P(t,T)\left[0.5 * Max(198{,}000 - 110{,}000{,}0) + (1-0.5) * Max(66{,}000 - 110{,}000{,}0)\right]$$
$$= 40{,}000.$$

In this IT investment example, the alternative strategies are to either exercise at t_0 or wait until t_1. The manager has the right to invest immediately in the IT investment at t_0, and a call option on delaying the IT investment with exercise at t_1. The DCF analysis implies that the project should proceed immediately as the discounted cash flows are positive. The decision to proceed with the IT investment at time t_0, however, gives up the right to exercise the call option at t_1. Therefore, the value of the decision to invest immediately at t_0 is \$20,000 (the value of investing immediately) minus \$40,000 (the value of the call option, or the option to defer), which equals (\$20,000).

Although the initial insight is that the best alternative based on the NPV analysis is to invest at t_0, it is clearly not the optimal decision for adding value to the firm when the option to defer is also considered. If the DCF analysis had given a negative NPV for the project at t_0, the right to defer the project may still have value due to the call option. Even though the project had a positive NPV at t_0, the firm may gain by delaying the project and proceeding with it in the future. In the DCF analysis, this right is worthless and adds no value to the firm. When the right to delay the IT investment is considered as a call option, however, this right does have value and should therefore be considered in the analysis.

7.8 Patent Options

Intellectual property (IP) is a form of property right that covers intangible creations of the human intellect—such as literary and artistic works, designs, symbols, names, images and inventions—that are used in commerce and can be protected under law. There are two types of IP property rights—copyright, the granting of exclusive rights to a creator of an original work, and industrial property rights, which encompass trademarks, designations of origin, industrial designs and models, and patents.

A patent gives its owner the right to exclude others from making, using, selling or importing an invention for a limited time period without the authorization of the patent owner. A patent is therefore a limited monopoly granted by the government for the term of the patent, after which other parties can make, use or sell the invention once the patent expires. Patent rights are granted by government in exchange for the public disclosure of the invention, typically fall within the jurisdiction of civil law, and form an essential component of competitive advantage in some industries.

Patents can be used by firms as options—either internally or through licensing—that are exercised at some future stage to create and sell commercial products. The research and development (R&D) costs for a patent are typically significantly less than the actual product development for the majority of new technologies. A firm can, therefore, have an incentive to patent new technologies and not bring those patents to market. New technologies can be patented and either not used (or are sleeping), or licensed to others, and therefore used to maintain a dominant position. Sleeping patents can be used by a firm to either deter entry by investing in the R&D required to obtain a patent and letting the patent sleep as an option, or to license the patent to third parties, especially when there are network effects.

The sleeping patent can be viewed as a call option. Expenditure on R&D today that leads to a patent is effectively buying a call option on the subsequent technology that can be called in the future. A patent grants the right but not the obligation to either further invest and commercialize the technology or license the patent. The asset value is the NPV of the cash flows from either commercializing the technology or from its licensing. The R&D expenditures and commercialization investments—or annual patent fees in the case of licensing—are equivalent to the net exercise price on the patent option.

A significant percentage of the market valuation of high-tech firms is the optionality associated with their technologies and patents, and is a function of the firm's stock price volatility and option value. The stock price volatility can reflect the volatility of a patent value, and used as a proxy for the volatility of the patent returns and as a parameter in the valuation of patent call options.

Firms in R&D intensive industries can, therefore, hold patent portfolios and allow the majority to sleep, and either exercise and commercialize a small number of patents, or license the patent to other firms that do not have the capabilities for such R&D investments. The patent holder has two options embedded in a patent—the option to use it exclusively or to license it to third parties. The structure of patent licenses from the licensee's standpoint can also be framed as an option. A licensee can pay an initial fee, or premium, to buy an option when entering into a licensing contract, for the rights to develop and commercialize the technology covered by the patent.

The patent holder can also use it as an asset to establish standards in markets that have strong network effects, and grant the use of the patent by third parties. A firm can use a patent to support the adoption of a platform, encourage third-party developers to use the intellectual property to invest in the sunk costs required for the adoption of the platform, and therefore, use the asset to gain an edge for other parties.

A simple example follows to illustrate the value from licensing a sleeping patent. The sleeping patent is valued as a three-year European call option on the patent to attract third-party developers to a platform, with potential license fees from three to ten years. The intrinsic value for the patent call option is:

$$\text{Patent Call} = \text{Max}\left(S_t - X, 0\right) \tag{7.5}$$

and has the following parameters:

S_t = NPV of the potential licensing fees from three to ten years,
X = the patent capitalized R&D costs (assumed to be constant),
σ = volatility of the underlying asset,
r = the risk-free interest rate,
T = the call on the patent in three years.

The value of the patent call option with S_t = \$75 million, X = \$50 million, σ = 20%, r = 5% and T = 3 is therefore \$25.98 million. As the strategy is to use the patent to attract third-party developers to the platform, the patent call option value of \$25.98 million is capitalized, as the value is in the network effects of allowing the third parties to commit to the platform's functionality. In practice, an American binomial model would be used over the simple European Black-Scholes illustration used here to capture early exercise.

Bibliography

Baldwin, C. Y. and Clark, K. B. Design Rules: The Power of Modularity. MIT Press, Cambridge, MA, 2000.

Baldwin, C.Y. Design Rules Volume 2 - Chapter 14 - how technology shapes organizations - introducing open platforms and business ecosystems - HBS Working Paper 19-035 2018a

Baldwin, C.Y. (2018) "Introducing Open Platforms and Business Ecosystems", HBS Working Paper (October 2018b)
Baldwin, C.Y. Design Rules Volume 2, How Technology Shapes Organizations. Chapter 13, Platform Systems vs. Step Processes, The Value of Options and the Power of Modularity, HBS Working Paper 19-073 2018c.
Baldwin, C.Y. From step processes to platforms: A real options view of organizational design. Real Options Conference Keynote speech. Boston June 29 2017.
Baldwin, C.Y. Platform systems v Step processes - options and modularity. Chapter 13 Platforms vs Steps 6-15-18
Bridgwater, A. What's the difference between a software product and a platform? Forbes. March 17, 2015.
Clewlow, L. and Strickland, C. Implementing Derivative Models, Wiley, 1998.
Cotropia, C.A. Describing patents as real options, 34 J. Corp. L. 1127 (2009)
European Parliament. Intellectual, industrial and commercial property | Fact Sheets on the European Union. 2019.
Financial Times, Waters, R. Microsoft and Amazon face challengers from on edge of cloud, February 1, 2018a
Financial Times, Waters, R. Tech trends for 2018: the big will get bigger, Richard Waters, December 28, 2017.
Financial Times, Waters, R. The future of computing is at the edge, June 5, 2018b
Financial Times, Waters, R. Time to invest in skills for quantum computing revolution, May 17, 2018c.
Financial Times. Waters, R. Thornhill, J. Quantum computing the power to think outside the box, September 3, 2018.
From step processes to platforms: A real Options View of Org Design, C Baldwin RO Conference Keynote Speech, 6/29/17
Gilbert, R. J, Newbery, D. M. Preemptive patenting and the persistence of monopoly. American Economic Review, June 1982.
Greene, W.H. Econometric Analysis 8th Edition, Pearson, 2017.
James, G. Witten, D. Hastie, T. Tibshirani, R. An Introduction to Statistical Learning: with Applications in R (Springer Texts in Statistics), 7th Edition, 2017.
Kester, W. C. "Today's Options for Tomorrow's Growth." Harvard Business Review 62, no. 2 (March–April 1984): 153–160.
Kulatilaka, N., Balasubramanian, P. and Storck, J. Using real options to frame the IT investment problem, in Trigeorgis, L. (ed.) Real Options and Business Strategy, Risk Books, 1999.
Martin, K, Zysman, J. The Rise of the Platform Economy. Issues in Science and Technology 32, no. 3 (Spring 2016).
Pearson, N. An efficient approach for pricing spread options, The Journal of Derivatives, Fall, 1997.
Pitkethly, R. The valuation of patents - A review of patent valuation methods with consideration of option based methods and the potential for further research. Judge Institute Working Paper WP 21/97, The Judge Institute of Management Studies. March 1997.

Reiss, P, C. Wolak, F, A. Structural Econometric Modeling: Rationales and Examples from Industrial Organization, Handbook of Econometrics, chapter 64, Volume 6, Part A, 2007.
Sullivan, K., Chalasani, P., Jha, S. and Sazawal, V. Software design as an investment activity: a real options perspective, in Trigeorgis, L. (ed.) Real Options and Business Strategy, Risk Books, 1999.
The Economist, Can Sonos beat back the tech giants? Oct 20 2018a
The Economist. Google leads in the race to dominate artificial intelligence - Battle of the brains, December 7, 2017.
The Economist. The race is on to dominate quantum computing, August 18, 2018b.
The Rise of the Platform Economy, Kenney, Martin, and John Zysman, Issues in Science and Technology 32, no. 3 (Spring 2016).
Thomsett, R. Third Wave Project Management, Prentice Hall – Yourdon Press, 1993.
Vakulenko, M. Singh, S. The 9 Types of Software Platforms. Medium.com. June 12 2016.
World Intellectual Property Organization. Understanding Industrial Property. 2016

8

Energy

8.1 The Global Energy Markets

The energy market is the largest market in the world after currencies. The two most significant drivers of energy demand are population and income growth. Since the start of the twentieth century, the world's population has more than quadrupled, with real income and primary energy consumption growing by factors of 25 and 22.5, respectively. Over the twentieth century, the global annual average 3% GDP growth rate was sustained by an annual 2% growth rate in the energy supply, while energy consumption increased from an annual equivalent of 4 barrels (23.2 MM Btu) to 13 barrels (75.4 MM Btu) of oil per person.

Figure 8.1 illustrates the global total, urban and rural population trends from 1950 to 2050. Over the last 60 years, the world has seen a rapid urbanization, with more people living in urban areas today than in rural areas. In 2007, the world's urban population exceeded the rural population historically for the first time, with 55% of the global population living in urban areas as at 2018. The percentage of people living in urban areas ranked by regions in 2018 was North America (82%), Latin America and the Caribbean (81%), Europe (74%) and Oceania (68%). In Asia, the level of urbanization has reached approximately 50% of the population, while Africa continues to have a rural majority, with 43% living in urban areas.

The growth in urban populations is being driven by the expansion in the general population and people increasingly living in urban areas. These two factors are estimated to add an additional 2.5 billion to the world's urban

J. Rogers, *Strategy, Value and Risk*, Global Financial Markets,
https://doi.org/10.1007/978-3-030-21978-9_8

population, with the total global population estimated to reach 9.7 billion by mid-century.

Urban population growth is therefore expected to continue, with close to two-thirds of the global population living in urban areas by 2050. Urbanization levels will, however, vary significantly across regions. Close to 90% of the urban growth is expected to occur in Asia and Africa, with urbanization by the mid-century reaching 64% and 56%, respectively.

Figure 8.2 illustrates the long-term real GDP forecasts by region from 2020 to 2050. Global total long-term real GDP is forecast to rise from $103 trillion

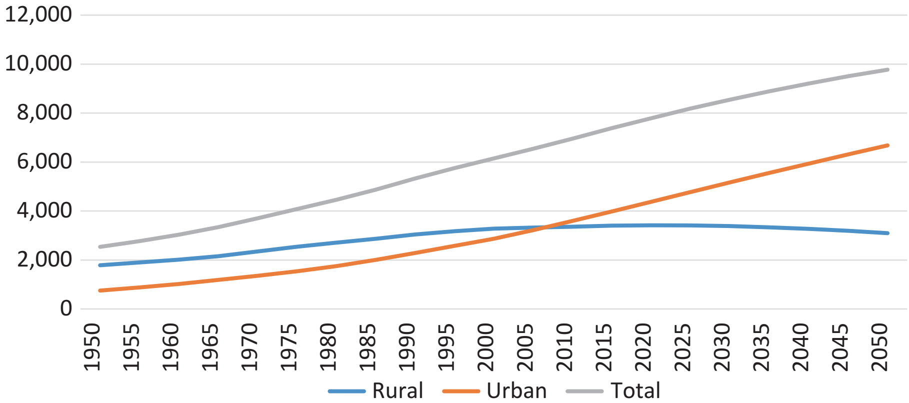

Fig. 8.1 Rural, urban and total population trends, 1950–2050 (M). Source: United Nations, Department of Economic and Social Affairs, Population Division (2018). World Urbanization Prospects: The 2018 Revision, Online Edition

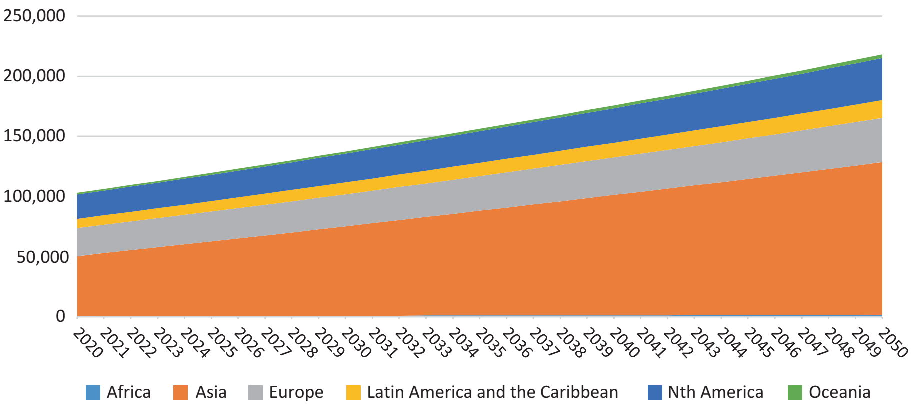

Fig. 8.2 Real long-term GDP forecast by region (billions US dollars). Source: data.oecd.org (2018), GDP long-term forecast (indicator) measured in US dollars at 2010 purchasing power parities, https://data.oecd.org/gdp/gdp-long-term-forecast.htm

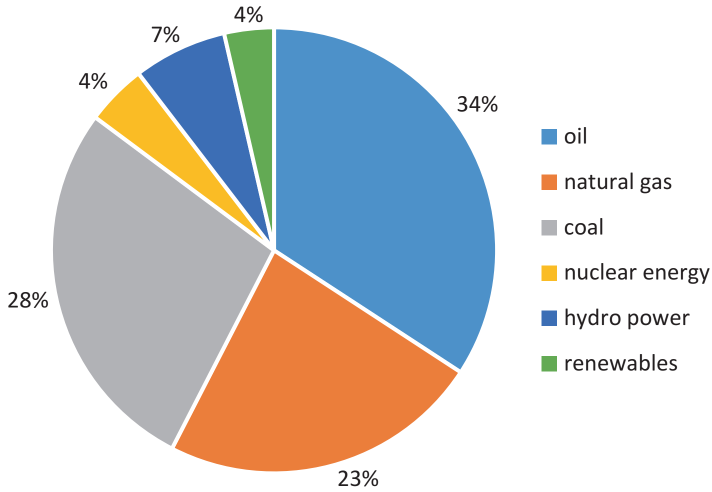

Fig. 8.3 Total world primary energy consumption by fuel. (13,511.2 million total tonnes of oil equivalent). Source: BP Statistical Review of World Energy, June 2018

to $218 trillion; in Europe, from $23.4 to $36.6 trillion; and in North America, from $20.3 to $34.9 trillion. The real GDP forecast for Asia is $49.7 to $126.9 trillion, with China ($24.9 to $54.4 trillion, a 118% increase) and India ($10.1 to $41.5 trillion, a 311% increase) a significant component of Asian growth. Real GDP growth in Latin America and the Caribbean is projected to rise from $7.6 to $15.0 trillion; in Africa, from $740 billion to $1.7 trillion; and in Oceania, from $1.4 to $3.1 trillion.

Figure 8.3 illustrates the total world primary energy consumption by fuel as at 2017, with nuclear 4%, renewables 4%, hydro power 7%, natural gas (NG) 23%, coal 28% and oil 34%.

Figure 8.4 shows the forecast for the world energy consumption by fuel from 2020 to 2040. The forecast for petroleum and other liquids (including biofuels) is 202.2 to 229.5 quadrillion British Thermal Units (quad Btu), a 13.5% increase; natural gas, from 132.2 to 181.6 quad Btu, a 37.4% increase; coal, from 162.3 to 160.9 quad Btu, a 0.9% decrease; nuclear, from 28.5 to 37.9 quad Btu, a 33% increase; and renewable energy (excluding biofuels), from 84.7 to 128.8 quad Btu, a 52% increase.

Figure 8.5 shows the forecast for global energy consumption by region from 2020 to 2040. By 2016, the world was consuming a total of 572.8 quad Btu of energy per year, and is forecast to rise to 736 quad Btu in 2040. Non-OECD (Organisation for Economic Co-operation and Development) regions are expected to account for most of the global growth in energy consumption,

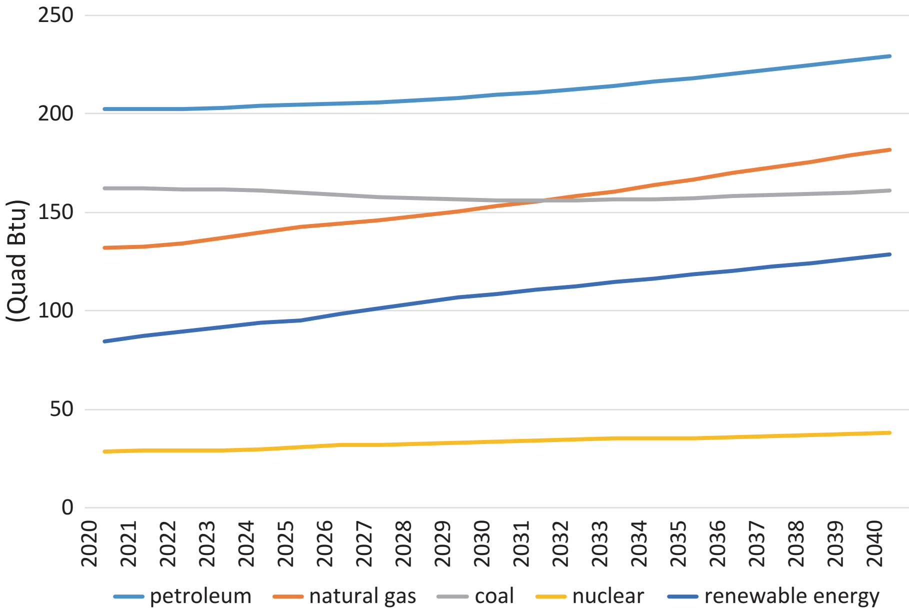

Fig. 8.4 World energy consumption by energy source, 2020–2040 (Quad Btu). Source: U.S. Energy Information Administration, International Energy Outlook (2017). Note: Petroleum and other liquids includes biofuels; renewable energy excludes biofuels

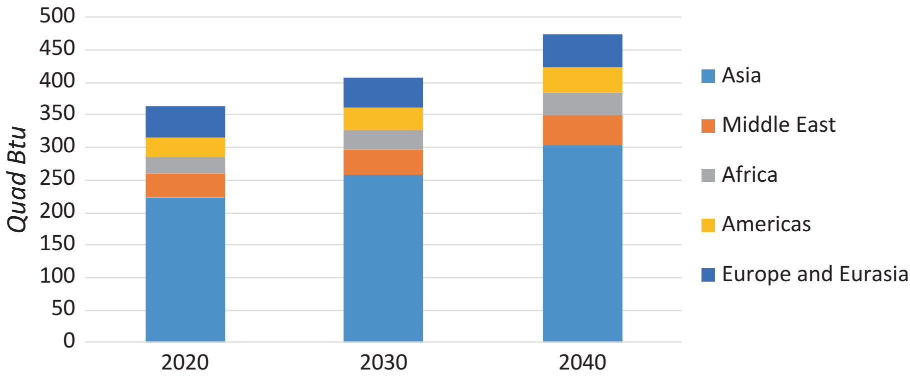

Fig. 8.5 Non-OECD energy consumption by region, 2020–2040 (Quad Btu). Source: IEA Annual Energy Outlook 2018

with increasing energy demand driven by long-term economic growth. Greater than 50% of the global increase in energy consumption will occur in Asia, predominantly in China and India, and by 2040, will exceed OECD energy use by 41 quad Btu.

Efficiency improvements are also likely to be a factor in energy demand. Energy efficiency—generally defined as the energy to GDP ratio or energy per unit of GDP—is forecast to accelerate and factor into income growth over the next 20 years, with more rapid efficiency gains seen in non-OECD economies.

8.2 The Transformation of the Energy Sector

The global energy sector is currently undergoing a fundamental transformation. Driving this process are the growth and cost competiveness of fossil fuel alternatives that include renewables, the rise of electric vehicles, the globalization of natural gas markets, gains in energy efficiency, battery developments and greenhouse gas emission issues.

These trends are creating unparalleled pressure within the oil industry. As the global energy market more than doubled in size since 1971, fossil fuels remained relatively stable in the energy mix at 80–85% of the total energy market. In 2016, however, the beginnings of a structural change emerged in the energy markets, with the global share of electricity produced by wind and solar rising from 4.5% to 5.2% (International Energy Agency [IEA] figures).

The future global demand for oil is a fundamental component of this structural change. The peak demand for oil is predicted to arrive as early as 2025 and any time up to the late 2040s. This peak will not be the result of the typical oil price cycle, and instead, signify a structural change in energy consumption, in which oil prices begin a slow permanent decline to a level where oil investments are uneconomic.

The International Energy Agency's (IEA) New Policies Scenario (NPS) includes current and planned policies that will have an impact on future energy demand and supply, distribution, carbon emissions, air pollution and all fuels and technologies within the energy system in general. Figure 8.6 shows the oil demand in mb/d under the NPS from 2025 to 2040. The growth in energy demand in the NPS is forecast to be more than 25% to 2040, and require greater than $2 trillion in new energy supply investment per annum. Renewables will constitute more than 60% of gross additional capacity to 2040, and reach 50% of global power generation capacity in most regions by 2035.

One of the fastest growing energy technologies in the NPS is solar photovoltaic (PV), which is forecast to have the second largest installed capacity behind hydro generation, surpassing wind capacity in the near future and coal before 2040. China and India will be the main drivers of the growth in global solar PV, with greater than 50% in additional global solar PV capacity.

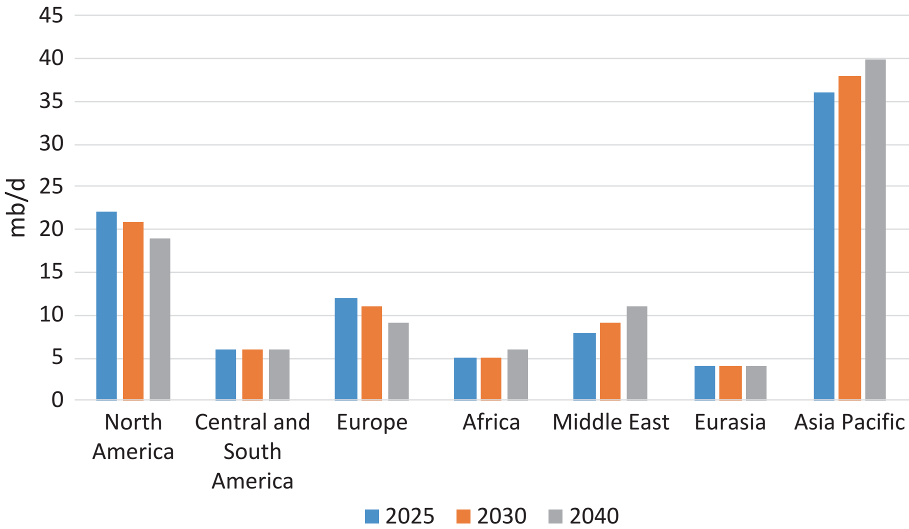

Fig. 8.6 Oil demand under the IEA New Policies Scenario, 2025–2040 (mb/d). Source: IEA World Energy Outlook 2018

Installed wind power will also see rapid growth, reaching approximately 1700 GW or 14% of global capacity by 2040.

While global oil demand growth does slow in the NPS, peak oil does not occur before 2040, and there is no peak in global energy-related CO2 emissions. The NPS oil demand from 2025 to 2040 for North America is 22 to 19 mb/d; Central and South America, flat at 6 mb/d; Eurasia, flat at 4 mb/d; Europe, from 12 to 9 mb/d; Africa, 5 to 6 mb/d; the Middle East, 8 to 11 mb/d; and the Asia Pacific, from 36 to 40 mb/d. China under the NPS is predicted to be the globe's single largest oil consumer and net oil importer by the 2030s, importing more than 13 mb/d by 2040.

The IEA Sustainable Development Scenario (SDS) provides a sustainable energy benchmark, compared to the NPS current and planned policies, which combines three critical policy goals—climate issues, air quality and energy access. Figure 8.7 illustrates the oil demand in mb/d under the SDS from 2025 to 2040. The SDS oil demand forecasts from 2025 to 2040 are declines in North America, from 20 to 12 mb/d; Central and South America, from 5 to 4 mb/d; Europe, from 11 to 5 mb/d; the Middle East, from 8 to 7 mb/d; Eurasia, from 4 to 3 mb/d; the Asia Pacific, from 33 to 27 mb/d; and Africa, flat at 5 mb/d.

Scenarios for the future energy mix will be a function of future global energy demand and the diffusion of new energy technologies. High energy demand and low technology diffusion will see global oil demand peak around the late 2040s, with oil, and natural gas and coal approximately 25% and

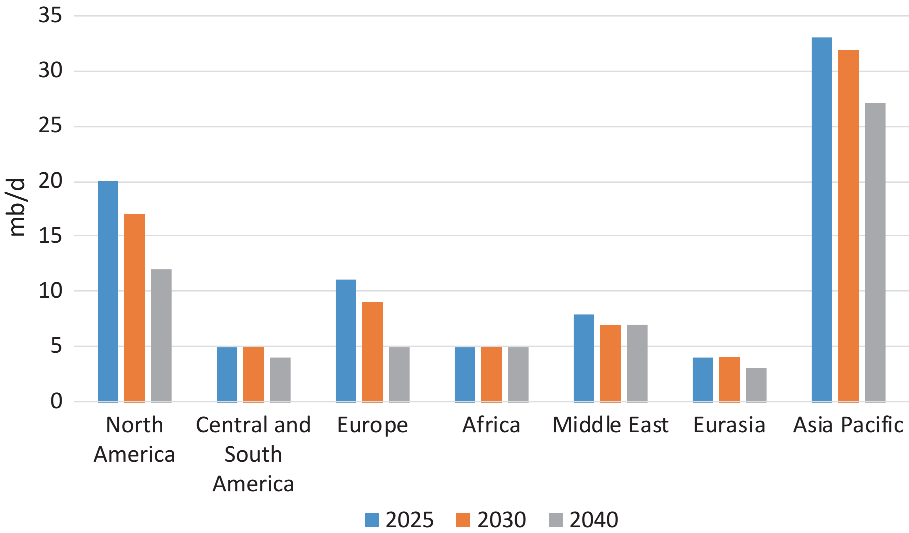

Fig. 8.7 Oil demand under the IEA Sustainable Development Scenario, 2025–2040 (mb/d). Source: IEA World Energy Outlook 2018

wind and solar 5% of total demand. Low energy demand driven by energy efficiency and high technology diffusion will potentially see global demand for oil peak as early as 2025. Under this scenario growth in global energy, demand will be significantly less by 2040, with oil and natural gas each making up approximately 25%; coal 20%; and solar and wind, 15% of total demand.

Over the twenty-first century, the trends in the energy mix will potentially see renewables reach one-half to two-thirds of the energy system, with the majority being solar energy, and wind energy in niche regional markets. Global demand for natural gas is expected to continue rising, and is viewed as a bridge fuel in the transition from fossil fuels to renewable energy. The peak natural gas forecast, however, is a function of the trends in the rapidly falling costs and continued investments in renewable energy.

These scenarios will dramatically alter energy business models, as the growth in alternative energy technologies and their declining costs transform industries across the global economy. Energy firms will therefore require the resources and capabilities to adapt to the evolving energy markets and maximize value at every stage of the energy value chain. Firms will need to decarbonize their portfolios as the energy system transforms from oil and gas to electricity and global clean energy. Those firms with capitalized oil and gas reserve exploration costs also face the risk of being unable to monetize the asset values, as downstream sales demand softens due to the structural decline in peak oil prices, and therefore, need to reduce their exposures to stranded assets.

Investment in renewable energy will therefore significantly increase in the future, as will the capabilities required to produce, sell and trade energy. Building and managing a new energy complex will require management technologies that include portfolio trading and risk management, data and analytics, and advanced methods for the analysis of energy asset investments, divestments and value.

8.3 Natural Gas and Renewables

One consequence of the long life of energy assets is that change in the energy mix is slow; however, gas and non-fossil fuels are expected to increase their share at the expense of oil and coal. The fastest growing energy source is renewable energy, with 40% of the increase in primary energy, and natural gas demand expected to see continued growth globally. By 2040, the energy mix is expected to be the most diversified in the history of the industry, as the growth of other new energy technologies develops.

Figure 8.8 illustrates the natural gas (NG) demand under the IEA Sustainable Development Scenario (SDS) in billion cubic metres (bcm) from 2020 to 2040. The globalization and growth in natural gas, supported by the US shale revolution and growth in liquefied natural gas (LNG), will continue

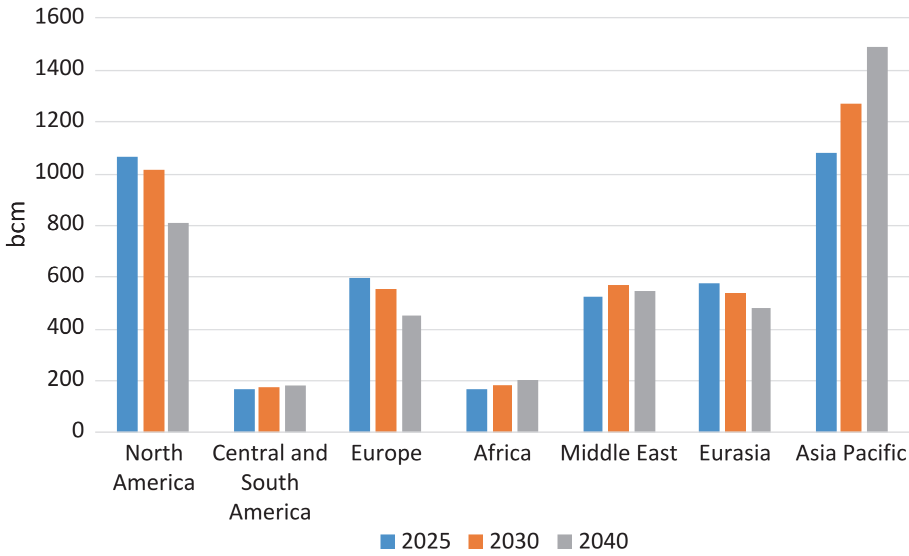

Fig. 8.8 NG demand under the IEA Sustainable Development Scenario, 2025–2040 (bcm). Source: IEA World Energy Outlook 2018

to transform the global gas markets. Gas is more responsive to short-term supply and demand dynamics, and has greater availability in markets across regions as a result. A key driver of natural gas demand growth in the Asian emerging economies are policies aimed at addressing air pollution.

NG demand under the SDS from 2025 to 2040 in North America is a decline from 1066 to 814 bcm; growth in Central and South America from 170 to 184 bcm; in Europe, a decline from 596 to 450 bcm; growth in Africa from 166 to 201 bcm; the Middle East, from 528 to 545 bcm; and in the Asia Pacific, from 1081 to 1491 bcm; and a decline in Eurasia from 574 to 485 bcm.

Figure 8.9 shows the renewables demand under the SDS from 2025 to 2040. Renewables demand under the SDS sustainable policies in terawatt hours (TWh) in North America is 1841 to 3719 TWh; Central and South America, 1155 to 1773 TWh; Europe, 1569 to 2537 TWh; Africa, 410 to 1528 TWh; the Middle East, 125 to 956 TWh; Eurasia, from 321 to 725 TWh; and Asia Pacific, 5013 to 12,481 TWh.

Wind energy is a renewable energy technology that uses the physics of wind to drive turbines that generate electricity, and is one of the fastest growing energy technologies in the world today.

Wind turbines are combined into wind farms that provide scale, and can either be connected to electricity transmission grids and networks for distribution or provide electricity to meet off-grid demand. The advantages of wind

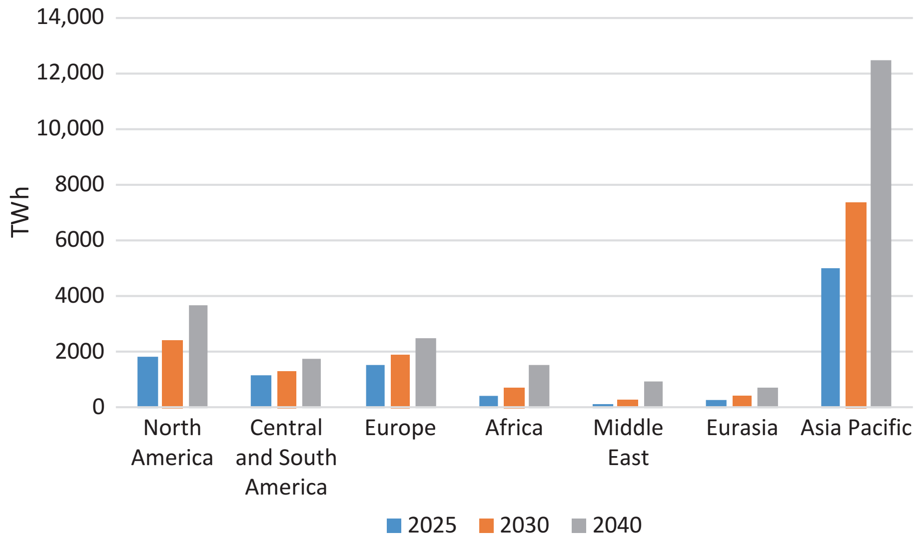

Fig. 8.9 Renewables demand under the IEA Sustainable Development Scenario, 2025–2040 (TWh). Source: IEA World Energy Outlook 2018

energy as a substitute for fossil fuels are its abundance and distribution; it is renewable, and has no greenhouse gas emissions. Although wind power is capital intensive and therefore requires upfront investment, it does not have any fuel costs, and therefore, like solar energy, has low-to-zero marginal costs.

The majority of installed wind power today are horizontal axis wind turbines (HAWT), with increases in average turbine capacity, rotor diameter and hub height the long-term trend. Wind turbines can be located either onshore or offshore, and currently reach up to 10 MW of onshore generation, and 12 MW for offshore generation. Offshore wind can be more consistent and have a higher velocity while having relatively higher construction and maintenance costs. One issue with wind energy is its consistency and variation over short time frames, and is therefore used in combination with other electricity generation assets and power management methods to meet demand schedules.

By the end of 2017, the total global installed wind turbine capacity was 539 GW, with an additional 52 GW of capacity added during the year, and total global wind turbine capacity able to meet greater than 5% of global electricity demand. Wind turbine energy capacity is widely distributed in Europe, with other regions increasingly migrating to wind energy power to meet global electricity demand.

Solar energy, or power from the sun, is a resource that is larger than every other energy source available on the planet, with approximately 174,000 terawatts (TW) of power provided constantly through solar radiation to the atmosphere's higher levels. Global power consumption was approximately 22,015 TWh as at 2017, and therefore, the solar energy that radiates to the planet is more than sufficient to supply total energy requirements. The solar resource is also freely accessible, and globally distributed relative to other energy resources.

Solar energy is a renewable energy that sources the Sun's radiant light and heat for a range of energy technologies that include photovoltaics, solar heat, solar thermal electricity and solar cells. These technologies can address energy security and climate change issues, and integrate into electricity systems. Solar energy is generally more available in countries with warm and sunny climates, which are also the regions that will see the majority of the global population and economic growth in the coming decades. The global population in warm and sunny climates is estimated to be 7 billion, compared to 2 billion in cold and temperate climates, by 2050.

Light, in general, concerns a particular type of electromagnetic radiation, that consists of oscillating electric and magnetic fields that vibrate at a given frequency and wave length, and disseminate linearly. The electricity produced by solar radiation to a semiconductor surface, directed at the sun on a clear day at noon, is approximately one kilowatt per square meter. Silicon is a semiconductor solid comprised of electrons, a subatomic particle that has an electrical charge. A six-inch square silicon solar cell generates approximately 0.5–0.6 volts and 4–5 watts under direct sunlight with an intensity of one kilowatt per square meter. Solar cells are linked in a PV module in series to boost the combined output voltage. A PV module will typically be composed of 60–96 solar cells, and produce 30–48 volts and 260–320 watts of power. PV modules also include additional operational mechanical components, and are linked together as solar arrays, either in series to increase output voltage, or in parallel to increase output current.

Solar energy electricity technologies have the potential for large-scale growth, with global expansion reaching terawatt generation capacities. The recent rapid growth in solar generating capacity, technology, performance and price improvements, combined with innovative business model developments, have driven investment in commercial and residential solar systems. Two solar energy technologies that can scale for the generation of grid electricity are concentrated solar power (CSP) and photovoltaic (PV) systems.

Each technology has significant differences. CSP technologies offer large-scale installations, a generating capacity of 100 MW or greater and the ability to store thermal energy to generate electricity outside of periods with little or no sunlight. CSP systems use direct irradiance only, and are therefore more sensitive to the influence of cloud, haze and dust. CSP systems currently offer large-scale installations without the potential for materials availability bottlenecks, and the inclusion of thermal energy storage in CSP systems also provides a dispatchable electricity resource.

PV systems can use all solar radiation incident, and can be configured as utility plants that have greater than a 1 MW capacity, to residential installations under a capacity of 10 kW. The output of PV systems is sensitive to changes in solar radiation. PV costs by convention are split into solar module costs and balance-of-system (BOS) costs. BOS costs, or all PV system components other than the PV panels, include inverters, hardware, labour, financing, marketing and regulatory costs. Continuing developments in PV technology can also have an influence on solar module and BOS costs.

The competitiveness of solar energy in relation to other generation technologies is a function of the revenue and cost structures within a specific electricity market. Most installed solar electricity generation globally today is PV. There are three issues, however, that need to be addressed in regards to solar energy having a major role in the future energy mix. Although solar electricity costs have significantly declined in recent years, solar power is still relatively more expensive than the current fossil fuel technologies in many regions, although this would be offset by carbon pricing. Currently, BOS costs account for approximately two-thirds of the price of utility-scale PV installations.

Second, the solar resource also fundamentally differs from other energy resources due to its intermittency at any location on the planet's surface. Predicting the solar resource is a function of both stochastic processes, with uncertainty occurring at frequencies that span minutes to days due to cloud cover and weather, and deterministic processes, with oscillations that have frequencies from days to months and are a function of the planet's daily rotations and seasons.

The third issue is scaling. Forecasting solar power at any location due to its intermittency can be a significant impediment to building large-scale solar generation in many regions. Matching generation with demand in electrical power systems is essentially a real-time process, with demand variations not entirely predictable. Solar generation within a power system will incrementally add volatility and reduce the net load predictability due to its intermittency.

The intermittency in renewable generation has been a key obstacle for contracts in the forward wholesale market, which are used to manage risk for generators and retailers in volatile electricity markets. Renewable generation is typically dispatched first as the electricity output cannot be regulated, has zero marginal generation costs and is paid the prevailing spot price. Solar firming contracts offer the ability to 'shape' contracts that match the load of buyers. A contract can replicate the shape of solar generation, for example, when there is no solar generation due to the time of day or the weather, and match a buyer's demand with supply contracts in the wholesale market, and therefore, provide a flat or fixed price. Developments in solar contracts include a solar shape and a solar firming, or inverse solar shape.

8.4 Energy Statistics

8.4.1 The Schwartz Single Factor Model

Energies such as electricity and natural gas exhibit the property of mean reversion. Modelling these price series using Black–Scholes-type models can produce unrealistic spreads between the two related energy commodities. Mean reversion can be captured in a more realistic single factor model introduced by

Schwartz (1997), which assumes that the spot price follows a mean reverting process:

$$dS = \alpha\left(\mu - \lambda - \ln S\right)Sdt + \sigma Sdz \tag{8.1}$$

where α is the mean reversion rate, which is the speed of adjustment of the spot price back towards its long-term level μ, σ is the spot price volatility and λ is the market price of energy risk. By defining $x = ln\ S$ and applying Itô's lemma to Eq. (8.1), the log price can be characterized by the Ornstein–Uhlenbeck process:

$$dx = \alpha\left(\hat{\mu} - x\right)dt + \sigma dz \tag{8.2}$$

where,

$$\hat{\mu} = \mu - \lambda - \frac{\sigma^2}{2\alpha}.$$

8.4.2 Schwartz Single Factor Futures and Forward Pricing

Futures and forward prices with maturity s in the Schwartz single factor model are equal with the appropriate boundary conditions and are given by:

$$F\left(t,s\right) = \exp\left[e^{-\alpha(s-t)}\ln S + \left(1 - e^{-\alpha(s-t)}\right)\left(\mu - \lambda - \frac{\sigma^2}{2\alpha}\right) + \frac{\sigma^2}{4\alpha}\left(1 - e^{-2\alpha(s-t)}\right)\right] \tag{8.3}$$

The mean reversion rate α determines how quickly forward prices revert to the long-term level. Figure 8.10 illustrates the sensitivity of the futures price defined in Eq. (8.3) to the mean reversion parameter, or the speed of mean reversion. The parameters used in the illustration are $S = 110$, $\sigma = 0.3$, $\lambda = 0$, $\mu = \ln(100)$ and $\alpha = 0.1$, 1 and 10. The long-term level of the futures curve does not equal $\exp(\mu)$, as it is adjusted by an amount that depends on the relative size of α and σ:

$$F\left(t,\infty\right) = \exp\left[\mu - \lambda - \frac{\sigma^2}{4\alpha}\right] \tag{8.4}$$

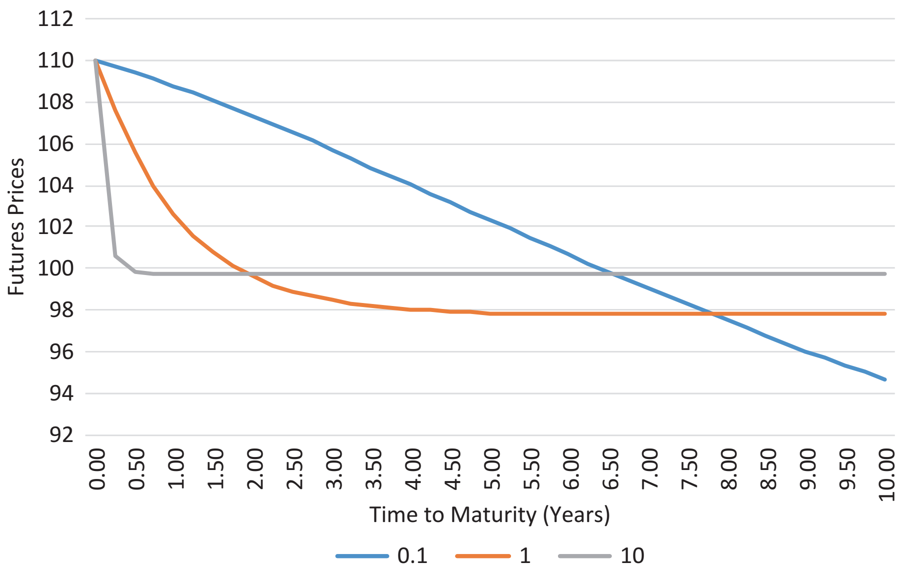

Fig. 8.10 Schwartz single factor model futures prices. Source: Clewlow and Strickland 2000

8.4.3 Volatility

The volatility measure used in energy pricing models should be estimated in the context of the specific stochastic price process that captures the key features of the energy markets, such as mean reversion. The constant volatility assumption used in the Black–Scholes model is not consistent with the empirical observation that long-dated energy forwards have less volatility than short-dated energy forwards.

Itô's lemma can be applied to Eq. (8.3) to provide the term structure of proportional futures volatilities in the single factor model:

$$\sigma_F(t,s) = \sigma e^{-\alpha(s-t)} \tag{8.5}$$

Figure 8.11 illustrates the effect of the speed of mean reversion, α in Eq. (8.5), on the term structure of volatility of futures prices. Volatility parameters of 0.3 and α = 0.1, 1 and 10 are used in the illustration. Increasing the speed of mean reversion, for example, increases the attenuation of the volatility curve. As the maturity of the forward increases, the volatility also tends to zero.

While the volatility term structure based on the Schwartz single factor model is a more accurate representation than the Black–Scholes model, its shape is still relatively simple. Even though a volatility function of this type

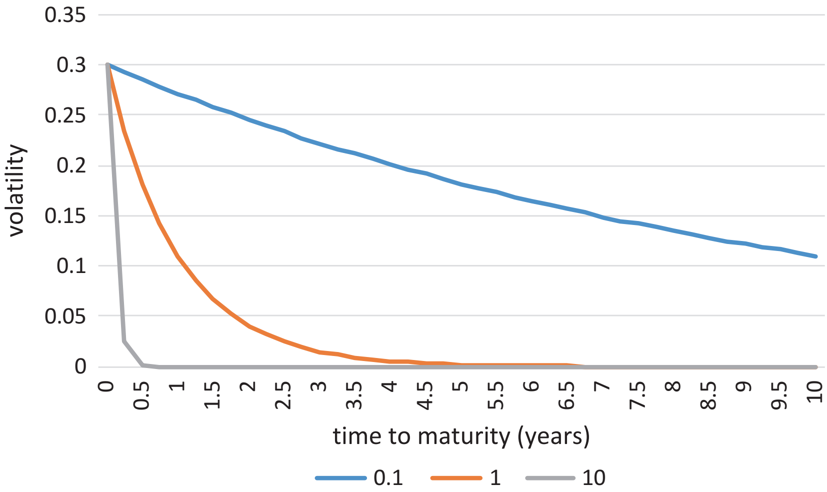

Fig. 8.11 Volatility of futures prices in the Schwartz single factor model. Source: Clewlow and Strickland 2000

describes the attenuation typical of market forward volatility term structures, the volatility parameters tend to zero for longer dated maturities. While market volatilities of forward energy prices do decrease as maturities increase, they typically do not approach zero, and therefore, the Schwartz model has a potential drawback when pricing options on long-dated maturity forward contracts.

This attribute obviously is not correct, and is a function of the simple assumptions in the form of mean reversion in Eqs. (8.2) and (8.5). This issue can be addressed in the representation of Eqs. (8.2) and (8.5) by directly specifying the volatility function, and adding a constant long-term level of forward price volatility to the simple negative exponential specification.

8.4.4 Correlation

The volatility of a spread is less than the sum of the volatilities of the individual components, and should be considered in the pricing of spread options. The correlation between two assets is captured by:

$$\xi_{1,t} = \varepsilon_{1,t} \tag{8.6}$$

$$\xi_{2,t} = \rho\varepsilon_{1,t} + \varepsilon_{2,t}\sqrt{1-\rho^2} \tag{8.7}$$

where $\xi_{1,t}$ and $\xi_{2,t}$ are two independent random numbers from a standard normal distribution, and ρ is the correlation between the two assets.

8.4.5 Simulating Mean Reversion

A Monte Carlo simulation simulates possible future values of an underlying asset using a stochastic process based on assumptions of the behaviour of the relevant market variables. The advantages of Monte Carlo simulation are that it can facilitate the accuracy in the modelling of market price behaviour by including factors such as jumps, seasonality, stochastic volatility and possible future structural changes in the market. The disadvantages are its relative complexity in implementation and the resources required for computation.

The mean reverting spot price model in Eq. (8.2) was specified in terms of the natural logarithm of the spot price $x = \ln(S)$:

$$dx = \left[\alpha\left(\mu - x\right) - \frac{1}{2}\sigma^2\right]dt + \sigma dz \tag{8.8}$$

which can be discretized as:

$$\Delta x_i = \left[\alpha\left(\mu - x_i\right) - \frac{1}{2}\sigma^2\right]\Delta t + \sigma\sqrt{\Delta t}\varepsilon_i \tag{8.9}$$

In contrast to the GBM (geometric Brownian motion) model, the discretization in this specification is only correct in the limit of the time step tending to zero, as the drift term is dependent on the variable x. Time steps that are relatively small to the speed of mean reversion should therefore be chosen. To simulate the path of the spot price, the parameters α, μ, σ and Δt are estimated, normally distributed random numbers ε_i are repeatedly generated and new values of Δx are calculated, from which a new spot price at each time step is then derived.

8.4.6 Estimating the Mean Reversion Rates

Two methods can be used to estimate the mean reversion rate α, either through linear regression using spot price data or by fitting the single factor volatility function to the empirical volatility term structure. The simple mean reverting process for the natural logarithm of the energy spot price:

$$dx = \alpha\left(\overline{x} - x\right)dt + \sigma dz \tag{8.10}$$

is essentially the same as Eq. (8.8), but with the $-\frac{1}{2}\sigma^2$ included in $\bar{x}$. This can be discretized as:

$$\Delta x_t = \alpha_0 - \alpha_1 x_t + \sigma\varepsilon_t \tag{8.11}$$

where $\alpha_0 = \alpha\bar{x}\Delta t$ and $\alpha_1 = \alpha\Delta t$. Observations of the spot price through time imply the linear relationship between Δx_t and x_t with the noise term $\sigma\varepsilon_t$. Regressing observations of Δx_t against x_t obtains $\alpha_0 = \alpha\bar{x}\Delta t$ and $\alpha_1 = \alpha\Delta t$ as estimates of the intercept and slope of this linear relationship. As the time interval between observations Δt is known, estimates of α and $\bar{x}$ can be obtained.

An alternative is to estimate α, the mean reversion rate, from the term structure of volatility. The volatility of short-term energy forward contracts is typically more volatile that long-term contracts, with the volatility declining as the maturity increases. Equation (8.12) represents this decline in the volatility as t gets large:

$$\sigma_1(t,T) = \sigma e^{-\alpha(T-t)} \tag{8.12}$$

where $n = 1$. The mean reversion rate of the spot energy price can be estimated through the relationship between the spot price process and single factor model described in Sect. 8.4.1.

8.4.7 The Jump Parameters

The combination of mean reversion and jumps into the same model can be represented by the stochastic differential equation (SDE):

$$dS = \alpha(\mu - \ln S)Sdt + \sigma Sdz + \kappa Sdq \tag{8.13}$$

The mean reversion jump-diffusion model can be discretized as:

$$\Delta x_t = \left[\alpha(\mu - x_t) - \frac{\sigma^2}{2}\right]\Delta t + \sigma\sqrt{\Delta t\varepsilon_{1i}} + (\bar{\kappa} + \gamma\varepsilon_{2i})(u_i < \phi\Delta t) \tag{8.14}$$

where u_i is a uniform (0,1) random sample, and ε_{1i} and ε_{2i} are independent standard normal random variables. The term $(u_i < \phi\Delta t)$ returns one if true and zero if false, and generates the random jumps with the average frequency.

The jump size when it occurs is the mean jump size, represented by $\bar{\kappa}$, plus the jump standard deviation γ scaled by ε_2 to derive a normally distributed random shock. The time step Δt should be small in relation to the jump frequency, where $\phi\Delta t << 1$, so that the simulation of the jump frequency is accurate. dS is sum of the GBM mean reversion process and $\left(\bar{\kappa} + \gamma\varepsilon_{2i}\right)$ if a jump occurs. The Weiner and Poisson processes are assumed to be independent and not correlated, and therefore, the jump process is independent of the mean reversion process.

Estimating energy price jump parameters is complex in that the observed jumps are a subset of the time series, which also includes the non-jump price behaviour. The exact arrival of a jump is unknown, and the probability of these large price spikes within a GBM is effectively zero. The following jump parameters can be estimated using the recursive filter method:

ϕ = number of jump returns divided by the number of jumps within the sample
$\bar{\kappa}$ = average jump size of returns
Γ = standard deviation of jump returns

The standard deviation of the price returns is used to derive a probability for the identification of outliers that are greater than the chosen probability for actual jumps. The diffusion volatility is then re-estimated by deriving the standard deviation of the price returns with the jumps removed. The new diffusion volatility is then used to identify those jumps that exceed the probability limit, with the process repeated to where the estimates converge and no new jumps are identified.

8.4.8 The Half-Life of a Mean Reverting Process

A key property of a mean reverting process is the half-life. This is the time taken for the price to revert half-way back to its long-term level from its current level if no more random shocks arrive. Ignoring the randomness allows a focus on the mean reverting behaviour itself. The half-life, denoted by $t_{1/2}$, can be derived as:

$$t_{1/2} = \ln\left(2\right)/\alpha \tag{8.15}$$

The half-life is an average over a long time period, representing the time that shocks to the spot price take to decay to half their deviation from the long-term level. Table 8.1 illustrates the range of half-lives for various values of alpha.

8.4.9 Energy Spot Price Model Simulation

Figure 8.12 illustrates a simulation using Eq. 8.14 of a mean reversion jump diffusion path for an electricity spot price, where $\bar{S}$ is the long run mean, ϕ is the average number of jumps per dt, $\bar{\kappa}$ the average jump size of returns is set to zero, and γ is the standard deviation of jump returns. The function $\left(u_i < \phi\Delta t\right)$ takes a value of one if true and zero otherwise, and generates jumps at a random frequency as defined by ϕdt. If a random number is generated below the average jump frequency, a jump is simulated in a random direction, while if above the frequency, then no jump is generated. The jump size when

Table 8.1 Mean reversion rates and the corresponding half-lives

α	$t_{1/2}$
1	8 months
10	25 days
100	2.53 days
1000	6 hours

Source: Clewlow and Strickland (2000)

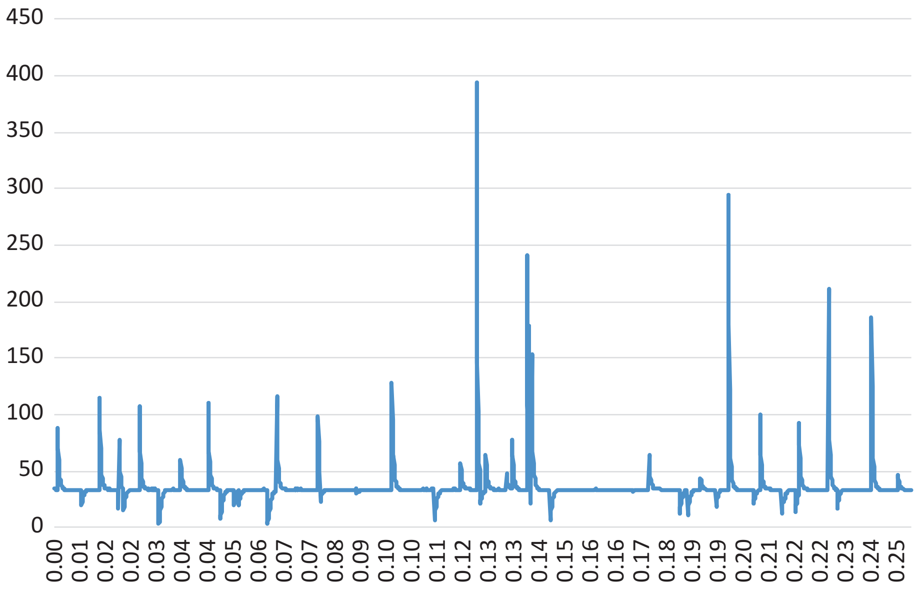

Fig. 8.12 A mean reversion jump diffusion electricity spot price simulation

it occurs is the average of the historical jump returns plus the jump standard deviation γ multiplied by a uniform normally distributed random variable, with the probability of either positive and negative jumps occurring. As the mean jump size is typically problematic in the derivation of robust estimates, it is set to zero. The half-life is 3 hours, which is derived as $\ln(2)/2000 \times 24 \times 365$.

$$\text{Price} = 35,\ \alpha = 2000,\ \bar{S} = 35, \sigma = 30\%,\ \bar{\kappa} = 0,\ \phi = 250,\ \gamma = 1.2,\ \Delta t = 1/\left(24 * 365\right)$$

The single factor model is a relatively simple model for the forward curve. While the volatility structure under the Schwartz single factor model is more realistic than the Black–Scholes model, it still has a relatively simple shape, and the volatilities tend to zero for longer maturities. The single factor model in Sect. 8.4.1 can be generalized and modified, as illustrated in Sect. 6.4, to include multiple sources of uncertainty in the forward curves. Clewlow and Strickland (2000) value energy options in a general multi-factor model for the forward curve, which can capture multiple sources of uncertainty. As spread options depend simultaneously on forwards related to separate energies, Clewlow and Strickland extend the multi-factor model to a specification that can simultaneously model a number of different energy forward curves, and capture the multiple dynamics of the forward curves in the valuation of spread options.

8.4.10 Wind Power Statistics

The conventional methodology for utilizing wind resources for the generation of electricity is the use of power curves. Wind turbine power curves provide values for the production of electricity as a function of wind speed at the turbine hub height and turbine type. Wind is measured in metres per second (m/s), with 10 (m/s) an established measure. A wind power curve is described by variables on two axes, v(m/s), the hub height wind speed on the horizontal axis, and power (MW) on the vertical axis, and has three significant values on the wind speed axis:

- The cut-in speed—the minimum wind speed at which the turbine delivers power.
- The rated output speed—the initial point on the wind speed axis at which the maximum turbine rated power is generated.
- The cut-out speed—the maximum wind speed for the turbine power generation, above which generation is zero.

Table 8.2 Typical features of a power curve

Characteristic	Value (m/s)	Power output (MW)
Cut-in wind speed	3	$P_{min} = 0.0533$
Rated wind speed	11.5	$P_{rated} = 4.5$
Cut-out wind speed	25	–

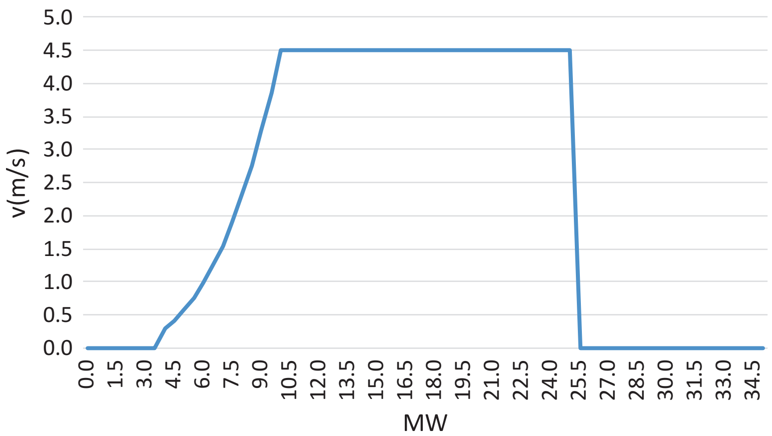

Fig. 8.13 The assumed wind turbine power curve

Table 8.2 illustrates the three significant values for a power curve example with a hub height of 100 m and a rated power of 4.5 MW.

While wind turbine model types have unique power curves that are specific to the manufacturer, the power curve in Fig. 8.13 is a common representation based on the data in Table 8.2.

The power obtained from wind is described by a cubic function:

$$P_{wind}(v) = \left(\frac{1}{2}\right) pA^3 \tag{8.16}$$

where,

$P_{wind}(v)$ = power (Watts)
P = 1.225 kg/m^3 (air density)
A = the rotor swept area (m^2)
v = hub height wind speed (m/s)

The rotor swept area is derived as:

$$A = \pi \left(\frac{rotor\ diameter(m)}{2} \right)^2 \tag{8.17}$$

For example, for a 100 m rotor diameter, the swept area is 7854 m^2. Annual potential wind energy is calculated as wind power (kW) multiplied by 365 × 24.

The power generated by the wind turbine between the cut-in speed and the rated speed is scaled by an efficiency parameter C_p:

$$P_{turbine}(v) = \left(\frac{1}{2} \right) pAC_p v^3 \tag{8.18}$$

The power coefficient *Cp* is the Betz limit, or the maximum power that can be generated from wind at the site, and is the ratio of 16/27 or 59%, the Betz coefficient. The potential maximum wind power in Watts is therefore:

$$59\% * P_{wind}(v). \tag{8.19}$$

A power curve is, therefore, described by the following, with $P(v)$, the power output, a function of v, the hub height wind speed (m/s):

0, where $v \leq v_{cut\text{-}in\ speed}$

$\left(\frac{1}{2} \right) pAC_p v^3$, where $v_{cut\text{-}in\ speed} < v \leq v_{rated\ power}$

$P_{rated\ power}$, where $v_{rated\ power} < v \leq v_{cut\text{-}out\ speed}$

0, where $v_{cut\text{-}out\ speed} \leq v$

The Weibull probability distribution is typically used to describe a wind turbine power output, and is characterized by two parameters for 10-minute average wind speeds—v, the hub height wind speed (m/s), and c, a scale parameter. A value of 2 for the Weibull shape parameter, k, is consistent with annual wind speed distributions. The parameter c can therefore be derived for an assumed wind speed:

$$c = \frac{\overline{v}}{\Gamma(1 + 1/k)} \tag{8.20}$$

where,

$\overline{v}$ = the average wind speed for $1 \leq k < 10$
$\Gamma(x)$ = the gamma function.

For an average wind speed of 10 (m/s), c is equal to 11.28. The Weibull probability density function is therefore:

$$f(v) = \frac{k}{c}\left(\frac{v}{c}\right)^{k-1} e^{-(v/c)^k} \text{ for } v \geq 0 \tag{8.21}$$

and 0 for $v < 0$.

The shape of the Weibull distribution is a function of the turbine location and time frequency, with the parameters k and c estimated and fitted according to the wind turbine location. Figure 8.14 illustrates the Weibull distribution that describes the yearly wind speed for k = 2 and c = 11.28.

The shape parameter k can also be estimated based on the assumptions in regards to the average wind speed and the wind volatility:

$$k = \left(\frac{\sigma_v}{\overline{v}}\right)^{-1.086} \tag{8.22}$$

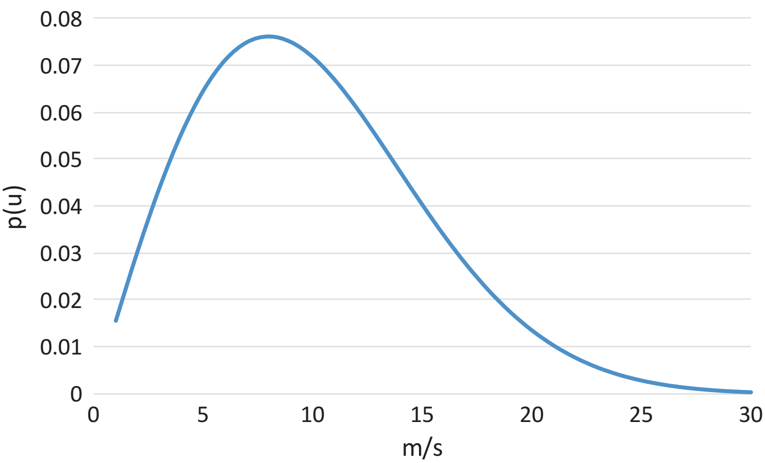

Fig. 8.14 Weibull distribution for k = 2, c = 11.28

where $\overline{v}$ is the average wind speed, and σ_v is the standard deviation of wind speed, derived as:

$$\sigma_v = \overline{v}^2 \left[\frac{\Gamma(1+2/k)}{\Gamma^2(1+2/k)} - 1 \right] \tag{8.23}$$

The empirical relationship between the three parameters, k, c and $\overline{v}$ is specified as:

$$\frac{c}{\overline{v}} = \left(0.568 + \frac{0.433}{k} \right)^{-1/k} \tag{8.24}$$

Using the underlying assumptions in regards to the wind turbine power curve and the wind speed probability distribution, a wind farm capacity factor can then be derived. The capacity factor is the ratio of:

$$\textit{Capacity factor} = \frac{\textit{average annual production MWh}}{\textit{nameplate capacity}\,(MW) \times 365\,\textit{days} \times 24\,\textit{hrs}} \tag{8.25}$$

where the nameplate capacity is the wind turbine maximum output rating. The average power generated $E[P(v)]$ is the power curve multiplied by the Weibull probability distribution:

$$E\left[P(v)\right] = \int_{v_{\min}}^{v_{\max}} P(v) f(v) dv \tag{8.26}$$

An additional parameter is required for the wake effects produced by a wind turbine within a wind farm. The clustering of wind turbines can reduce the total energy converted to electricity relative to the energy generated by the individual turbines operating under the same wind flow conditions. A turbine's downwind wind speed is inevitably less than its upwind wind speed, and therefore, upstream turbines 'shadow' downwind turbines. The downwind and cross spacing of wind turbines at the site can, however, maintain losses from wake effects at less than 10% of the wind farm power generation.

Table 8.3 summarizes the variables and parameters for a wind turbine and wind farm.

Table 8.3 Summary of the variables and parameters

Wind resource		Wind turbine		Wind farm	
v	10 m/s	Rated power	4.5 MW	Loss factor	10%
Shape factor	2	Cut-in speed	3 m/s		
Scale factor	11.28	Rated speed	11.5 m/s		
		Cut-out speed	25 m/s		
		Cp	59%		

8.5 Natural Gas Generation Valuation

8.5.1 Overview

Natural gas (NG) is a fossil fuel in which the principal element, methane, is a hydrocarbon, a compound that consists of hydrogen and carbon. Technologies for the generation of electricity through the combustion of natural gas include simple cycle turbines, conventional steam and combined cycle gas turbines (CCGT). Simple cycle combustion turbines are typically used for peak electricity demand, while conventional steam electricity turbines account for the majority of electricity generation. A CCGT plant combines a gas and steam turbine to generate electricity, and can produce significantly more power than simple cycle plants from the equivalent fuel. The growth in renewable capacity has increased the intermittency in the electricity supply, creating a structural change in the generation load and the need for increased flexibility in the power system. CCGT generation assets, by design, are well suited to respond to this requirement with their ability to ramp up within minutes and meet peak or unscheduled demand loads. Natural gas power generation generally produces relatively less emissions that other fossil fuels such as oil and coal, and is considered a bridge fuel as the energy mix transitions to clean energy.

The valuation of power plants is typically conducted for M&A (mergers and acquisitions) transactions, business entity going concerns and value analysis, and by convention, uses the intrinsic value, where value is a function of the plant dispatch relative to electricity prices in the forward market. The dispatch model values plants by determining the marginal cost-based clearing price and calculating cash flows based on the intrinsic spread, the spark spread, between the electricity price and the cost of fuel for generation. The margin received, or the spark spread, is derived over the life of the plant and discounted to the present.

The problem with the industry standard dispatch model is that it takes no account of volatility. Although the intrinsic discounted cash flow (DCF) model recognizes that a peaking plant has intrinsic value, that is, electricity prices can increase to a point where there is value in switching on the plant, it does not capture value derived from the volatility in the spark spread. The extrinsic value method is therefore increasingly being used in valuation, as flexibility now has a significant role in the use of NG power generation assets. Although CCGT generators are able to hedge some of the intrinsic value during peak loads, a greater percentage of the asset value is extrinsic.

A power plant's extrinsic value can account for the value derived from the volatility within the electricity and NG spark spread, and valued as a portfolio of European electricity and NG spread options. Power peaking plants valued as options on the fuel cost and power price spread provides the owner, or option holder, with the right to operate the unit when electricity prices are higher than the cost of the fuel used to generate. A portfolio of peaking units can increase marginal value significantly.

The operating characteristics of a NG power generation plant can therefore be defined as the equivalent to a spark spread option. A peaking plant can choose to only run when the power price exceeds the marginal fuel cost. The real option in the peaking plant is the ability to choose whether to generate or not at a given power price. While the peaking plant may have available capacity, there is no obligation to generate, regardless of the electricity price.

A rational generator operator would choose to generate when power prices are above the fuel cost and any start-up costs. These real options are switching options, defined in the following valuation example as a series of European call options on the spark spread between power and NG prices. Valuing the power plant as a real option illustrates the value in the flexibility to call the plant when the energy spread is positive, which can be used to optimize operations—and therefore, value.

8.5.2 Energy Forward Curves

Capturing the significant features of the energy markets is important in applications of energy pricing models. Although a number of energy derivative models use forward curves for pricing, there are some associated problems. Energy forward curves are typically composed of discrete monthly futures contracts, and therefore, are not continuous as assumed in the pricing model. Some energy markets can be in *backwardation* (where futures prices are lower than spot prices) while others might be in *contango* (where futures prices are higher than spot prices), which gives the spread its own forward curve. The spread can also become negative as a consequence of these properties. Another issue is that seasonality can also exist in the spreads.

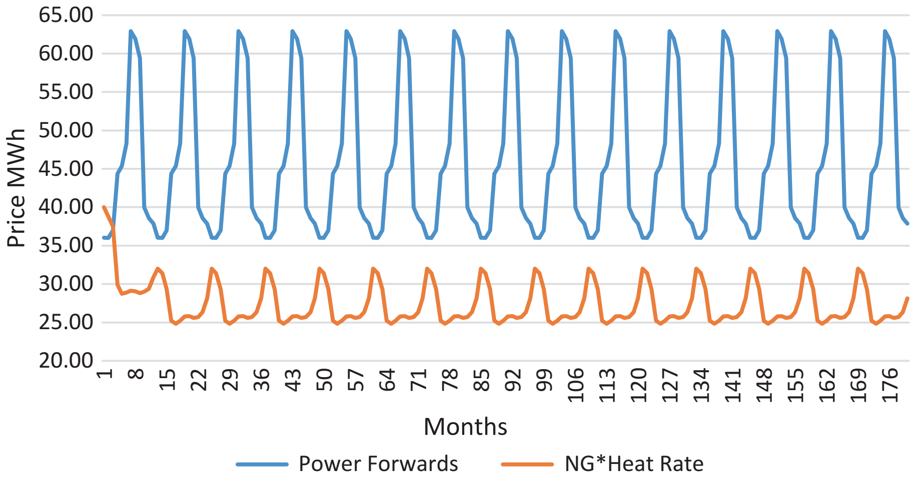

Fig. 8.15 The spark spread forward curves

Figure 8.15 illustrates the forward curves used in the valuation example. The forward curves were derived for the following valuation example, with annual seasonal patterns and negative spark spreads from January to March for each year.

Although electricity cannot be easily stored, the fuels used to generate the electricity can be stored, and the link between the two implies that the forward curve for electricity should be related to the input fuels. An arbitrage pricing approach takes this into account by considering the conversion process.

One of the key steps in the conversion process is the generation process itself, and this depends on the efficiency of generation expressed as the heat rate, the number of British Thermal Units (Btu) required to generate one kWh of electricity. A basic electricity forward curve can be obtained for the fuel forward curve via the following relationship:

$$Cost_{electricity} = Heat\ Rate \times Price_{fuel} \tag{8.27}$$

A constant value of the heat rate implies that the shape of the electricity forward curve should resemble the forward curve of the input fuel. The cost of electricity can be converted into a forward price after taking into account costs associated with fixed assets, transmission and tolling charges and others, such as fuel storage and fuel transportation. These costs obviously change through time.

Forward energy curves can be created as composite curves that consist of market data such as futures, forward prices and curve modelling. One feature exhibited by energy prices is the high level of seasonality, a repetitive cyclical pattern in the price over time. Seasonality in the power markets is driven typically by demand caused by weather factors such as hot summer months.

8.5.3 Energy Derivatives

There are a number of practical problems associated with derivative modelling in energy markets. Some of the important issues associated with energy derivative pricing that were often overlooked in early modelling approaches are:

- Energy prices tend to be drawn to production costs. The geometric Brownian motion (GBM) assumption permits price series to drift to unrealistic levels when applied to energy markets. In the short run, divergence from the cost of production can be possible under abnormal market conditions; however, in the long run, supply will adjust to the anomaly and prices will move to the level determined by the cost of production. This property is described as mean reversion.
- Energy prices display seasonality. Seasonality in energy prices and volatility may correspond to the time of year, such as winter or summer, and also can result from regular demand patterns due to factors such as the weather.
- Energy commodities cannot be treated solely as financial assets, as energy commodities are inputs to production processes and/or consumption goods. Models based on an automatic extension of those developed for financial markets may therefore break down when applied to energy markets.
- Another problem with applying the GBM assumption to energy prices is that their market price behaviour is often not consistent with the assumption of price continuity over time. Commodity and energy prices often display jump behaviour, determined in many cases by fluctuations in demand and supply. The frequency of these extreme values is often larger than the probability implied by GBM models.

In some markets, such as energy, the concept of being able to perfectly replicate options by continuously trading the underlying asset can be unrealistic. Many energy derivatives, however, actually rely on futures prices rather than the spot price, with the prompt futures contract a proxy for the spot price, and therefore, futures can be used to replicate options positions and permit the application of the risk-neutral pricing approach.

8.5.4 Energy Spread Options

The payoff in a spread option is derived from the price differential between two underlying assets. These types of exotic options can be used either to take a position on, or to hedge the risk associated with, the relative performance of two underlying assets. The payoff of a European spread call option at maturity T is:

$$c = \max\left[S_{1T} - S_{2T} - K, 0\right] \tag{8.28}$$

where S_{1T} and S_{2T} are the spot prices of the two underlying assets, and K is the exercise price.

Most option pricing models have the underlying assumption that the risk-neutral price distribution of the underlying asset is lognormal. A spread option is priced as the discounted double integral of the option payoffs over the risk-neutral distribution of the two underlying assets at maturity T (Pearson, 1997).

Analytical Black–Scholes-type models for valuing spread call and put options that include a strike are not known, and therefore, the Kirk approximation method is used in the following valuation illustration.

The Kirk approximation formula for the pricing of European call spread options on futures or forwards is:

$$C \approx \left(F_2 + K\right)e^{-r(T-t)}\left(FN\left(d_1\right) - N\left(d_2\right)\right) \tag{8.29}$$

which can be rewritten as:

$$C \approx e^{-r(T-t)}\left(F_1 N\left(d_1\right) - \left(F_2 + K\right)\right)N\left(d_2\right) \tag{8.30}$$

where,

$$\sigma \approx \sqrt{\sigma_1^2 + \left(\sigma_2^2 \frac{F_2}{\left(F_2 + K\right)}\right)^2 + 2\rho\sigma_1\sigma_2\left(\frac{F_2}{\left(F_2 + K\right)}\right)}$$

$$F = \frac{F_1}{\left(F_2 + K\right)}$$

$$d_1 = \frac{\ln(F) + \left(\frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}$$

$$d_2 = d_1 - \sigma\sqrt{T-t}$$

and F_1 is the power forward price, F_2 is the natural gas forward price times the heat rate, and K is the strike.

If the futures contracts underlying the option are written on two separate energies—in this case, natural gas and electricity—then the option is referred to as a *spark spread option.* Firms exposed to the price differences between two different energies often use options of this type. In this case of a natural gas-fired power generator, energy is an input into a process that produces another type of energy. If $F_a(t, T)$ represents the price of a T maturity futures contract on energy a—in this case, power—$F_b(t, T)$ represents the price of a T maturity futures or forward contract on energy b—in this case, NG times the heat rate—and K represents the start-up costs, then the payoff for a European call option with maturity T and strike K on the spread between the two forward contracts is:

$$c = Max\left(F_a(t,T) - F_b(t,T) - K, 0\right) \tag{8.31}$$

and therefore, the value of the call option at time t can be written generally as:

$$= P(t,T)\mathrm{E}_t\left[Max\left(F_a(t,T) - F_b(t,T) - K, 0\right)\right] \tag{8.32}$$

where $P(t, T)$ is the continuously compounded discount factor.

8.5.5 Natural Gas Peaking Power Plant Valuation

The assumptions for the valuation analysis are:

- The plant output is 300 MWh.
- The heat rate is 10,000—to derive the spark spread, the heat rate is divided by 1000.
- The generation plant has a total remaining life of 15 years.

- The power and natural gas forward curves consist of a series of one-month forward prices that represent the hourly average over each month.
- The plant will run 16 hours per weekday, except in the summer months of July, August and September, when it will run 24-hours per weekday.
- The start-up cost is $5000 per start-up.
- The MWh start-up cost was derived per month as the number of start-ups per month times $5000, which is divided by the total operating hours per month times the number of MWs.
- The operations and maintenance (O&M) costs are $1.50 per MWh.
- The discount rate is 6.50%, with the maturity for each discount factor the middle of each month.
- The power and natural gas monthly volatility curves represent the average volatility of the hourly average forwards for each month.
- The middle of each month is used as the maturity or expiry date for each spread option.
- The strike K represents the start-up costs, which is reflected in the per MWh start-up costs for each month.
- The 10% correlation (rho) between the two assets was estimated from spot price returns. Although there is typically a term structure of correlation similar to the term structures for forward prices and volatility, the correlation (or rho) between the volatilities is defined as a constant term in the valuation example.

Figure 8.16 illustrates the power and natural gas volatility term structure.

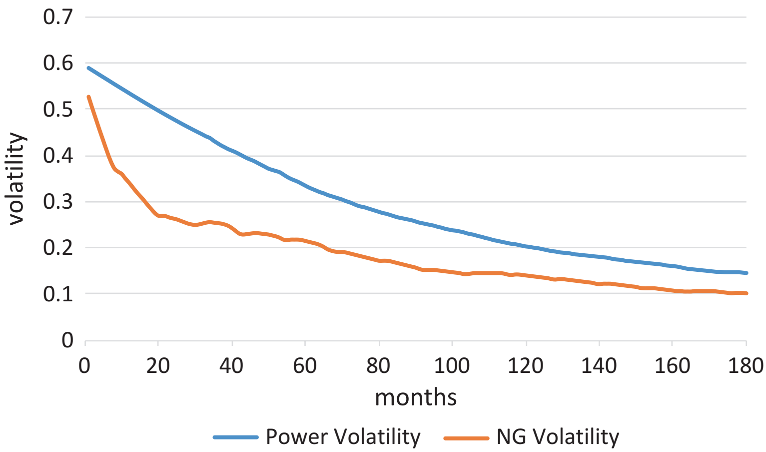

Fig. 8.16 The power and natural gas volatility curves

8.5.6 Energy Analysis and Valuation

Table 8.4 illustrates the intrinsic and extrinsic valuation of the natural gas peaking power plant. The intrinsic net present value (NPV) of the power plant is the sum of the present values of the net cash flows for each month. The spark spread options are priced as a series of European call spread options, with payoffs as specified in Eq. 8.32 from t_0 to T_i, the maturity or expiry for each European call spread option (in this case, the middle of each month) for each of the 180 European spread options. The extrinsic value equals the sum of spark spread options for each maturity. The intrinsic DCF valuation is \$236.6 M while the spread option valuation is \$301.8 M. While the DCF method does capture the intrinsic value, it does not account for the value in the volatility in the spark spread.

Figure 8.17 compares the spread option values and the DCF intrinsic values by month. The peaking plant is essentially an out-of-the-money option for at least part of the year. Although the spark spread can be negative, as is the case from January to March, the relatively large volatility associated with

Table 8.4 NG power generation intrinsic and extrinsic valuation

Month	1	2	3	4	....	179	180
Power forwards	36.02	36.00	37.00	44.36	....	38.57	37.88
Gas forwards	4.00	3.88	3.75	2.99	....	2.63	2.81
NG∗heat rate	40.00	38.80	37.50	29.90	....	26.32	28.12
Spark spread – \$/MWh	−3.98	−2.8	−0.5	14.46	....	12.25	9.76
Start-up costs	0	0	0	14.46	....	12.25	9.76
Total operating hours	384	384	384	384	....	384	384
Total revenue \$	0	0	0	1,545,792	....	1,291,200	1,004,352
Total O&M costs \$	172,800	172,800	172,800	172,800	....	172,800	172,800
Net cash flows \$	(172,800)	(172,800)	(172,800)	1,372,992	....	1,118,400	831,552
Discount factor	0.9963	0.9927	0.9891	0.9854	....	0.5186	0.5167
NPV \$	(172,167)	(171,537)	(170,909)	1,352,994	....	580,024	429,680
Time to maturity (*T*)	0.08	0.17	0.25	0.33	....	14.92	15.00
Strike (*K*)	1.04	1.04	1.04	1.04	....	1.04	1.04
Power σ	59.1%	58.6%	58.1%	57.6%	....	14.8%	14.7%
NG σ	52.8%	50.3%	47.9%	45.5%	....	10.1%	10.0%
Net option value[a]	(14,923)	148,385	349,212	1,516,525	....	799,240	704,388
Kirk unit price	1.36	2.78	4.51	14.64	....	7.72	6.89
σ	74.37%	72.51%	70.74%	68.86%	....	16.83%	16.73%
F	0.8776	0.9036	0.9600	1.4337	....	1.4096	1.2990
$d1$	−0.5006	−0.1945	0.0614	1.1048	....	0.8532	0.7277
$d2$	−0.7153	−0.4905	−0.2923	0.7073	....	0.2031	0.0797

[a]Note: Option value net of total O&M costs

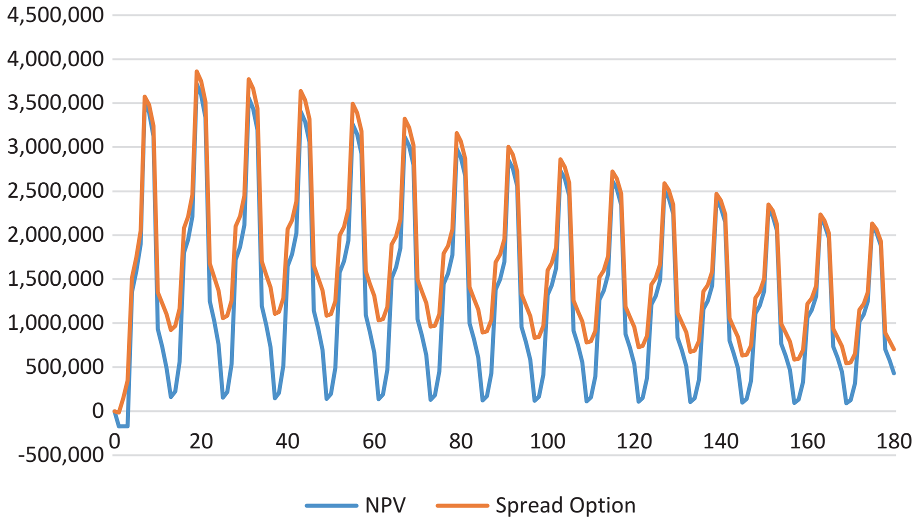

Fig. 8.17 Spread option and DCF intrinsic values by month

the energy markets implies that there is time value in the option, as there is some probability that the spark spread can be positive in these months. Consequently, the intrinsic DCF valuation of the peaking plant is likely to understate the value of the NG power generation plant. Value can also be maximized by considering the natural gas power generator as an asset that can be traded through its option value.

A number of factors can have an influence on the value of natural gas power generator. The 15-year life used for the power plant in this case study is arbitrary, and can feasibly be extended to value a new power plant. A benchmark for the cost of building a natural gas power plant is US$500,000 per MW. For the 300 MW NG generator illustrated in this case study, therefore, the cost to build would be US$150,000,000 million.

The value captured by including the volatility in the valuation will therefore have a significant impact on management decisions such as whether to build, divest or shut down a generator. Another driver of value that will influence decisions is the heat rate, the efficiency at which natural gas is converted into electricity. New natural gas generators will typically have lower heat rates than older plants, which will produce a relatively wider spark spread, and therefore have a competitive advantage over older generators.

Bibliography

BP Energy Outlook 2018.

Challa, R. Discovering multiple interacting options, Energy and Power Risk Management, July 2000.

Chignell, S, Robert Gross, R. Not locked-in? The overlooked impact of new gas-fired generation investment on long-term decarbonisation – A case study of lock-in to new CCGT in the UK, Imperial College Centre for Energy Policy and Technology Working Paper Ref: ICEPT/WP/2010/012, May 2012.

Clancy, H. Microsoft has figured out a way to reduce risks associated with PPAs, Greenbiz.com, October 16 2018.

Clewlow, L. and Strickland, C. Energy Derivatives, Pricing and Risk Management, Lacima, 2000.

Clewlow, L. and Strickland, C. Implementing Derivative Models, Wiley, 1998.

Financial Times. Sheppard D, Raval, A. Oil producers face their life or death question, June 19 2018.

Financial Times. Sheppard, D. Peak Oil forecast for 2036, July 2018.

Fortune. Ball, J. Shell Faces Lower Forever, February 1 2018.

IEA Key World Energy Statistics 2018.

IEA New Policies/Sustainable Policies.

IEA World Energy Outlook 2018.

Johnson, K, Thomas, N. Wind farm control addressing the aerodynamic interaction among wind turbines, 2009 American Control Conference, St Louis, Missouri, 2009.

Kirk, E, Correlation in the energy markets, Managing Energy Price Risk, Risk Publications (1995).

MIT Energy Initiative. The future of solar energy, 2015.

Parkinson, G. New solar "firming" contracts to boost corporate demand for big solar farms. reneweconomy.com, April 23 2018.

Schwartz, E.S. The stochastic behaviour of commodity price: implications for valuation and hedging, The Journal of Finance, LII(3): 923–73, 1997.

Shell energy transition report 2018.

U.S. Energy Information Administration Outlook 2017.

Wind Energy Explained: Theory, Design and Application, J F Manwell, J G McGowan and A L Rogers, 2nd edition, John Wiley and Sons, 2009.

World Wind Energy Report 2019.

9

Pharmaceuticals and Biotechnology

9.1 Overview

The pharmaceutical (pharma) industry dominates the healthcare sector with a market capitalization of $2.7 trillion, while the biotechnology (biotech) industry is a significant component of the sector with a market capitalization of $863 billion. Both industries focus on the discovery, development, manufacturing and sale of prescription drugs, and have drivers that are both similar and unique to the industries. These include a focus on innovation and patents, drug licensing and a significant regulatory and political environment.

There are, however, some significant differences between the two industries. Pharmaceutical drugs are generally based on chemicals defined as small molecules, while biotech drugs are derived from living organisms such as proteins and are defined as large molecules. The biotech industry also typically focuses on diseases that have no known cures. While both industries are involved in many types of innovation, the discovery and development of new chemical and biotech drugs is generally the most significant.

The drug industry is research and development (R&D) intensive, as it is essential for firms to have a pipeline of new drugs in development due to the revenue constraint imposed by the life of a drug's patent. A drug patent typically has a duration of 20 years, during which a drug firm has exclusive rights on the patent. With an average of ten years in development, drug firms therefore have around ten years to recover a drug investment and make a profit. Once the patent expires, other firms can sell the drug, described as a generic, with off-patent drugs potentially losing up to 80% or more in revenues.

J. Rogers, *Strategy, Value and Risk*, Global Financial Markets,
https://doi.org/10.1007/978-3-030-21978-9_9

The drug value chain follows a discovery, development and commercialization process, with all healthcare products regulated to some extent in their manufacture, testing and use. Alliances have become a fundamental component of pharmaceutical and biotechnology business models to minimize the risks associated with bringing new drugs to the market. Firms now rely on these partnerships at multiple stages within their value chains. For large pharmaceutical firms, these partnerships provide a source of innovation and the ability to outsource R&D, while providing access to capital, technical expertise and global reach for biotech companies. Managing a portfolio of strategic alliances is therefore a critical component of the industry innovation processes and value strategies.

A drug has two life cycles, with a development life cycle before the drug's market launch, and the product life cycle once the drug is released for commercial sale. The development life cycle can last up to 10–15 years, and is divided into a number of stages that a drug follows before entering the market. The commercialization process starts with the product approval, with marketing expenses—typically, the largest component of a drug's commercialization costs—spread over the life of the drug's patent.

9.2 New Drug Development

The drug development process follows a number of stages over which drug firms organize evidence to present to regulators that a drug can be safely, effectively and reliably manufactured for the intended medical condition. Drug firms also have to make decisions as to whether to abandon or to continue a drug's development using the technical and market information available at the start of each stage. In general, drugs go through the following stages in the US before reaching the market:

1. *Discovery:* significant resources are applied to the development of new molecular entities (NME), with many abandoned at this stage.
2. *Pre-clinical:* NMEs are screened for activity and toxicity. An Investigational New Drug application is then filed with the Food and Drug Administration (FDA) by the drug firm if the NME has the potential for further development. Once the Investigational New Drug application is approved, the drug firm can continue development through the testing of the drug in clinical trials.

3. *Clinical trials*: these generally have three consecutive phases:

    - *Phase I*—tests are performed on a small number of volunteers to gather information on the drug's toxicity and to establish safe dosages. Data on the compound's absorption, distribution, metabolic effects, toxicity and excretion is also compiled.
    - *Phase II*—trials are performed on subjects with the disease or condition for which the drug is intended, and are designed to obtain evidence and data on safety and efficacy.
    - *Phase III*—large-scale trials designed to further establish effectiveness are conducted on patients. The large sample size of these trials is designed to approximate the drug's intended use and uncover any potential side effects.

4. *FDA filing and review:* once the clinical development phases are finalized and sufficient evidence has been compiled for approval, the drug firm submits either a New Drug Application (NDA) or a Biological License Application (BLA) to the FDA for review. Marketing can start for approved uses once the notification is received from the FDA.
5. *Post-approval:* the drug firm performs additional research in support of marketing and product extensions as it collects revenues from sales of the new drug.

9.3 Valuation

The pharma and biotech industries' unique value drivers present a range of issues in regards to valuation. The three most common valuation approaches used in drug valuations are the Net Present Value (NPV), Expected Net Present Value (ENPV) and real options methods, and all require assumptions in regards to revenues, development costs, probabilities of success, development phase durations, the cost of capital and other cash flow parameters.

The trends in these factors over the last twenty years include the increasing uncertainty in revenues, the increasing development costs and the decrease in the success probabilities for drug development. DiMasi et al. (2003) observed that average clinical phase costs have increased by a factor of five in real terms since the authors' previous study (1993). These trends are having an influence on the cost of capital for the pharma and biotech industries, and on investments in R&D generally.

Table 9.1 Revenues by quality category ($000)

Drug quality	Launch 1	Launch 2	Launch 3	Launch 4	Launch 5	Probability
Dog	110	3310	6620	6620	6620	10%
Below average	100	3720	7440	7440	7440	10%
Average	1100	33,100	66,200	66,200	66,200	60%
Above average	44,130	209,620	413,720	661,960	661,960	10%
Breakthrough	44,130	275,820	772,290	1,323,920	1,323,920	10%

Source: Myers and Howe (1997)

The following data is drawn from various sources to illustrate the valuation methods.

9.3.1 The Revenues

Myers and Howe (1997) defined five categories for the quality of drugs that reached the marketing stage, with probabilities for each drug quality of 10% with the exception of 60% for an average quality. Myers and Howe also found the revenues for each quality category to be highly skewed.

Table 9.1 illustrates the five drug quality categories, the annual revenue assumptions for each category and the quality probabilities. The Myers and Howe revenues ramp up after launch, with assumed annual peak revenues for breakthrough drugs of $1.3 billion a year, which compares to $7.4 million for below average drugs and $6.6 million for dogs.

The Myers and Howe revenues are in 1994 dollars, and were inflated to 2010 for the valuation illustrations using GDP data. Figure 9.1 shows the timeline for new drug revenues, with Myers and Howe (1997) data for the first 13 years and the following years sourced from the Office of Technology Assessment 1993 report (US Congress), with both adjusted for inflation.

A drug's product life cycle typically peaks just before the expiry of its patent. Generic versions of the drug can be marketed once the patent expires, with revenues for the off-patent drug generally falling.

9.3.2 The Research and Development Costs

New drug discovery and development is a long and expensive process. Information on the costs of drug innovation is therefore essential for the valuation of drug investments. As with any investment analysis, both the amounts and timing of expenditures need to be identified, as R&D outlays have both an opportunity and a direct cost. The probabilities of a drug's success are factors that account for the risk associated with each development stage. These success probabilities are used to scale the cash flow estimates for each stage.

Table 9.2 shows for each drug development and commercialization stage:

- the pre-tax cost assumptions
- the duration of each stage, and
- the probability of success for the completion of each stage conditional on completion of the previous stages

Total costs are out-of-pocket and full-time employee combined. The cost for each stage is equally weighted to each year when a stage has a duration

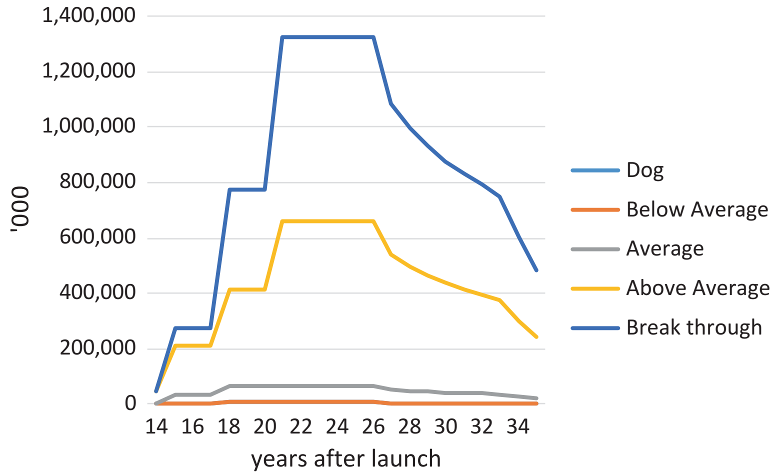

Fig. 9.1 New drug revenue streams by quality category ('000). (Source: Years 14 to 26: Myers and Howe (1997); years 27 to 36: US Congress, Office of Technology Assessment (1993), GDP inflation index)

Table 9.2 R&D stage (pre-tax) costs, durations and conditional success probabilities

Stage	Cost ($000)	Duration (yrs)	Probability
Discovery	13,500	4.5	51%
Pre-clinical	5000	1	69%
Clinical:			
Phase I	15,000	1.5	54%
Phase II	40,000	2.5	34%
Phase III	150,000	2.5	70%
FDA filing	40,000	1.5	91%
Launch	44,275	9	100%

Source: Paul et al. (2010); launch costs: Myers and Howe (1997) inflated to 2010

Table 9.3 Additional cash flow parameters

Parameter	% revenue assumption	Source
Cost of sales	20%	Industry estimate[a]
Marketing expense:		Myers and Howe
Year 1 post launch	100%	
Year 2 post launch	50%	
Year 3–4 post launch	25%	
Year 5–13 post launch	20%	
General and admin expenses	11.1%	US Congress
Working capital	17%	US Congress
Tax rate (% operating profit)	35%	Myers and Howe
Revenue inflation adjustment	–	–

[a]From public financial reports

longer than a year, and weighted within the relevant year for the half-year durations.

Table 9.3 shows the additional cash flow assumptions.

9.3.3 The Cost of Capital

Most of the weight in the pharma industry's cost of capital consists of equity, with debt typically making up less than 10% of market valuations.

DiMasi et al. (2003) found an average real rate of 11.1% and an average nominal rate of 15.1% for the pharmaceutical industry's cost of capital from 1985 to 2000. Myers and Shyam-Sunder (1996) analysed the cost of capital for smaller biotech firms. These firms typically have fewer commercialized products and more early stage R&D projects than large pharma firms, which was reflected in the study results, with an average cost of capital nominal rate of 19% and a real rate of 14%. DiMasi et al. (2007) found a downward trend in the cost of capital for biopharmaceutical firms, with real and nominal costs of capital of 12.5% and 17% in 1994 to 10% and 13% by 2004.

Drug valuation can also be framed as a compound option, where a firm has a series of call options during drug development, with the exercise price equivalent to the future R&D costs. Myers and Howe (1997) reduced this valuation problem to two discount rates, with a higher discount rate for the net revenues and a lower discount rate for future costs. The two real discount rates were specified as 6.0% and 9.0% for R&D costs and revenues, respectively. Adding a five-year GDP inflation average estimate of 1.9% to the real discount rates yields nominal discount rates of r_d = 7.9% for the R&D development stage, and r_c = 10.9% for the commercialisation or revenues stage.

Large pharma firms typically bundle their R&D projects into portfolios of later stage marketed products. This implies that the cost of capital, and therefore, the hurdle rate or required return on R&D investments should actually be actually higher. An alternative source for the market's cost of capital expectations for individual R&D investments are smaller biotech firms or the rates of return used by venture capitalists. Stewart (2001) argues that a 20% discount rate should be for biotechnology firms, as this is the general benchmark hurdle rate used by venture capital and pharmaceutical firms, the main sources of biotech capital.

9.4 NPV Valuation

The NPV method is a standard valuation technique used in drug valuations. The NPV method nets the present value of revenues and costs, and if the value is positive, then the drug is profitable and adds value. One issue with pharma and biotech valuations is that the revenue and cost cash flows are not the same in terms of timing or likelihood. In an NPV analysis, all risk is reflected in the discount rate. In the past, a number of researchers observed that the net present value of new drug development investment was approximately zero.

Table 9.4 illustrates an NPV valuation of an NME drug in the discovery stage using the data in Tables 9.1, 9.2 and 9.3. The cash flow assumptions include:

- a cost of capital of 11%, which reflects all risks in the example
- the present value of the R&D costs, which assume a 100% probability for each development stage, and

Table 9.4 NPV of a drug at the discovery stage ($000)

Drug phase:		R&D NPV	
	Discovery	(10,494)	
	Pre-clinical	(13,386)	
	Phase I	(21,005)	
	Phase II	(37,525)	
	Phase III	(85,247)	
	FDA filing	(95,557)	
Quality of drug:		Quality prob	Revenue NPV
	Dog	10%	4168
	Below average	10%	4687
	Average	60%	37,148
	Above average	10%	344,353
	Breakthrough	10%	664,169
	NPV:		124,026

- the NPV for the launch stage, which is derived from:
  - the revenues for each drug quality from year 14 minus taxes, marketing, general and administration costs and other expenses, and
  - the sum of the product of the net revenue for each drug quality scaled by the probability for that drug quality

The NPV of the drug is therefore the sum of the present value of the (negative) R&D costs and the NPV of the launch net revenues:

$$(95{,}557{,}000) + 124{,}026{,}000 = 28{,}470{,}000.$$

9.5 The Expected NPV Method

The Expected Net Present Value (ENPV) or Risk-adjusted Net Present Value (rNPV) method is used extensively in the pharma and biotech industries for drug valuations. The method includes risk as a probability in the valuation of future cash flows. Multiplying the discounted cash flows with their success probability estimates derives the expected values.

ENPV can be specified for drug valuations as:

$$\text{ENPV} = \sum_{i=1}^{7} \rho_i \sum_{t=1}^{T} \frac{\text{DCF}_{it}}{(1+r)^t} + \rho_7 \sum_{j=1}^{5} q_j \sum_{t=1}^{T} \frac{\text{CCF}_{jt}}{(1+r)^t} \tag{9.1}$$

where:

- $i = 1,\ldots n$ is an index for each drug stage (discovery to launch)
- ρ_i is the probability for each drug end stage i,
- DCF_{it} is the expected development stage cash flow at t
- $j = 1,\ldots 5$ is the index for the drug quality
- q_i is the probability for drug quality j
- CCF_{jt} is the expected commercialization cash flow at t for drug quality j
- r is the discount rate for the development and commercialization cash flows
- T is the maturity for all future cash flows

Table 9.5 illustrates the ENPV results for a discovery phase NME drug using a cost of capital of 11%.

Table 9.5 ENPV of an NME drug in the discovery phase ($000)

			1.	2.	3.	4.	5.
Phase:	*i*	*j*	ρ_i	q_j			
Discovery	1		49.0%		(10,494)		(5142)
Pre-clinical	2		15.8%		(13,386)		(2116)
Clinical phase I	3		16.2%		(21,005)		(3400)
Clinical phase II	4		12.5%		(37,525)		(4706)
Clinical phase III	5		1.9%		(85,247)		(1652)
NDA submission	6		0.4%		(95,557)		(389)
Launch:	7						
Dog		1	4.12%	10%	(95,557)	4168	(376)
Below average		2	4.12%	10%	(95,557)	4687	(374)
Average		3	4.12%	60%	(95,557)	37,148	(1442)
Above average		4	4.12%	10%	(95,557)	344,353	1024
Breakthrough		5	4.12%	10%	(95,557)	664,169	2340
ENPV:							(16,235)

The calculations are:

1. ρ_i represents:

   – the probability at the R&D phase, derived as one minus the probability of success, or probability of failure, times the product of the success probabilities for the previous stages, and
   – for the approval stage, the product of the probabilities of success from Table 9.2.

2. q_j is the probability for each of the drug qualities.
3. Column '3' is the cumulative sum of the present value of the development costs at each stage.
4. Column '4' is the commercialization phase NPV, or revenues net of cost of sales, general and administration costs, marketing expenses and taxes for each drug quality revenue stream.
5. Column '5' is the sum of columns '3' and '4' multiplied by column '1' and for revenues columns '1' and '2', the total of which derives the ENPV of ($16,235,000).

9.6 The Real Options Method

Biotech firms often have a significant premium over book value, or high discounted cash flow valuations well before a drug finally reaches the market. This premium can be defined as a growth option, where a biotech firm's R&D

expenditure is an asset that, although has no current cash flows, does have the probability of growth.

The market value of a firm can also be framed as having two components, the assets in place and the growth options. A firm's value is derived from both the expected profitability from current products and the growth potential from drugs in development.

The concept of growth options can be applied to the optimal capital structure for pharma firms. Well over half the value of a large pharma firm is derived from growth options, with the remainder consisting of the current products or the assets in place. As R&D is the driver of value or a substantial source of expected future cash flows, a pharma firm needs either sufficient equity or access to equity capital to maintain its R&D programmes. Major pharmaceutical firms maintain low debt levels as a result, with biotech firms carrying almost no debt, as R&D expenditure is the principal driver of value.

The value of an NME drug project is illustrated using a real options approach. Kellogg and Charnes (2000) specified a binomial tree to derive a biotech firm's growth option. The value of the growth option was derived by adding a second binomial tree that represents a research phase NME, with the second option value added to the end branch of the initial NME binomial tree.

The variables used in the binomial tree are:

- the asset value A at t_0
- the asset's standard deviation σ
- the exercise price amount and timing
- the probability of moving to the next development phase, and
- the risk-free rate r

The binomial tree's asset value A is derived as the expected commercialisation cash-flow values discounted at an 11% cost of capital at t_0:

$$A = \sum_{j=1}^{5} q_i \sum_{i=1}^{T} \frac{\mathrm{CCF}_{jt}}{\left(1+r_{coc}\right)^t} \tag{9.2}$$

The asset volatility is specified as the standard deviation $\sigma = \ln\left(h/A\right)^{1/l}$, with h representing the maximum value at node (i,j) for the value of the asset, where j = asset level and $i = l = T$.

$$h = \max_j \left\{ \sum_{t=1}^{T} \frac{\mathrm{CCF}_{jt}}{\left(1+r_{\mathrm{coc}}\right)^t} \left(1+r_{\mathrm{COC}}\right)^l \right\} \tag{9.3}$$

The underlying asset A follows a binomial process, and at any time can only change in either one of two possible values. For a time period Δt, the asset A can move to up to uA or down to dA, with u and d determining the volatility and average behaviour of the asset. The parameters u and d are derived as:

$$u = e^{\sigma\sqrt{\Delta t}} \quad and \quad d = \frac{1}{u} \tag{9.4}$$

The payoffs for each node within the binomial tree are:

$$P_k = \max\left[E_k\left(\theta_t\right) - \mathrm{DCF}_t, 0\right] \tag{9.5}$$

where θ_t is the probability at t of continuing to the following t, and DCF_t represents the R&D cost at t. The binomial risk-neutral probabilities are derived as:

$$p = \frac{e^{r\Delta t} - d}{u - d} \tag{9.6}$$

where r is the risk-free rate.

The value $V_{t,k}$ represents the option value for each t and k within the tree, with each payoff value adjusted by the probability of success θ_t at each development phase, and DCF_t the equivalent to the exercise price for each Δt.

$$V_{t,k} = \max\left\{\left[V_{t+1,k}\,\rho + V_{t+1,k+1}\left(1-\rho\right)\right]e^{-r\sqrt{\Delta t}}\theta_t - \mathrm{DCF}_t, 0\right\} \tag{9.7}$$

Two scenarios for the cash flow assumptions in the valuation are illustrated:

- 100% probabilities of success in the development stage cash flows for the R&D pre-commercialization costs, and
- the actual probabilities of success from Table 9.2

Table 9.6 shows the binomial tree and parameters for the commercialisation cash flow asset. Table 9.7 shows NME drug call option payoffs with the success probabilities equal to 1, and Table 9.8, the NME drug call option payoffs with the probabilities of success.

For the 100% probabilities of success in the R&D stages, the value of the option on the NME drug is $30,652,000. If the probabilities of success are included, the option on the NME drug has a zero value, or is out-of-the-money. The tree is most sensitive at the phase II probability of 34%.

Table 9.6 The 14-step binomial tree for the commercialization cash flow asset

A at t_0	*h*	*T*	σ	*r*	*df*	*dt*	*u*	*d*	*p*	1 − *p*
188,281	3,482,109	9	32%	1.89%	0.9813	1	1.383	0.723	0.449	0.551
Year	0	1	2	3	4	5	6	7	8	9
R&D costs		5000	10,000	13,000	16,000	16,000	60,000	60,000	43,333	26,667
Asset tree	188,281	260,368	360,056	497,911	688,547	952,171	1,316,731	1,820,869	2,518,028	3,482,109
		136,152	188,281	260,368	360,056	497,911	688,547	952,171	1,316,731	1,820,869
			98,456	136,152	188,281	260,368	360,056	497,911	688,547	952,171
				71,197	98,456	136,152	188,281	260,368	360,056	497,911
					51,485	71,197	98,456	136,152	188,281	260,368
						37,230	51,485	71,197	98,456	136,152
							26,922	37,230	51,485	71,197
								19,469	26,922	37,230
									14,078	19,469
										10,181

Table 9.7 The binomial tree NME drug call option payoffs with success probabilities equal to 1

Probabilities	1	1	1	1	1	1	1	1	1	1
Option tree	0	1	2	3	4	5	6	7	8	9
	30,652	67,161	144,353	278,246	477,939	753,851	1,130,931	1,692,670	2,448,527	3,455,442
		2012	15,929	58,906	149,448	299,590	502,747	823,972	1,247,230	1,794,202
			–	–	11,312	62,048	174,256	369,711	619,046	925,505
				–	–	–	2482	132,169	290,555	471,244
					–	–	–	7953	118,780	233,702
						–	–	–	28,955	109,485
							–	–	–	44,530
								–	–	10,564
									–	–
										–

Table 9.8 The binomial tree NME drug call option payoffs with the probabilities of success

Probabilities	1	0.69	1	0.54	1	0.34	1	1	0.7	0.91
Option tree	0	1	2	3	4	5	6	7	8	9
	–	–	–	1041	59,071	130,754	660,517	1,039,397	1,542,333	3,142,052
		–	–	–	–	32,370	260,364	486,037	777,107	1,630,324
			–	–	–	–	51,115	196,673	376,954	839,809
				–	–	–	–	45,358	167,705	426,432
					–	–	–	–	58,284	210,268
						–	–	–	1066	97,232
							–	–	–	38,122
								–	–	7213
									–	–
										–

Bibliography

US Congress, Office of Technology Assessment, Pharmaceutical R&D: Costs Risks and Rewards, OTA-M-522, Washington, DC: US Government Printing Office, February 1993.

DiMasi, J and Grabowski, H. The Cost of Biopharmaceutical R&D: Is Biotech Different? Managerial And Decision Economics 28: 469–479, 2007.

DiMasi, J, Hansen, R, and Grabowski, H. The price of innovation: new estimates of drug development costs, Journal of Health Economics, 22, 151–185, 2003.

DiMasi, J.A., Hansen, R.W., Grabowski, H.G., Lasagna, L., 1991. Cost of innovation in the pharmaceutical industry. Journal of Health Economics 10, 107–142.

Stewart, J. Allison, P. and Johnson, R. Putting a price on biotechnology, nature biotechnology, Volume 19 Sept 2001.

MIT Roundtable on Corporate Risk Management, Journal of Applied Corporate Finance, Volume 20 Number 4 Fall 2008.

Myers, S. and Shyam-Sunder, L. Measuring pharmaceutical industry risk and the cost of capital, in Helms, R.B. (ed.) Competitive Strategies in the Pharmaceutical Industry, Washington, DC, American Enterprise Institute Press, 1996.

Lütolf-Carroll, C, From Innovation to Cash Flows, Wiley, 2009.

Kellogg, D. and Charnes, J.M. Real options valuation for a biotechnology company, Financial Analysts Journal, May/June 2000.

Kelly, M, Teufel, S, Fisher Investments on Health Care, Wiley, 2011.

Myers, S. C., Howe, C. D., A Life Cycle Financial Model of Pharmaceutical R&D, Program On the Pharmaceutical Industry, MIT, 1997.

Paul, S, Mytelka, D, Dunwiddie, C, Persinger, C, Munos, B, Lindborg, S and Schacht, A. How to improve R&D productivity: the pharmaceutical industry's grand challenge, Nature Reviews Drug Discovery 9, 203–214, 2010.

10

A Growth Firm

10.1 Overview

Industry sectors are generally defined through the products they create or the services on offer. Industries are also identified by their response to economic and business cycles, typically defined as growth, defensive or cyclical. The industry life cycle considers an industry's viability over time, with the four stages of pioneer, growth, mature and decline indicating a phase in an industry's evolution.

The term 'growth' as used in investment analysis has a range of interpretations. While growth is typically applied to sales, earnings and assets, growth generally describes the capability to create value. A growth industry is typically identified with new technologies and products, while a growth firm is able to grow sales and earnings, regardless of its stage in the business cycle. Growth firms can have annual growth rates that exceed 12–15%, and while earnings will be typically steady or growing, these firms will generally consume cash, with investments exceeding cash generated internally.

Growth fundamentals are not the same for all firms. Three attributes can be identified that differentiate growth firms. The classic growth firm typically has innovative products or technologies in markets where demand is yet to be determined. The market share growth firm is located within a mature industry where growth is derived from increasing market share. A consolidating firm grows through acquisitions, typically in mature and fragmented industries, where cost synergies and valuations are the focus.

J. Rogers, *Strategy, Value and Risk*, Global Financial Markets,
https://doi.org/10.1007/978-3-030-21978-9_10

10.2 The Valuation of Growth

The value of a firm is equal to the sum of the value of its debt and equity, or the value to all claimants on the firm. Summing the claims on the firm therefore represents the firm's value. The valuation of equity can be based on earnings, cash flow or assets. Earnings, or the income statement's net income, is the fundamental bottom-line metric for shareholder value. Growth is typically framed as a firm's capacity to grow earnings. The issue with earnings growth for valuation is that a firm can create growth in earnings though investment and accounting techniques that do not add value.

The valuation of equity based on the present value of future free cash flows focuses on the cash flow from operations with capital expenditure deducted. Cash flow from operations is the total cash flow derived from operations. Free cash flows will generally fluctuate over the life cycle of a product or firm, and are analysed in three life cycle phases. During the growth phase, cash flow from operations is usually down while new products are being established and investments consume cash. The maturity phase will have positive free cash flow if the new venture is successful and investment requirements are reduced. Finally, the decline stage is reached if product demand or cash flow from operations diminishes. Free cash flow at this point should continue while investment decreases.

The cash flow life cycle illustrates why free cash flow can be an imperfect metric for valuing a firm. Free cash flow does not measure the value added from operations over a period in time. Profitable firms can potentially have negative cash flows, while those with significant free cash flow can be barely profitable. A negative cash flow from investing more cash into operations than being received will give a negative Net Present Value (NPV). This investment could produce a return from cash flow from operations at some later stage. The three life cycle phases are also not consistent for all firms. Some firms can maintain growth consistently for extended periods, while others are unsuccessful at the initial phase. Firms have also been able to react to the risk of decline and continue in the mature phase.

A more appropriate metric for measuring growth is the residual earnings model. A firm's net assets quantify the book value of equity, which represents the total shareholder investment in the firm. The net assets are used in the firm's operations to generate shareholder value. The residual earnings measure captures the value in excess of book value over the cost of capital. Residual income models distinguish the excess earnings for a time period over an investor's required rate of return on equity. Residual earnings therefore represent the differential between earnings and investors' expected return on book value:

$$\text{Residual Earnings}_t\left(\text{RE}_t\right) = \text{Earnings}_t - \left[\left(p_{\text{E}} - 1\right) \times \text{common shareholders equity}_{t-1}\right]$$

where $p_{\text{E}} - 1$ is the required return on equity. The firm value added for shareholders can therefore be defined in terms of the opportunity cost to investors.

10.3 A Growth Case Study

Wave Nouveau Materials Inc. (Wave Nouveau) manufactures industrial materials, and focuses on using science, engineering and chemistry to conceive, develop and commercialize new industrial products. Wave Nouveau has also developed proprietary innovative processes for the manufacture of materials, equipment and various components. In addition, the firm has created an intensive research and development (R&D) programme and partnered with commercial firms and government for the development of new technologies.

Over time, Wave Nouveau has developed markets worldwide, creating a reputation for performance and innovation with its products. The firm framed its growth strategy as leveraging its materials and manufacturing technologies to compete in the growing environmental and alternative energy markets. To take advantage of these trends, the firm expanded its operations by adding a production facility for the development and production of batteries. Wave Nouveau's strategy is to leverage its existing manufacturing and customer base, and use its new facility for the manufacturing of batteries. The condensed financial statements for Wave Nouveau are illustrated in Table 10.1.

Table 10.1 Condensed financial statements for Wave Nouveau ($000)

Year	20X1	20X2	20X3	20X4
Balance sheet:				
Assets	400,000	432,000	488,000	528,000
Liabilities	200,000	200,000	220,000	220,000
Shareholders' equity	200,000	232,000	268,000	308,000
Income statement:				
Revenues	80,000	90,000	100,000	120,000
Expenses	48,000	54,000	60,000	72,000
Net income	32,000	36,000	40,000	48,000

Table 10.2 Residual earnings trend for Wave Nouveau ($000)

Year	20X1	20X2	20X3	20X4
Earnings	32,000	36,000	40,000	48,000
Shareholders' equity	200,000	232,000	268,000	308,000
Cost of equity	24,000	27,840	32,160	36,960
Residual earnings	–	12,000	12,160	15,840

The residual earnings with the required return on equity of 12% follow in Table 10.2.

Wave Nouveau grew fixed assets by investing in plant and equipment for the manufacturing of batteries, and financed this expansion through debt in 20X3. The firm also leveraged sales through its existing customer base, and therefore maintained the trend in sales and marketing costs. As a result, expenses remained constant at 60% of revenues. As there were no dividends, the total earnings for each year flowed into shareholders' equity. The growth in residual earnings can be attributed to the increase in revenues from battery sales and maintaining the ratio of expenses to revenues.

10.4 Growth Options

Defining a firm's future growth investments as call options on real assets provides a framework for approximating the value of the firm's growth options. A firm's value (V) can be broken down into the current earnings stream, or the value of assets-in-place (VAIP), and the value within the firm's potential future investment choices, or the value of growth options (VGO):

$$\mathrm{V} = \mathrm{VAIP} + \mathrm{VGO} \tag{10.1}$$

The value of growth options VGO is a function of the value that is ultimately derived from a firm's investment choices, while the assets-in-place VAIP represents assets whose value is independent of these investments. This relationship can be extended by quantifying the firm's growth option value (GOV) as the differential between its market value and capitalized current earnings stream. The capitalized current earnings stream symbolizes the firm's VAIP with no growth factored into the value. The percentage of a firm's value that can be sourced to growth options, the growth option value (GOV), is therefore represented by:

$$\mathrm{GOV} = \mathrm{VGO} / V = \left[V - \text{current earnings} / \text{discount rate}\right] / V \tag{10.2}$$

This relationship can be extended through the use of the economic profit concept, which considers the firm's equity cost of capital. Earnings only take into account the cost of debt capital, whereas economic profit considers a firm's earnings or profit net of all capital costs, including the cost of equity:

$$\mathrm{EP} = \mathrm{NOPAT} - \left[\mathrm{CI} * \mathrm{WACC}\right] \tag{10.3}$$

where,

CI = capital invested
EP = economic profit
NOPAT = net operating profits after tax
WACC = the weighted average cost of capital

The firm value (V) includes the book value of all equity, debt and other capital invested (CI) plus a residual factor over CI, the market value added (MVA):

$$V = \mathrm{CI} + \mathrm{MVA} \tag{10.4}$$

The firm's MVA represents the total NPV of all investments, which is the equivalent to the present value (PV) of the expected economic profit:

$$\mathrm{MVA} = \mathrm{PV\ of\ expected\ EP} \tag{10.5}$$

Expected EP therefore is composed of a current EP with no growth plus a future residual factor that represents EP growth, which can be either positive or negative due to declining operations or unsuccessful investment outcomes.

$$\mathrm{PV\ of\ expected\ EP} = \mathrm{PV\ of\ current} - \mathrm{level\ EP} + \mathrm{PV\ of\ EP\ growth} \tag{10.6}$$

The firm's value (V) can therefore be redefined using the above equations:

$$V = \mathrm{CI} + \mathrm{PV\ of\ current} - \mathrm{level\ EP} + \mathrm{PV\ of\ EP\ growth} \tag{10.7}$$

Combining the first two terms CI and PV of current-level EP is the equivalent to the firm's value of assets-in-place VAIP in Eq. (10.1), while the firm's value of growth options is established through the PV of EP growth or VGO in Eq. (10.1). The firm's growth option value (GOV) is derived by solving Eq. (10.7), and then, dividing the result with the firm value (*V*):

$$\mathrm{GOV} = \left[V - \mathrm{CI} - \mathrm{PV\ of\ current-level\ EP}\right]/V \tag{10.8}$$

where,

V = the firm's market value
GOV = the firm's growth option value as a %
CI = capital invested, the total book value of debt and equity
PV of current-level EP = EP perpetuity discounted by the firm's WACC

10.5 A Growth Option Case Study

The percentage of Wave Nouveau's value that can be sourced to growth options is examined using Eq. (10.8).

The book value of debt is not a sufficient proxy for the market value of debt in deriving a firm's enterprise value. The market value of a firm's debt can be implied from stock option prices. Geske and Zhou (2007a) imply the market value of a firm's debt from two simultaneous liquid at-the-money option prices that also account for a firm's capital structure.

A firm's stock return volatility is a function of its stock price level which, in itself, is a function of the firm's value, and is not assumed to be constant as in the Black and Scholes model. A firm's leverage ratio increases as the value of its stock declines, and therefore, the equity's risk, and consequently, its volatility increase. Geske's approach therefore considers bankruptcy effects and includes Modigliani and Miller theorems.

The market value of the firm *(V)* was derived using Geske's compound option model. The firm value *(V)* was found by solving the three Eqs. (6.19), (6.20) and (6.21) simultaneously for three variables—the total firm market value *(V)*; the total instantaneous volatility of V, σ_V; and V^*, the value of *(V)* at a specific time in the future (T_1) where S_{T1} is equal to the strike price K of an option expiring at t_1. The option price C represents a quoted bid/ask average, near term at-the-money option price. At-the-money options are usually the deepest and most liquid.

The value for the firm's market value *(V)* was solved using the following variables for Wave Nouveau:

S = 400 million, the total market capitalization
M = 220 million, weighted duration of total debt
$K = S_T$
t_1 = option expiry (1 month)
T_2 = debt duration (5 years)
S = \$80.00
K = \$80.33
r = 5%
σ_s = 15%
C = \$1.38

Wave Nouveau has 5 million shares outstanding, and as the firm's current stock price per share is \$80, the total market capitalization (S) is \$400 million. The firm's market value (V) was derived using Eqs. (6.19), (6.20) and (6.21) as \$571 million. The capital invested (CI) for Wave Nouveau is equal to the sum of the book value of debt (\$220 million) and the book value of equity (\$308 million), which equals \$528 million. Table 10.3 summarizes the compound variables.

The total value of a firm is equivalent to the sum of the original investment, or the book value of debt and equity, and the MVA, or the NPV of the economic profit cash flows discounted at the weighted cost of capital. This is the equivalent to the total value of the firm as the sum of the original investment and the NPV of the sum of the residual earnings discounted at the cost of equity.

The residual earnings number in Sect. 10.3 is substituted for the economic profit parameter to arrive at Wave Nouveau's growth option value. A WACC of 5% is used for illustration. The perpetuity for the PV of current-level EP is derived as the 20×4 residual earnings of \$15,840 divided by the WACC of 5%, which equals \$316,800.

Table 10.3 Wave Nouveau's compound option variables for firm market value (V) (\$M)

Compound option model variable	Description	Variable
Asset price (S)	Total market capitalization	400
Strike underlying option (X_1)	Face value of debt (M)	220
Strike option on option (X_2)	ATM option price × total equity	402
Time to maturity—option on option (t_1)	Option maturity	0.08
Time to maturity—underlying option (T_2)	Duration of M	5.00
Discount rate (r)	Interest rate	5.00%
Volatility (σ_V)	Volatility	10.50%

Wave Nouveau's growth option value, or the firm's growth option value as a percentage of firm value is therefore:

$$\begin{aligned} \text{GOV} &= \left[571 - 528 - 0.317\right] / 571 \\ &= 7.48\% \end{aligned}$$

Bibliography

Geske, R. The valuation of compound options, Journal of Financial Economics, 1978.
Geske, R. and Zhou, Y. Capital Structure Effects on Prices of Firm Stock Options: Tests Using Implied Market Values of Corporate Debt, UCLA Working Paper, 2007a.
Geske, R. and Zhou, Y. Predicting Risk and Return of the S&P 500: Evidence from Index Options, UCLA Working Paper, 2007.
Tong, T. and Reuer, J. Corporate Investment Decisions and the Value of Growth Options, 2004.
Zhou, Y. Pricing Individual Stock Options on Firms with Leverage, Anderson School of Management at UCLA, 2007.
Penman, S. Financial Statement Analysis and Security Valuation, McGraw Hill, 2012.

11

Firm Transformation and Divesture

11.1 Industry Decline

Many industries follow life cycles that are defined by convention as pioneer, growth, mature and decline. An industry in decline either has negative growth or is not expanding at the general economic growth rate. These industries are typically identified by extreme price competition, excess capacity, a lack of innovation, a dwindling number of competitors and buyouts.

Firms within a declining industry can transition from the mature to the decline stage as a consequence of foreign competition, shifts in consumer behaviour and changes in technology and demographics. Declining revenues, margins, operating cash flows and earnings are all features of the final phase of a firm's existence. Other factors identified with declining firms include commoditized products, diminishing resources or substitutes appearing due to innovation.

Industry decline and falling market demand leads to uncertainty around the sustainability of product demand, which could either stabilize at some significantly lower level, or in the worst case, fade to where operations are no longer justified. Declining demand leads to industry-wide pressure to reduce capacity, as excess capacity diminishes profitability. Reducing capacity has significant costs, as both fixed and sunk costs can be a comparatively large percentage of an industry's total costs. These large sunk costs also create high barriers to exit and prolonged excess capacity in the industry, which can lead to lengthy periods of industry downsizing. Firms in a declining industry are also motivated to reduce investments in products when decreasing sale volumes lower profitability. This leads to a fall in product quality, and ultimately, firms exiting the industry.

J. Rogers, *Strategy, Value and Risk*, Global Financial Markets,
https://doi.org/10.1007/978-3-030-21978-9_11

11.2 Strategies in Declining Industries

The conventional strategies in declining industries are either to maximize cash flows from existing investments with no further investment or to divest. These strategies can be extended by dominating the market through increased investment, staying at these investment levels, then reducing investments, milking the cash cow, and finally, divesting.

Continuing to invest in the industry through research and development or marketing programmes has significant risks, and requires forecasts of declining product demand. In some cases, a firm can be successful in a turnaround and return to the maturity phase, or reposition itself to take advantage of new market trends. Many firms in a declining industry will, however, ultimately face liquidation. If the future for a firm appears bleak, then the best management alternative may be divesting the business and selling assets before the industry decline gains momentum.

The market for a declining industry's assets is dynamic and can diminish rapidly. A firm would focus on resale asset values to establish potential realized proceeds, and maintain the flexibility to sell before asset values significantly decline. The firm also has value if the assets have the flexibility to switch to operations for other products. Both the industry and a firm's exit barriers are also factors that will influence a firm's decision on the timing of an exit.

11.3 Industry Decline and Transformation

The newspaper industry is inundated with problems. Consumers are switching to other media such as streaming media, Internet news sites and blogs for news, while revenues from classifieds have also declined due to online classified websites. This has led to uncertainty in the industry's growth and future profit margins, with declines of 15–20% in the level of earnings. Many newspapers are shutting down operations as a result.

Attempts at increasing revenue through newspaper price increases have a significant negative influence on circulation, and therefore, the advertising charge rates that flow into revenue. There are growth areas in the newspaper industry, however, which include diversifying into niche newspaper distribution and investments in online media. Newspapers are now offsetting losses from print editions with online editions. A critical issue in this migration is the offset on the loss on revenue from traditional print media with the gains in revenue when a consumer switches to the newspaper's online content.

The Sunset Examiner is a morning newspaper based in a provincial city with a daily circulation of 180,000 on weekdays and 225,000 for a weekend edition. The newspaper is published from an editorial operations centre and printing plant in the provincial city. The publisher of *The Sunset Examiner*, Sunset Inc., is a single business newspaper firm. The circulation area covers the provincial city and surrounding counties. The firm has 1500 employees, over 1000 independent newspaper carriers, and operates on a 15% pre-tax profit margin.

The newspaper had enjoyed a monopoly on local news content for decades, with growth in circulation flowing into profit margins. In the late twentieth century, however, alternative sources of news became available—initially with cable news, and then, with Internet news sites. The newspaper's circulation started to decline as a result. While print advertising revenue had grown through to the end of the twentieth century, advertisers were aware of the declining circulation numbers and growth in advertising and classifieds revenue slowed considerably. By the start of the twenty-first century, the newspaper's advertising revenues from the print edition were declining year-on-year.

Offsetting the decline in revenue was difficult as most of the firm's costs were fixed costs. In an attempt to reverse the declining circulation, Sunset Inc. invested in marketing and redesigns to retain readers. These measures proved to be unsuccessful in reversing the decline in print circulation. Sunset Inc. was successful in establishing an online edition strategy to offset the decline in print revenues.

A firm as a going concern assumes that it will continue operations indefinitely, and firm value is based on this assumption. Managers and investors can, however, consider other values. A firm's liquidation also has value if the firm ceases operations and the assets are sold off. The firm as a going concern is one strategy while selling off the assets is another. Firm value is then either the maximum of the value as a going concern or the value in liquidation.

Sunset Inc.'s management and shareholders examined the value of the newspaper in liquidation while there was still a market for the newspaper's physical assets. The printing plant assets have the potential to be switched into alternative print media such as direct marketing and niche media markets.

A liquidation analysis involves valuing each asset in a sell off. Percentages for assets in liquidation founded on comparable firms are established. The sum of these asset values is netted against the liability redemptions and associated liquidation costs. A positive result provides a maximum value from which the investors' required return is deducted.

Table 11.1 illustrates Sunset Inc.'s liquidation analysis using balance sheet data. Percentages of asset values in liquidation were determined and the net

Table 11.1 Sunset Inc. liquidation analysis ($000)

Assets		Liquidation (%)	Values
Cash	50,000	100%	50,000
Accounts receivable	22,000	72%	15,840
Inventory	2000	55%	1100
	74,000		66,940
Plant & equipment	125,000	54%	67,500
Goodwill	(1000)	0%	–
	198,000		134,440
Liabilities & shareholder equity			
Short-term debt	50,000	100%	(50,000)
Other liabilities	50,000	100%	(50,000)
	100,000		(100,000)
Shareholder equity	100,000	Cost of shutdown	(5000)
	200,000		(105,000)
		Net liquidation value	29,440
		Shares outstanding	3,000,000
		Value per share	9.81

liquidation value calculated. The investors' required rate of return would consider the range in the percentage estimates for the assets in liquidation, the rate at which the firm is consuming cash and the probability of other firms also analysing the value in an acquisition.

The discount-to-book value of $29.4 million in liquidation value compared to the equity of $100 million reflects the value of the firm as a going concern. The book value per share is $33, while the value to investors is $9.81 per share. If the assets take six months to liquidate, then the present value at a 15% internal rate of return is $9.13 per share.

11.4 Investor Valuation of the Abandonment Option

The value to investors in a firm's liquidation can be defined as an option to abandon the firm for its exit value. This option can be valued as an American put on a dividend paying stock, where the stock price is the equivalent to a firm's cash flows and the strike is an uncertain exit value. The value of this option increases along with the firm's exit value. Structural change within an industry may therefore persuade investors to value and even exercise this option. The data in a firm's balance sheet can be used to value the option to abandon the firm for the exit value of its assets.

Berger, Ofck and Swary (BO&S) (1996) present a method for assessing the value of the abandonment option. The following equation defines the relationship between firm value and those attributes that determine the value of the abandonment option:

$$\text{VALUE} = \text{PVCF} + \text{P}\left(\text{PVCF},\text{EXIT},\text{SDEV}\right) \tag{11.1}$$

where,

VALUE = the firm's market value, or market capitalization
PVCF = the firm's present value of expected operating cash flows
P = American put option
EXIT = the firm's assets exit value
SDEV = standard deviation of the PVCF/EXIT ratio

The investors' decision to sell the firm is illustrated by Eq. (11.1), where firm market value is equal to the total of the value of the expected operating cash flows and the abandonment option value. Equation (11.1) can be redefined as a function of the option's parameters, where the value of the abandonment option is expressed as the percentage by which firm value exceeds the PVCF:

$$\left(\text{VALUE} / \text{PVCF} - 1\right) = \text{P}\left(1,\text{EXIT} / \text{PVCF},\text{SDEV}\right) \tag{11.2}$$

The value of the abandonment option is now the percentage by which the value of equity is greater than the after-interest cash flows for a firm as a going concern. There is no closed form solution available for the abandonment option; however, the relationships in Eq. (11.1) can be used to determine the option payoff. This option value is a function of the EXIT/PVCF ratio, which is defined as the excess exit value. The excess exit value in Eq. (11.2) is a put option's stochastic strike with the underlying stock value normalized to one.

The abandonment option value is at-the-money when the excess exit value is equal to one, in which case, a firm's exit value is equal to its expected cash flows. The option moves to in-the-money with an increase in the exit value, and out-of-the-money with an increase in the value of the expected cash flows. There is therefore a positive relationship between a firm's excess value and the value of the abandonment option.

Observable exit values for the assets of firms as going concerns are typically unavailable. The focus, however, is on the relationship between balance sheet data and the value of the abandonment option. BO&S derive estimates for the relationships between the book value of assets and the exit value, and find, on average, exit values of \$0.72 on the dollar for receivables, \$0.55 for inventory and \$0.54 for fixed assets.

BO&S derive the cash flow present value (PVCF) proxy through analysts' earnings forecasts. The use of earnings forecasts instead of cash flows has advantages and disadvantages. Earnings are measured for firms as going concerns, and therefore, a PVCF proxy based on earnings forecasts does not include the value of the abandonment option. Forecasts of distributed cash flows would include the expected cash flows from occurrences outside the going concern assumption such as a firm's exit.

The main disadvantage, however, is that earnings are not the same as cash flows, and therefore, the present value of earnings from a going concern should be modified to arrive at the PVCF as a going concern. These adjustments are required as capital expenditures may not match depreciation and working capital growth is not deducted from earnings. No adjustments are made for changes in the capital structure as these are unknown.

The specification of the PVCF proxy is:

$$PVCF = \sum_{t=1}^{n} \frac{EARN_t}{(1+r)^t} + \sum_{t=n+1}^{10} \frac{EARN_2 * (1+gr)^{t-2}}{(1+r)^t} + \frac{EARN_2 * (1+gr)^9}{(r-tg)} * \frac{1}{(1+r)^{10}} - CAPEX\_ADJ - WC\_ADJ \tag{11.3}$$

where,

PVCF = present value of analysts' forecasted going concern cash flows
$EARN_t$ = the forecast at year *t* of the analyst's after-interest earnings
r = the expected CAPM (Capital Asset Pricing Model) return
gr = the five-year earnings growth consensus forecast
tg = the terminal growth rate for earnings
n = number of years for the earnings forecast
t = year index
CAPEX ADJUST = present value of analysts' earnings forecasts adjustment down for the capital expenditures and depreciation difference
WC ADJUST = present value of analysts' earnings forecast adjustment down for working capital growth

The *CAPEX ADJUST* is specified as:

$$CAPEX\ ADJUST = \frac{(CAPEX_0 - DEPN_0)/(r-g)}{EQUITY_0}$$

where,

$CAPEX_0$ = the year 0 capital expenditures
$DEPN_0$ = the year 0 depreciation expense
$EQUITY_0$ = the year 0 market value of equity
g = the growth rate of excess capital expenditures

The *WC ADJUST* is specified as:

$$WC\ ADUST = \frac{\left(\left[0.5(gr) + 0.5(tg)\right] * 9.5\left[NETWC_0\right]\right)/r}{EQUITY_0}$$

11.5 An Abandonment Option Case Study

The estimates in Sect. 11.3 provided *The Sunset Examiner*'s exit value. The BO&S dollar book exit value estimates for receivables, inventory and fixed assets were used for each book value component and the book values of debt and payables were deducted. A ratio of one was used for cash and short-term marketable securities in non-inventory current assets. The exit value is therefore equal to the sum of $29.44 million.

The BO&S PVCF equation as specified in Eq. (11.3) has five components. The first is the sum of the analysts' discounted expected earnings forecasts up to two years. The second term represents the sum of the discounted earnings forecasts from three to ten years, based on analysts' long-term five-year earnings growth projections. The third variable is the present value of a perpetuity for earnings greater than ten years, with the assumption of a constant nominal terminal growth rate *(tg)*.

The fourth component deducts an adjustment for the net present value of future capital expenditure after subtracting future depreciation. Earnings forecasts are not cash flows that can be distributed, and growth in earnings usually requires capital investments that are greater than depreciation. BO&S arrive at a value of 12% for the capital expenditure adjustment. The fifth item, the working capital adjustment, is a similar adjustment representing the present value of the growth in working capital. BO&S derive a value of 5.5% for the working capital adjustment.

Table 11.2 illustrates Sunset Inc.'s condensed income statements. The decline in circulation has created a downward trend in net income, while advertising was stable due to Sunset Inc.'s online venture.

Table 11.2 Condensed income statements for Sunset Inc. ($000)

Revenues	2XX1	2XX2	2XX3	2XX4
Advertising	120,000	120,000	120,000	120,000
Circulation	35,000	33,000	32,000	30,000
Total revenue	155,000	153,000	152,000	150,000
Total operating costs	140,000	140,000	140,000	140,000
Net income	15,000	13,000	12,000	10,000

Table 11.3 PVCF earnings projections

Year	First term	df	Second term	gr
2XX5	8,571,429	0.9524	–	1.0200
2XX6	8,163,265	0.9070	–	1.0404
2XX7	–	0.8638	11,056,328	1.0612
2XX8	–	0.8227	11,841,328	1.0824
2XX9	–	0.7835	12,682,062	1.1041
2X10	–	0.7462	13,582,488	1.1262
2X11	–	0.7107	14,546,845	1.1487
2X12	–	0.6768	15,579,671	1.1717
2X13	–	0.6446	16,685,827	1.1951
2X14	–	0.6139	17,870,521	1.2190
2X15	–	0.5847	–	1.2434

Note: The third term in Eq. (11.3) is the perpetuity which is derived as $17,520,119

Table 11.3 illustrates the calculation of the first two terms in Eq. (11.3) using an earnings forecast of $9 million, which is based on a decline in circulation revenues from $30 million to $29 million for 2XX5. The earnings forecast to ten years and perpetuity for earnings from year 11 assume a constant 2% nominal terminal growth rate. The present value of future cash flows is then derived by adjusting the present value of future earnings with the deduction of the present value of future capital expenditures minus future depreciation and working capital growth. A CAPM discount rate of 5% is used for illustration.

The total of the sum for each of the first two terms, $16.7 million and $113.9 million, and the third term perpetuity of $17.5 million results in the unadjusted PVCF of $148.1 million. Deducting the capital expenditure adjustment of $17.8 million and the working capital adjustment of $8.2 million provides an adjusted PVCF of $122.2 million. Table 11.3 illustrates the PVCF earnings projections.

BO&S's research provides results for the variables in the valuation of the abandonment option. The median sample value of the abandonment option is 11.5% with an excess exit value of −0.761. Using the Sunset Inc. example, the market capitalization value of the firm is \$136.2 million and the excess exit value is −0.759. As the exit value of \$29.4 million is less than the PVCF of \$122.2 million, the value of the abandonment option is out-of-the-money for investors.

The factors that can influence the value of the abandonment option include the degree of specialization in a firm's assets. The option has greater value when the assets are less specialized. Current assets are less specialized than fixed assets, and therefore, provide greater option value, while land is less specialized than other fixed assets and also provides greater option value.

The model in Eq. (11.3) was stressed tested to identify the sensitivity of the model variables. The at-the-money value of the abandonment option was derived by adjusting the advertising revenue forecast, the discount rate and growth rate forecast. The earnings forecast for 2XX5 was adjusted down to \$2.2 million based on a decline in 2XX5 advertising revenues of approximately 5.7%. The discount rate was increased to 8% and the growth forecast lowered to minus 1%.

Table 11.4 illustrates the calculation of the stressed test first two terms in Eq. (11.3) using an earnings forecast of \$2.2 million, the 8% discount rate and growth forecast of minus 1%. The sum of the first two terms, \$3.9 million and \$27.4 million, and the third term perpetuity of \$4.3 million resulted in an unadjusted PVCF of \$35.6 million. Subtracting the \$4.3 million capital expenditure adjustment and \$2 million working capital adjustment resulted in an adjusted PVCF of \$29.4 million. As the exit value is \$29.4 million, the stressed abandonment option is therefore at-the-money. The PVCF was found to be most sensitive to declines in advertising revenue, and therefore had a significant influence on the value of the abandonment option.

Table 11.4 At-the-money abandonment option PVCF earnings projections

Year	First term	df	Second term	gr
1	2,027,778	0.9259	–	0.9900
2	1,877,572	0.8573	–	0.9801
3	–	0.7938	2,676,831	0.9703
4	–	0.7350	2,862,068	0.9606
5	–	0.6806	3,060,123	0.9510
6	–	0.6302	3,271,883	0.9415
7	–	0.5835	3,498,298	0.9321
8	–	0.5403	3,740,380	0.9227
9	–	0.5002	3,999,214	0.9135
10	–	0.4632	4,275,960	0.9044
11	–	0.4289	–	0.8953

Note: The third term perpetuity value from Eq. (11.3) is \$4,319,151

Bibliography

Berger, P., Ofck, E. and Swary, I. Investor valuation of the abandonment option, Journal of Financial Economics, 42: 257–87 1996.

Hooke, J. Security Analysis on Wall Street, Wiley, 1998.

Jones, J. What You Should Know About The Newspaper Industry, Frontline, PBS, February 2007.

Kohut, A, Doherty, C, Dimock, M, Keeter, S. Online Papers Modestly Boost Newspaper Readership, The Pew Research Center, 2006.

Rudie Harrigan K, Declining Demand, Divestiture, and Corporate Strategy, Beard, 2005.

Index

J. Rogers, *Strategy, Value and Risk*, Global Financial Markets,
https://doi.org/10.1007/978-3-030-21978-9