BUBBLES & CONTAGION

in Financial Markets

VOLUME 1: AN INTEGRATIVE VIEW

EVA R. PORRAS

Bubbles and Contagion in Financial Markets, Volume 1

Bubbles and Contagion in Financial Markets, Volume 1

An Integrative View

Eva R. Porras
Independent scholar, Spain

palgrave
macmillan

© Eva R. Porras 2016

All rights reserved. No reproduction, copy or transmission of this publication may be made without written permission.

No portion of this publication may be reproduced, copied or transmitted save with written permission or in accordance with the provisions of the Copyright, Designs and Patents Act 1988, or under the terms of any licence permitting limited copying issued by the Copyright Licensing Agency, Saffron House, 6-10 Kirby Street, London EC1N 8TS.

Any person who does any unauthorized act in relation to this publication may be liable to criminal prosecution and civil claims for damages.

The author has asserted her right to be identified as the author of this work in accordance with the Copyright, Designs and Patents Act 1988.

First published 2016 by
PALGRAVE MACMILLAN

Palgrave Macmillan in the UK is an imprint of Macmillan Publishers Limited, registered in England, company number 785998, of Houndmills, Basingstoke, Hampshire RG21 6XS.

Palgrave Macmillan in the US is a division of St Martin's Press LLC, 175 Fifth Avenue, New York, NY 10010.

Palgrave Macmillan is the global academic imprint of the above companies and has companies and representatives throughout the world.

Palgrave® and Macmillan® are registered trademarks in the United States, the United Kingdom, Europe and other countries.

ISBN 978-1-137-35875-2 ISBN 978-1-137-35876-9 (eBook)
DOI 10.1057/9781137358769

This book is printed on paper suitable for recycling and made from fully managed and sustained forest sources. Logging, pulping and manufacturing processes are expected to conform to the environmental regulations of the country of origin.

A catalogue record for this book is available from the British Library.

A catalog record for this book is available from the Library of Congress.

To Luis Javier

Contents

List of Tables and Figures

Tables

Figures

Preface

Between June and September 2015 the Shanghai Stock Exchange Composite Index lost around 40 percent of its value, and China, the world's second-largest economy after the USA, experienced a stock market crash. Earlier, while the bubble – the inflationary process – was still growing, the Chinese government had tried to take measures to moderate the bubble's progress; nonetheless, these measures failed. Later, when the bubble was about to burst, the government again attempted to control the situation by slowing the pace of its collapse. Once more, however, these desperate actions failed to achieve their objective. Fearing a slowdown of its economic growth China devalued its currency and then cut interest rates repeatedly, but to no avail. At the time of writing this book, the Chinese government is looking into the activities of individuals who expressed negative sentiments about the stock market, hence supposedly contributing to its demise – yet another effort to "set the record straight." China's continuous, and ultimately futile, struggle to contain the development and collapse of a bubble demonstrates the difficulty of dealing with these types of occurrences.

In retrospect, the bubble was in the making for years. However, it is difficult to detect clear cues as to when this process began gaining shape and speed. Given that the Chinese crash was preceded by unprecedented growth, the implications of this crisis were not immediately obvious from the start. But now it is becoming clear that this event has serious implications in terms of China's real output of goods, and the current public debate centers on whether a recession will strike this nation. Should this turn out to be the case, the world will have to brace itself for yet another global economic upheaval.

So what determines the price of a commodity, whether real estate or equity shares? And why is it that often these prices seem to develop in a manner totally unrelated to their fundamental economic parameters, defying "logical" reasoning? Well, as we have experienced time and again in the last decade, the powerful forces fueling these events are the so-called bubbles: inflationary processes that burst, sending shockwaves throughout different markets and unsettling financial stability.

Thanks to the recent subprime mortgage crisis in the USA and the ensuing worldwide economic crisis, everyone is now familiar with the occurrence of bubbles. However, what do you really know about them? Are you aware that there are many kinds of bubbles and that some can actually become contagious? Do you know there are specific conditions where bubbles form and that there are methods to detect the growth of a bubble, even at a very early stage?

Bubbles are fascinating phenomena. In my university days economists often used this term to refer to significant inflationary runs in specific assets which ended abruptly and for which they had no precise explanation. However, in 1997 during the Asian crisis this colorful, exciting, and emotionally charged word took a more definite shape for me. And "contagion" – the spread of market changes or disturbances from one regional market to others – went along with it, as the crisis which began in Thailand soon spread to other countries near and far.

At the outset of the Asian crisis, my attention was initially focused on the specific mechanisms of contagion. However, soon thereafter I dedicated an equal amount of attention to the bubble formation process. While studying these matters, I learnt a lot from books such as Malkiel's *A Random Walk Down Wall Street*[1] and Galbraith's *A Short History of Financial Euphoria*.[2] However, it was Keynes's *General Theory*[3] that I remember as being groundbreaking for me, especially the chapters dedicated to the "workings" of the capital markets as well as investor psychology and behavior.

Eight decades have passed since Keynes first wrote his masterpiece, and during this time a sequence of bubble episodes has taken place in various markets around the world, most recently in China as already mentioned. However, even though the amount of research and analysis dedicated to these subjects is flabbergasting, still no uniform economic theory exists to explain stock market bubbles, or contagion for that matter. Furthermore, the key questions posed today are the same Keynes used to introduce his study: How and why do price bubbles form and burst? And what are the necessary and sufficient conditions for these events to take place?

This two-volume work approaches these questions by providing a well-rounded synthesis of the different aspects of bubbles. In addition, this outlook is extended to contagion and the infection mechanisms that work to extend these crises beyond their initial epicenters.

These pages explore the existing main models and their conclusions: issues such as share price development in the presence of symmetric and asymmetric information in the context of rational expectations, fundamental value, and herding; key aspects related to behavioral finance; and the empirical findings pertinent to decision-making or behavioral patterns that trigger market price and volume changes.

The results of empirical economics, carried out through simulations, add valuable insights. But no less relevant is the speculative behavior of not entirely rational noise traders and chartists, and the feedback and learning mechanisms that surge within the markets and which help transmit crises. In addition to exposing the most common trading techniques followed by speculators and their impacts on the bubble formation processes, typical biases such as overconfidence, accessibility, and other psychological mechanisms and traits which influence decision-making in trading are also considered.

A rational bubble occurs when the differences between the market price of an asset and the fundamental value of that asset are justified on the bases of the rational expectations of the market players. However, in the event of speculative bubbles, the market price and the fundamental value differ to a point that no dividend income that could be realistically expected can support the current market price of an asset. Consequently, some chapters are dedicated to the issues of valuation and value growth, including related aspects of technical trading and fundamental valuation principles.

Given that the sufficient and necessary conditions for bubbles to form and contagion to occur escape a narrow exploration of financial markets, we look beyond into macroeconomics, monetary policy, risk aggregation, psychology, incentive structures, and many more subjects which are in part co-responsible for these events.

Thus, in these volumes the concepts, intuition, theory, models, mathematical and statistical background, and alternative thoughts related to bubbles and contagion in financial markets are explored. The aim is to give readers the conceptual and information background to provide them with a command of the theory and practice in all matters related to the subjects addressed within these pages. The key objective is to ensure a comprehensive understanding of the aspects that can potentially create the conditions for the formation of bubbles, the mechanisms that make a bubble burst, and the inner workings of the aftermath of such an event: the contagion of macroeconomic processes and the ensuing recession.

Within this volume, Chapter 1 summarizes the events experienced as a result of the recent housing crisis and those of other historically relevant bubbles, presenting well-defined scenarios where patterns begin to emerge. In addition, formal definitions for these processes are proposed and the "life cycle of a bubble" is examined; appropriate policy responses to the challenges presented at different stages of this cycle are explained.

Chapter 2 analyzes the key macro players in the bubble and contagion formation processes. Issues such as monetary and fiscal policy, credit and global cash flows resulting in excess liquidity, and the connectivity system and risk sharing of the modern financial world, together with systemic risk and transmission mechanisms, and feedback effects between financial sector risk and sovereign risk and the real economy, are some of the aspects developed in this chapter.

Chapter 3 investigates the idiosyncrasies of the markets and investors' psychology which are vital to the bubble and contagion formation processes. The relevance of asymmetric information between the various parties to a negotiation is highlighted. However, other mechanisms of primary importance, such as self-fulfilling expectations and reflexivity, and the role of perverse incentive structures in the reward systems of top management and traders, are also scrutinized and debated along with a number of biases in the thought processes of market players. Additional market failures as well as policies and regulation are also analyzed and thoroughly discussed.

In a “rational expectations” framework, the price of a financial asset contains a bubble when the expected rents derived from holding the asset cannot be “sensibly” expected to justify its market price. Hence, valuation techniques as well as the concept of economic value creation are useful in assessing the bubble component of prices. Chapter 4 addresses these matters and helps clarify the issue of value while establishing a framework for the variables that can be affected by the bubble.

The investment horizon of market participants differs and with it the range of tools and strategies they use to trade. These disparate approaches impact prices and contribute to the creation of bubbles and the contagion mechanisms. Chapter 4 explores the scenario from the perspective of long-term investors, whereas Chapter 5 investigates the approach taken by short-term investors and speculators, looking into technical trading and chartism in financial markets. Here the basics of technical analysis and the impact that some of these techniques and strategies, such as positive feedback trading, have on prices are exposed.

Chapter 6 is dedicated to contagion and views this phenomenon from two different angles. The first meaning refers to the transmission of crises across borders or markets and the channels through which this occurs. The second is the transmission of opinion, information, and behavior among market participants. First, the chapter looks into contagion within the context of prior financial crises, analyzing the channels of propagation. Second, it examines “social learning”, exposing how informational cascades and herding occur within this context giving rise to bubbles and accelerating their implosion. These pages also introduce various theories and models of contagion, herding, and cascades, as well as noise trading and behavioral models. Finally, some of the most relevant studies within the contagion literature are reviewed to uncover numerous meaningful details relevant to the understanding of these multifaceted and complex issues.

Chapter 7 is dedicated to exploring bubbles using frames such as rationality, information, value, and terminal life of the bubbled asset to structure their analysis. The chapter starts with an overview of rational and near-rational growing bubble models like “sunspots,” and then discusses others such as “fads” and “information bubbles.” A partial history of the classical literature on bubbles is also presented along with the findings of bubble modelling experiments and the related accounting literature. The last section of this chapter summarizes the findings with respect to the most frequently asked questions about bubbles: How are bubbles started? Why do bubbles implode? What are the consequences? Should the government intervene?

Given the breadth of subjects discussed, it is my hope that anyone interested in learning more about bubbles and contagion will find this volume enlightening, including undergraduate, postgraduate, and PhD students in business administration, as well as those specializing in economics, finance, and accounting. Students in areas as diverse as mathematics, physics, statistics, and computer engineering may also find it of value. It goes without saying that I hope to attract the interest of the financial industry itself: the practitioners, analysts, and researchers with an academic interest in investment banking, hedge funds, and risk management institutions and organizations.

Achieving a better understanding of the formation of bubbles and the impact of contagion will no doubt determine the stability of future economies. Perhaps these two volumes will help provide a rational approach to mastering these seemingly irrational phenomena.

Eva R. Porras

Acknowledgments

This book is the product of the efforts and experience of many people who in different ways have contributed toward its development.

First, I would like to thank all the professors who devoted their energy and time to my education. It was their persistence and dedication which instilled in me the love for the subject of finance and the insatiable curiosity I have for all matters related to bubble events and contagion in financial markets.

Second, I thank my parents whose motto was "the more you know, the more you will relish life." It was their guidance in intellectual matters which has endowed my existence with endless sources of inspiration, gratitude, and joy.

Finally, I would like to thank my family and closest friends for their unfathomable patience with my need for time while I was developing this work.

I also wish to give special thanks to Noelia Camilla, my assistant, and particularly to Peter Baker and Josephine Taylor of Palgrave Macmillan. They already know the many reasons for my gratefulness. Of course, any errors remain my sole responsibility.

List of Acronyms

ABS	Asset-backed security
AIG	American International Group
Alt-A	Alternative A-paper
ARM	Adjustable rate mortgages
CDO	Collateralized debt obligations
CDS	Credit default swaps
CED	Committee for Economic Development
CEO	Chief executive officer
CF	Cash flows
CFTC	US Commodity Futures Trading Commission
CLO	Collateralized loan obligations
DCF	Discounted cash flow
DJIA	Dow Jones Industrial Average
DLRE	Dynamic linear rational expectations
EAT	Earnings after tax
EBIT	Earnings before interest and taxes
EBT	Earnings before tax
ECF	Equity cash flows
EMH	Efficient market hypothesis
EMT	Efficient market theory
EPS	Earnings per share
EU	European Union
EVA	Economic value added
EWP	Elliott Wave Principle
FCF	Free cash flows
FCIC	Financial Crisis Inquiry Commission
Fed	Federal Reserve System (USA)
FSA	Financial Services Authority
FX	Foreign exchange
GDP	Gross domestic product
GE	General Electric
HAM	Heterogeneous agent model
IMF	International Monetary Fund
IPO	Initial public offering
IPT	Informational price theory
LU	Lucent Technologies
MBS	Mortgage-backed securities

Nasdaq	National Association of Securities Dealers Automated Quotations
NPV	Net present value
NYSE	New York Stock Exchange
OBV	On-balance volume
P/E	Price-to-earnings ratio
P&F	Point and figure chart
PV	Present value
RE	Rational expectations
REE	Rational expectations equilibrium
ROE	Return on equity
RSI	Relative strength index
RV	Residual value
S&P	Standard & Poor's
S&P500	Standard & Poor's 500 Index
SEC	Securities and Exchange Commission
VNM	Von Neumann–Morgenstern utility theorem

1
Introduction to Bubbles and Contagion

1.1. Current situation

The harm caused by the bursting of financial asset bubbles can have a devastating impact on investors' wealth and the welfare of society. For instance, at the end of the 1990s, the rise and fall of Internet stock prices during the dot-com bubble destroyed about $8 trillion worth of shareholders' wealth.[1] More recently, the bursting of the housing bubble created a global financial crisis that affected nations around the world, its impacts likely to be felt for generations to come with many people and communities irreparably harmed.

Given its size, it is difficult to get a comprehensive idea of the housing bubble wreck. Nonetheless, we can get partial information from reports such as "The Financial Crisis Response in Charts"[2] which highlights that $19.2 trillion in household wealth was lost between 2007 and 2009 during the financial crisis, peak-to-trough. For its part, the US Financial Crisis Inquiry Commission (FCIC) reported that more than 26 million Americans lost their jobs, and about 8.5 million either lost their homes to foreclosure, having slipped into the foreclosure process, or fell badly behind on their mortgage payments as of 2011.[3] Thus, 1 in 20 families lost their homes and livelihood in the USA, an impact usually associated with major natural disasters or war.

The data above refers to the USA alone, but the burst of the housing bubble triggered a worldwide crisis that many countries are still trying to overcome. In Spain, for example, as of the first quarter of 2013, unemployment reached six million, close to one-third of the total working population and double the average of the European Union (EU). As of August 2015, joblessness afflicts approximately over four million, representing 22.37 percent of the workforce.[4] These figures can be contrasted with those of 2006–2007, when the unemployment rate in Spain was at 8 percent.[5]

Moreover, a large number of European countries have seen their sovereign debt cost skyrocket and their economies slump. This ongoing eurozone financial crisis has hampered the ability of some of these countries to repay or refinance their government debt without the assistance of third parties, and has badly impaired their recent economic growth.

Ireland, Italy, Greece, Portugal, Slovenia, Slovakia, Spain, Cyprus and Malta have been particularly hurt. For instance, prior to the crisis, Spain had a comparatively low debt level among the advanced economies, and enjoyed a triple-A credit rating.[6] In 2010, its public debt relative to gross domestic product (GDP) was 60 percent, some 20 to 60 points less than Germany, France, Italy, Ireland, Greece, or the USA.[7] However, when the bubble burst, Spain had to spend large amounts of money on bank bailouts which, together with the economic downturn, increased the country's deficit and debt levels and led to a substantial downgrade of its credit rating.[8] By 25 July 2002, Spain had a BBB– rating and was paying 7.753 percent on its ten-year bonds, a major hike from the 3.3 to 4 percent range pre-crisis level.[9]

For a second example of the effects of this crisis we can look at Greece whose sovereign debt was downgraded by Standard & Poor's (S&P) to junk status on April 2010, after its government requested a €45 billion loan from the EU and the International Monetary Fund (IMF).[10, 11] This downgrade sent ripples across countries, as investors were set to lose some 30 to 50 percent of their stake and fears of default drove international stock markets down and caused the euro to decline.[12] Since then, austerity measures have helped Greece reduce its primary deficit from 10.6 percent of GDP in 2009 to 2.1 percent of GDP in 2013, although GDP has contracted by more than 25 percent since 2010.[13]

The social cost of these events has been horrendous. In January 2013, the seasonally adjusted unemployment rate recorded an all-time high of 27.2 percent, up from 7.5 percent in September 2008, while the youth unemployment rate reached 59.3 percent, up from 22.0 percent in the same year.[14, 15, 16] As of August 2015, unemployment is still at 26 percent, and Greece's total debt, amounting to €320 billion, represents 177 percent of the nation's GDP. Discontent prevails throughout this country and Europe in general, and Greece's exit from the euro is debated on a daily basis.[17]

Volatility in stock exchanges and bond markets, contagion among markets, and re-allocation of resources are important consequences of the bursting of financial asset bubbles. The full costs resulting from additional uncertainty, greater business risks, and social unrest driven by these events are impossible to quantify. The reason is that they range from increased financial costs which hamper economic growth for years to come (e.g. during the first quarter of 2012, Greece was paying close to 30 percent for long-term debt) to personal tragedies such as the deaths of citizens protesting against governmental actions[18, 19, 20] or the suicides of those confronted with personal losses and shame. These are just some of the pernicious effects nations had to confront as a result of the bursting of the last large financial asset bubble. Needless to say, given the gravity of these events, it is in the common interest to prevent them. However, some people question whether bubbles can be forecast and dealt with in advance, while others attempt to duck responsibilities. Differences in opinion also stem from the fact that the interests of all parties are not aligned, as some sectors of society benefit from these collapses, while others bear the losses that result.

With reference to the 2008 housing bubble collapse, the Majority Report of the FCIC in 2011 concluded the following:

> Some on Wall Street and in Washington with a stake in the status quo may be tempted to wipe from memory the events of this crisis, or to suggest that no one could have foreseen or prevented them.[21][. . .] The crisis was the result of human action and inaction, not of Mother Nature or computer models gone haywire. The captains of finance and the public stewards of our financial system ignored warnings and failed to question, understand, and manage evolving risks within a system essential to the well-being of the American public. Theirs was a big miss, not a stumble. While the business cycle cannot be repealed, a crisis of this magnitude need not have occurred. To paraphrase Shakespeare, the fault lies not in the stars, but in us.[22]

As concluded by the FCIC, it is clear that some individuals and organizations recognized the bubble process and acted to prevent its negative impact. Some of these people had information which the average investor could not access on his/her own, such as cumulative shorting contracts or the quality of the underlying loans, while others were helped by their own intuition. Nonetheless, many more were just immersed in the process, either failing to anticipate the housing market crash or, on recognizing it, tried to take advantage of the situation.

While the existence of bubbles was frequently questioned in the past, it is now undisputable that understanding developments in the techniques used for identifying asset bubbles and their consequences and, more basically, grasping the intuition behind the concept and its processes, is an important first step in preventing a recurrence of these events. This knowledge is of particular importance for researchers and policymakers as well as for those institutions responsible for monitoring the economy and others working in risk management. Sensitivity and understanding can help individuals take immediate action and develop pre-emptive policies and other measures to ameliorate the negative impacts of speculative bubbles before they grow too big and collapse.

1.2. Definitions

1.2.1. Contagion definition

When a bubble bursts, that is, when there is great discontinuity in the market-clearing price of the asset, as a consequence of excess supply, high price volatility results. Under certain circumstances, the impact of this event can be devastating because contagion among markets and assets affects both the rise of the price of the financial asset in question and, upon the bursting of the bubble, the downfall of the asset elsewhere, spreading the crisis beyond its original epicenter. Contagion spreads because the global economic system operates in a series of interdependencies which facilitates the transfer of risks, unless firewalls are put in place.

In this context, financial contagion refers to the phenomenon that occurs when one asset or basket of assets is affected by changes in prices in other markets of this asset or basket of assets. For instance, during the housing bubble, US policymakers were afraid that the sudden and disorderly failure of large firms would trigger balance-sheet losses in counterparties. The direct financial link between firms puts at risk the wellbeing of a *second* company when a *first* is threatened. This is contagion as related to the condition of "too big to fail"; if financial firm *X* is a large counterparty to other firms, *X*'s sudden bankruptcy might weaken the finances of the others and cause them to fail as well. A financial firm *X* is too big to fail when policymakers fear contagion cannot be assumed by the market. This judgment is based on how much counterparty risk other firms have to the failing firm, and on the likelihood and possible damage of contagion. If a firm is considered too big to fail, authorities will decide how to "bail it out." The fear of contagion through the "too-big-to-fail" mechanism explains some actions taken by US policymakers during 2008. Two examples are when the US Federal Reserve (the Fed) facilitated JPMorgan's purchase of Bear Stearns by providing a bridge loan and loss protection on a pool of Bear's assets, and when, with support from the Treasury, it "bailed out" American International Group (AIG).

A second way in which contagion occurs is through a common factor that affects a number of firms in the same manner. The common factor in this crisis was concentrated losses on housing-related assets in large and mid-size financial firms in the USA and Europe. Unconnected financial firms were failing at the same time for the same reason. Since they had made similar failed bets on housing, they shared the problem of large housing losses. Policymakers were not just dealing with a single insolvent firm that might transmit its failure to others, they were dealing with a scenario in which many large, mid-size, and small financial institutions took large losses at the same time. These losses wiped out capital throughout the financial sector. In a common shock, the failure of one firm may inform us about the breadth or depth of the problem but does not cause the failure of another.

Usually, the term "contagion" takes on multiple meanings. It is therefore useful to clarify that in our context, contagion is an episode which has significant *immediate* impacts. This is in contrast to instances where these effects are gradual, regardless of whether they may, cumulatively, have major economic consequences. We refer to the latter cases as "spillovers."

1.2.2. Bubble definition

The expression "bubble" was coined in the 1720s in reference to the events concerning the South Sea Company. Economic bubbles have existed since the birth of currency. There is a long recorded history of financial bubbles, starting with the Tulip Mania, the first and probably most famous of all bubble events. However, even though much has been written about "bubbles" since then, there is no exact definition of the word in this context. In general, though, it is used to refer to asset prices that are not justified by the assets' "fundamentals" or intrinsic value.

The value of a company rests on its capacity to create wealth over time. In a corporation, each of the company's owners will share in the profits in a manner directly proportional to their investment in the company. That is, when you buy a share of stock, you buy a piece of the company and your share of its expected "growth" is your return. Thus, the determinants of the fundamental value of a company are those factors that ensure a sustainable growth and the sharing of it among the various corporations' owners. Overall, these determinants can be captured by earnings and dividends growth, the dividends-to-net-earnings ratio, the risk of the cash flows (CF) generated by the firm, and the cost of financial capital.

An asset bubble occurs when a financial asset is traded in the market at a price higher than the level its economic fundamentals can sustain, such as when the price of the share grows in the exchange markets for a sustained period of time at a rate much greater than its earnings.

To illustrate the idea, we can think of Tirole's[23] model in which the value of the fundamentals of the asset in the market is the discounted present value (PV) of its future payoffs, proxied by expected dividend payments. Tirole's proposal was that if the asset's price in the markets is above what can be justified by its fundamentals, then there is a bubble. In general if,

$$x_t = F_t + B_t \tag{1.1}$$

where

x_t is the price of the asset today

F_t is the part of the price that corresponds to the fundamentals

B_t is the part of the price that corresponds to the bubble (what we cannot justify according to the firm's fundamentals)

When $x_t = F_t$ there is no bubble component in the price of the asset.

The problem is that determining F is not a simple matter. The main difficulties in financial asset valuation lie in forming correct expectations about the future, as specific prices cannot be estimated with uncertain data, and also in ascertaining whether it is the proposed model or the specific values assigned to the various variables that contain errors. Thus, the basic complexity involved in testing for the existence of rational bubbles is that the contribution of hypothetical rational bubbles to the asset price would not be directly distinguishable from the contribution to market fundamentals of variables the researcher cannot observe.[24, 25]

As stated, in the context of this book, the term "bubble" refers to the mispricing of financial assets. However, not every temporary mispricing should concern us. Rather, the bubbles of interest to us have a negative impact on the economy after a long period of sustained significant mispricing and higher-than-average volatility in financial markets. Ultimately, bubbles are important because they drain resources from the system and the resulting prices affect the real allocation of resources in the economy. For example, the presence of bubbles may distort agents' investment incentives, and the bursting of bubbles may affect the balance sheets of firms, financial institutions, and households, reducing the overall economic activity of the country. It is because of these serious repercussions that it is important to understand the circumstances under which bubbles can arise and why market asset prices can deviate systematically from the assets' fundamental value.

Below, we summarize the stories of some of the most legendary bubbles.

1.3. Brief history and analysis of some bubbles

1.3.1. Tulip Mania

Toward the middle of the sixteenth century, the Ottoman Empire began to export tulips. These flowers differed greatly from others known at the time, particularly due to their bright, vibrant colors, and the fact that they proved quite resilient to adverse weather. They also had a distinct, and quite extraordinary, trait: they could be afflicted by the mosaic virus which determining results in spectacular, intricate lines and multicolored effects. This bizarre quality made the flowers particularly interesting and tulips became fashionable as well as a desirable luxury item. Rare bulbs that gave rise to a profusion of new varieties with remarkable patterns and colors were introduced to the market every year. Consequently, demand for tulips began to grow exponentially, and bulb wholesalers began to fill their inventories.

Even though tulips became wildly popular in many countries, it was in the Netherlands that the passion for these flowers reached its height as rich merchants who traded with the East Indies chose to exhibit their wealth by designing sprawling flower gardens. In time though, what had once been reserved for the elite trickled down and, by 1634, owning and trading tulips involved all ranks of society. And as the flowers grew in popularity, professional growers paid higher and higher prices for the bulbs with the virus.

The spot market where tulips were traded took place between June and September, the plant's dormant phase when the bulbs could be uprooted. For the rest of the year, tulip traders signed notarized contracts to purchase bulbs at the end of the season. This was called the wind trade market because no bulbs were physically exchanged.

By 1636, tulips were the fourth leading export from the Netherlands and bulb trading was done on the exchanges of numerous Dutch cities. Various accounts help us assess the circumstances. For instance, according to Mackay[26] people were selling their possessions to speculate in the tulip market. As an example, he mentions an offer of 12 acres (49,000 m^2) of land for one of the two existing Semper Augustus bulbs. He also lists the basket of goods used to purchase a single bulb of Viceroy (see Table 1.1).

At the beginning of 1637, 70 tulips were auctioned for 53,000 guilders. It is difficult to get a precise idea of what this money means in modern terms, but it helps to keep in mind that in those days, the annual salary of a skilled craftsman was around 150 guilders. A few days after the auction, for reasons that are unclear, buyers did not show up at a regular bulb fair that took place in Haarlem, an important city in the north of Holland. This event unveiled the underlying nervousness already spreading through the market, and the fear that the interest in bulbs had passed its prime spread quickly after someone refused to pay for the tulips he had bought at a later auction.

By February 1637, when tulip traders could no longer find buyers willing to pay increasingly inflated prices for their bulbs, the demand for tulips collapsed, and prices plummeted. Some were left holding contracts to purchase tulips at

Table 1.1 Viceroy bulb trade (1637)

Two lasts[a] of wheat	448*f*	Four tuns[c] of beer	32*f*
Four lasts of rye	558*f*	Two tons of butter	192*f*
Four fat oxen	480*f*	1,000 lb of cheese	120*f*
Eight fat swine	240*f*	A complete bed	100*f*
Twelve fat sheep	120*f*	A suit of clothes	80*f*
Two hogsheads of wine[b]	70*f*	A silver drinking cup	60*f*
Total			2500*f*

Notes: Depending on the source of the information and the size of the bulb, its worth varied between 2,500 and 4,200 guilders (florins). According to Chapter 3 of Mackay[85] (1841) and Schama[86] (1987) this basket of goods was actually exchanged for a bulb, but Krelage[87] (1942) and Garber[88] (2000, pp. 81–83) dispute this interpretation of the original source, an anonymous pamphlet, stating that the commodity bundle was clearly given only to demonstrate the value of the florin at the time.

[a] A "last" in the Dutch East India Company in the seventeenth century was approximately 1,250 kg, becoming later as much as 2,000 kg, http://www.historici.nl/Onderzoek/Projecten/VocGlossarium/vocoutp

[b] A "wine hogshead" was equal to approximately 232 to 240 liters.

[c] The "tun" is an English unit of liquid volume equivalent to 252 wine gallons, approximately 2,240 pounds or close to 1,000 liters, http://en.wikipedia.org/wiki/English_wine_cask_units

prices that were ten times greater than those on the open market. Others found themselves in possession of bulbs that were worth a fraction of the price they had paid just a short time ago.

Some contemporary researchers suggest Mackay's account is not particularly accurate. For instance, Goldgar[27] claims that the trading was an urban phenomenon limited to a fairly small group of wealthy merchants and skilled craftsmen, while Peter Garber[28] argues that the bubble "was no more than a meaningless winter drinking game, played by a plague-ridden population that made use of the vibrant tulip market." Either way, the story illustrates quite well the essence of a bubble run and burst, and many of its typical connotations, including the psychological ones.

1.3.2. South Sea Bubble (1719–1720)

The term "bubble" originates from the South Sea Corporation's inflated stock prices.[29, 30] The *causa remota* of this bubble was the British government's need to manage £10 million of debt it had acquired during a war with Spain. To resolve this situation, the government set up a deal by which a company would take up the nation's debt in return for a 6 percent interest to be paid over a certain period. The company, which was granted exclusive rights to trade with the Spanish colonies in South America (the South Seas) and to supply the colonies with slaves for 30 years, took the name South Sea after becoming incorporated by an Act of Parliament.

The deal aroused much public expectation. The popular "self-serving" thought was that South America would be eager to trade gold and jewels in exchange for British wool and clothing. This perception was based on the rumor, spread by the company's managers, that Spain would provide a permit for free trade with ports in Peru, Chile, and Mexico. Thus, to finance the project, the South Sea Company issued stocks which immediately became sought after by wealthy speculators and later on by anyone who could afford them.

Several reasons explain this interest. For one, the South Sea Company was one of the largest and better-publicized business ventures England had ever embarked on. Also, at that time very few companies offered shares and those which did presented scant returns. In comparison, the South Sea Company was perceived as the most potentially lucrative monopoly on earth. Thus, with many impressed by the company's imminent returns and the project's image, it became fashionable both in Great Britain and the Continent to own shares in the South Sea Company.

Nonetheless, by 1718 when the first vessel set sail, a new war with Spain disrupted the possibility of any effective maritime traffic, although this fact did not deter people from building dreams of riches. When it eventually became clear that the company could not generate income from its stated operations, the management decided to put their efforts into alternative speculative ventures and the South Sea Company managed to stay afloat as a financial institution.

Responding to a request from the king that national debt should be decreased, the South Sea Company proposed their capital stock be enlarged by £2 million in exchange for a reduced interest rate of 5 percent. After this proposal was accepted in the South Sea Act, the company's directors began to concoct new ways of extending their influence. This fact was reflected in a new scheme presented to Parliament: the South Sea Company would take on the whole debt of the state, some £30,981,712 in exchange for a 5 percent interest to be received until June 1724, at which time the debt would become redeemable and the interest reduced to 4 percent.

In 1720, toward the end of January, deliberations in the House of Commons had brought the market's hysteria to a height and the company's stock value was raised from £130 to £400. Rumors spread continuously: peace and merchant treaties between Spain and England were being negotiated; Spain was granting free access to all of its colonies; the more productive of the Spanish mines were being sold to Britain; commerce was going to be reactivated immediately; piles of South American gold and silver were waiting to be exported to England to be exchanged for wool; and so on. Accordingly, unrealistic expectations were the norm among the South Sea's investors and speculators, and the management team was helping increase the company's lure by creating illusions of grandeur in investors' minds.

When the Bill was passed in both Houses, the speculating frenzy was so high that Exchange Alley was blocked by crowds on a daily basis. As the time was then propitious, other schemes named bubbles were started with countless joint-stock companies popping up all over. Some of these schemes were plausible enough, but they were established with the purpose of raising shares in the market and, in many instances, selling them once the price had increased.

The situation was so alarming that on June 11, the king published a proclamation declaring these projects unlawful. According to Mackay's account (1814–1889[31]) the following were included in the list of charges against the companies:

a. For supplying London with sea-coal.
b. For extracting silver from lead.
c. For carrying on an undertaking of great advantage, but nobody to know what it is.
d. For insuring of horses.
e. For a wheel for perpetual motion.
f. For insuring to all masters and mistresses the losses they may sustain by servants.

The shares of the South Sea Company were now valued at £890. However, by that time the general opinion was that the price could no longer increase. For instance, part of the nobility traveling to Hanover with the king became anxious to sell before their departure. On June 3 supply was totally outweighed by those selling and the price fell to £640. This was quickly met with company directors' orders to buy. Their agents helped the price climb again to £750, and by the beginning of August it reached £1,000: the bubble's peak.

As in many other bubble events, corruption was prevalent among insiders, politicians, and regulators. When this news became widely known it began to unsettle the market. First, there were allegations against the directors who stood accused of partiality in making out the lists for shares in each subscription. Then it became known that the chairman and others within the company had sold out their entire stake in a clandestine manner. As these secrets leaked, the public panicked and investors began to sell immediately, losing fortunes in the process. Public excitement was so high and petitions from all parts of the kingdom, crying for justice, so numerous that George I had to return to England to attend to the alarming state of affairs.

After an audit proved that large quantities of stock had been transferred to the chancellor of the exchequer, the treasurer of the company made an escape to Calais, from where he was later retrieved. Numerous members of the House and South Sea officers were summoned to answer for their corrupt practices. When it all came tumbling down, many of the executives were arrested; thousands of investors, including hundreds of members of government, saw their fortunes evaporate; and suicides became a regular occurrence. At this point, the Bank of England decided to step in as a "lender of last resort," helping stabilize the banking industry.

Despite the government's efforts, Britain's economy was in a shambles after the South Sea bubble and did not fully recover until a century later. In addition to the long-term crisis, another consequence was the Bubble Act of 1720, by which the British government restricted the ability to create new business ventures and limited joint-stock companies (synonymous with incorporation) to prevent future bubbles.

1.3.3. Railway Mania[32]

The world's first modern intercity railway for transporting both passengers and freight opened between Liverpool and Manchester in 1830. During that decade, the British economy was suffering and interest rates were raised to channel funds into government bonds. However, once the situation began to improve by the mid-1840s, the Bank of England decided to cut interest rates. This decision shifted the attention of the investors to new financial instruments which, comparatively, offered a better return. With government bonds selling at reduced rates and railway companies booming as a consequence of new industrial demand, people began to consider investing in these ventures.

The Industrial Revolution had created a new, wealthy middle class who, together with the banks and nobility, were ready to invest their savings in British businesses. Improvements in various sectors were also easing the transfer of information and funds. For instance, by this time, newspapers were strong enough to help companies market themselves, and the modern stock market organization facilitated public investment. Furthermore, by 1825 the government had already repealed the Bubble Act, which meant anyone with means could invest in a new corporation.

Railway companies took advantage of these new technologies and promoted themselves as rock-solid ventures. The usual offer was that shares could be purchased with a 10 percent deposit and the company could call in the remainder as needed. The deal offered was de facto a credit to prospective investors, therefore enlarging the company's lenders' pool. This fact, together with a solid marketing campaign promoting the railways as sound investments, resulted in thousands of investors, including citizens with very limited savings, purchasing large quantities of shares.

The British railways regulation of those times could be described along laissez-faire lines. To obtain the right to purchase land, companies had to present a Bill to Parliament that outlined the proposed railway route. However, no genuine analysis of financial viability was ever performed. This was a market open to any company who wished to put itself forward. Thus, around 272 bills were passed, amounting to 9,500 miles of railway. Not surprisingly, one reason for such Parliamentary largesse was the fact that many Members of Parliament were also railway investors.

Over time, as numerous companies began to operate, it became obvious that many of the routes were not commercially viable: contrary to earlier suppositions, railways were not simple projects with certain growth and return patterns. By 1845, the idea that the Railway Mania was the result of a self-promoting scheme supported by over-optimistic speculation began to take root. And toward the end of that year, as the Bank of England began to increase rates, money began to flow from the railways into bonds.

The peak of the bubble occurred at the beginning of 1846. Thereafter, the increase in prices slowed down and finally leveled out before beginning to fall. As soon as that happened, all investment was halted, leaving many companies with insufficient funds, and investors with no prospects of getting their money back.

When the speculation collapsed, a representative portion of the middle class was ruined, and by the early 1950s, all but a few of the largest railway companies were gone. Around a third of the railways that had been authorized were never built.

1.3.4. Dot-com Bubble[33]

The Railway Mania can be compared to the fever which affected the stock of telecom companies in the 1990s. The first part of the telecom Mania culminated in the installation and exploitation of a huge fiber-optic telecommunications infrastructure. This resulted from the recognition that the same railway "rights-of-way" could make inexpensive conduits for fiber optics. A second boom occurred during the expansion of the Internet, when countless companies were created to provide services on the growing network. The final burst took place in 2000 when, after an all-time high of 5,133 on March 10, the National Association of Securities Dealers Automated Quotations (NASDAQ) Composite Index lost over 37 percent of its value, falling to a low of 3,227 by April 17. All in all, from the peak to the bottom, the loss amounted to 78 percent, a blast whose tremors lasted until October 2002.[34, 35]

The Internet was created by the US military long before there were any thoughts of its viability as a commercial network. However, by 1995 it already had 18 million users, and its potential began to surface, offering insights on the latent international market. A range of new buzzwords started to conjure up an exciting world of possibilities and a new euphoria swept the markets. The transition to a service-based economy around technological developments was baptized "the New Economy."

As the renewed energy and enthusiasm brought by the "New Economy" spread through the markets, a string of Internet-related initial public offerings (IPOs) began to surface.[36] In 1999 alone there were 457 IPOs mainly related to Internet and technology ventures, and this was just in the USA. Such was the excitement that about 25 percent of them doubled their price on the first day of trading.

Investors wanted "big ideas" but, unfortunately, it was soon to be discovered that many of these start-ups were not backed by solid business plans. The first signs of trouble became visible when numerous companies reported huge losses and folded just months after opening. In 2001, there were only 76 IPOs, none of which doubled its price on the first trading day.

The NASDAQ Composite Index was made up mainly by companies associated with the New Economy: computer software and hardware, Internet services, telecommunication, and so on. A key characteristic of these companies was that their price-earning-ratios (P/Es) were much higher than those of companies representing the "Old Economy." This had to do with the standard fundamental valuation formula in which the value of a share is a function of the CF *expected* to be generated by a company. New Economy companies compensated for their lack of current earnings by promising enormous growth potential. Thus, the bull market that drove the NASDAQ until the first quarter of 2000 was mainly the result of the expected growth of potential earnings. Just to give an example, the price-to-dividend ratio of Lucent Technologies (LU) prior to its crash on January 5, 2000 was over 900. At the same time, an Old Economy company such as DaimlerChrysler produced a return more than 30 times higher. Nevertheless, Lucent Technologies' shares rose by more than 40 percent during 1999, whereas those of DaimlerChrysler declined by more than 40 percent over the same period.[37]

The dot-com is one of the better-known bubbles in living memory and also one of the strangest. Many argue that it was a case of too much too fast. Companies that could not decide on their corporate vision were given millions of dollars and told to grow to Microsoft size by the next day. The novelty of the industry and the embedded difficulty of valuing these companies at the start led to some rather over-enthusiastic investments. But the growth in the technological sector proved deceptive. Poor business practices led to high-profile court cases, and the stock market began to tumble, along with hundreds of dot-coms.

There were many contributing factors to the dot-com bust, but overall, the key reason was the high growth expectations that never materialized. The long-term potential of the sector overshadowed the short-term viability of specific companies. The bubble finally burst on March 10, 2000, resulting in a mild but long-felt recession (Figure 1.1).

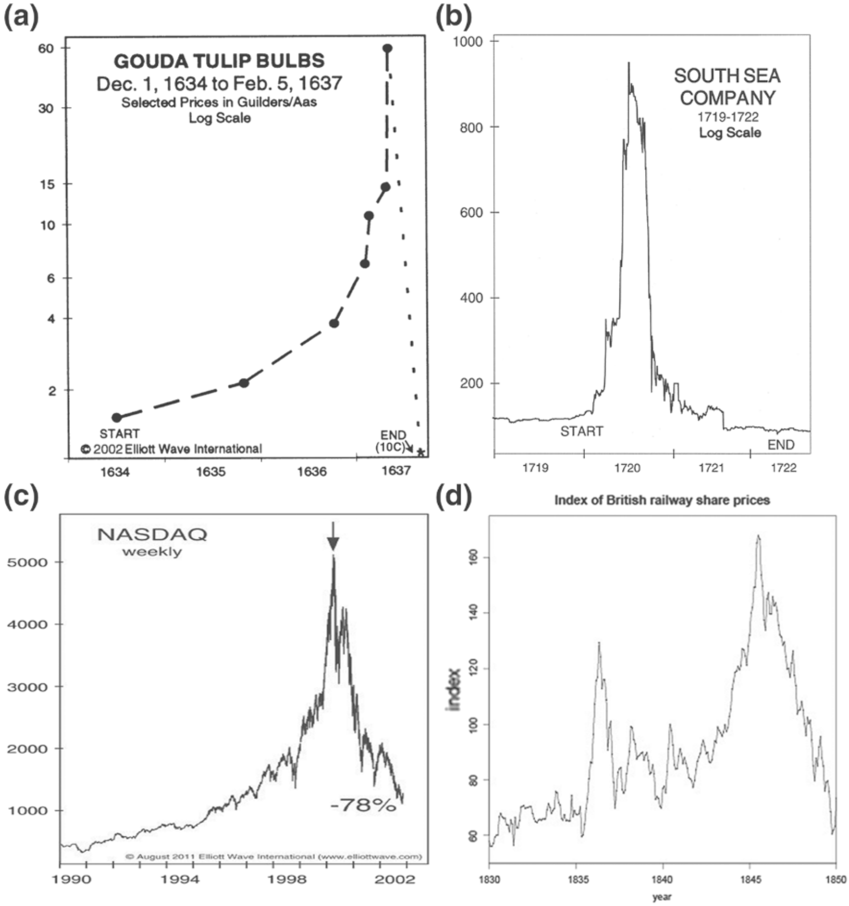

Figure 1.1 (a) Gouda Tulip Bulbs 1634–1637, (b) South Sea Company 1719–1722, (c) Nasdaq Composite 1990–2002, (d) Index of British Railway share prices 1830–1850

Source: Graphs (a)–(c): *Conquer the Crash* by Robert R. Prechter. Graph (d): Andrew Odlyzko, Collective hallucinations and inefficient markets: The British Railway Mania of the 1840s. Preliminary version, 15 January 2010, http://ssrn.com/abstract=1537338

1.3.5. The housing bubble

As of December 2010, the USA was still suffering the consequences of the crisis resulting from the collapse of the housing bubble. The primary features of this crisis were a financial shock in September 2008 and a related financial panic, which triggered a severe contraction in lending during the fourth quarter of 2008. The background to this crisis is as follows.

At the start of 2001, the US unemployment rate of 4 percent marked a 30-year low. Although the recession subsequent to the dot-com bust was almost over, the economy was slowing down. Thus, to stimulate borrowing and spending, the Fed decided to lower short-term interest rates. On January 3, 2001, the benchmark rate at which banks lend to each other overnight was cut by a half percentage point. Later that month, the rate was cut by another half point. This was done 11 more times throughout the year, all the way to 1.75 percent. With the recession over and mortgage rates at 40-year lows, the housing sector began to soar. In 2003 alone, builders started more than 1.8 million single-family dwellings, and between 2002 and 2005, residential construction contributed three times more to the economy than at any time since the 1990s.[38]

However, even with all this activity, employment in other sectors remained undersized. Thus, the Fed persevered in using monetary policy as worries of a "jobless recovery" (an increase in production with no marginal gain in employment) began to surface. During 2003, short-term interest rates were kept so low that large US firms could access short-term funds in the 90-day commercial paper market at an average of 1.1 percent, from 6.3 percent just three years earlier. Also, three-month Treasury bill rates had dropped under 1 percent from 6 percent in 2000.[39]

Given that interest rates for the typical 30-year fixed-rate mortgage are historically related to the overnight federal funds' rate, low rates cut the cost of home ownership. For instance, by 2003, creditworthy home buyers could get fixed-rate mortgages for 5.2 percent, three percentage points lower than three years earlier. That is, for the same monthly payment of $1,077, a homeowner could move up from a $180,000 home to a $245,000 one.[40]

As people jumped into the housing market, prices increased at an overall annual rate of 9.8 percent between 2000 and 2003. But that rate was much higher in the fastest-growing markets such as Florida and California. Household wealth rose to nearly six times income, up from five times a few years earlier, and many families benefited as home ownership peaked at 69.2 percent of households in 2004.[41]

Even though house prices rose, declining affordability that would have normally constrained demand was overridden by the use of nontraditional mortgage products. These included interest-only adjustable rate mortgages (ARMs), pay-option ARMs that gave borrowers flexibility on the size of early monthly payments, and negative amortization products in which the initial payment did not even cover interest costs. These products were created in order to encourage buyers to purchase above their means, as they often resulted in significant reductions of initial monthly payments compared to traditional mortgages. Not surprisingly, many who entered into these contracts could not maintain payments once all the charges were made effective to them.

Fed Chairman Bernanke summed up the situation:

> At some point, both lenders and borrowers became convinced that house prices would only go up. Borrowers chose, and were extended, mortgages that they could not be expected to service in the longer term. They were provided these loans on the expectation that accumulating home equity would soon allow refinancing into more sustainable mortgages. For a time, rising house prices became a self-fulfilling prophecy, but ultimately, further appreciation could not be sustained and house prices collapsed (p. 16).[42]

This explanation posits a relationship between the surge in housing prices and the surge in mortgage lending which appears to have been mutually reinforcing.

As house prices increased, consumers' spending outpaced incomes, resulting in a reduction in the personal savings' rate from 5.2 percent to 1.4 percent between 1998 and 2005. Furthermore, higher home prices coupled with low mortgage rates also resulted in massive refinancing within the prime mortgage market.[43] Between 2001 and 2003, cash-out refinancing amounted to $427 billion which, in addition to another $430 billion via home equity loans, was used by homeowners to cover medical bills, taxes, electronics, and vacations or to consolidate debt, make home improvements, and for other equivalent uses.[44]

However, by early 2007, it was already clear that home prices were falling, mortgage originators faltering, and a growing number of families could no longer afford their mortgage payments. As 2007 went by, an increasing volume of delinquencies and defaults forced rating agencies to downgrade mortgage-backed securities (MBSs)[45] and collateralized debt obligations (CDOs).[46] Inevitably, startled investors sent prices plummeting, while hedge funds facing margin calls began to sell at distressed prices, and banks began to write down the value of their holdings by tens of billions of dollars.

In addition, several securitization markets were brought to a halt after June. For example, $75 billion in subprime securitizations were issued in the second quarter of 2007. That figure dropped to $27 billion in the following quarter, and to $12 billion in the fourth. Alternative A-paper (Alt-A) issuance[47] earmarked $100 billion in the second quarter, but fell to $13 billion during the last. By the end of 2007, these previously booming markets were almost gone, with only $4 billion in subprime or Alt-A MBS issued in the first half of 2008, and almost none thereafter.[48]

Other structured products followed, with a reduction in CDOs from $90 billion during the first quarter to barely $5 billion in the fourth, and from over

$80 billion of collateralized loan obligations (CLOs) in 2007 to $10 billion during 2008. Also, the issuance of commercial real estate MBSs plummeted from $232 billion in 2007 to $12 billion in 2008.[49] Furthermore, those securitization markets that held up during 2007 eventually suffered the same fate in 2008 as the crisis evolved. And securitization of auto loans, credit cards, small business loans, and equipment leases all nearly ceased in the third and fourth quarters of 2008.

The collapse of securitization markets for these other kinds of debt further restricted access to financing for credit cards, car loans, student loans, and small business loans. This reduction of credit, together with the implosion of the housing bubble and wealth losses resulting from the declining stock market, led to a blunt contraction in consumption and an increase in unemployment. The inability to access funding, financial firm deleveraging, and macroeconomic weakness resulted in tighter credit for both consumers and businesses which caused a type of retro-feedback mechanism: with no credit or customers, companies had to trim costs and lay off employees, exacerbating the situation.

During the ensuing months, financial intermediation spread the crisis to other sectors of the economy and other countries. Some funding markets completely collapsed, while in others the crisis spread until the US government intervention began to stabilize them. For example, within the interbank lending market, even large banks were unable to get overnight loans, which added to the problem of funds unavailability elsewhere. During the last semester of 2008, the repo market presented significantly increased rates and a ballooning of financial haircuts. After the Lehman Brothers bankruptcy,[50] money market mutual funds withdrew from investing in the commercial paper market, further contributing to the hike in funding costs for all financial and nonfinancial firms.

Seventeen trillion dollars in household wealth had vanished in less than two years.[51] As housing prices kept declining, even more families were presented with the dilemma of how to deal with mortgages that exceeded property values. A most obvious choice was to cut spending and the immediate side effect of cumulatively taking this action was to put a halt on economic expansion, shedding further jobs. As the unemployment rate grew, the number of families who could not afford their mortgages and were now stuck with their houses grew in tandem. Thus, as time progressed, millions entered foreclosure and millions more fell behind on their mortgage payments, while others returned the keys to the banks. Ultimately, the increase in foreclosed and abandoned properties pushed prices further down, depressing the value of neighborhoods across the country, and affecting local budgets that relied on property taxes.

The recession began in December 2007 and its effects, as reflected by the speed and breadth of the rise in the unemployed, were the worst on record. In the USA alone, the economy shed 3.6 million jobs in 2008 and another 4.7 million jobs by December 2009.[52]

1.4. Causes of bubbles and contagion

In the classic study *Manias, Panics and Crashes: A History of Financial Crises*,[53] the authors make the following distinction:

> Causa remota of any crisis is the expansion of credit and speculation, while causa proxima is some incident its that saps the confidence of the system and induces investors to sell commodities, stocks, real estate, bills of exchange, or promissory notes and increase their money holdings.

In 2011 Report,[54] the FCIC provides, through evidence and insights, its own list of direct and indirect factors to explain the 2008 crisis and how structural situations, decision making, and exogenous events came together to deliver events as we have experienced them:

> While the vulnerabilities that created the potential for crisis were years in the making, it was the collapse of the housing bubble – fueled by low interest rates, easy and available credit, scant regulation, and toxic mortgages – that was the spark that ignited a string of events [. . .] When the bubble burst, hundreds of billions of dollars in losses in mortgages and mortgage-related securities shook markets as well as financial institutions that had significant exposures to those mortgages and had borrowed heavily against them. This happened not just in the United States but around the world. The losses were magnified by derivatives such as synthetic securities. The crisis reached seismic proportions in September 2008 with the failure of Lehman Brothers and the impending collapse of the insurance giant American International Group (AIG). Panic fanned by a lack of transparency of the balance sheets of major financial institutions, coupled with a tangle of interconnections among institutions perceived to be "too big to fail," caused the credit markets to seize up. Trading ground to a halt. The stock market plummeted. The economy plunged into a deep recession. (p. xvi)

The conclusions drawn by the Majority Report can be summarized as follows:[55]

- The crisis was avoidable.
- The widespread failures in financial regulation and supervision were devastating to the financial markets.
- The major malfunction of corporate governance and risk management at financial institutions were a key cause.
- Excessive borrowing, risky investments, and the lack of transparency shot down the financial system.
- There was a systemic breakdown in accountability and ethics.

- Collapsing mortgage-lending standards and the securitization pipeline lit the flame and spread the crisis.
- Over-the-counter derivatives contributed significantly to the crisis.
- The failures of credit rating agencies were essential cogs in the wheel of financial destruction.

The Minority Dissenting Statements further elaborated:

> Even absent market fundamentals driving up prices, shared expectations of future price increases can generate booms.[56]

The above accounts describe the outcome of defective leadership which resulted in excess liquidity due to governmental and private actions; lack of supervisory effectiveness by governmental agencies of private corporations and by agencies into both government and private activity; unethical attitudes and actions which affected governance, regulatory, and supervisory activities; and a general lack of transparency either because the necessary information was not made public or because the information provided was tainted.

Psychological and information-related manipulation played a big role in the process, giving rise to the illusion of control from the perspectives of both the regulators and the regulated. Even well-informed key market analysts and communicators such as Alan Greenspan, chairman of the US Federal Reserve for 20 years, alluded to the difficulties implicit in seeing through the various layers that veiled the appearance of a bubble and the complexity of asset markets. In his words:

> How do we know when irrational exuberance has unduly escalated asset values? [. . .]. But we should not underestimate or become complacent about the complexity of the interactions of asset markets and the economy.[57]

And then, years later, he told the FCIC:

> History tells us [regulators] cannot identify the timing of a crisis, or anticipate exactly where it will be located or how large the losses and spillovers will be.[58]

Notwithstanding the conclusions reached by those analyzing the 2008 crisis and the reasons for the start and spread of the housing bubble, more generally and over time a number of hypotheses have been proposed to explain the existence of bubbles and contagion under conditions which assume both rational and irrational behavior for at least one group of market agents, and models with misaligned incentives and nonstandard preferences. Overall, these theories encompass specific economic conditions and market imperfections, specific behaviors, and interconnectivities between the micro and macro sectors across assets and boundaries. We briefly introduce some of these here and elaborate on them more formally over the subsequent chapters.

1.4.1. Conditions

Certain conditions reduce the set of possible bubbles. For instance, it has been argued that bubbles cannot grow in assets with upper-bounded prices, such as those having close substitutes. The reason is that consumers will replace the expensive asset with the substitute once the former becomes too pricey. It has also been proposed that a bubble in an asset cannot exist if the asset's required rate of return is higher than the growth rate of the economy. The explanation for this is that the bubble in this asset would outgrow the aggregate wealth of the economy.[59]

In scenarios where agents are perfectly rational and all information is available, a third suggestion is that a bubble cannot exist in the price of a finitely lived asset. Given the premises of full information and finite life, the bubble will burst at the end of the asset's life (T) with 100 percent probability when it is liquidated at its fundamental value. As both its fair value and deadline are public knowledge, no one would be willing to buy the asset at a price above fundamentals one minute before T. The same reason extends to a period of two minutes before T, and so on, all the way up to the present.

On the other hand, if we have a finitely lived asset but private information exists and short sale constraints are binding, then an asset price bubble can exist.[60] In this scenario, investors know the asset is overvalued, but ignore it. Thus, agents would be willing to hold an overvalued asset in the hope of reselling it to an uninformed trader at an even higher price in the future.[61]

There is a special case of rational bubbles in which the bubble is a function of the asset's fundamentals instead of time.[62] When agents are over-optimistic and overvalue the CF derived from holding an asset, "intrinsic" bubbles arise. Their name derives from the fact that they are deterministic functions of the assets' fundamentals, and the model relates to empirical observations of bubbles, explaining why stock prices are more volatile than dividends.

1.4.2. Bounded rationality

To explain bubbles and contagion, there are behavioral models assuming the existence of some rational traders. These models can be grouped into four categories:[63]

a. *Differences of opinion and short sale constraints*. Within this group, optimistic investors disregard the fact that pessimistic-short-sale constrained investors imprint their views into prices. Given the differences in opinion and action regarding future outcomes resulting from personality traits and sale constraints, the market price of the asset will have a bubble component.

b. *Feedback trading*. Feedback trading behavior results in a trading strategy based on recent price movements. When an asset's price increases, feedback traders push the price even further by purchasing it. This attracts additional attention from other feedback traders, who, through their dealings, keep pushing prices upward to a point where they exceed fundamentals.
c. *Biased self-attribution*. Biased self-attribution refers to a cognitive bias by which people recognize signals that confirm their prior beliefs while disregarding others that contradict their earlier formed opinions.[64] A plausible scenario which can result in bubbles is one in which agents observe a noisy private signal and form initial valuations. Later on, a noisy public signal is revealed. Given that this second signal confirms the investor's earlier research, she grows overconfident and further revises her valuation in the direction of the private signal. However, when public signals contradict her initial private signal she chooses to ignore it and prices are not adjusted.[65]
d. *Representativeness heuristic and conservatism bias*. These cognitive biases represent departures from optimal Bayesian information processing. Rational behavior depends on the ability to process information effectively, despite ambiguity or uncertainty. An advantage of Bayesian models is that they are probabilistic, and probability theory provides an *optimal* calculus for representing and manipulating uncertain information, that is, it allows characterizing information and the uncertainty in that information. Consequently, an advantage of Bayesian modeling is that it gives cognitive scientists a tool for defining rationality. Using Bayes's rule, Bayesian models optimally combine information based on prior beliefs with information based on observations or data. Bayesian models use these combinations to choose actions that maximize the task performance. Of course, the performance of a model depends on how it represents prior beliefs, observations, and task goals. But for any specific probabilistic formalization of a task, a Bayesian model specifies optimal performance given the set of assumptions made by the model. Bayesian inference derives the posterior probability as a consequence of two antecedents, a prior probability and a "likelihood function" derived from a probability model for the data to be observed.

Problems arise when people depart from Bayesian rationality. For instance, under representativeness, heuristic investors overreact to salient news by putting too much weight (high probability) on such signals relative to their base probabilities; whereas the opposite happens under conservatism bias when investors underreact to less attention-grabbing signals that assign low probability weights. Both biases can lead to the formation of price bubbles.

An example presented by Barberis, Shleifer, and Vishny (1998)[66] proposes a scenario in which agents imagine an earnings trend or a mean-reverting process rather than a random walk. The salient signal that leads agents in this direction is a number of earnings innovations of the same sign. Even though these occurred by chance, agents assume there is a pattern to be extrapolated into the future.

In behavioral models, bubbles may begin when agents overreact to signals about fundamentals. For instance, solid arguments for a permanent change in valuation resulted in setting off some well-known bubbles. Such was the case in the increase in real estate prices which led to the 2008 bubble in the USA. These prices were initially justified by some observers who assumed that securitization would permanently reduce real estate financial costs by diversifying idiosyncratic risk. The dot-com and railroad bubbles were similarly explained on the basis that technological breakthroughs would result in permanent productivity improvements.

1.4.3. The financial accelerator

It has also been proposed that the "financial accelerator" can amplify positive shocks to the fundamentals.[67] This statement accords with a financial theory which asserts that a small change in financial markets can result in a large change in economic conditions and create a feedback loop. Thus, the term financial accelerator is used for the economic shocks amplification and propagation mechanism, which aims to explain how small economic shocks can have large and persistent effects on the aggregate economic activity due to market imperfections. In macroeconomics, the financial accelerator represents the idea that adverse shocks to the economy may be amplified by worsening financial market conditions and, more broadly, that adverse conditions in the real economy and in financial markets propagate the financial and macroeconomic downturn.

In addition to the financial accelerator, several channels have been proposed as playing a role in the diffusion of financial crises. Some models emphasize investor behavior that results in herding and fads, influencing the behavior of capital flows and financial markets and exacerbating booms and busts. Other models stress economic linkages through trade or finance. Thus financial linkages – cross-border capital flows and common creditors – and investor behavior figure most prominently in the theoretical explanations of contagion. This section as playing a role in the diffusion of presents a summary of these theories.

1.4.4. Herding

Herding behavior as a transmission mechanism for bubbles and crises has been analyzed in different scenarios. For instance, some models have focused on mass behavior resulting from informational cascades which occur when an individual, having observed the actions of others, ignores her own information to follow their actions. In these situations, a few agents can have a disproportionate effect on the whole.[68]

Others have proposed that constraints, such as costs of gathering information which result in information asymmetries, can result in the transmission of crises. For instance, when the marginal cost of gathering information exceeds the marginal gain, it is rational for investors to mimic the market portfolios of others whom they assume to be better informed, giving rise to herding behavior and the transmission of crises through the global diversification of financial portfolios.[69]

Private versus public information has been the focus of models attempting to examine the reasons for and implications of decisions made by individuals who are influenced by the actions of others. For instance, in settings with sequential decision making, people may disregard their own information to copy the observed behavior of others. This type of herding may result from the assumption that those actions reflect superior information. An additional effect is that the herd externality provokes a positive feedback reaction: as more agents join the crowd, more people are induced to follow. The signals perceived by the first few decision makers – irrespective of whether these are random or correct – determine where the first crowd forms, and from then on the path everyone follows as they join the herd. This model captures excess volatility in asset markets, and the unpredictable changes in fashions which result from this behavior.[70]

1.4.5. Trade linkages

Looking into trade linkages, some analysts dug into the effects of competitive devaluations in currency, commodities, or markets.[71] For instance, since devaluation in one country makes its goods cheaper, it puts pressure on others to devalue as well in order to remain competitive in the international trade space.[72] Thus, devaluation in a second country is a policy decision whose effect on output is expected to be beneficial and have impacts beyond borders or within other sectors as the relative competitiveness of the various options is changed.[73]

1.4.6. Financial linkages

Common creditors of various types and financial linkages fulfill an important role in contagion. With respect to arbitrageurs, at times when they are fully invested and prices significantly decrease, they might choose to bail out right when their participation is most needed. That is, when arbitrageurs face fund withdrawals they might opt to liquidate their positions out of fear that further adverse price movements might cause a drastic outflow of funds. In these instances, they are not very effective in betting against the mispricing,[74] and that very liquidation results in the transmission of the crises.

The role of liquidity itself needs to be stressed.[75] A leveraged investor who faces margin calls needs to sell her asset holdings. Often, given information asymmetries, the strategy would not be to sell the asset whose price has already collapsed, but other assets in the portfolio. However, in doing so, other prices fall, and the original disturbance spreads across assets and markets.

Another channel that has been proposed to help reduce the shocks is the actions of commercial banks. These can exacerbate a crisis by calling in loans and cutting credit lines.[76] In this sense, the need of those banks to rebalance their overall risk portfolio and to recapitalize following balance-sheet losses can lead to a marked reversal in commercial bank credit across those markets where they have exposure.

The "wake-up call hypothesis"[77] relies on investor irrationality or the costs of acquiring information. It refers to a situation in which a shock in a country reveals a weakness that wakes up investors who then proceed to terminate their investments in other markets with shared characteristics.[78] An example would be the Greek crisis which encouraged agents to avoid other southern European countries.

Another scenario picking up on the idea of transmission due to shared characteristics and lack of perfect information is one in which contagion occurs when "informed" investors respond to private information concerning a country-specific factor by rebalancing their portfolio's exposures to the shared macroeconomic risk factors in other countries' markets. When there is asymmetric information in the countries affected by the rebalancing, "uninformed" investors cannot identify the source of the change in asset demand. Therefore, they may respond as if the rebalancing is related to information on their own country. Thus, contagion can occur between two countries even when contagion via correlated information shocks, correlated liquidity shocks, and wealth effects ruled out by assumption, and even when the countries do not share common macroeconomic factors, provided that they share at least one underlying macroeconomic risk factor with a third country.[79, 80]

1.5. The life cycle of a bubble

Each of the big bubble events recorded in history has been the result of a number of specific circumstances. One cannot attempt to compare situations as different as those of the sophisticated and liquid US market, with its many interconnections in the financial sector at present, with anything that happened in Holland, France, or the UK in the past few centuries. Nonetheless, one can find some common elements both in aspects of the bubbles' life cycles and some necessary (although not sufficient) market conditions that need to transpire for real market bubbles to develop. Figure 1.2 provides a succinct summary of the phases of the bubble life cycle.

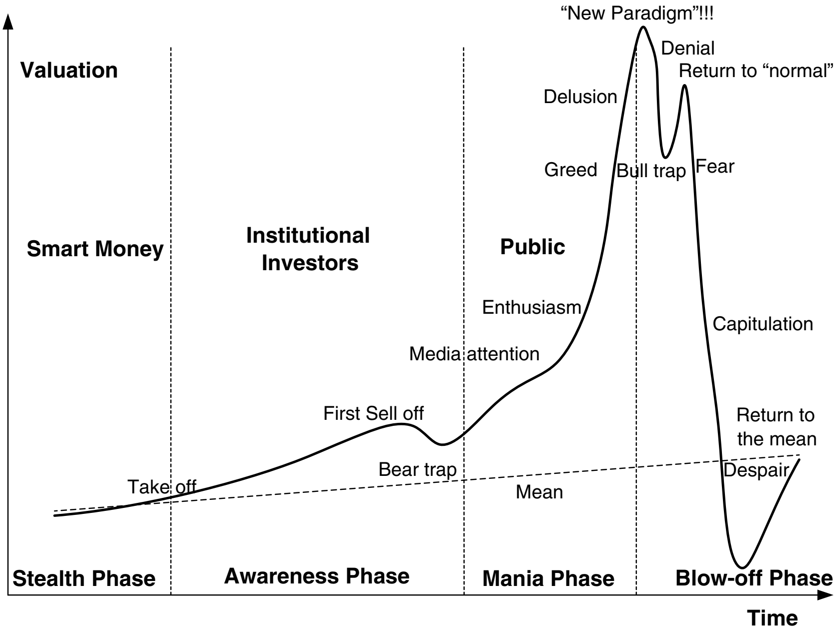

Figure 1.2 The bubble life cycle

Source: Reproduced from J. P. Rodrigue, 2009, http://people.hofstra.edu/jean-paul_rodrigue/jpr_blogs.html

1.5.1. Phases of a bubble's life cycle

The general condition prior to the bubble development is the existence of an economic environment where the "right mood" can be gauged. That is, speculative bubbles are driven by expectations of future earnings rather than current realizations. The potential of "important" wealth creation beyond other concurrently available alternatives needs to be "believed." In this context "important" is relative to the initial wealth of the investor. A second condition that must be met is sufficient credit availability to reach the market overall, as it is the aggregate speculative capacity of the market that drives the bubble. A third condition is the partial availability of information, rather than full transparency and knowledge. Imagining a limitless growth potential is more enticing than knowing that everyone else also knows that after a 5 percent growth, the price—value relation for the asset in question is "out of whack."

History provides numerous examples of bubbles that can be analyzed to find common factors. For instance, starting with a successful sector with strong fundamentals, markets go through a sequence of phases. Helped by the "strong sector," credit expands and money is accessed easily. As liquidity increases and gets used in purchasing the bubbled asset, prices rise, and as this happens more investors become immersed in the process with the expectation of additional price increases and quick profits. The bubble expands through the process of price inflation, and then bursts upon the realization that the rate in the price increase is no longer tenable.

Four phases of bubble life cycles can be identified, as shown in Figure 1.2:[81]

a. *Stealth*. During the stealth phase, informed market participants discover an opportunity for important wealth creation. Risk is attached to the opportunity as the future growth assumption is so far a hypothesis. At this time, it is sector experts who get into the market in a discreet manner and who, as prices gradually increase, purchase larger stakes in the business. This "smart money" has considerable opportunity for experiencing future revaluations as it understands the fundamentals and has entered the market early.
b. *Awareness*. In this second phase, investors start to realize the momentum, adding further capital and, in the process, increasing prices. Some players may also collect initial returns and there may be subsequent selling instances, every one starting at a higher level than the one before. During this phase, the smart money strengthens its position while the media begins to help more "unsophisticated traders" get into the market.
c. *Mania*. During this third phase, information regarding the business opportunity is generally known to all potential market participants. Increasing prices lure additional players into the market. Everyone has a stake in the growth potential of the bubble and incredible returns are "guaranteed" as past growth is transferred to the future. While prices keep increasing as a result of the money which pours in, the market "overall" has lost sight of the fundamental values. Also, these are ever more difficult to assess as they now appear intrinsically connected to the recent and expected growth. There is a retro-feedback mechanism by which beliefs are reinforced by increased prices, and prices increase further, pushed by the money invested in the bubble as a consequence of the actions resulting from these strengthened beliefs. In this phase, it is mainly the money of the general public – unsophisticated investors – that pushes the market up, while the experienced investors start to collect their returns and de-scale their investments. Because every Joe and Jane has made a little potential fortune, the market grows exponentially as more people, their families and friends, jump in to join the ride. To obtain a bigger share of the potential growth, an increasing portion of these investors access additional capital through leverage and loans. Again another feedback loop is blinding participants, as when the money pool enlarges and is used to purchase assets with overvalued growth expectations. Then there is a further appearance of wealth creation, which justifies additional investments to sustain prior debt-equity ratios. This tendency is halted when, for whichever reason, there comes the realization that fundamentals can no longer sustain

additional growth. Credit becomes unavailable, prices have reached a plateau, and the bubble is ready to collapse.

d. *Blow-off.* In this last phase, the mood has changed from festive to pensive. A general market realization has set in, changing expectations. While some sell, hesitation might trap others who are unable to distinguish a temporary setback from a permanent price collapse. By now the smart money is gone and it is the general public who are left holding the depreciated assets. There is a consequent discontinuity in prices which start to plunge at increasing rates. Now everyone is ready to cash-in and run off. Liquidity needs force sales. These new low prices compel further sales, particularly on the part of those who are most leveraged. The same retro-feedback mechanisms experienced during the price appreciation process are now running counter-clockwise and at a much higher speed. When prices reach a new low, the smart money returns once more to acquire the assets at bargain bottom prices.

Hyman Minsky[82, 83, 84] provided a characterization of the bubble cycle from birth to bust. His work distinguishes five phases: displacement, boom, euphoria, profit taking, and panic. We summarize them here although they closely resemble the ones described above. The process works as follows.

An initial *displacement* occurs as a new technology or financial innovation becomes available, leading to expectations of increased profits and economic growth. To illustrate this instance, we can think of the reduction in the US federal funds rate from 6.5 percent in May 2000 to 1 percent in June 2003 during which the interest rate on 30-year fixed-rate mortgages fell to a historic low, sowing the seeds for the housing bubble. A second example could be the moment of recognition of the uses of Internet technology and its impact on productivity.

During the ensuing *boom* phase, prices slowly gain momentum as new participants enter the market. This period is often marked by low volatility, credit expansion, and increases in investment. The asset also begins to attract media attention, as a result of which speculation draws in more participants. It is during this phase that prices start exceeding the fundamental improvements from the innovation.

The boom is followed by a period of *euphoria* during which prices increase exponentially and investors trade the overvalued asset in a frenzy. During this phase, there is often an associated high trading volume and price volatility. Valuations reach extreme levels and some investors become aware of the bubble. However, trading is sustained, given the hope that the assets can be sold to someone else in the near future. An example of this euphoria period was the peak of the 1989 Japanese real estate bubble, when land in Tokyo sold for as much as 350 times the value of Manhattan property. After the bubble burst, real estate lost about 80 percent of its inflated value, while stock prices declined by 70 percent.

After the peak of euphoria, smart money starts *taking profits* while demand from less sophisticated investors remains. Even when the bubble's existence is already known, it is not possible to predict the exact time of its collapse, so trading keeps unfolding. However, at some point, prices start to fall rapidly, leading to a *panic* phase. During this last phase, prices spiral down. If the run-up was financed by credit, amplification and spillover effects kick in. Faced with margin calls and plunging asset values, investors and speculators want to liquidate their holdings at any price. As supply overwhelms demand, asset prices slide sharply, often accelerated by margin calls and leveraged balance sheets. One example of global panic in financial markets happened in 2008, after Lehman Brothers declared bankruptcy and Fannie Mae, Freddie Mac, and AIG were close to collapse when, in a single month, global equity markets lost $9.3 trillion, adding up to 22 percent of their combined market capitalization.

As we shall see in ensuing chapters, much of the theoretical literature on financial bubbles tries to formalize this narrative. However, even though the literature on bubbles has made huge progress in the last 20 years, for the most part, models only work to explain portions of the Minsky framework. For instance, some may generate explosive price paths and others are quite incisive regarding trading volume, but a comprehensive model on the bubble cycle remains elusive.

1.5.2. Resulting misallocations

Bubbles create "fictitious" wealth and destroy real wealth. Even if these are inflationary occurrences, the final implosion brings about a deflationary period during which wealth vanishes from the overall economy. So the first lesson to learn is that the process from birth to implosion dislocates resources and can be very destructive, particularly to the most unsophisticated and less liquidity-wealthy participants or even innocent bystanders.

Figure 1.3 shows important resource misallocations resulting from bubble distortions, and a brief explanation is provided below.

a. *A bubble transfers wealth from the future to the present.* A share of the future demand is brought to the present, stealing expectations of future consumption. Demand that should have naturally occurred as part of an economic cycle is now squeezed into the current period. That is, demand is not creating sustainable wealth but rather shifting future consumption to the current time. One reason why future demand is altered is that the resources that should have been available to make future demand possible are consumed in the present, as bubbles are fueled by debt. Given that debt is consumed in the present and expended over the future, liquidity will be unavailable to pay for later consumption. Figure 1.3 shows that the amount for demand A is transferred into *bubble demand* B. The 2008 crisis allows us to ascertain this aspect with current inventories of housing available at depressed prices for years to come.

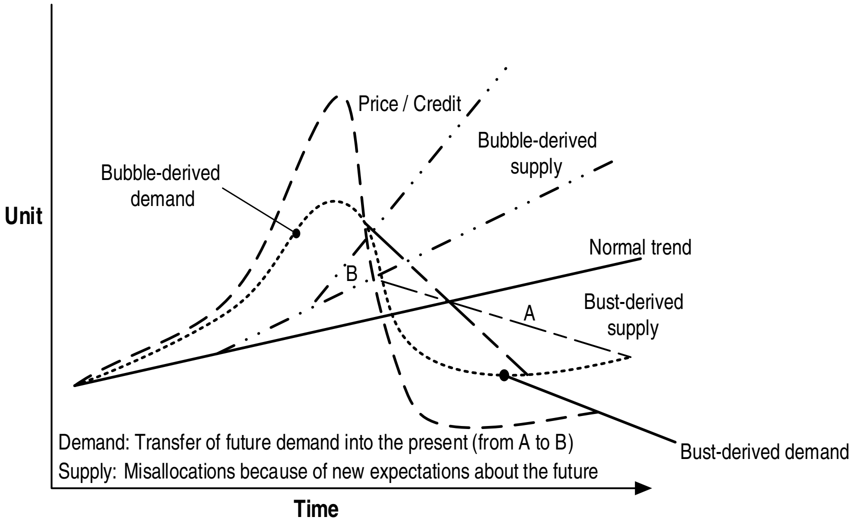

Figure 1.3 Bubble resource misallocations
Source: Reproduced from J. P. Rodrigue, 2009, http://people.hofstra.edu/jean-paul_rodrigue/jpr_blogs.html

b. *Bubbles reallocate wealth from some sectors of society to others*. Smart money will initiate the cycle by ripping most profits in capital revaluations and fees, while unsophisticated investors and latecomers will be left holding most of the undervalued assets (e.g. the loans to the devalued real state post-bubble implosion).
c. *Bubbles steal liquidity from the system*. Bubbles are fueled by easy credit. Given that liquidity is used to finance the asset bubble expansion, these resources are no longer available to finance growth in other sectors of the economy. Upon the bursting of the bubble, as assets are devalued and balance sheets require adjustments, credit simply disappears. Again, the 2008 crisis allows us to ascertain these aspects with easy credit being represented by "no down payment adjustable rate mortgages" and the evidence that credit is not reaching the street in sufficient amounts in many economies as of 2015.
d. *Bubbles induce misallocation of resources*. Given the spectacular returns produced during the initial phases of bubble growth, suppliers misallocate their resources and, while unable to cope with demand during the initial phase, they over-estimate future demand. The ensuing excess capacity results in a waste of capital, particularly worrisome in capital-intensive sectors, and eventually destroys know-how, production, and distribution capabilities. The reason is that suppliers have already scaled up their operations and, cutting production, in view of demand destruction, barely results in variable cost

reduction. The 2008 crisis illustrates this point as inventories of recently built houses deteriorated while waiting for owners, resulting in countless manufacturers supplying construction-related materials such as bricks, tiles, or housing sanitariums going bankrupt.

e. *Bubbles result in a post-implosion balance sheet induced recession*. After the bursting of a bubble, the affected assets will be revalued at their new, lower level, while the value of the liabilities will not change. Thus, from that time on, the corporations' main objective is to reduce debt rather than to maximize revenue. At the aggregate level, during the deleverage process, no one will borrow money independently of the interest rate. Consequently, money going into the local banking system cannot revert to the local economy. Furthermore, if both the government and the private sector deleverage at the same time, the economy is weakened. Balance-sheet recession can have long-term consequences as trauma resulting from the recent experience will prevent borrowing, even once the balance sheets are repaired. Fiscal stimulus will be needed to overcome this trend and get the economy back on "normal track."

1.5.3. Balance-sheet recession

Richard Koo, Chief Economist at the Nomura Research Institute, uses the arguments presented in the last point to advise a change in the way macroeconomic policy is conducted after a large bubble implodes, and when balance-sheet recession occurs. His main idea is that there are two differentiated phases to macroeconomics which he labels Yin and Yang. The Yang phase is identified with those periods of economic development when profit maximization is the overall goal of corporations. In this scenario, the "textbook" economic approach suffices as corporate balance sheets are healthy, monetary policy is desirable, and a less intrusive fiscal policy is recommended.

During the Yin phase, when debt minimization is the overall corporate objective, the effectiveness of these policies is reversed. Koo argues that during this latter scenario, low interest rates have no impact, as companies need to restore the health of their balance sheets before they embark on further capital investments. Given that all firms will be trying to minimize the value of their liabilities at the same time, the danger of economic contraction is ever present. His advice is to put in place measures to prevent falling into depression. Figures 1.4 and 1.5 summarize some of his proposals.

In summary, bubbles are damaging misallocation engines that transfer resources across time and segments of society. Some sectors benefit while others lose. There is no alignment of interests as it is more often than not a "non-zero sum game" where the interacting parties' aggregate gains and losses are either less than or more than zero, but whose shares of the gains are not proportional or equal to each other. That is to say, independently of the fact that overall

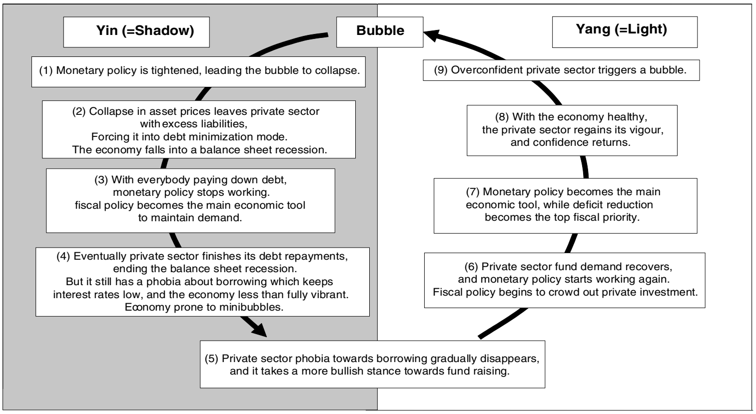

Figure 1.4 Yin-Yang cycle of bubbles and balance-sheet recessions

Source: Reproduced from R. Koo (2008), *The Holy Grail of Macroeconomics: Lessons from Japan's Great Recession* (Singapore: John Wiley & Sons), p. 160.

		Yang	***Yin***
1) Phenomenon		Textbook economy	Balance sheet recession
2) Fundamental driver		Adam Smith's "invisible hand"	Fallacy of composition
3) Corporate financial condition		Assets > Liabilities	Assets < Liabilities
4) Behavioral principle		Profit maximization	Debt minimization
5) Outcome		Greatest good for greatest number	Depression if left unattended
6) Monetary policy		Effective	Ineffective (liquidity trap)
7) Fiscal policy		Counter productive (crowding-out)	Effective
8) Prices		Inflationary	Deflationary
9) Interest rates		Normal	Very low
10) Savings		Virtue	Vice (paradox of thrift)
11) Remedy for Banking Crisis	a) Localized	Quick NPL disposal Pursue accountability	Normal NPL disposal Pursue accountability
	b) Systemic	Slow NPL disposal Fat spread	Slow NPL disposal Capital injection

Figure 1.5 Contrast between Yin and Yang phases of cycle

Source: Reproduced from R. Koo (2008), *The Holy Grail of Macroeconomics: Lessons from Japan's Great Recession* (Singapore: John Wiley & Sons), p. 161.

wealth is destroyed or created, some participants end up richer while others end up poorer. In the end, the temporary benefits of bubbles are more than compensated by the wealth destruction their bust triggers. Furthermore, usual monetary and fiscal policy mechanisms need to be reworked to fit the deleveraging needs of corporations after a major asset value adjustment event has taken place.

2
Macro "Players" in Bubble Formation and Contagion Processes

2.1. Monetary and fiscal policy

Business cycles set off periods of investment that result in economic expansion. Thereafter, as the technology behind this growth matures, the development rate decreases and often reaches a stage of recession during which market participants re-adjust and clear excess capacity. Mild recessions, therefore, are part of the normal cycle. However, since the 2008 housing crisis in particular, it has been alleged that actions initiated by central banks interfere with this normalizing mechanism, and contribute to bubble formation processes. This is the case when, for instance, monetary and fiscal policy or the micromanagement of markets through dealings and communications are used to manipulate rates and expectations. Thus, the impacts of monetary and fiscal policies on interest rates and credit availability have been proposed as key players in the bubble formation and contagion processes.

Monetary policy which is based on the connection between interest rates and total money supply affects outcomes such as economic growth, inflation, and unemployment. A specific monetary policy refers to the actions through which the amount of monetary assets, that is, currency in circulation and demand deposits, in commercial banks is controlled to achieve a certain objective. To implement a given policy, monetary authorities target the level of interest rates (the price at which money can be borrowed), the monetary base or amount of money in the economy, and the reserve requirements or minimum fraction of customer deposits and notes that each commercial bank must hold as reserves. The distinction between the various types of monetary policies lies primarily in the set of instruments and target variables that are used by the authorities to achieve their goals at any given time.

In practice though, the main tool used to implement any type of monetary policy is modifying the amount of base money in circulation. This is done by buying or selling financial government obligations in open market operations, impacting the amount of money, or its liquidity (the latter, if less liquid forms of money are bought or sold). Thus, regular market transactions by monetary authorities modify the supply of currency. The multiplier effect of fractional reserve banking, which depends on the percentage of deposits banks are required to hold as reserves, amplifies the effects of these actions.

The optimal monetary policy for an economy depends on the nation's economic situation. If the country is suffering unemployment and recession, authorities will use an expansionary policy and increase the total supply of money to encourage spending. On the other hand, if the country is undergoing an inflationary period, the authorities will decrease the rate of money supply to implement a contractionary policy to slow down consumption. In either case, there are a number of tools that can be used to strengthen the effect of such actions. For instance, when attempting to lower the unemployment rate or shorten a recession period, interest rates will be decreased to encourage business expansion and employment creation throughout the economy. These actions might also be coordinated with additional measures to ensure money supply travels to businesses and households in the form of different types of loans, mortgages, and credit. If, on the other hand, the government is attempting to control inflation, interest rates will be hiked to prevent spending and to stop the general increase of prices.

Governments use fiscal policy in conjunction with monetary policy to manage economic development and achieve price stability, full employment, and economic growth. The two main instruments of fiscal policy are changes in the level and composition of taxation and government spending in various sectors. These changes affect the aggregate demand and the level of economic activity, the distribution of income, and the pattern of resource allocation between the government and private sectors. Expansionary fiscal policy involves government spending that exceeds tax revenues, whereas a contractionary fiscal policy involves government spending that is lower than its tax revenues.

A broadly discussed example of the role of monetary and fiscal policy in a crisis was introduced in Chapter 1 in reference to the actions of the US Federal Reserve from 2001 to 2008. The US fiscal and monetary policy during the years preceding the housing bubble implosion had global ramifications and has been listed as a key culprit in the bubble development process.

As mentioned in Chapter 1, to stimulate the economy following the 2001 recession, short-term interest rates were kept low throughout the bubble growth period, starting on January 3, 2001, when the Fed lowered short-term interest rates. This action was repeated on 11 more occasions, until the rates reached 1.75 percent. In 2003, large US firms could access funds in the 90-day commercial paper market at an average of 1.1 percent, while the three-month Treasury bill rates stayed under 1 percent. From June 2003 to June 2004, the federal funds rates were kept at 1 percent while over the next two years, as deflation fears waned, they were raised to 5.25 percent in 17 quarter point increases. These prices had their mirror in the rates for the average 30-year fixed-rate mortgage which, at 40-year lows, helped the housing sector soar. The short-term monetary authority's objectives were achieved. However, in the view of many, the Fed simply kept rates too low for too long. These critics lay the blame for the 2008 recession primarily on this action.[1]

Excessive monetary liquidity, such as easy credit, potentially occurs when the practice of fractional reserve banking is used to implement an expansionary monetary policy, that is, lowering interest rates and flooding the financial system with money supply. In the words of Axel A. Weber, European Central Bank council member and former president of the Deutsche Bundesbank:

> the past has shown that an overly generous provision of liquidity in global financial markets in connection with a very low level of interest rates promotes the formation of asset-price bubbles.[2]

Some believe that money supply is controlled exogenously by central banks or an equivalent governing body and attribute the expansionary monetary policy to these organizations. However, others contend that money supply is produced endogenously within the banking sector and consequently attribute such policy to the actions of that sector. These alternative paths explain in part how efficiently policies such as the fractional reserve banking work to carry out central banks' efforts to raise or lower short-term interest rates. This is essential given their role in the creation, inflation, and ultimate implosion of bubbles in the economy.

Whichever mechanism is used to transmit these policies, the general finding is that when interest rates are excessively low, investors avoid putting capital into savings accounts. Instead, they borrow from banks and invest the leveraged capital in financial assets. It was this risky behavior prevalent during the 2008 crisis that lead to an increasingly fragile economy, contributing to the rise and growth of the bubble.

When too much liquidity floods the markets, to reverse the inflationary trend, monetary authorities may try to remove some of it from the system by raising interest rates. Investors respond by becoming more risk-averse and avoiding the now-too-expensive debt. Additional mechanisms to slow growth may include putting in place countermeasures such as increasing capital reserve requirements and imposing regulations that inhibit processes leading to over-expansion and excessive leveraging. These measures lighten the negative impact of a downturn, if taken while the economy is strong.

Figure 2.1 provides anecdotal evidence for the argument that low interest rate policies by the US Federal Reserve exacerbated the housing and commodities bubbles and that these bubbles only ended when the excess money was curtailed, raising federal funds interest rates.

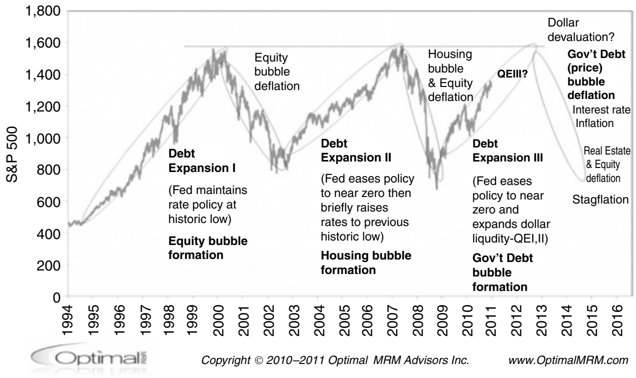

Figure 2.1 A recent history of debt expansion, bubble formation, and bubble deflation, as narrated by the Standard & Poor's 500 Index

Source: Reproduced from G. Z. Hajjar, CFA, Optimal MRM Advisors Inc., *Debt Expansion and Bubble Formation*, http://www.optimalmrm.com/debt-expansion-and-bubble-formation/ Downloaded 4 November 2013.

2.2. Credit, global flows, and the repricing of risk

Under normal circumstances, supply and demand tend to be closely related and evolve according to factors associated with demographics, technology, productivity, and income. However, when a bubble emerges, the market mechanisms that would have normally corrected undue asset inflation, such as higher interest rates or a decline in demand due to high prices, are disrupted. The most common source of disruption is lax credit, which can be the direct outcome of the policies of central banks.

Nonetheless, excess liquidity can also result from other factors, such as inappropriate lending standards or foreign flows. Hence, flows such as those resulting from the increased saving rates in China, oil-producing countries, and other big developing nations which make their way to the USA and Europe have also contributed to the decrease in interest rates. "The system was awash with liquidity, which helped lower long-term interest rates."[3] As credit spreads narrowed significantly, the costs of borrowing to finance risky investments declined relative to safe assets, such as US Treasury securities.

The residential housing bubble was the most visible effect of the credit bubble, but not the only one. Commercial real estate, high-yield debt, and leveraged loans were all boosted by the surplus of inexpensive credit. In this respect, 2008 Nobel Prize winner in Economics, Paul Krugman, declared: "It's hard to envisage us having had this crisis without considering international monetary capital movements. The USA housing bubble was financed by large capital inflows. So were Spanish and Irish and Baltic bubbles. It's a combination of, in the narrow sense, a less regulated financial system and a world that was increasingly wide open for big international capital movements."[4]

For his part, Fed Chairman Ben Bernanke described the strong relationship between financial account surplus growth and house price appreciation this way: "Countries in which current accounts worsened and capital inflows rose . . . had greater house price appreciation [from 2001 to 2006] . . . The relationship is highly significant, both statistically and economically, and about 31 per cent of the variability in house price appreciation across countries is explained."[5]

Monetary and fiscal policies are used to drive the economy toward the nation's desired objectives. However, these activities can have unintended consequences. There are at least two reasons for this. The first is that even though, in economics, there is the assumption that a complex system of determinants will tend toward equilibrium, this tendency may actually be absent. Instead, unstable patterns of events such as virtuous and vicious circles appear, both showing positive feedback loops and staying in the direction of their momentum, until an exogenous dynamic force intervenes and stops the trend or cycle.

The second reason is that there exist interconnections between asset prices and economic fundamentals which interact via a two-way causality relation: the economy drives asset prices and asset prices drive the economy. These interactions are difficult to disentangle because the firms' business activities are embedded in a general economic environment that impacts the organizations' development through variables beyond their control, such as inflation, the cost of credit, or tax expenditures. At the same time, the general economic background is set and responds to overall conditions, including those within the business environment. The result is that the practices through which investors are expected to maintain the optimal equilibrium are undermined.

Every economy needs a well-functioning ecosystem of finance to manage systemic risks and cyclical bubble growth. Financial markets perform many functions. For instance, they transfer value over time, across nations and sectors, and supply information on the expectations and value of the assets. Another key role of financial markets is the allocation of risk among market participants. This function is completed through a number of privately negotiated derivative instruments, such as swaps, which assist in shifting risks across various economic units. The concern is that, as their use expands, these instruments might increase the vulnerability of the financial system rather than contribute to a better allocation of risks.

We dedicated the previous section to discussing government monetary policy, interest rates, and credit in the context of their role within a bubble environment. In this section, we examine their unintended consequences, the risks that these present, and the costs resulting from interconnections between different segments of the market economy. To see how this might happen, we can analyze some of the documents investors use to build their asset value estimations. In times of financial stability, but even more so in times of bubble formation, balance sheets, income statements, and other information relative to economic fundamentals are all affected by financial markets in a manner that hinges on the investors' ability to obtain objective exogenous measures of value.

For example, with respect to the balance sheet, if in times of credit expansion one compares the amounts of loans received by any firm against the market value of the assets purchased, the values are a match. Credit gets approved because the expansion of these loans corresponds to the value of the assets being purchased. At the aggregate level, the accumulation of debt in the economy increases in proportion to the valuation of the economy's assets.

However, given that during a bubble formation period there is a fast increase in the value of the assets, it is often the case that in time, the matching pile of debt seems inadequate. That is, the leverage ratios are too low and companies are prompted to increase them. The reverse is also true when the cycle turns around and asset prices fall. When a bubble implodes, asset prices suffer discontinuities and the stock of outstanding debt immediately becomes disproportionate. This refers to the instance presented in Chapter 1, resulting in the firms' need to deleverage as a result of decreased asset values while debt balances are kept constant.

When debt ratios change, credit downgrades follow, and the cost of capital rises, leaving firms in weak positions. This is always undesirable, but much more so when the effects impact a whole sector. In these instances, mark-to-market accounting to balance-sheet analysis provides destabilizing positive feedback effects.

Accounting rules involve a trade-off between the "loss spiral" and the "margin spiral." When a negative shock lowers asset prices and erodes the financial institution's capital, worsening the balance sheet, two liquidity spirals come into effect. The "loss spiral" refers to the instance when in order to maintain the leverage ratio the financial institution must sell its assets, precisely at a time when their price is low. Thus, assuming that asset prices depend on the health of investors' balance sheets, the erosion of the investors' net worth further reduces asset prices, which feed back into their balance sheets. At the same time, lending standards and margins tighten, leading to the "margin spiral."

If all positions are mark-to-market, the "loss spiral" is more pronounced. However, allowing institutions to hide losses does not stabilize markets. This is because when mark-to-market rules are not followed, asymmetric information exacerbates the "margin spiral." The less liquid the stock, the more pronounced the price drop. For some structured products with very little market liquidity no reliable price might exist. Selling some of these assets in a financial crisis would establish a low price and force the holder to mark down the remaining holdings. Hence, investors prefer to sell assets with higher market liquidity first.

When assessing the quality of a loan or the equity of a company, analysts focus on the individual financial statements of the firm, given that its assets are the collateral for its credits. Nonetheless, within a bubble period, what once held true at the micro level no longer works at the macro level. A single home can always be sold to repay its mortgage, at least in theory, but when 40 percent of all homes within a neighborhood are up for sale at the same time during a period of devaluation, the market will not clear.

The income statement is also affected in situations when, all else being equal, additional easy credit is available and savings are discouraged, introducing further credit and liquidity into the system. Once borrowing is permitted and workers can use extra money from the bank (or from home equity, as during the housing crisis) to spend on more goods, corporate profits will be boosted without an increase in the corresponding wage bill. Thus, price-to-earnings ratios (P/Es) or any other measure dependent on the flow of money around an economy will be affected.

The opposite is also true, and decreasing the borrowing capacity or increasing the savings rate will tend to depress overall corporate profits. Consequently, central banks will try to reduce the savings rate of the economy when their priority is to boost economic activity. If needed, they will associate this measure with decreased interest rates to further push spending. The unintended consequences of reducing the savings rate is that it leaves the economy in a weak position, making it difficult to deal with long-term investment and growth, and sets the economic environment dangerously positioned to deal with further adverse shocks.

Keynes's "paradox of thrift"[6] and his recommendation for fiscal stimulus summarize the process by which borrowing and saving drive economic activity. In this paradox, Keynes explains how when one sector of the economy increases its saving rate, another sector's income is reduced. This will retro-feed back to undermine the income of the original savers, leading them to further reduce their spending, and causing a self-reinforcing cycle of declining activity. If, for any reason, most economic agents increase their risk-aversion and respond by further escalating their savings rates, a self-fulfilling economic contraction may arise.

According to Keynes, the way out of the paradox is to let one agent (the government) increase spending to boost profits, thus promoting additional borrowing to generate further profits, ultimately leading to a virtuous cycle of economic expansion. In reference to this paradox, Minsky[7] argued that borrowing may lead to a self-reinforcing positive spiral in which increased borrowing leads to higher profits, endorsing the decision to further borrow and spend. Keynes's and Minsky's proposals are relevant because they show how a credit expansionary process can drive asset market instability which, in turn, can hinge on the investors' ability to form objective judgments about asset values.

In summary, there is a connection between asset inflation, credit creation, and profit formation. Higher borrowing results in both higher profits and asset prices, while falling levels of borrowing reduces them. In these situations, investors can be misled as, at the aggregate level, corporate earnings do not always provide a dependable measure of the true "value" of the stock market, given that P/Es, revenue growth, and other such variables do not present investors with objective external measures of asset values.

Furthermore, dependence on these figures can lead to self-reinforcing positive and negative asset price and credit cycle patterns, which feed back into the real economy. As profits and asset prices start to increase, companies feel more disposed toward making new investments, which is a key driver of corporate profits, with higher investment spending from one company flowing through to higher corporate profits of another. Increased investment spending also implies higher employment, stronger consumer confidence, greater willingness to borrow, and stronger retail sales and economic growth figures. The upshot of these linkages is that asset inflation and credit expansion flow back into the real economy, generating self-ratifying and stronger economic data. The reverse is also true: credit contraction undermines profits, reduces investment spending, weakens employment, cuts consumption, and creates conditions ripe for still more credit contraction. Thus, credit formation (borrowing for either consumption or investment) lies at the heart of the financial market's fundamental instability.

Booms and busts create an economic climate which prevents the formulation of objective measures of the assets' fair value, and hinges on the capacity of those operating within the bubble to ensure a sustainable level of credit. Thus, once incomplete information is taken into account, financial markets may form bubbles without the investors ever behaving irrationally.

Interconnections between credit and data became quite clear during the events surrounding the 2008 housing crisis, when the collapsing mortgage market led to reduced corporate profits, deteriorating labor markets and weakening consumer demand. As has been argued, increased money supply, low interest rates, and credit availability are all contributors to this state. However, their existence "alone" does not constitute sufficient conditions for bubbles to form. In addition to these factors, an important sector of the market has to run the demand for an "asset" to the point where a bubble can be started.

2.3. Connectivity, systemic risk sharing, and transmission mechanisms

2.3.1. Connectivity and exposures

Assessments and exposures that allow us to understand how the unintended accumulation of large risks travel through the system are complicated by partially available data on the positions of counterparties and insufficient information deriving from accounting rules. Financial guarantees can also transfer risk across different sectors in the economy and produce negative feedback loops. For instance, implicit guarantees extended by governments to banks resulted in the accumulation of unanticipated risks in the public sector's balance sheet.[8]

Contagion played a major role in the 2008 crisis. The reason is that over the last 20 years, the financial system has become more complex as the separation between hedge and investment funds, insurance companies, banks and broker/dealers has become blurry due to financial innovation and deregulation. This additional complexity entails an increased interdependence among holders of different financial instruments within the same and other markets around the world. For instance, during the housing crisis, the financial system was vulnerable because the sudden failures of large firms could trigger balance-sheet losses in their counterparties. Policymakers deemed these institutions too-big-to-fail as they held important interconnections to other large financial parties.

It is necessary to recognize how bubble implosions spread throughout the financial system, and to recognize the interdependencies that exist among sectors, markets, and asset classes. These contagion mechanisms will highlight the need to have an "integrative view" of monetary, fiscal, and financial stability policies which consider their effects within the system and across countries. Thus, while appraising a structure for assessing macrofinancial risk, particularly financial system credit risk and sovereign credit risk, one should be aware of three key facts[9] as detailed below.

First, the degree of connectedness across different types of entities such as banks, corporations, and governments, changes over time.[10] Accordingly, in monitoring the connectedness of the system, one needs to consider these dynamic structures.

Second, macro risks accumulate in a nonlinear fashion. That is, when the value of the underlying assets falls, the impact on the accumulated risks increases, exhibiting convexity. A first shock has a smaller impact on the assets than a second equal shock. Consequently, risk parameters need to be adjusted when the circumstances deteriorate.

Third, the feedback loops of these risks arise to a greater degree when each party – banks, governments, and corporations – is effectively guaranteeing the other. Governments are hit by weakening banks of which they are guarantors, and the banks (which hold public debt) suffer more when the sovereign debt weakens. The same mechanisms work with respect to corporations and banks, banks which have interactions with other banks, and the sovereign debt of their own and other countries. Another consideration is the interactions among sovereigns. Thus, negative feedback loops link sovereign tensions to the problems of the banking sector, which is linked to the corporate and household sectors, and vice versa. Simply put, the weakness of one spreads to the other. These points can be better clarified with an example.

Let us think of a credit instrument such as a loan, a corporate bond, or a mortgage. These instruments are risky, but if we add to them a full faith and credit guarantee of the US government, what we get is "risk-free" debt. The US government guarantee is risk free because it (the government) will pay if the issuer does not. Thus, without the guarantee component, we have a risky loan, whereas once we have the guarantee, the risk disappears and the loan becomes "risk free." That is why the addition of a risky loan and a loan guarantee is for all practical purposes equal to a default-free loan. So a "risky" loan is equal to the "default-free" loan minus the loan guarantee.

Imagine a scenario where the guarantor of the debt and the lender are two different entities. In this instance, the borrower would first buy the guarantee at a cost of $5, and then take the guarantee to the lender from whom she would borrow $100 at a default-free (risk-free) interest rate of 7 percent. Given the numbers in our example, the borrower ends up receiving a net amount of $95 ($100–$5) in return for a promise to pay back $107 in a year. The interest rate on the loan is then stated as 12.63 percent, ($107–$95)/$95. On the asset side of the balance sheet, we have a "risky asset" with a value of $95, and a "guarantee" worth $5. On the liability side we have a "guarantee debt" for $100.

On the other hand, in case the lender and the guarantor are the same entity – such as in the case of a commercial bank – the borrower would simply get the net $95 from the bank in return for a promise to repay $107 in a year. Effectively, this would amount to the same interest of 12.63 percent. This is risk-free because there is a guarantee. Thus, this *promised* rate reflects both the risk-free interest rate and the charge for the guarantee. To see that the two are separable activities, note that the holder of the risky debt could have bought a third-party guarantee for $5.

The point of the example is to show that when an investor acquires a risky bond, loan, or mortgage, she is performing two actions: risk-free lending, similar to the time value of money, and writing a guarantee, similar to selling insurance. The same extends to the acquisition of any real-world loan. The relative value of the guarantee with respect to the value of the default-free component of the loan depends on the specific characteristics of the security. A triple-A high-grade bond would have a very small guarantee component, whereas a "junk" bond would contain a very large one.

Functionally, this also applies to banks as their lending equally consists of these two distinct activities: pure default-free lending and the bearing of default risk. More generally, this is valid for other forms of debt obligations: whenever a lender makes a loan, she is implicitly selling a guarantee.

Table 2.1 illustrates a very simple example of a corporation with Assets (A), one debt issue (D) with face value (B), and Equity (E).[11] Note that the assets of the balance sheet equal the sum of the debt and equity of the corporation.

Table 2.1 Guarantor of debt when there is risk of default

Risky Debt + Guarantee of Debt = Risk-Free Debt Risky Debt = Risk-Free Debt – Guarantee of Debt	
Corporation Balance Sheet	
Operating Assets (*A*)	Debt (*D*) with face value *B* Common Stock (*E*)
$A = D + E$	

In default, the holder of the guarantee receives the promised value of the debt minus the value of the assets recovered from the defaulting entity, that is the MAX [0, $B - A$]. The value of guarantee is then equal to a put option on the assets of the borrower. Credit default swaps (CDS) are guarantees of debt and, therefore, put options on the assets of the borrower.

In this example we have a zero-coupon corporate loan[12] with a face value of *B* dollars which is issued together with a government guarantee. Upon maturity, either the creditor gets paid what she is owed or she does not. *Ex post*, if the creditor is paid, the value of the guarantee is zero. This is just as if you were to purchase health insurance but not make any use of it. On the other hand, if the lender is not repaid, after the liquidation process, she receives the company's assets, whatever these are worth. However, given that in this instance there is a guarantee, she would then get the face value of the loan, *B*. As a result, the *ex post* value of the guarantee is equal to the promised amount, *B*, minus the value of the assets that she is also getting, *A*. The payoff function of the guarantee is the maximum of $0, or the loan face value minus the value of the assets, that is the MAX [0, $B - A$].

Guarantees are contingent liabilities for their issuers with required future payoffs dependent on the values of other assets. The correspondence between a risk-free loan and a combined risky loan plus guarantee suggests the analogy between a loan guarantee and a put option[13] on that loan. The put option is a binding commitment to buy the loan for V_0, the promised value of the loan, in case the borrower is unable to repay in part or fully. Thus, the position of the holders of the corporation's debt is analogous to a default-free debt combined with a short position in a put option on corporate assets. Using this analogy, and given that a government guarantees the liabilities of the banking system, Merton[14] suggested the use of "delta" to measure the risk of exposure resulting from changes in the value of corporate assets (hence, the assets of the banking system).

Option theory proposes a number of ways to measure exposure to risk but the most common one, delta,[15] measures the change in the value of an option as the value of the underlying asset changes. In our example delta can be used to measure the extent to which the government's exposure to its guarantee varies, as the value of corporate assets, and thus of the assets of the banking system, vary.

To get an idea of how this works, Figure 2.2 shows two relations. The straight line is the *payoff* to the holder of the guarantee (the option) at maturity, which is a function of the mark-to-market price of the underlying asset (depending on the case, the bank's loan or the capital of the corporation, etc.). The payoff is positive for $V < V_0$, where V_0 is the promised value of the loan, and zero for $V \geq V_0$. The other relation displays the *value* of the guarantee (the option) as a function of the mark-to-market price of the underlying asset. This is inversely and nonlinearly related to the price of the underlying stock.

Figure 2.2, shows that the "payoff" function of the guarantee resembles the payoff function of a put option on the underlying assets of the borrower. The specified strike price is the promised payment on the debt, and the expiration date is the maturity of the debt. For a home mortgage bond, the put option is on the value of the house; for a corporate bond, the put option is on the value of the corporate assets; and for a sovereign bond, the put option is on the value of whatever sovereign assets the creditor lays claim to, including taxing power.[16]

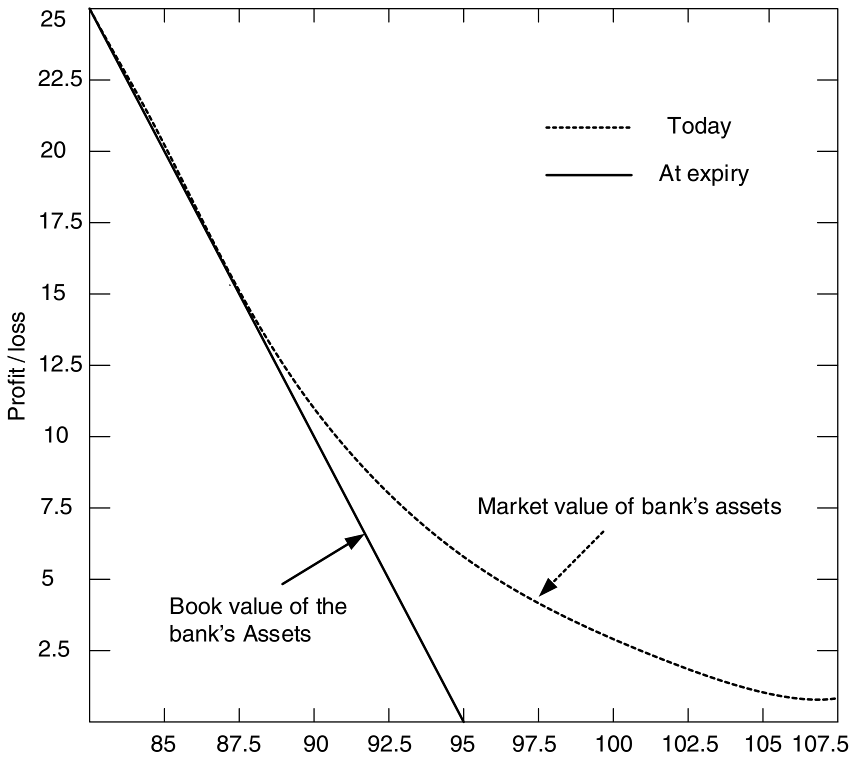

Figure 2.2 Value and payoff of a guarantee

If we examine the guarantees banks write to their borrowers through loans or mortgages, we realize the value of the guarantee/put is a nonlinear inverse function of the value of the underlying asset. Thus, as the value of the assets decrease, the value of the guarantee/put increases and vice versa. The nonlinear curvature is the generic shape of a put option.

In Figure 2.3, panels A, B, and C illustrate the relationship between the banking sector's debt and assets, the value of the guarantees, and the nonlinear function of macro risk aggregation across all productive sectors. Let us assume that in a first instance the value of the corporate or household sector's asset is worth A_{t0} in panel A with corresponding debt level D_{t0}, the value of the debt at time t_0 and G_{t0} in panel B, the value of the guarantee at time t_0. Suppose that an exogenous shock impacts the value of the assets, which now falls to A_{t1}, the value of the asset at the new time t_1. At t_1, the value of the underlying debt is devalued to D_{t1},

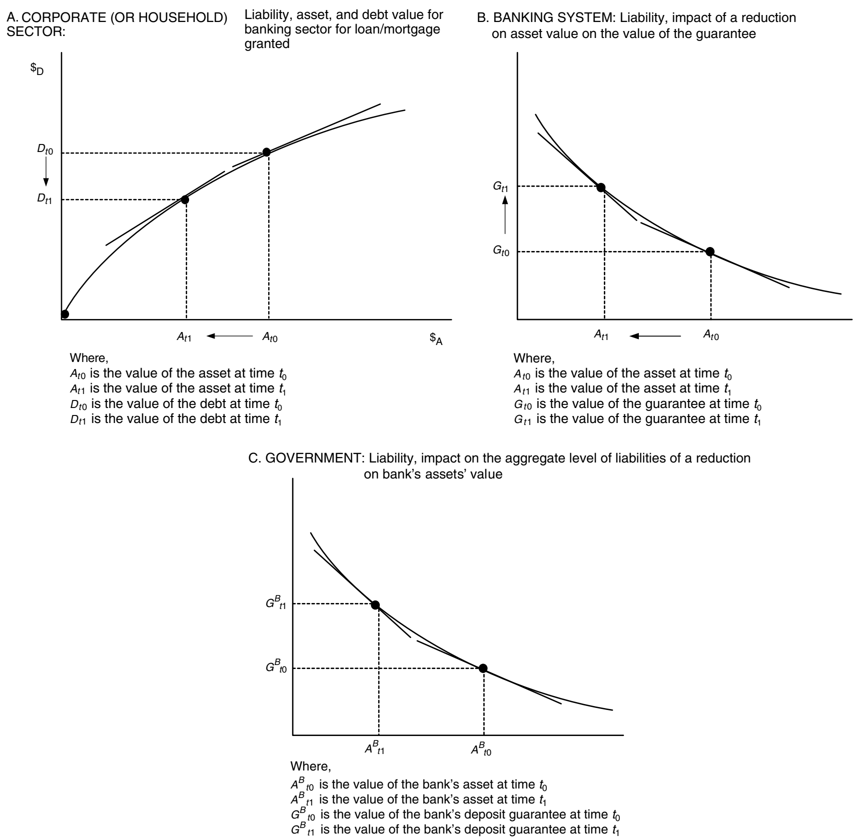

Figure 2.3 Nonlinear function of macro risk aggregation

the value of the debt at time t_1, whereas the value of the guarantee/put increases to G_{t1}, the value of the guarantee at time t_1. When the corporate/housing assets' value decrease to A_{t1} the bank's assets' value in panel C, A^B_{t0} decrease to A^B_{t1}, and the value of the bank's deposit guarantee G^B_{t0}, increases to G^B_{t1}.

The key consideration here lies in the "nonlinear" sensitivity of the loan value to the movement in the value of the underlying assets. In a crisis, the problem the bank needs to resolve is how much would the value of the guarantee change as a function of changes in the value of the underlying assets. And as already said, the bank would have to estimate the option's delta, the slope of the straight tangent lines in Figure 2.3. If the slope is –0.15, then the value of the guarantee/put would increase by \$0.15 for each \$1 decrease in the value of the bank's assets. The slopes of the tangent lines change from the original position of the value of the assets to the second position once their value has been decreased.

Given the obvious nonlinearity of the function, we would need to ask the question of what impact a second shock would have on the value of the guarantee. As we are now in a different, steeper section of the curve, let us say we obtain –0.45. Hence, a \$1 decrease in the asset's value will have three times the impact of the first shock. Thus, a second shock of equivalent magnitude will have a much larger impact than the first. The reason is that the shape of the curve is convex; hence this risk property can be described as *convexity*.

The effect of this relationship appears counterintuitive to those who are not familiar with it, but it easily explains some events observed during the aftermath of the housing crisis. Even when banks were no longer taking new risks, a second shock of equal magnitude to the same book of loans would lead, during the crisis, to an even larger loss. The balance sheet had not changed, but the risk had increased. That is because market prices tend to look "forward," whereas banking practice suggests that accounting values are not marked down when asset values decrease, at least not as long as the loans are performing. Rather, the losses in loan value from a series of asset value declines are recognized cumulatively and appear as a single large "unexpected" shock following a period of relative stability.

A related aspect is that, regardless of changes in asset values, a rise in the volatility of the asset returns also increases the value of the guarantee/put. A key characteristic of financial markets in times of distress is increased volatility, one of the drivers of the price of options. The underlying connotation is that the risk profile of an asset can change and this change does not necessarily immediately show up in related asset values. However, increased volatility has a significant impact on the value of those guarantees.

Governments guarantee banks in several ways, but the foremost method is through deposit insurance. De facto governments write guarantees on bank assets, which are effectively short put options. Hence, governments are guaranteeing puts. That is, they are writing puts on short puts which result in double convexity.[17, 18] Consequently, government guarantees (via the banks) are shaped by assets in the corporate sector, as shown in our example. For instance, we could propose that during the last crisis, the US government had a put on a put on the residential housing sector. If we focus on Figure 2.3 and dig through bank assets to the underlying corporate and household assets, we would see a curve of larger convexity than that of the banks' guarantees. When the asset value is high, this curve would be close to flat, but it would become very steep as soon as the asset value decreases.

The idea to keep in mind is that when corporate or household sectors suffer a shock, the newly created risk travels through the banking sector into the government sector. However, this is just an example, as these shocks can occur in any sector, individually or simultaneously, and then move on to others. For instance, the shock could start with the banking sector and then travel to corporations, households, and governments.[19]

The propagation of risk leads to the concept of feedback loops. For instance, rating agencies may downgrade the sovereign bonds of a country after learning that the value of its guarantees has increased. The downgrade will raise the cost of the public debt and lower the PV of future budget surpluses. This will lower the value of the government's assets precisely when the value of its liabilities has increased.

A deteriorating economy impacts the government's balance on the liability side, as the value of the guarantees increases. In addition, given that worsening economies receive fewer tax revenues, the asset side of the balance sheet is also affected as it diminishes. These two effects interact, as the debt-to-asset ratio changes more dramatically when the numerator increases and the denominator decreases at the same time.

Nonetheless, this is not the whole story, given that banks often hold their own government's debt as well. Thus, the government is guaranteeing its banks, and these same banks are holding the government's bonds. At the beginning of this section, we said that by holding bonds one is writing a guarantee, whoever the issuer is; thus, the banks are guaranteeing the government in return. The end result is that there are two parties, each of whom is guaranteeing the solvency of the other. So if the value of the government debt declines, the value of the put option that banks have written for the government through holding those bonds rises. Consequently, the value of the banks' assets falls, and the banks become weaker. Given that the government has guaranteed the banks, when the banks become weaker the value of the government guarantee increases which, in turn, means the government becomes even weaker, which feeds back to the banks' becoming weaker as well.[20]

The contagion of risks becomes more interesting when we realize that banks located in different countries interact with one another as well. For instance,

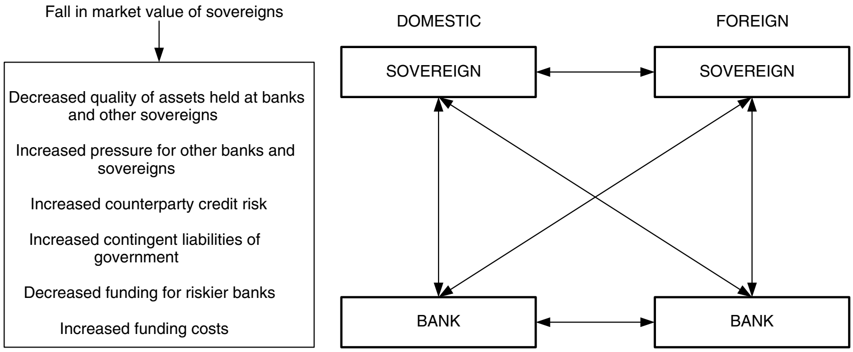

Figure 2.4 Feedback loops from explicit and implicit guarantees

a European bank may hold the sovereign debt of an Asian bank. Thus, when the situation of a given bank worsens, banks all around the world may be affected, and not only those which do business directly with the troubled entity. As an example, Figure 2.4 shows the feedback loops that tie domestic to foreign institutions, banks, and sovereigns.

In addition to feedback loops among banks and sovereigns, there are also linkages between sovereign tensions, the banking sector and the corporate sector, and lastly the household sector. Simply put, the weakness of one spreads to the others.

A number of studies have analyzed "connectedness" in terms of the impact that banks, insurance companies, and sovereigns have on each other's credit.[21, 22] The findings conclude that banks, insurance companies, and sovereigns are dynamically interconnected, with one country disseminating risk to another, and vice versa. A second finding is that the degree of connectedness between the different types of entities varies over time, and has increased in the post-crisis environment. These dynamic changes in risk exposure are expected in response to changing asset values and volatilities. Merton et al. (2013)[23] suggest that over time, the degree of connectedness between the following network connections changes substantially: banks to sovereigns, sovereigns to banks, and sovereigns to sovereigns:[24] "At this time, the sovereigns are having a big impact on banks, whereas before, the banks were having a major impact on the sovereigns."

2.3.2. The financial accelerator's role in the transmission of crises

In Chapter 1 we suggested that the "financial accelerator" amplifies shocks to the fundamentals.[25] In the original work by Bernanke, Gertler, and Gilchrist (1989),[26] the accelerator results from the interaction between *economic agents'* *net worth* and the *external finance premium* that emerges due to differences in the information held by lenders and borrowers.

The key to the financial accelerator is the relationship between the external finance premium (the difference between the cost of external funds and the opportunity costs internal to the firm), and the agents' net worth (the sum of liquid assets plus collateral value of illiquid assets' less outstanding obligations).[27] This relationship rests on two assumptions. First, the smaller the contribution of own wealth by the borrower, the greater the departure of her interests from those of the supplier of external funds. Second, when borrowers have superior information about the projects' characteristics, or greater capacity to understand the impacts on the projects' returns, the greater the incompatibility of interests that arise between borrowers and lenders, increasing agency costs.[28] Given the additional resources needed for monitoring, selecting projects, and so on (and/or) the larger the reduction of credit available in the economy, the greater the incompatibility of interests, the higher the interest rate, and the lower the availability of credit.

The financial accelerator effect works as follows. A change in aggregate economic activity affects the economic agents' net worth: if markets go down, so does their wealth. The terms under which economic agents raise external funds (the external finance premium) are inversely proportional to their net worth. The implication is that their ability to access money and the cost of external finance is countercyclical. Thus, funds are more difficult to obtain during bust periods than during expansionary phases. This results in swings in investment, spending, and production over business cycles. The higher costs and limited access to funds reduces the ability to borrow and hence the overall level of investments, spending, and production, depressing the economy even further (Figure 2.5).

The bind between the real economy and the financial markets derives from the corporations' need for external funds to exploit new business opportunities. However, the firms' capacity to borrow rests on the market value of their net worth. Given that lenders need to secure their returns, they require some sort of insurance, often in the form of collateral. This collateral is diminished when the value of the assets in the firm's balance sheet falls. The consequence is a deterioration of the firm's ability to borrow. Of course, the lack of necessary funds has a negative impact on the firms' investments, which on the aggregate decreases the overall economic activity further. This lower level of economic activity drives prices down, deteriorating balance sheets again, further tightening financing conditions and reducing economic activity even more. This vicious cycle is the financial accelerator. What starts as a minor change in the markets through this feedback loop results in a large change in economic conditions.[29] Frictions, such as information asymmetries, have an important bearing on the transmission mechanism of shocks.

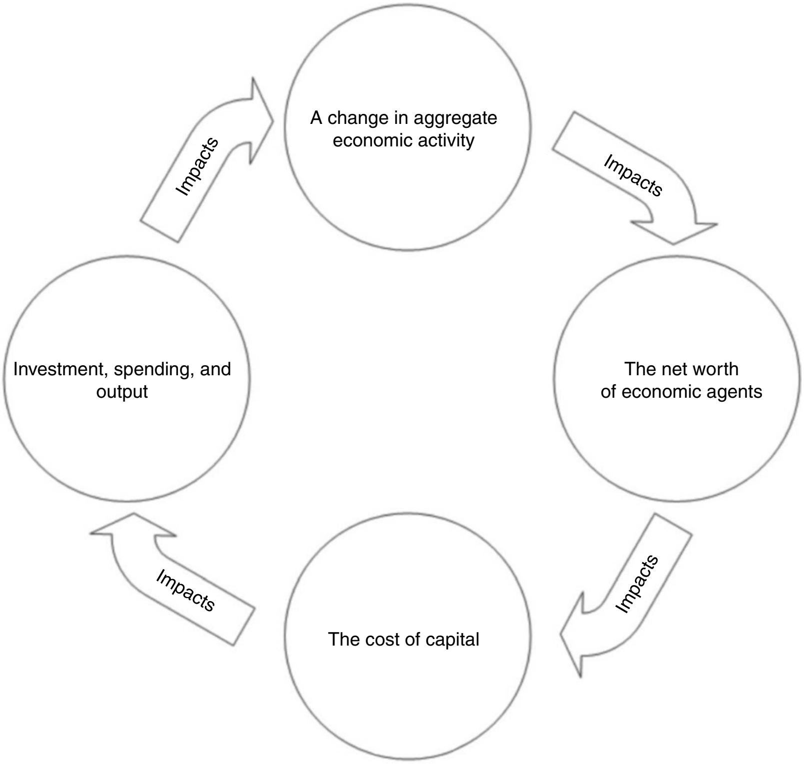

Figure 2.5 The financial accelerator effect

2.3.3. Economic instability and financial risk

The literature on financial instability is based on Keynes's "General Theory"[30] as well as the views of Joseph Schumpeter[31]. However, Minsky[32,33] has been credited with the key works on this subject. The financial instability hypothesis formulated by Minsky[34] has both empirical and theoretical aspects. The observation that the capitalist economic system does not seek equilibrium and stability but exhibits cycles of serious inflations and deflations resulting from changes in the real economy amplified by the economic system itself has motivated the study of disequilibrating processes and financial instability. A review of the main findings will help us achieve a better understanding of the processes of crises.

According to Minsky, the theoretical argument of the financial instability hypothesis characterizes a capitalist economy with pricey capital assets and a complex financial system. Individuals become leveraged, mainly through the banking system, to own those assets. This system provides a guarantee between the lenders – depositors of funds – and the borrowers – customers – who purchase the assets. Money is traded across time: in expectations of future profits, current money finances the resources needed to create future investment output, whereas the returns that accrue to those owning the capital assets are future CF. The balance sheet then presents a series of liabilities which are the prior payment commitments, and some hypothetical future CF derived from the productive capacity of the assets. Money flows through time and between depositors, banks, firms, and back to depositors. Initially these exchanges are done to finance the assets, and subsequently to repay the commitments. The past, the present, and the future are linked by the labor force, the capital assets, and the financial relations between them. But it is the expectation of profit which determines the flow of funding to firms and the market price of such agreements. In addition to businesses, modern financial relations have to consider households which also access debt, and the local and international governments which function as refinancing agents. It is the structure of aggregate demand relying on profits which determines the financial stability of a nation.

In his theory, Minsky identifies three income-debt relations which affect the stability of the system: hedge financing, speculative financing, and Ponzi. Hedge financing refers to those instances when obligations can be fulfilled with CF owned by the borrowers. Equity financing is key in this scenario as the more equity there is, the greater the likelihood that this is a hedge financing unit. The speculative units are those which expect to make payments out of liabilities, even by rolling them over, whereas Ponzi units rest on selling assets or borrowing further to repay principle and interest. The larger the proportion of speculative and Ponzi finance in a system, the greater the probability that the economy is a deviant system.

The first theorem of the financial instability hypothesis proposes that an economy is only stable under certain financing regimes. The second posits that during times of prosperity, the economy transits from financial relations that make the system stable, such as hedge finance, to others that make the system unstable, such as Ponzi. Thus, the financial instability hypothesis proposes that a capitalist economy generates business cycles from its own internal dynamics.

2.3.4. Systemic risk and transmission mechanisms

"Systemic risk" is used to describe events at the firm level which can trigger a collapse in a certain industry or economy. Brunnermier et al. characterized systemic risk as: "the risk that a shock affects the financial sector and triggers an endogenous adverse feedback significantly amplifying these shocks, causing further deterioration in the financial sector, and leading to significant output losses."[35] Systemic risk needs to be differentiated from systematic risk, which refers to overall market risk.

Systemic risk typically builds up in times of low volatility and materializes when it becomes apparent that the accumulated imbalances are not sustainable.

The subsequent fallout involves amplification mechanisms with spillover effects across the financial sector and the real economy of one or more countries. Thus, in most financial crises we need to differentiate the run-up phase during which bubbles and said imbalances form, and the crisis phase, when risk that has built up in the background materializes and the crisis erupts.

During the run-up phase of the crises, imbalances build up slowly and volatility is low. These imbalances are hard to detect, particularly because the price growth is not exorbitant and might be initially justified on an innovation. However, in time, it becomes evident that the fundamentals' improvements are surpassed by changes in prices and a bubble has formed. It is also during this phase that incentive distortions appear either due to rational behavior responding to incentives such as over-leveraging, or behavioral belief distortions, given the insufficient information available.

Given that run-up phases grow in low volatility environments, the resulting low-risk premium available during these periods makes financing easy. Speculators can get leverage, lowering the return differential between risky and less risky securities. As each individual speculator is exclusively concerned with herself and not the externalities her debt acquisition causes to the system, there may be, overall, an excessive leverage and maturity mismatch. For example, when levering up with short-term debt, each speculator may take into account that she might not be able to roll over her debt, forcing her to sell off assets at fire-sale prices. However, this same investor might not consider that her selling will depress prices, potentially forcing others to follow suit, exacerbating the fire sale.

After the gradual buildup of a bubble and the associated imbalances, a trigger event, maybe at the firm level, leads to the bursting. A sudden transition, a "Minsky moment," often occurs long after most market participants have become aware that a bubble has formed, the imbalances suddenly appear, and a time of high volatility follows. The fact that the price correction occurs too late and that the amplification mechanisms transfer the effects of the burst to other sectors of the economy can turn it into a major crisis. The amplification mechanisms increase the magnitude of the correction and spread the effects via direct contractual links, or spillovers or externalities due to common exposures, or the endogenous response of market participants to the event.

Indirect methods of transmission, such as spillovers, take place when the effect is transmitted via a third variable. For instance, when an institution is liquidating its assets at a loss, these new prices impact the market value of other portfolios holding the same asset. This reduction in the value of the portfolio of a second institution might result in forced fire selling. This action will again impact prices and an increased number of additional portfolios.

Amplification also arises due to the *self-reinforcing nature of market participants' actions*. For example, if one financial institution sells and depresses prices, others are more likely to follow suit and destabilize prices even further. Thus, *adverse feedback loops and liquidity spirals* may arise and amplify the crisis even more. This is often the case when the liquidity mismatch of market participants is high. An example of liquidity mismatch is when real investment on the asset side of the balance sheet is irreversible due to technological illiquidity, or the assets can only be sold with a large discount in times of crisis due to market illiquidity. At the same time, on the liability side of the balance sheet, the firm's maturity structure is very short term due to, for example, low funding liquidity. In times of crises, market liquidity is especially high for flight-to-safety assets, such as US Treasuries. However, the flight-to-safety status can disappear if agents stop agreeing on the flight-to-safety asset. Liquidity mismatch should not be confused with maturity mismatch. For example, holding 30-year US Treasuries funded with short-term paper has very little liquidity mismatch but a large maturity mismatch. Below we list some of the most relevant mechanisms of financial risk transfer in times of crises.

2.3.4.1. Creditor/depositor runs

Given the liquidity mismatch inherent in the financial system, a potential amplification mechanism during financial crises are creditor or depositor runs. These occur as depositors run on banks, as creditors run on unsecured short-term credit of financial institutions, and as margins run on secured credit. The rationale is that if consumers start withdrawing money, banks would need to liquidate some of their long-term investments to cover these withdrawals. This early liquidation of investments will reduce the bank's ability to service future payments and thus increase the incentive for more consumers to withdraw their money early. If the bank has to service its customers in a sequential order, depositors are incentivized to withdraw their funds as soon as possible. This is particularly the case if customers fear withdrawals will render the bank insolvent.

2.3.4.2. Information-induced bank runs

Empirical studies show that bank runs are related to negative news about the health of a financial institution or system. As bank runs usually occur when a fundamental variable crosses a given threshold, a small change in the information environment can result in large changes in behavior, leading to actions with the potential for amplification.

Another amplifying mechanism is Knightian uncertainty. "There is a fundamental distinction between the reward for taking a known risk and that for assuming a risk whose value is not known," Knight wrote.[36] A known risk is "easily converted into an effective certainty, while true uncertainty is not susceptible to measurement." Ricardo Caballero has invoked Knightian uncertainty to explain the behavior of investors in times of financial panic: "When investors realize that their assumptions about risk are no longer valid and that conditions of Knightian uncertainty apply, markets can witness destructive flights to quality in which participants rid their portfolios of everything but the safest of investments, such as USA Treasury bonds."[37] During the recent housing crisis, Knight's uncertainty realization drove investment banks to cut back in trade and investments, further slowing the economy.

2.3.4.3. Collateral/margin runs

Banks are subject to counterparty bank runs via increased margin requirements. Given that collateralized lending is secured by the value of the collateral asset, price reductions depressing the financial institutions' net worth would lead to losses which can instigate runs. When banks are forced to fire-sell assets, these sales further depress prices and increase losses, leading to a second round of selling, and so on. In this context, two liquidity spirals emerge: a loss spiral driven by the loss of net worth and the margin (haircut) spiral which works through increased volatility. The reason for the latter is that when volatility increases, margins increase, reducing the maximum leverage investors can take on. The resulting fire sales lead to higher volatility, which exacerbates the initial shock and leads to a further tightening. The two liquidity spirals often co-exist and reinforce each other.

2.3.4.4. Loss spiral

The loss spiral arises because a decline in asset values erodes the net worth of levered financial institutions, increasing their leverage ratios. To return to the target ratio, the financial institution might be forced to sell some assets at a discount. The lower the entrepreneurs' net worth, the lower the overall economic activity, profits, and retained earnings. The size of the loss spiral is determined by the liquidity mismatch and aggregate selling of all institutions. In addition to the persistent effect of losses on net worth, dynamic amplification mechanisms occur when entrepreneurs face a debt constraint and cannot issue equity because the borrowing capacity is limited by the collateral value of the physical capital.[38] A period's temporary shock not only adversely impacts future periods, but the cutback on future investment also feeds back to the current period. Reduced future investment depresses future asset prices as well as current prices, which again lowers collateral values, borrowing, and investment.

2.3.4.5. Margin/haircut or leverage spiral

A drop in asset prices leads to higher margins, haircuts, and cautious lending. There are several reasons for this. First, asymmetric information problems worsen as financiers become careful about accepting assets as collateral.[39] Unexpected price shocks might be precursors of higher volatility, and because margins are set to protect financiers, these increase with volatility. For instance, in a scenario where bad news leads to higher fundamental volatility in the future, there appear endogenous collateral/margin constraints as no payments in future periods can be credibly promised.[40] Thus, margins are destabilizing and the more sensitive they are to a change in the price level, the lower the experts' funding liquidity. If margins were to jump to 100 percent, margin funding would dry up. This is equivalent to not being able to roll over debt, because the firm becomes unable to use its assets as a basis for raising funds.[41]

In addition, liquid investors may hesitate to acquire collateral assets given expectations of future funding constraints. This is particularly the case when using short-term debt to purchase assets and becoming exposed to rollover risk. In this situation, relatively small changes in the fundamental value of the collateral asset can lead to large changes in its debt capacity.[42] Shocks to agents' funding conditions can also cause liquidity spirals, further deteriorating liquidity positions, spreading the crises across markets through funding liquidity and falling prices.

2.3.4.6. Contagion and flight to safety

In a setting with multiple assets, risky asset prices might move at the same time, since they are exposed to the same funding liquidity constraint. Losses can also generate contagion between assets when these are held by the same investors. In times when the experts' net worth is depressed, flight to quality takes place when the difference in market liquidity between high-margin and low-margin assets increases. This in turn can lead to larger endogenous margins, exacerbating the price difference even further. Relatedly, there is the flight to safety as less informed financiers, who set margins, cut back their funding to leveraged expert investors and invest their funds in nearly risk-free assets. The prices of risky assets with low market liquidity fall, while prices of flight-to-safety assets increase.

2.3.4.7. Lenders' limitations

When lenders have limited capital, they may restrict their lending as their own financial situation worsens, thus amplifying shocks. We can distinguish two main mechanisms through which this happens: (a) moral hazard in monitoring and (b) precautionary hoarding. Precautionary hoarding happens if lenders are afraid that they might suffer from interim shocks, such that they would lack funds for their own projects or trading strategies. The anticipation of those future shocks may lead to a cut in current lending.

2.3.4.8. Network externalities

Some studies have proposed the most direct way for the propagation of losses through a financial network is via direct linkages between the balance sheets of financial institutions or firms. The defaulting of one financial institution leads to losses on the balance sheets of other financial institutions, which may lead to further defaults. This is known as the domino model of contagion. If the initial default affects both direct losses on obligations and leads to decreasing asset prices, it provides an additional channel through which contagion can spread through the system. The two channels of contagion reinforce each other.

2.3.4.9. Feedback effects between financial sector risk and sovereign risk

In the previous sections of this chapter, we discussed how a feedback effect transfers through national and international networks when sovereign debt becomes risky.[43, 44] If financial institutions rely on sovereign debt for risk and liquidity management purposes, this introduces an interdependence of sovereign and financial sector risk through two main channels. First, an increase in the riskiness of government debt impairs financial institutions that have large exposures to sovereign risk. This raises the probability that the sovereign would have to bail out the banking sector, which further compromises the former's fiscal position. This impacts yields on sovereign debt and hence makes refinancing more challenging for the sovereign. Second, banks that suffer losses on their holdings of sovereign debt may reduce their lending to the real economy. The resulting decrease in credit slows down economic growth and thus reduces the sovereign's tax revenue, which again increases the riskiness of sovereign debt. In the context of the European debt crisis, this feedback mechanism has been referred to as the "diabolic loop" between sovereign risk and banking risk.[45] The following facts have been documented to underline these links:[46] banking crises often follow increases in external debt; banking crises usually precede or accompany sovereign debt crises; and public borrowing usually increases prior to external sovereign debt crises.[47]

3
Contributors to the Bubble Formation and Contagion Process

There is no agreed-upon, finite, and comprehensive list of "culprit" variables responsible for the creation of bubbles; instead, evidence suggests there are many contributing factors. For instance, in addition to factors pertaining to particular economic conditions and monetary policy (the low interest rates, easy credit, and transmission mechanisms discussed in Chapter 2), there are variables that relate to market imperfections as well as the collective behavior that prevails during such events.

3.1. Market imperfections

Within economic theory, market failures or imperfections describe the inefficient allocation of goods and services in a free market. Efficiency refers to a manner of using resources that maximizes the production of goods and services. A situation can be called "economically efficient" if no one can be made better off without leading to someone else becoming worse off (Pareto efficiency). That is, nothing more can be achieved given the resources available.

Market failures can be viewed as scenarios where individuals' pursuit of pure self-interest leads to results that are not efficient from the societal point of view. These failures are often associated with time-inconsistent preferences, information asymmetries, noncompetitive markets, principal–agent problems, externalities, or public goods. Some markets fail due to the nature of their exchanges, as these may suffer significant transaction costs, agency problems, or informational asymmetries. Such incomplete markets could culminate in economic inefficiencies which nonetheless can be ameliorated through market, legal, and other regulatory remedies.[1, 2, 3, 4, 5, 6, 7]

Market failure is frequently the reason for government intervention in a particular market.[8, 9] However, some government interventions, such as taxes, subsidies, bailouts, wage and price controls, and regulations, including attempts to correct market failure, may also lead to inefficient allocation of resources.[10] Thus, the choice is often posed between imperfect market outcomes, with or without government interventions.

The main feature of financial markets leading to imperfection is the *information asymmetry* between borrowers and lenders. Information asymmetry refers to decisions on economic transactions that are made in situations where one party has more information than the other. *Adverse selection* and *moral hazard* are two examples of market failures caused by this.

Adverse selection occurs when the seller values the good more than the buyer, because she has a better understanding of the asset's value. The seller is unwilling to settle for a price below the one she considers fair, while the buyer, unsure of the true value of the good, wants to insure herself against the possibility of purchasing a "lemon," or a bad asset. This asymmetry prevents the transaction from taking place.

Moral hazard describes a situation where one of the parties takes unnecessary risks because someone else bears the potential costs of her careless actions. One example is when a party insured against burglary leaves the windows of her home open. In this scenario, the potential for market failure is not the result of possessing prior information, but rather the inability of one party to control the increased risk-taking behavior of the second party. Also, while in adverse selection the seller is the one with "full" information, in moral hazard it is the buyer who has an unbalanced proportion of control.

3.2. Asymmetric information

Entrepreneurs and managers typically have better information than savers regarding the value of business investment opportunities. Consequently, savers face an "information dilemma" when they attempt to make investments in business ventures. The agency problem arises because savers do not intend to play an active role in the administration of the businesses in which they invest. That responsibility is delegated to the management who, thereafter, has an incentive to make decisions that expropriate savers' funds. For example, if savers acquire an equity stake in a firm, the management can use those funds to acquire perquisites, pay excessive compensation, or make investment or operating decisions that are harmful to the interests of the outside investors.[11] The *agency problem* arises because of the difficulties involved in motivating the "agent/manager" to act in the best interest of the "investor/principal," rather than in self-interest. This imbalance leads to agency costs incurred from monitoring the agent's actions, and may damage the interests of the party with less information, resulting in market failure.

The agency problem can be lessened through the appropriate design and use of contracts, disclosure rules, corporate governance, information intermediaries, and corporate control contests. Nonetheless, a variety of economic and institutional factors determine the effectiveness of these actions. The ability to write and enforce optimal contracts aligning the interests of management with those of equity and debt holders, the potential incentive problems for corporate boards and intermediaries, and the nature of the corporate control market are some examples. Additionally, financial analysts and rating agencies engaging in private information production to uncover any manager's misuse of firm resources are also of special importance. In conclusion, information and incentive problems impede the efficient allocation of resources, although required disclosures and the institutions created to ensure transparency may mitigate some of these problems.

During the 2008 crisis, information asymmetry between mortgage sellers, intermediaries, and final mortgage holders played a major role in the bubble formation process. This happened because mortgages sold to homeowners were resold to financial intermediaries who, after repackaging and charging their commissions, sold them once again to investment funds. Consequently, local mortgage issuers and financial intermediaries did not pay the cost of granting loans to high-risk holders. As a result, thousands of mortgages were sold to people who had no means to repay them and hordes of unaware investors were left footing the bill.

A large portion of the work analyzing the impact of intermediaries indicate that financial analysts add value to markets and play an important role in improving market efficiency. There is also evidence to suggest that at least some of the disclosures made by them, the business press, and the bond-rating agencies affect stock prices.[12] However, the fact that analysts' earnings forecasts and recommendations affect stock prices creates a whole new set of problems.[13] This is particularly the case when the same analysts are rewarded for providing information that generates trading volume and investment banking fees for their own brokerage houses. In these instances, analysts have incentives to make optimistic forecasts and to provide purchase recommendations. This fact is confirmed by empirical evidence showing analyst bias toward buy recommendations.[14, 15]

The management's reporting decisions that are related to voluntary disclosure and the role of financial reporting in capital markets are also of key importance in reducing information asymmetry.[16] Researchers discuss six forces that affect managers' disclosure decisions: capital market transactions, corporate control disputes, stock compensation, litigation, proprietary costs, and management talent signaling.

With respect to corporate control disputes, this is motivated by evidence that boards of directors and investors hold managers accountable for current stock performance. For instance, Chief Executive Officer (CEO) turnover is associated with bad stock performance.[17] In this respect, voluntary disclosure theory hypothesizes that managers use corporate disclosures to reduce the likelihood of losing their jobs.

Notwithstanding their salaries, managers are also rewarded with a variety of stock-based compensation plans, such as stock option grants and stock appreciation rights, which provide incentives to engage in voluntary disclosures. There are several reasons for this. First, managers who wish to trade their shares have to meet restrictions imposed by insider trading rules. In this respect, they would also want to increase the firm's stock liquidity, and correct any publicly perceived undervaluation before their options expire. Should there be no insider trading restrictions, managers could purchase shares to take advantage of the undervaluation of their stock rather than make disclosures to increase the value of their options.

Second, managers have incentives to provide voluntary disclosures to reduce contracting costs for new employees. If stock prices are not fair estimations of the firm's value, new management will demand additional compensation for bearing risks tied to the probability of misvaluation. Misvaluation results from information asymmetry between managers and investors and refers to the gap between the value of the firm, conditional on the managers' information set and the investors' information set.

Consistent with these ideas, the evidence shows that management forecasts are positively associated with insider trading.[18] For instance, firms delay disclosure of good news and accelerate the release of bad news prior to stock option award periods to increase their stock-based compensation.[19] Also, managers of firms in turnaround situations are more open to providing earnings forecasts if they hold higher stock options compensation at risk.

The extent to which voluntary disclosure mitigates resource misallocation depends on the credibility of the information released by the firm. Thus, given that managers have incentives to make self-serving voluntary disclosures, much of this evidence focuses on its accuracy and the effects of management forecasts. For instance, positive price reactions have been shown to follow management forecasts of earnings increases, while forecasts of decreases are followed by negative reactions.[20] Other findings demonstrate that the market's reaction to unexpected management earnings forecasts is similar in magnitude to its reaction to unexpected earnings announcements.[21]

A number of studies examine the economic consequences of voluntary disclosure. They argue that there are three potential types of capital market effects for firms that make extensive voluntary disclosures: improved liquidity for their stock in the capital market, reductions in their cost of capital, and increased following by financial analysts.

3.2.1. Improved stock liquidity

Institutional ownership increases with higher disclosure and higher stock liquidity. This may be due to the fact that voluntary disclosures reduce information asymmetry among informed and uninformed investors[22] as these affect the speed with which information is impounded in prices, increasing the stocks' liquidity.[23] In this respect, a significant negative relation has been found between analysts' ratings of firms' disclosures and bid-ask spreads.[24] Consequently, investors can be confident that transactions occur at "fair prices" in firms with high levels of disclosure.

3.2.2. Reduced cost of capital

When disclosure is imperfect, investors bear the risk of forecasting future payoffs, consequently requiring an additional premium. As a result, other things being equal, firms with high levels of disclosure and little information risk will have a lower cost of capital.[25] Some evidence in this respect includes (a) firms with a small analyst following have a negative relation between the cost of equity and the extent of their voluntary disclosures, (b) firms providing additional segment disclosures have a contemporaneous increase in the market's capitalization of their earnings, and (c) there is a negative cross-sectional relation between the cost of capital and analyst rankings of annual report disclosures.[26]

3.2.3. Increased information intermediation

The argument in this case is that if management's private information is not fully revealed through required disclosures, voluntary divulgence lowers the cost of information acquisition for analysts. In this respect, firms with more informative disclosures have larger analyst following, less dispersion in analyst forecasts, and less volatility in forecast revisions.[27]

Information asymmetry models assume that at least one party to a transaction has relevant information that the others do not possess. For instance, in adverse selection models, at least one party lacks information while negotiating a transaction. In a moral hazard situation, the missing information pertains to the performance of the agreed-upon transaction by one of the parties who may also lack the ability to retaliate in case of a breach of agreement.

In 1970, George Akerlof's *The Market for Lemons*[28] brought information issues to the forefront of economic theory and proposed two main remedies: signaling and screening. In a situation of information asymmetry, signaling suggests that people can transfer information to the other party to resolve the asymmetry by "signaling their type." A classic example given by Michael Spence[29] is that of an employer interested in hiring someone "skilled in learning." Because this ability is required, most applicants will claim to have it. Information asymmetry arises because the employer cannot tell which candidates are truly capable of learning. Spence's suggestion is that having a college degree works as a credible "signal" of this ability.

3.3. Self-fulfilling expectations and reflexivity

Sociologists William I. Thomas and Dorothy S. Thomas (1928) first articulated the principle of *reflexivity* by declaring:[30] "If men define situations as real, they are real in their consequences." This statement is known as the "Thomas Theorem." The existence of an objectively correct interpretation is not relevant to guiding individuals' behavior: the interpretation of a situation causes the action, but this interpretation is not necessarily objective. Hence, actions are affected by subjective perceptions of situations.

The Thomas Theorem was thereafter built upon by Robert K. Merton (1948, p. 195) to define the "self-fulfilling prophecy":[31] "The self-fulfilling prophecy is, in the beginning, a false definition of the situation evoking a new behavior which makes the original false conception come 'true'. This specious validity of the self-fulfilling prophecy perpetuates a reign of error. For the prophet will cite the actual course of events as proof that he was right from the very beginning." To sum it up: once a prediction has been announced, agents adapt their behavior to make the prediction come true. As a result of positive feedback between belief and behavior, an announcement that might otherwise have been false, is made true through its own prediction.

In economic science, *reflexivity* has been associated with "reflexive prediction,"[32, 33] proposed as an important aspect related to the Lucas Critique, and questioned as a methodological concern. The Lucas Critique claims it is naïve to forecast the effects of changes in economic policy on the basis of relationships observed in historical data.[34] Recently, economic philosopher and philanthropist George Soros has been an active promoter of the relevance of reflexivity to economics, first propounding it publicly in his 1987 book.[35] In 2008, in *The New Paradigm for Financial Markets*,[36] Soros presented a current role for reflexivity in the economic reality by observing how thinking and reality interact in that market prices affect the fundamentals they are supposed to reflect.

Within finance, reflexivity and self-fulfilling expectations can refer to self-reinforcing effects, such as feedback loops. Feedback is a process in which output or information about an event or phenomenon in the past or the present influences the same phenomenon in the present or the future. Feedback loops, mechanisms, and signals are present in reflexive relations. For instance, rising prices attract buyers whose purchase orders drive prices even higher. The opposite could also be true, with lowering prices attracting sellers whose actions depress prices even further. Both selling during market declines and buying during market advances are called "positive feedback trading," an investment rule used by market participants. There is ample evidence for this in developed stock markets. Positive feedback trading increases volatility and can provoke autocorrelation in stock returns. Furthermore, if a large number of traders engage in positive feedback trading strategies, asset prices will start to diverge increasingly and persistently from fundamental values.

Positive feedback generated by optimistic investors pushing prices into bubble territory has links with the selling waves generated by pessimistic investors. The competition between imitation and contrarian behavior, and between value investors and technical analysts, or even the main mechanisms leading to positive feedbacks, provide fuel for the development of speculative bubbles, creating instability. These patterns have been documented in all crashes.

Positive feedback leads to collective behavior, such as herding during sales that take place during a financial crash. This collective behavior does not require the coordination of people, but results from the convergence of interests and interactions between them, made possible through various networks. Complex system theory argues that such collective behavior may be robust in spite of external intervention as long as the "selfish" individualistic nature of the utility function dominates.

However, beyond some obvious relations, reflexivity is a far-reaching phenomenon in finance with wide-ranging implications for financial markets, corporate finance, financial regulation, central bank and public finance, behavioral finance, interest rates and bond markets, contagion risks, networks, and system stability, among others. Reflexivity and reflexive relations in finance are also important because they introduce nonlinear relationships, obscuring the role of micro/macro variables and the mechanisms of cause and effect in financial relations. Furthermore, reflexivity presents an interesting problem because a prediction leads to changes in the system about which the prediction is being made. Accordingly, when assessing scientific hypotheses, one might be confronted by the question whether observed events are themselves the results of the prediction which, by its own formulation, contaminates the experiment. Reflexivity complicates all three of the traditional roles played by a classical science: explanation, prediction, and control.[37]

Reflexive relationships affect asset valuation processes and decision-making (and thus the economic lives of nations) through multiple processes. Some of these processes are linked to emotional decision-making, trading rules, and group imitative behavior, resulting in financial contagion. Others are related to its impact on financial statements of corporations and countries. In either case, when undisclosed, these relationships can contribute to macroeconomic instability and uncertainty, with serious implications for sustainability. Furthermore, during bubble formation, these effects can contribute to shifts in income and wealth distribution. Consequently, these mechanisms and their implications for market stability should be analyzed.

3.3.1. Reflexive relations

The essence of reflexivity as applied to finance is that a belief can affect reality and be made real simply through its formulation. This can lead to feedback loops between thoughts and reality. The term "reflexivity" has been used in other areas with analogous meanings. For instance, in sociology, it is defined as "an act of self-reference where examination or action 'bends back on,' refers to and affects the entity instigating the action or examination" (i.e. a circular relationship between cause and effect). From an epistemological point of view, "it is considered to occur when the theory being formulated or disseminated affects the behavior of the individuals or systems the theory is meant to be objectively modeling."[38]

In economics, the literature on reflexivity focuses mainly on the validity of economic theories under reflexivity from an epistemological view.[39] Contributors across different areas develop equivalent arguments. Of special interest is the work by Davis and Klaes (2003),[40] who define three levels of reflexivity: "immanent" (between agents and economic events), "epistemic" (between economists and their theories), and "transcendent" (between social forces and economists).

A second niche centers on reflexivity in predictability.[41] Soros (1987, 2008)[42, 43] addresses the rational expectations equilibrium (REE) and the efficient markets theory (EMT)[44] and several lines of research have been developed based on these works. For instance, Calandro (2004)[45] proposes an explanation for business cycles based on Soros's ideas, while Sandri (2009)[46] devotes the chapter "Reflexivity of Finance" to evaluating his proposals. Besides these direct analyses, several papers on agent-based simulation of financial markets cite Soros's thoughts as a source of inspiration.[47]

Within financial economics, additional frontier work includes modeling feedback effects from stock prices to fundamentals. Research has shown how small changes in stock prices can trigger permanent changes in fundamental corporate values[48] and how the information content of stock prices affects CF when managers make real investment decisions, even when these are random movements.[49, 50, 51, 52] An equivalent effect can also be found in lenders' decisions.[53] Ozdenoren and Yuan (2008)[54] develop an asset price formation model that incorporates the effects of feedback from prices into fundamentals. The implications of the model are higher excess volatility, self-fulfilling beliefs, and multiple equilibria.

Angeletos et al. (2008)[55] provide a comprehensive analysis of the two-way causation between investment and asset prices. The ideas of "self-reinforcing leverage cycles" and "first and second round effects of the interaction between real sector and the banking system" have been discussed,[56] with a number of papers focusing on the two-way interaction between credit and property prices[57] and bank lending and macroeconomic activity.[58, 59]

Uribe and Yue (2005)[60] also document two-way causation between country spreads and macroeconomic activity in emerging market economies. Studies on developed economies usually find causality running from property prices to bank lending,[61] a typical example being one where market prices affect real economic activity, while studies on emerging markets find causality running from bank lending to property prices.

An understanding gained from all these works is that feedback effects from asset prices to fundamentals enhance short-lag positive and long-lag negative autocorrelation in asset returns. Correlation will be analyzed at length in subsequent chapters, including later sections of this one and Chapters 5, 6, and 7; however, for our purposes here it refers to statistical relationships involving dependence. Autocorrelation, also called serial correlation, refers to the cross-correlation of a signal with itself. Informally, it is the similarity between observations as a function of the time lag between them. One example is the relationship between the change in price in the current period and that of the prior period.

Behavioral finance theories also attempt to explain these feedback effects and their impact on the autocorrelation of returns.[62] However, these theories are based on cognitive and behavioral biases in human decision-making, and hence not proof against the argument that sophisticated investors can learn and modify their behavior and that the documented behavioral biases are only a temporary phenomenon and not laws of nature. Here, reflexivity can complement behavioral biases as the profitability of momentum strategies can be explained by a combination of feedback effects and behavioral biases, rather than behavioral finance theories alone.[63] Ultimately, reflexivity complements behavioral theories to make an important modification to the EMT.

Examples of self-fulfilling hypotheses and related reflexivity effects also abound in studies of cognitive dissonance theory and self-perception theory: people are often seen to change their attitudes in order to align them with what they express publicly. The theory of cognitive dissonance argues that individuals have a motivational compulsion to reduce dissonance by changing or deducing the relevance of existing cognitions to form a consistent belief system. Two key assumptions here are that individuals need to maintain equilibrium where their expectations meet reality and that they will avoid situations and information sources that result in feelings of discomfort or dissonance.[64] However, during a period of mental stress, they suffer a dichotomy between things they do and their individually held beliefs; thus, individuals will ignore information that challenges their pre-existing beliefs and/or they will change their beliefs in order to justify their actions.[65] This is highly relevant to the case when traders are commanded to follow purchase and sale orders as part of company strategy. According to Festinger,[66] people engage in "dissonance reduction," a process that lowers the relevance of discordant sources, adds consonant elements, or changes dissonant factors.[67]

3.3.2. Feedback, collective behaviors, and herding

Research in the field of complex systems advocates that both the economy and stock markets function under the opposing tensions of positive and negative feedback mechanisms. When positive feedback forces prevail, divergence from equilibrium can lead to crises. These instabilities are inherent to the dynamics of such an organizational system and thus, under the theory of collective behavior, speculative phases resulting in both growing bubbles and decreasing prices can coexist.[68]

In accordance with this theory, large market crashes are analogues to the critical points the physics community studies in relation to the phase transformation of solids, liquids, and gas. Underlying this theory is the existence of a cooperative behavior of agents acting alike and progressively escalating the build-up of market cooperativity, or their effectual exchanges which often result in price increases for long periods before the bubble implosion. The core notions of these mechanisms that lead to the development of endogenous instabilities are imitation, herding, self-organized cooperativity, and positive feedbacks.

3.4. Executive compensation and bad incentives

When herd behavior prevails, investors tend to buy or sell in the direction of the market trend. This mechanism is sometimes helped by technical analysts who work to detect these tendencies in order to respond to them, thus contributing to the creation of a self-fulfilling prophecy.

As mentioned, investment managers are compensated and retained due, partly, to their performance relative to peers. Thus, the typical short-term focus of many equity markets exacerbates the risk for investment managers who do not participate in the building phase of a bubble, particularly one that builds over a longer period of time. Taking a conservative or contrarian position as a bubble builds will result in poor performance and may cause clients to go elsewhere, thus affecting the manager's own employment and/or compensation level. Consequently, in attempting to maximize returns for clients and maintaining their own employment, managers may rationally participate in a bubble they believe to be forming, as the risks of not doing so outweigh any potential benefits for themselves.

Thus, another form of self-collective behavior is the general agreement of a class of individuals that results in a pattern of executive compensation. This pattern has been alleged to have played a major role in the housing bubble (Chapter 1). The observation that, for years, executives of publicly held companies have been sacrificing long-term economic stability for short-term financial goals[69] supports this claim.

One mechanism for corporate governance is a compensation package that aligns the interests of shareholders and management. In addition to salary, management compensation typically includes bonus, stocks, options, severance packages, and performance-based termination. The incentives can also lead to value destruction. The *bad incentives* view propounds that the 2008 crisis could have been avoided had there been appropriately designed incentives in place.

For instance, excessive focus on short-term objectives can drive executives to pass up promising long-term investments. Also, when shareholders of levered organizations gain from disproportionate risk-taking, compensating executives with stock or options might result in management taking additional risks. In addition, if the firm is overvalued, stock-based compensation may lead managers to overinvest or manipulate earnings to justify the firm's current stock price. These effects have been identified as *short-termist* managerial behavior.

There is abundant evidence for the distortions that stock options create in management decision-making and the disconnect between performance-based compensation and actual value added by high-ranking executives. This has been a matter of discussion for decades. A recent curious example is that of Bear Stearns' CEO James Cayne, who in 2006, at the height of the housing bubble, received compensation amounting to more than $40 million.[70] This amount included a $17 million bonus, approximately $15 million in restricted stock awards, and $1.6 million in stock options. The firm's rationale for such an extraordinary sum was elucidated in a proxy statement dated March 27, 2007: "[t]he Company's performance as measured by profit margins remained strong and earnings per share increased over the prior year. In addition, return on common equity was among the highest of the Company's key competitors. The compensation paid to the Company's executive officers for fiscal 2006 reflects the strength of this performance."[71] There is no need to remind everyone that a year later Bear Stearns was sold to JPMorgan, making the obvious even more obvious: the stratospheric performance-based bonuses granted were totally unrelated to any long-term measure of value added to the corporation.

In 2007, when the effects of the mortgage crisis were already evident, Wall Street bonuses came to $33.2 billion.[72] For example, Goldman Sachs's Chief Executive, Lloyd C. Blankfein, received a $67.9 million bonus. Furthermore, the 2007 bonuses for seven of Wall Street's top firms (Merrill Lynch, Citigroup, Bear Stearns, Morgan Stanley, JPMorgan Chase, Lehman Brothers, and Goldman Sachs) were up 10 percent from the year before, leading to a combined total of $122 billion. This happened despite these companies experiencing in the same year mortgage-related losses of $55 billion and a collective net revenue fall of 6 percent, resulting in a $200 billion loss in shareholder value over the preceding year. Employee compensation at those firms was equal to 47 percent of net revenue in 2007, compared to 40 percent the year before.[73]

Examples of outrageous compensation packages are easy to find, even outside of the bubble scenario. Overall, the average total CEO pay for S&P 500 firms has increased substantially over that received by nonexecutive employees. Published figures report $850,000 in 1970 up to $10.5 million in 2008 after the crisis, peaking at $14 million in 2000. While CEO pay increased by an average 34 percent, the salaries of other employees did not follow the same trend. Thus, during the same years, the salary of top management went from 40 times that of the rest of the employees to 320 times, peaking at 400 in 2000.[74]

With respect to compensation, an essential perspective is presented by the idea of *agency costs* which center on the conflicts of interest arising among various corporate stakeholders. For instance, Jensen and Meckling (1976)[75] show how the incentive structure of top management with less than 100 percent firm ownership can drive actions that diminish corporate value. The reason for this is that a manager with, let us say, a 2 percent stake would get 100 percent of the benefits and incur only 2 percent of the costs from consuming a dollar of perks. The sensitivity of the manager's wealth to that of the shareholders has been proposed as an index of the degree of alignment achieved by the compensation structure. This widespread view results in the extensive use of employee stock and option plans.[76, 77]

In this respect, Jensen and Murphy (1990)[78] estimated that for the period 1974–1986, top executive pay increased by about $3.25 for every $1,000 increase in shareholder wealth. This pay–performance sensitivity measure led them to conclude that executive pay was not sensitive to shareholder wealth.[79] However, a series of recent studies suggest that, in time, compensation has grown to align more closely with performance. For instance, in Hall and Liebman's 1998 study,[80] 95 percent of the 1996 pay–performance sensitivity for CEOs in the manufacturing sector results from changes in the value of *existing* grants of stock options and stock.

Option grants have also become the norm for executives below the top tier, rising to 90 percent by 2004 from 85 percent in the mid-1990s.[81, 82, 83] Fahlenbrach and Stulz's (2009)[84] findings from a sampling of bank CEOs show that their wealth increases by an approximate $24 for every $1,000 of shareholder value created,[85] a very significant change from the earlier estimates reported by Jensen and Murphy.[86]

However, correlation and causation are two different things. The upward trend, mainly in the nonsalary portion of the CEO compensation packages, has resulted in a larger correlation between total pay and firm performance. This fact goes along with the basic tenants of agency theory. Nonetheless, it is also correlated with the above-mentioned undesirable collateral damage resulting from short-termist behavior, which has been listed as a major contributor to the bubble creation process. Stock options have been viewed as particularly culpable in this regard as they can be blamed for the creation of incentives leading to the manipulation of corporate statements to "artificially" increase stock prices, thereby contributing to the scandals of the post-dot-com era and the housing crisis.

The 2008 report by the Research and Development Committee of the Committee for Economic Development (CED) attests to this idea. It states: "[d]ecision making based primarily on short-term considerations damages the ability of public companies, and therefore, of the USA economy to sustain superior long-term performance."[87] The report further argued that the focus on short-term results was largely the product of compensation packages which correlated bonuses and perks with the meeting of short-term objectives. This problem is compounded by the high turnover of executives as a consequence of using quarterly results as a measure of performance, which links decision-making and the avoidance of the decision's long-term consequences. As per the CED panel, other collateral consequences of these compensation policies are "prominent shortfalls of ethical performance [. . .] distortion of short-term financial results and speculative trading at least partly caused by pressure their managers felt to meet the financial market's quarterly earnings expectations."[88]

3.5. Speculative trading

The US Commodity Futures Trading Commission (CFTC) definition of a speculator is "a trader who does not hedge, but who trades with the objective of achieving profits through the successful anticipation of price movements."[89] Speculators try to benefit from short- or medium-term changes in the market price of assets such as stocks, bonds, and commodities by engaging in risky financial transactions. Thus, rather than focusing on the fundamental aspects of financial instruments – such as dividends or interest – they concentrate on following the assets' price movements.

Along with investors, hedgers, and arbitrageurs, speculators play a key role as they provide liquidity by trading when other participants are not available, absorbing risks other agents reject, and helping improve market efficiency overall. For instance, speculators in competition with other speculators enlarge the market and reduce spreads between bid and ask prices. Also, by taking futures contracts and their corresponding risks, they are thought to impact the production ability of others.

Though positive aspects characterize speculators' roles, excessive speculation can be harmful to the proper functioning of futures markets.[90] For example, when larger numbers of speculators participate, the real, underlying demand and supply can become diminishingly small as compared to trading volume, and prices can be distorted. Also, speculation is tied to financial bubbles, even if not all bubbles are caused by speculation.[91] Additionally, the presence of speculators contributing to short-term volatility is also a matter of debate.

The US CFTC has proposed a regulation to limit speculation in futures markets. The three basic elements of the regulatory framework refer to the "the size (or levels) of the limits themselves, the exemptions from the limits (for example, hedged positions) and the policy on aggregating accounts for purposes of applying the limits."[92] The proposed limits apply to 28 physical commodities traded in various exchanges across the USA.[93] A second aspect, the Volcker Rule, passed on January 21, 2010, has dealt with speculative investments of banks that are not geared to benefit their customers. The Volcker Rule states that these investments played a key role in the financial crisis of 2008.[94]

3.6. Aspects of behavioral finance

In the following pages, we review some aspects of behavioral economics relevant to the analysis of bubble formation/implosion and contagion mechanisms. Recognizing the main concepts used in behavioral economics is the first step to identifying behavioral insights in practice. Therefore, we present the key biases and ideas germane to analyzing problems in financial markets.

3.6.1. Rational choice and bounded rationality

Economics is "behavioral" in the sense that it makes predictions about the way people behave. In a hypothetical "ideal world," such as the one described by Gary S. Becker,[95] decisions would result from the weighing of costs and benefits and be informed by existing preferences. Consequently, consumers would consistently make optimal decisions. Becker's ideas have come to be accepted as the foundation of "rational choice theory," a theory that infers that human actors have stable preferences and engage in maximizing behavior. Ever since this work was presented, rational choice theory has been the postulate accepted by most mainstream economists.

Nonetheless, and regardless of the merits of this theory, models of behavior used by standard economists do not explain everything. For instance, according to the preference approach in standard normative economics, what people choose is the same as what maximizes their utility. Under this assumption, the appropriate policy is to provide as many options as possible, that is, to increase the budget allowance. However, this remedy will not work if agents make mistakes in their statistical reasoning or repeatedly make bad choices. Hence, behavioral economics complements standard economics in providing models to answer questions about the way the economy works both from positive (descriptive role) and normative (prescriptive function) points of view.

In this context, Herbert Simon's works (1955, 1979) are of particular relevance, given that they advanced the idea of "bounded rationality" as an alternative in modeling decision-making.[96, 97] Bounded rationality argues that in decision-making, the rationality of agents is bounded by their information set, their cognitive limitations, and the time available for making decisions. Hence, "bounded rationality" complements "rationality" which views decision-making as the rational process of finding an optimal choice given the available information set. In essence, bounded rationality states that given the limited capabilities and resources decision-makers have to reach optimal solutions, they simplify their choices before applying "rationality." Thus, decision-makers seek a *satisfactory* solution rather than an *optimal* one.[98]

3.6.2. A dual-system theory: intuition and accessibility

For years, Tversky and Kahneman (1971) observed persistent discrepancies between statistical intuition and statistical knowledge in sophisticated researchers. The fact that these mistakes led to systematic errors of judgment motivated the authors to create a system to differentiate between intuition and reasoning.[99]

Stanovich and West (2000) identified two types of cognitive processes that characterize operations: *System 1* and *System 2*. System 1 is fast, automatic, effortless, associative, implicit, emotionally charged, governed by habit, and difficult to control or modify. These effortless processes do not cause or suffer great interferences when combined with other tasks. On the other hand, System 2 processes are slower, serial, effortful, consciously monitored, and deliberately controlled. They are relatively flexible and potentially governed by rules.[100] Given that capacity for mental effort is limited, effortful processes tend to disrupt each other.[101, 102, 103]

Intuition allows us to make impressions and automatic judgments effortlessly, thereby helping us perform complicated functions quickly and accurately. Reasoning is much slower, and must be deliberately controlled (see Table 3.1[104]).

We use both systems in decision-making. However, most choices made in daily life are the result of these intuitive, fast, automatic processes. The problem arises from the fact that though intuition is efficient in performing many complicated tasks, it is not always correct. Biases arise and intuitive processes can lead people astray in systematic and predictable ways. For instance, intuition can automatically replace complex questions our brain cannot respond

Table 3.1 Dual-system theory, two modes of thought

Perception and process	
Intuition	**Reasoning**
Fast, swift	Slow
Parallel	Serial, sequential
Automatic, instinctive	Controlled, consciously monitored
Effortless	Effortful
Associative	Rule governed
Slow-learning	Flexible
Emotional	Neutral
Latent, driven by routine	Conscious
Hard to restraint or change	Willfully controlled

to quickly with easier questions it can answer. We provide an example of such a mechanism in Table 3.2. Even though the answers to the two sets of questions are only weakly related, we might not be aware of the subconscious replacement process and hence rarely check whether we are, in fact, correct.

As Tversky and Kahneman observed, most people cannot recognize their biases and thus follow their mistaken intuitions. This includes people acquainted with various types of biases, such as professional investors. Two well-known biases in this sector are over-extrapolation from limited information regarding past stock returns and *overconfidence* which drives investors to overestimate future returns.

Intuitive judgments occupy a position between the automatic operations of perception and the deliberate mechanisms of reasoning. A central feature of intuitive thought is that they come to mind instinctively. *Accessibility* refers to the ease with which specific mental contents come to us, and is determined by the characteristics of the cognitive mechanisms that produce a thought as well as the events that evoke it. The determinants of accessibility include stimulus salience, selective attention, specific training, associative activation, and priming.

What becomes "accessible" in any particular situation is determined by the actual properties of the object of judgment. For instance, physical *salience* determines accessibility. If we see a street sign with a red uppercase letter and a black lowercase letter, thanks to *salience*, the red comes to mind first. However, this can be corrected by instructing the observers to search for the lowercase letter, thus enhancing the accessibility of its features. Thus, specific training and selective attention can help in the masking or highlighting of salient features.

These effects also apply to more abstract stimuli. For instance, "Real Madrid defeated Barcelona" and "Barcelona lost to Real Madrid" provide the same information. Nonetheless, the two versions of the same information set make accessible different thoughts. Accessibility also spurs temporary states of *priming* and *associative activation*. Mentioning a place of birth, such as a city or country, temporarily increases the "accessibility" of the traits associated with the stereotype.[105, 106] Furthermore, *priming* shows that exposure to a "stimulus" influences a response to a later stimulus. If a person reads a list of nouns, including the word *lamp*, and she is later asked to complete a word starting

Table 3.2 Question substitution by intuition

Intended question	**Replacement question**
Can I trust this adviser?	Is she nice?
What is the probability that a bad event will occur?	Have I ever experienced such an event in the past?
How should fraud be penalized?	I hate people who commit fraud.

with "lam," the probability that she will use "lamp" is greater than if she had not been "primed." The effects of priming can be very salient and long-lasting. Knowledge of these biases is often used in marketing.

The effect of context as related to accessibility is also important. An ambiguous stimulus that is perceived as a letter in the context of letters can be seen as a number in the context of numbers. Just think of the digit one and "l" or capital "I." If these are shown in close proximity to each other, the reader will become aware of the ambiguity. However, if they are presented separately, this conundrum will not be obvious, as observers would not become aware of alternative interpretations at the same time. They "see" the interpretation that is the most likely in "a" context, but generally have no awareness that the same cues could be understood differently in different scenarios.

The central finding in studies of intuitive decisions is that experienced decision-makers who work under pressure, such as a policeman in the middle of a chase or a trader in the midst of a debacle, rarely choose between options. This is because in most cases, only one option comes to mind.[107] Thus, uncertainty is hardly represented in intuition and perception. The options that were rejected are not represented. Hence, doubt is a phenomenon in System 2.

3.6.3. Framing effects

The assumption that preferences are not affected by variations of the irrelevant features of options or outcomes (*extensionality* and *invariance*) is an essential aspect of the concept of rationality in classical economic theory. *Extensionality* refers to principles that judge objects to be equal if they have the same external properties. It stands in contrast to the concept of *intensionality*, concerned with whether the internal definitions of objects are the same.[108, 109] However, *invariance* is violated in demonstrations of *framing effects*.

The *framing effect* is a cognitive bias in which people react differently to a particular choice depending on whether it is presented as a loss or a gain. Framing is an inevitable process of selective influence over the individual's perception of the meanings attributed to words or phrases which affect both decision-making and problem-solving.[110] For instance, people tend to avoid risk when a positive frame is presented, but seek risk when a negative frame is presented.

In finance, an investor is said to suffer from *narrow framing* if he is making investment decisions without considering the context of his total portfolio. Narrow framing occurs when people who are offered a new gamble evaluate it in isolation, and separately from their other risks. That is, they ignore the gamble as just one of many that determine their overall wealth risk, and act as if they get utility *directly* from the outcome of the gamble. This behavior is at odds with traditional utility functions, under which the agent only gets utility from the outcome of a new gamble *indirectly*, via its contribution to his total wealth. The foundation of Prospect Theory,[111] developed further below, is that the carriers of utility are variations in wealth rather than asset positions. This entails that decisions made by agents account for gains and losses rather than the final states.

Thus, one unrealistic inference of the rational agent model is that investors make choices considering all aspects of their current situations, plus those of future opportunities and risks. On the contrary, evidence suggests that the agent's thoughts are contextualized by *narrow framing*[112] and *mental accounting*.[113, 114]

A second unreasonable framing is to think of problems as isolated instances of decision-making, each worthy of a separate decision. In reality, decision-makers handle many problems at once and one decision is not necessarily made independent of another. Furthermore, decision-makers expect to live long enough to make many other decisions about accepting gambles. On the other hand, a restrictive concern with the long-term or the broad view may disregard the fact that life is lived in the present and that changes trigger emotions. A theory of choice that ignores fear, pain of losses, and regret is impractical and results in prescriptions that do not maximize the utility of outcomes as and when they are felt.[115, 116]

3.6.4. Attribute substitution: a model of judgment by heuristic principles

In order to reduce the complexity of decision-making when evaluating probabilities and assessing values, agents rely on certain heuristic principles. These heuristics are very handy but can also lead to systematic errors.[117]

Tversky and Kahneman's (1974) work identified three heuristics of judgment: *representativeness*, *availability*, and *anchoring*. In the same article, the authors presented a list of observed systematic biases, including nonregressive prediction, neglect of base-rate information, overconfidence, and overestimation of the frequency of events that are easily recalled. Some of these biases were recognized by systematic errors in estimates of known quantities and statistical facts, while others resulted from systematic discrepancies between the regularities of intuitive judgments and the principles of probability theory, Bayesian inference, or regression analysis. This 1974 work launched a new branch of literature looking into the so-called heuristics and biases approach to the study of intuitive judgment[118, 119] and has been the source of much controversy.

In financial markets, extrapolation refers to projecting historical data into the future. More specifically, it refers to the general expectation that prices which have risen at a certain rate in the past will keep rising at the same rate in the future. The contention is that investors tend to extrapolate past extraordinary returns into the future, overbidding in order to obtain those same rates. In addition to extrapolation, over-optimism also results in overbidding.

During the 2008 crisis, Wall Street employees sold securities backed by dubious-quality mortgage loans and took massive housing market risks for their firms. This was because they used bad models and extrapolated past growth of home prices;[120] psychological biases and cognitive dissonance led them to ignore risk and warning signs;[121] and optimistic shareholders used short-term stock-price-based compensation to select and motivate managers.[122] According to this "bad models" view, distorted beliefs and over-optimism on Wall Street resulted in individuals, even those properly incentivized, failing to anticipate the housing market crash.

3.6.5. Changes or states: standard utility versus Prospect Theory

As mentioned at the beginning of the chapter, traditional normative theories of choice in economics typically assume that people are rational and that they have clear and stable preferences based on some *utility* function. A utility function specifies the wellbeing of a consumer for all combinations of goods consumed, representing both the person's welfare as well as her preferences.

Choice theory assumes *description invariance* (the manner in which the information is presented to the decision-maker should not change the decision) and *procedure invariance* (the method of eliciting preference should not change the decision). Accordingly, if X, a hypothetical decision-maker, were to be presented a choice, she would be guided by her utility function, since the choice would be presented devoid of background.[123, 124]

However, real-world decisions bring other variables into play. For instance, X might consider the worst-case scenario (earning nothing) as a yardstick and treat all other outcomes as an improvement on this benchmark. Alternatively, she could treat the best fallout as a benchmark and everything else as an unsatisfactory outcome. Then again, she might view this choice as a one shot or as a choice within a series of decisions she needs to make, or she might undertake the task of deciding which of the two options to select, or to reject.

The use of different *frames* results in different choices. By *framing*, that is, contextualizing the information in part from extrinsic manipulation of the options offered and in part from compulsions intrinsic to the decision-makers, such as norms, habits, and temperament, people can process data and reduce ambiguity. The choices people make are influenced by the creation of a frame: different representations of fallouts highlight some features of the situation and mask others. Thus, *framing* can affect the outcome of choice problems to the extent that several of the classic precepts of rational choice do not hold true.[125]

Because of lack of a system that generates appropriate canonical representations, intuitive decisions are shaped by the factors that determine the accessibility of different features of the situation. Consequently, highly accessible features influence decisions, whereas those with low accessibility are usually ignored. However, the most accessible features may not be the ones most relevant to making a good decision.

In economics, the standard model of utility is *reference independent* in that it assumes the utility assigned to a given state of wealth is *invariant* to the decision-maker's initial state of wealth. This assumption conflicts with the principle of *perception*, where the effectual stimulus is not the new level but the difference between the new and the existing one. That is, one would care about gains and losses rather than the final states of wealth, and so it is "reference dependent," the reference being the status prior to effecting the change by selecting an option.

The expected utility hypothesis used in economics, game theory, and decision theory is one in which the "betting preferences" of people with regard to uncertain outcomes are represented by a function of payouts, the probabilities of occurrence, risk aversion, and the different utility of the same payout to people with unequal personal preferences. Daniel Bernoulli initiated this theory in 1738, prompted by his uncle's description of the St. Petersburg paradox.

The St. Petersburg paradox put forth by Nicolas Bernoulli arises when there is no upper limit to the potential rewards of events with very low probability. Because some probability distribution functions have an infinite expected value, a person interested in maximizing expected wealth would pay an infinite amount to take this gamble. Given that people cannot behave this way in real life, Daniel Bernoulli hypothesized diminishing marginal utility of increasingly larger amounts of money. That is, the expected utility of the gamble is finite, even if its expected value is infinite.

The von Neumann–Morgenstern utility theorem (commonly referred to as VNM) provides the necessary and sufficient "rationality" axioms under which the expected utility hypothesis holds.[126] John von Neumann and Oskar Morgenstern proved that an agent is VNM-rational if, and only if, there exists a real-value function u defined on possible outcomes such that every preference of the agent is characterized by maximizing the expected value of u, which can then be defined as the agent's VNM-utility. VNM-utility is a decision utility in that it is used to describe decision preferences.

Let us consider an agent who is given two options with mutually exclusive outcomes and probabilities adding to one: $L = 0.20A + 0.80B$. That is, either A will happen or B will happen. In this case, the probability of A happening is $P(A) = 20$ percent, whereas the probability of B occurring is $P(B) = 1 - A$. This idea is extended, as the outcomes of a lottery can represent the outcomes of other lotteries. For instance, $0.75(0.6A + 0.4B) + 0.25C = 0.45A + 0.30B + 0.25C$. For a lottery with many possible outcomes (A_i) we can write:

$$L = \sum p_i A_i, \tag{3.1}$$

where the sum of the p_i's = 1.

The four axioms of VNM-rationality are *completeness*, *transitivity*, *continuity*, and *independence*. Completeness assumes that an individual has well-defined preferences. Transitivity presupposes that preference is consistent among any three lotteries. Continuity tells us that the preference relation is continuous and small changes in probabilities do not affect the ordering between two lotteries. Continuity presumes that there is a "critical point" between being *better than* and *worse than* a given middle option. Finally, independence of irrelevant alternatives proposes that a preference holds irrespective of the chance of a different fallout.

Any agent attempting to maximize the expectation of a function u will fulfill all axioms. Thus, for any VNM-rational agent (i.e. one satisfying all axioms), there exists a function u assigning to each fallout A a real number $u(A)$ such that

$$L \prec M \quad \text{iff} \quad E(u(L)) < E(u(M)), \tag{3.2}$$

where $\prec$ indicates preference and $E(u(L))$ or $E(u(M))$ the expected value of u in L or M lottery.

$$Eu(p_1A_1 + \cdots + p_nA_n) = p_1u(A_1) + \cdots + p_nu(A_n). \tag{3.3}$$

Not long after the formulation of this theory (1947),[127] questions were raised regarding its value as a descriptive model.[128] In their critique of the theory, Kahneman and Tversky (1979, 1981)[129] addressed some of these issues and built a parsimonious theory to fit a number of violations of classical rationality. Their work, *Prospect Theory*, deals with the way in which we frame decisions, the different ways we label – or code – outcomes, and how they affect our attitude toward risk.

In their 1979 paper, Kahneman and Tversky[130] explained how choices among risky alternatives show effects inconsistent with the creeds of utility theory. For instance, people underplay outcomes that are simply probable in contrast to outcomes that are certain, and discard components that are shared by all prospects under consideration. These *certainty* and *isolation effects* contribute to risk aversion and lead to inconsistent preferences when the same choices are presented in different scenarios.

"Prospect Theory," the alternative to the rational choice presented by the same authors, replaced the notion of "utility" with that of "value" assigned to gains and losses, rather than to final states of wealth. That is, whereas utility is typically defined in terms of net wealth, value is defined in terms of deviations from a reference point.

Prospect Theory is based on the notion that people evaluate gains or losses from "the reference," some neutral point or status quo. This assumption is consistent with the adaptation-level findings that occur in real experience. Unlike Expected Utility Theory, which concerns itself with how decisions made in uncertainty should be handled (a prescriptive approach), Prospect Theory concerns itself with how decisions are actually made (a descriptive approach). Also, unlike Expected Utility Theory, Prospect Theory predicts that preferences depend on how a problem is framed. If the reference point is defined such that an outcome is understood as a gain, the resulting value function will be concave and decision-makers will tend to be risk-averse. However, if the reference point is defined such that an outcome is viewed as a loss, then the value function will be convex and decision-makers will be risk-seeking.[131]

Prospect Theory and Expected Utility Theory deal differently with the probabilities attached to specific outcomes. Whereas classical utility theory presumes decision-makers value a 60 percent chance of winning as precisely a 60 percent chance of winning, Prospect Theory manages preferences as a function of "decision weights" where these weights do not necessarily correspond to probabilities. That is, a probability of 60 percent can be assigned a different weight. In particular, Prospect Theory proposes that decision weights tend to overweight small probabilities and underweight moderate and high probabilities. In addition, the value function for losses is different from the value function for gains. The first (the curve lying below the horizontal axis in Figure 3.1) is convex and relatively steep. But in the second, the value function for gains (above the horizontal axis) is less steep and concave. These differences lead to several relevant results. For instance, given that the value function for losses is steeper than that for gains, losses cause more hurt than gains bring happiness, and thus a loss of \$600 is felt more than a gain of \$600.[132]

In 1992, Tversky and Kahneman[133] extended their earlier work to create a new version of Prospect Theory that employs cumulative rather than separable decision weights. This "cumulative Prospect Theory" applies to uncertain as well as to risky prospects with any number of outcomes, and allows different weighting functions for gains and losses.

According to Prospect Theory, people make choices with respect to a reference point to which the outcome can be compared. Thus, choices depend on the reference point rather than on the material outcome of the choice alone, and options are judged considering whether they represent gains or losses relative to that point. Hence, rather than modeling $C(A)$, choices made from a set A, one would model $C(A; x)$, the choices made from set A considering reference point x.

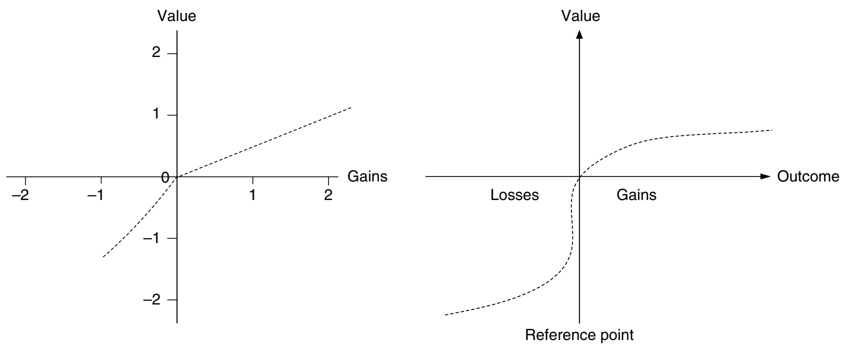

Figure 3.1 The value function of Prospect Theory

We present a simple example of how cumulative Prospect Theory replaces the *u* utility function in the cumulative probability weighting model with a *v* value function. That is, whereas *u* maps outcomes into utilities, *v* maps the changes in the fallouts into utilities. For instance, let us assume someone with a total net worth of $1,000 is offered a lottery *p* which would leave him with a 50 percent chance of losing $900 and a 50 percent chance of winning $1,100. The cumulative probability weighting model would rank this lottery using the function

$$U(p) = \psi(0.5)\,u(1{,}100) + (1 - \psi(0.5))\,u(900) \tag{3.4}$$

while cumulative Prospect Theory would rank it using the function

$$U'(p) = \psi(0.5)v(100) + (1 - \psi(0.5))v(-100) \tag{3.5}$$

Kahneman and Tversky reported the following main properties[134] of the function *v*:[135]

a. Individuals regard monetary outcomes as a function of variations from a reference level (the status quo of the decision-maker) and the values of the outcomes have diminishing returns. That is, once there is a gain of *x*, the impact of an additional gain of $1 falls as *x* increases. The same goes for losses. The impact of an additional loss of $1 on top of a loss of *y* falls as *y* increases. Thus, a central aspect of this theory is that the law of diminishing returns applies to both good and bad choices with economic impact.
b. The value function is steeper for losses than for gains. This entails loss aversion: equally sized gains and losses have asymmetric repercussions on the decision, with losses being about twice as painful as gains are gratifying. The impact of the gain of an amount *x* is less than the impact of a loss of *x*.
c. The curve is concave for gains and convex for losses, implying that decision-makers will be risk-averse when choosing between gains and risk-seeking when choosing between losses.

Thus, this process has three major characteristics: reference level dependence, gain and loss satiation, and loss aversion.

The shape of the value function defined on gains and losses describes the predictions of Prospect Theory and has the following characteristics: (a) it is concave in the domain of gains, favoring risk aversion; (b) it is convex in the domain of losses, favoring risk-seeking; (c) it is kinked at the reference point and *loss averse* – that is, it is steeper for losses than for gains by a factor of about 2–2.5;[136, 137] and (d) the functions in the two domains are approximated by power functions with similar exponents, both less than unity.[138] This means that people are risk-seeking over low-probability gains, risk-averse over high-probability gains, risk-averse over low-probability losses, and risk-seeking over high-probability losses. For these reasons, the probability weighting function is usually assumed to be *S* shaped: initially concave, then convex. Thus, small probabilities of winning are given too much weight, while large probabilities are underweighted. *S*-shaped functions also exemplify the principle of diminishing sensitivity.

Reference dependence and loss aversion help explain several phenomena of choice and related behaviors:

a. Utility comes from returns, given the fact that under Prospect Theory, investors' utility functions depend on the *changes* in the value of their portfolios.
b. Choices are compared to the current situation and judged as favorable or unfavorable, given that the status quo is the usual reference point, and the weaknesses of the options matter more than their advantages.[139]
c. Investors' aversion to holding stocks is known as the "equity premium puzzle," a phenomenon that explains the historical real returns of stocks over "risk-free" government bonds; it is read as a reflection of risk aversion. The puzzle occurs because this average premium, 6 percent in the USA, is understood to be high, pointing to an equally elevated degree of risk aversion among investors.[140, 141]
d. Investors withdrawing money from the market when prices fall. This action may be explained by loss aversion and the inability to ignore sunk costs, even though these are costs that cannot be reversed.
e. Investors stressing more about out-of-pocket losses, than about uncertain gains.
f. Investors tendency to avoid the sale of shares that lose value[142] while selling those which have increased in value.

3.6.6. Applying behavioral economics to financial markets[143]

People do not always make "rational" choices. As proposed earlier, much human decision-making is done in an automatic and intuitive manner instead of following deliberative and controlled processes. Accordingly, academic works have proposed a number of ways in which everyday thought departs from rationality. Furthermore, using insights from psychology, behavioral economists help us detect, understand, and remedy problems that arise from biases that lead people to misjudge situations or be inconsistent in their choices. Understanding these biases is of particular relevance in financial decision-making for the following reasons.

It is difficult to learn about financial products from personal experience. This is obvious considering many financial decisions (such as taking a mortgage or making a retirement plan) are made sporadically and their consequences only learnt after some time. Also, choices may hinge on economic circumstances and investors have little chance to master them as the next time the same issues pop up, the overall scenario has changed. In addition, cultural taboos may impede open and frank discussions of financial issues and outcomes, precluding one from learning about the experience of others. Furthermore, complexity is inherent to many financial products. Informed decisions require time to evaluate options and, occasionally, specific training. Also, economic scenarios might be unclear, and many financial products are difficult to understand given both the limited cognitive capabilities of most people and the instruments' intricacies. Faced with complexity, people can simplify decisions in ways that lead to errors.

Above, we proposed one such instance: replacing a complex question with an easy one. But other examples include ignoring dissonant information and selecting partial information to scrutinize a problem. Focusing only on returns is an example of the latter. Risk and uncertainty regarding the present and future need to be considered in financial decision-making. Most people are bad intuitive statisticians and thus prone to making systematic errors when decisions involve uncertainty. Additionally, assessments of risk and uncertainty might be obscured further by partial and opaque information provided by the industry. Consequently, consumers will often misjudge probabilities and make poor investment decisions. Furthermore, one needs to make compromises between the present and the future when assessing financial options. The fact that people regularly focus on the present and overestimate their future resources may likewise lead to making decisions that go against their best long-term interests.

To complicate matters, emotions are often involved in financial decision-making. Rather than the analysis of costs and benefits derived from the given options, financial decision-making may be driven by strain, apprehension, fear of losses, social pressures, greed and regret, and conflict of interest. Most often, it is difficult to get complete insight into our motivations when emotions are being played. In addition, consumer biases might be manipulated by firms marketing their financial products. Accordingly problems often arise because firms' product design and sales processes may accentuate, rather than ameliorate, the effects of consumer biases.

Given the large number of biases identified over time by behavioral economists, it is hard to provide a comprehensive list of those most relevant to financial markets. Thus, in Table 3.3 we follow Stefano DellaVigna (2009),[144] who grouped the biases according to which component they affect: *preferences*, *beliefs*, or *decision-making processes*, and explain each of these biases thereafter.[145]

3.6.6.1. Preferences

Preferences are things consumers want and value. In economics, preferences are implied when people choose among alternatives, and when arranging options in terms of expected levels of happiness, gratification, utility, and so on.[146] Biases in preferences may arise when, for example, consumers do not realize

Table 3.3 Ten behavioral biases and effects in financial services

Preferences	Beliefs	Decision-making
Time discounting and present bias	Overconfidence	Mental accounting and narrow framing
Reference dependence and loss aversion	Over-extrapolation	Framing, salience and limited attention
Regret and other emotions	Projection bias	Decision-making rules of thumb
		Persuasion and social influence

how much their decision to purchase a product or service is influenced by emotions and psychological experiences. One example is the "hypothetical bias," which occurs when declared preferences are different from those expressed through actual choices.[147] Within this category we list and describe three key biases and their main implications:

> *Time discounting and present bias*:[148] Time discounting theories scrutinize differences in the relative valuation placed on rewards at different points in time, and propose that present events are given much more weight than future ones. They also suggest that the value of occurrences farther in the future decreases more slowly than the value of those closer to the present.[149] Time discounting research also shows that once rewards are seen as being very far in the future, they stop being valuable. Tests of this theory show that discounting is not time consistent as it is neither linear nor does it occur at a constant rate.

Relatedly, the present bias,[150] also known as time-inconsistent preferences or preferences for immediate gratification, refers to the tendency to give stronger weight to payoffs that are closer to the present time when considering trade-offs between two future moments. For instance, people would rather get $100 now than $110 in a month's time. But this preference will not be sustained when the same choice is transferred to a year and a month from the current period. Present bias can also lead to other self-control problems such as procrastination and inertia.[151]

The overall idea is that people react to compulsions for immediate gratification. These self-control problems result in the overvaluation of the present over the future, and thus making choices that may be regretted later. Keeping insufficient savings for retirement and not adapting the portfolio investment strategy to what is appropriate to different needs across an individual's life are just two examples. Also, procrastination drives consumers to postpone doing tasks that demand effort, such as switching investment managers when the current one does not work. Of course, these biases provide firms with an opportunity to take advantage of such typical errors. For instance, since consumers will not make the time to shop for the best deals or bother to cancel unnecessary services, firms adopt strategies to profit from these attitudes. Also, firms may offer prices that look attractive should the buyer make use of all the contracted services, or they may set low charges that are later hiked once the services are truly used. Thus, consumers focus on the full range of services they "could use" and pay for them, or investors fixate on the low charges they initially expect to incur but not the higher ones they do not think to contract (such as interest rates).[152]

Pricing strategies designed to convince individuals to acquire mortgages too expensive for the holders is one example of how these biases were manipulated during the housing crisis. One way this was achieved was by offering cheap interest rates initially. These were raised after a few months. This, coupled with the implicit promise that the properties would soon increase in value (thereby insinuating that home equity could be used to repay mortgage charges), worked to convince many people that acquiring such debt and taking advantage of the revaluations in real estate was a good opportunity.

> *Loss aversion and reference dependence in Prospect Theory*:[153] Loss aversion highlights that individuals prefer avoiding losses to acquiring gains, and explains these preferences on the basis of deeply felt emotions related to the fear of loss. Loss aversion has several consequences, such as the endowment effect, which implies that we value an asset more just because it belongs to us; a preference for the *status quo*, even to the point of ignoring the chance of obtaining small net gains because "they are not worth the effort"; and distortions in attitudes toward risk. The effects of these distortions depend on whether the events relate to gains or losses and the probabilities involved. Very often, these attitudes drive individuals to respond to events in a range of ways that are not rational. Some examples include paying for "peace of mind" (such as when consumers pay for unneeded insurance); seeking risks hoping to avoiding losses (such as when individuals deep in debt take crazy gambles); pursuing risks and overpaying for the chance of a large gain (like overpriced lotteries); and avoiding risks for the fear of a loss (such as over-insuring for very small risks). Furthermore, since people are more willing to take risks to avoid losses, loss aversion can explain differences in risk-seeking.[154]

The reference point was introduced earlier in Prospect Theory and can be summarized by the realization that investors assess outcomes in terms of gains and losses from a reference point. For instance, the price at which a house was bought serves as the reference point for its selling price later, even when this price is above the market value. The reference point could be the status quo, expectations, other products available in the market, or the publication of recent research, among other things. Preference for a particular product may change when the reference point changes. However, depending on the reference chosen, the same outcome may be framed as a gain or a loss. Given that, psychologically, losses are felt about twice as much as gains, investors downplay gains and overweigh losses. Ergo, consumer choice can be erratic and vary based on the chosen reference.

Having said all of this and keeping in mind that, when evaluating a product or future prospects, people's choices are affected by both loss aversion and reference dependence, firms design strategies to exploit these errors. For instance, firms might manipulate the reference points and exploit endowment effects and loss aversion.

> *Regret and other emotions*:[155] People make choices to prevent negative feelings, so their actions (including inertia) are distorted by these temporary emotions. A common example is to shy away from fixing debt problems, or not purchasing financial products just to avoid stress and anxiety. Obviously, firms design processes to profit from these feelings. Our housing crisis example also works to illustrate this instance. Prospective home owners were induced to enter into debt in order to avoid the future regret of not having taken advantage of the expected increase in real estate prices.

3.6.6.2. Beliefs

Beliefs are what people think about the facts of a situation, the options available, and the likely outcomes. Often, these beliefs are formed by implicit, unreliable rules of thumb regarding the likelihood of events occurring, or the individual's own abilities to succeed, even when these contradict objective facts.

> *Overconfidence*[156] refers to individuals who overestimate the precision of their judgments and are too optimistic concerning the probability of good events occurring to them. Overconfidence is similar to *optimism* bias. Individuals' overconfidence has been explained on the basis of hindsight bias – everything can be elucidated when looking back – and self-attribution bias – taking responsibility for successes and blaming others for failures. High rates of entrepreneurs entering a market despite low chances of success[157] and people thinking they are better than others in their ability to select winning stocks are some instances of overconfidence.

Overconfidence also explains investors' underestimating uncertainty and taking on excessively risky assets. Thus, firms may profit by catering to these misperceptions. An example could be the well-publicized legend used before the housing crisis, that real estate prices had, historically, always increased. This was used as a forecast for the future.

Over-extrapolation from small samples, also known as the "Law of Small Numbers"[158] refers to forecasting events on the basis of few observations that are not truly representative. By pretending that these observations are sufficient to assess future trends, people also underestimate uncertainty. One case would be selecting a small sample of historical returns that may just reflect specific past circumstances in order to assess future performance and using these to make investment decisions. Another example would be to consider financial advice good on the basis of just a few experiences, when success could have been the result of good luck rather than exemplary strategy. One way in which these effects are used to take advantage of investors is in advertising, such as when firms declare the effectiveness of a product based only on a sample test of a few people.

Projection bias[159] indicates people's tendencies to believe their current preferences will be long-lasting, ignoring the possibility of change and, for this reason, being unprepared for the future. For instance, investors' underappreciation of habit formation may lead to projection bias in planning for the future. An example would be binding capital in long-term contracts and overlooking the likelihood of different future needs before the contract lapses.

3.6.6.3. Decision-making process

Decision-making process refers to "unreliable" shortcuts used in selecting the preferred option. The following are some of the most typical shortcuts.

Mental accounting[160] depicts the manner of handling money, dependent on factors such as the money's origin (salary or sale of an asset) or intended use (school money, funds for staples, etc.), rather than considering it all equal. According to this theory, people treat money as not-fungible and frame assets as belonging to current wealth, current income, or future income. An instance when this can become a problem is when individuals keep "mental accounts" for saving and borrowing, the first at a low rate and the latter at a higher one. Another example of mental counting is credit card payments, which are treated differently than cash.

Mental counting and narrow bracketing or framing describes how investors may regard choices in isolation rather than combining them with other decisions related to their overall level of wealth and risk-taking. One example is making investment decisions on an asset-by-asset basis instead of planning the whole portfolio.

> *Framing, salience, and limited attention*[161] have been extensively covered earlier. In summary, framing describes how various presentations of the same information can result in disparate investor choices. This occurs because choices can be worded in a way that highlights the positive or negative aspects of the same decision, leading to changes in their relative attractiveness. Thus, frames work because they trigger a given bias (loss aversion, reference dependence, regret, a rule of thumb, etc.). For instance, by appealing to loss and regret, firms may include irrelevant information to distort the assessment of meaningful features, to provoke an emotional response, or to set off a specific rule of thumb. Some examples include risky choice framing (the risk of closing 20 out of 100 businesses versus the chance of saving 80 out of 100), attribute framing (beef being 90 percent lean versus having 10 percent fat content), and goal framing (offering people a reward to motivate them versus imposing a penalty).[162] Another example is to create complex pricing structures so that assessing total cost is a Herculean task. Overall, framing effects are difficult to uncover because of the elusive nature of what the "neutral" frame could be.

As mentioned earlier, salience affects how and what information is processed by highlighting some of the features of the object or situation. For example, a medicine can be framed as being 98 percent reliable versus having only a two percent failure rate, thereby emphasizing either positive or negative information. Salience also underlies heuristic judgments that rely on external cues such as "packing" or other forms of presentation from which perspective customers can infer quality.[163]

> *Decision-making rules of thumb or heuristics*[164] are used by investors unconsciously to simplify activities when problems are complex. Heuristics are cognitive shortcuts that can also lead to cognitive biases. Examples include choosing the most familiar option from a number of alternatives to reduce uncertainty, and, when estimating unknown quantities, anchoring to some figure and adjusting from there so as to apportion equal amounts of assets within a portfolio instead of studying the correct allocation.[165] Other common heuristics are avoiding the most ambiguous options and choosing what draws the most attention.

Different rules of thumb affect how consumers search for or evaluate services, or how they view a situation. For instance, investors may only consider the financial products provided by the three largest companies within a sector, or having looked at three of these firms, they might conclude that the services of all others will be equal. Obviously, firms purposefully use their knowledge about which situations trigger which heuristics and use it to sell their services.

> *Persuasion, social influences, and norms*[166] are relevant as well. Social norms are behavioral expectations or rules that apply to a group of individuals.[167] There are different types of norms and these vary across cultures and contexts, but they always affect our preferences. For example, while professional norms require payment for services, social norms modify these expectations (thus, the practice of getting free advice from a trusted consultant). One obvious example of "social influence" is peer pressure, which played a major role during the housing crisis and is a common factor in portfolio investment selection. A second important aspect is to remain aware of the automatic character of System 1, and realize it is also reflected in an aversion to change. We have a default preference for things to stay the same and not change our behavior, not unless adequate incentives are set in motion.[168, 169]

3.7. Biases, strategies, competition, market failures, policy, and regulation

Behavioral biases interact with information asymmetries, externalities, and other market failures to such an extent that markets that appear competitive might not actually be so. Interactions between behavioral biases and competition have been the focus of extensive research.[170] Among other factors, studies report that traditional competition policies, such as increasing the number of competitors or introducing simple products, may be ineffective as long as other products exacerbating the biases exist in the same markets or the underlying biases are not addressed. The following summarizes the main implications of these works:

a. Behavioral biases create and strengthen market power. One reason for this is that consumers insist on sticking to their current products and services to, among other reasons, preserve the status quo and inertia biases.
b. Investors make mistakes when they encounter complex pricing schemes, such as teaser rates.
c. Investors' needs are not served by bogusly differentiated products. For instance, firms may offer attractive rewards that have no real value and complicate the investors' ability to compare among services offered by various suppliers.
d. Biases may worsen with the entry of more firms if they compete on an exploitative basis. That would be the case when, for instance, there is collusion in

the price/service-level ratio offered by the various market participants, however well-disguised under layers of incomprehensible jargon.

e. The presence of consumers less prone to bias-induced mistakes may not improve market transparency. On the contrary, it is often the case that these investors will benefit at the expense of the less sophisticated ones.[171]
f. Biased behavior also applies to the firms' decision-making. That is, firms might not be the rational agents.[172, 173]

In addition to market power, information asymmetry and externality are the main categories of market failures that interact with behavioral biases to produce inefficient competition. In this respect, the theory of the "second best" proposes that when a market suffers from multiple failures, actions to fix one problem might make matters worse. We outline some of the key interactions among these failures below.

3.7.1. Information asymmetries

As stated earlier, information asymmetry arises when a party to a transaction (a firm) has access to more relevant information or can process this information better than the other party (the investor). Some information asymmetry problems may be resolved by requiring that firms disclose all relevant information. However, behavioral biases affect how consumers understand and act upon these data. For instance, biased investors may simply disregard or misinterpret the available information. There is also evidence that too much information can confuse investors by distracting them from the key data or causing them to react emotionally.[174, 175] In these cases, investors might just focus on a few headline rates, ignoring the additional information about features, risks, or charges provided to them.[176] Also, if information is provided in a convoluted way, such as using lengthy and complicated legal terms, it might be too costly to access anyhow, leading to the same result. In these instances, consumers may rely on advisers. Nonetheless, this may not solve the problem either as the expert may be affected by her own biases or misaligned incentives.

3.7.2. Externalities and cross-subsidies

Externalities occur when consuming or producing a financial product or service affects a third party who is not part of the negotiations. For instance, overconfidence may lead consumers to overborrow, thus increasing the chance of default. Mortgage defaults bring externalities, given that home repossessions lower the value of neighboring properties. The additional risk thereby created might push other homeowners into negative equity, and consequently into foreclosing on their loans.[177]

Cross-subsidies are similar to externalities in that they also affect third parties not directly involved in a transaction. However, in this instance, the effects are felt through market responses. One example is that of sophisticated investors taking advantage of pricing structures designed to trap consumers while avoiding the additional charges that less sophisticated investors will have to defray.

3.7.3. Root causes and remedies for problems

A way to assess risks resulting from consumer mistakes is to reflect upon the difference between the intended use of a financial product and its real use. For instance, step-up mortgages were designed to allow those with bad credit or low income a time window so that once these situations were resolved, they could refinance at more reasonable rates. However, during the onset of the housing crisis these mortgages were offered to individuals who had no real chance of overcoming their financial problems. Thus, once the low rate period elapsed, the inadequacy of expectations resulted in massive defaults among "subprime mortgage holders."

With this evidence at hand, it would be reasonable to hold that in attempting to identify relevant explanations for the root cause of problems, one should deconstruct proposals on how well their predictions fit the evidence.[178] For instance, one should be watchful for early warning indicators and ask whether the necessary features for each proposed explanation are present in the market. Equally, one should assess which predictions are less likely to occur, especially if they do not completely fit the evidence.

However, determining the real causes of behavioral problems is complex. The first reason for this is that it is often unclear whether consumers are making mistakes, as biases might be interacting with competition problems and information asymmetries. Furthermore, even after concluding investors are making mistakes thanks to behavioral biases, it can be difficult to establish their true preferences, given that we can only observe their biased actions.[179]

Thus, to develop effective interventions, these interactions should be disentangled and well understood. For example, despite high interest rates, some consumers continue to use payday credit. The reason for this is that high-risk borrowers may have no other source of credit. Therefore, any action to restrict their ability to borrow will immediately and effectively negatively impact this class of borrowers. In this respect, usury laws have been denounced as a "regulatory failure" driven by the regulators' own behavioral biases.[180]

Thus, when structural or systematic problems exist in addition to consumer biases, various market-wide measures need to be considered. It is clear that behavioral problems may not be resolved with solutions that are "behavioral" themselves but rather by imposing stronger governance, banning certain services, channels of distribution, or customer types. Likewise, when addressing nonbehavioral problems, consideration should be given to matters such as investors' reactions to product design and so on. Hence, overall, the policy development process requires an analysis of potential market-wide problems, and when analyzing problems leading to poor consumer/investor outcomes, one should consider all the following variables:

a. Information asymmetries, externalities, and misaligned incentives.
b. Regulatory failures and consequences of prior interventions.
c. Behavioral biases of consumers, firms or regulatory organizations.
d. Competition failures.
e. Macroeconomic environment and other "out of reach" variables.

A number of actions could be taken, such as providing more and/or better product related information and education to investors, changing the choice environment, and controlling the product and its distribution.

With respect to the first action, if consumers cannot properly understand or evaluate the features, costs, returns, and risks of a product or service, one solution is to provide "better" information. For instance, corporations could be asked to publish their material in a way that does not exacerbate consumer weakness, and certain marketing practices that target behavioral biases could be forbidden.[181] In the context of our payday loan example, what this means is that we may disclose the cash charges loan holders would incur through time should they take the loan, rather than just publishing a percentage interest expense they may not fully understand.[182]

The other aspect under this headline relates to education. This would encompass not only education with respect to the financial product or service itself, which so far has brought mixed results when used to tackle behavioral weaknesses, but also in reference to consumer biases. Here the idea is to equip investors with the necessary knowledge to become sensitive to their own biases, and provide them with a set of skills that would allow them to make better decisions.[183] This is a step in the right direction, although so far providing information to educate consumers on their own biases has also had minimal positive effect on how they make decisions.

The second category of possible actions refers to the choice environment. Issues include the amount of information provided, how simple it is to read, how prominently numbers are displayed and in what order, how information and marketing is framed, and whether salience will affect the choices made by investors.[184] Here we address how biases are stimulated by presenting alternatives in a manner that highlights or hides certain product features. A key example is the typical "default" option most consumers stick with, and the rules of thumb used to choose between alternatives, given that just by changing the default option or the first choice offered within a list, we push investors toward making specific decisions.[185] One way to get around this practice is to request that firms ensure that what is more relevant is more salient as well. However, without specific testing, it is impossible to assess how effective each proposed change could be, or what could work for investors. Thus, in this respect, writing market-wide policies is difficult given that case-specific instances could easily be distorted by firms.[186]

The third category of actions refers to controlling the product or service and/or its distribution. That is, one could ban specific products or features that appear to exploit consumer mistakes. Alternatively, one could require that products contain specific features or that these are promoted or sold exclusively through specific channels, or impose restrictions in relation to the kinds of clients these products might reach. These measures could include total bans on products that are inordinately likely to result in consumer harm because of design, use, or important and persistent problems with how it is promoted or sold. In addition to the prohibitions, product controls also ensure that they contain specific features. One example is setting limits on credit card spending in accordance with the product being purchased (such as staples, casinos, etc.). A different example refers to a Financial Services Authority (FSA) intervention regarding small businesses benefiting from interest rate swaps. These sales depend to a great extent on exploiting information asymmetry and a mix of optimism and persuasion/trust biases.

3.7.4. Practical challenges for interventions

The effort to impose effective and fair restrictions encounters serious challenges. For instance, an important task is to ensure the above-mentioned controls do not affect innovation and competition.[187] Other types of challenges refer to evidence from behavioral economics that suggests that consumers do not always act in their own best interests, no matter how educated with respect to financial products or their own behavioral biases. This brings to the table issues regarding the limits of consumer responsibility and the value of choice and product variety. If consumer mistakes are predictable, there is scope for a regulator to intervene. The crucial point is that the option selected must correct the relevant market failures. Alternatively, if correction is impossible at an acceptable cost, the intervention must attempt to counterbalance the effects of the failures identified.

In this section, we have discussed specific interventions targeting behavioral biases and their impact on investor and consumer choices. However, consumer psychology is subtle. Despite a high-level understanding of the problem, the specific intervention chosen might succeed or fail thanks to small details and factors. This matters particularly when applying behavioral insights in information and design of choice environment. So, another consideration that could have a profound impact on the effectiveness of the proposed measures is whether these recommendations are performed in a vacuum. If the right environment is created, so that consumers are educated with respect to financial products and self-biases, and policy is also crafted to help institutions become more transparent, chances are that the warnings coming from different angles will have a greater impact on investor decision-making.

4
Bubbles versus the Valuation of Fundamentals

4.1. The role of "expectations"

Bubbles in financial markets exist when assets do not trade at their fundamental values. The bubble in the price of an asset, let us say 100 percent in the case of payment for a service which does not exist, is in its entirety the opposite of the case in which this same service is 100 percent delivered at an agreed-upon price representing the fundamental value that was supposed to be delivered to those who purchased it.

Let us propose an example and assume this service is the result of hiring online a four-door luxury sedan with a chauffeur to drive us to a destination at the prepaid cost of $0.25 per mile. The case in which the price paid represents the service fundamentals is that in which we do get picked up at the prearranged time by an adequately uniformed chauffeur in an Audi A6, Mercedes-Benz E350, or Lexus GS 350. After a polite greeting, opening the door for us and fixing the car heating mechanisms to ensure our comfort, the chauffeur drives at the legal speed, considering road, traffic, and climate conditions, all the way to the desired destination. Upon arrival, the chauffeur opens the door, helps us out of the car, and with a cheerful, friendly, and respectful voice wishes us a very nice day.

In the above instance, we feel we received 100 percent of the value we had purchased and paid for and, in consequence, the price of $0.25 per mile corresponds 100 percent to the fundamental value of the service. The opposite situation would be that in which we found ourselves waiting for a car and a driver that never showed up, and upon attempting to call the "help" line received no answer and realized after a 30-minute wait that we have paid for a scam. In the latter case, 100 percent of the cost of the prepaid service would correspond to the bubble.

Should we confront this last scenario, though, we would not call our experience "a bubble"; we would just acknowledge that we were victims of a scam. Let us then modify the story so we can get closer to a very simple bubble experience. For this purpose, we go back in time to the moment we read the online advertisement for the driving service before we had either of the two experiences related above. Thus, at the point at which we are looking into the future, the only thing we have are expectations of what this service could be like should we decide to purchase it. We can then assign probabilities to those "expectations." However, given that the expectations we form about anything are driven by certain exogenous factors, we can adopt three different personalities to see how our specific circumstances might affect these expectations.

As person A, we are optimists and know nothing about the cost of cars, gas, salaries, or taxes. As person B, we are moderate optimists and knowledgeable about the cost of cars, gas, salaries, and taxes. As person C, we are pessimists and also have knowledge about the cost of cars, gas, salaries, and taxes. Under these assumed personalities, our expected probability distributions to receiving the service exactly, or in a certain proportion, as described in the advertisement could look as described in Table 4.1.

The first column indicates the person assigning probabilities to the level of completion of the services promised; the first row indicates the percentage of completion of the services promised; 100 percent means that the services will be delivered exactly as promised, while 0 percent refers to a scenario where no services are delivered. For instance, 75 percent refers to a scenario where the car is not a luxury sedan but is still a somewhat new and presentable four-door car, while 25 percent might be a scenario where you are collected late in a pick-up truck.

Considering each person's expectations, we could say the following. A, having no information of any kind and possessing a very optimistic personality, sees no problem assigning a probability of 90 percent to the chance that she will enjoy 100 percent of the promised benefits, and 10 percent to the "rare" occurrence that she will only enjoy 75 percent of these. She allows for the possibility that the driver might be late, or that the company's concept of a luxury sedan might not match her own. B, knowing that the cost of these cars ranges between \$50,000 and \$75,000, that they will depreciate over four years, and being fully aware of the costs of gas, salaries, and taxes among others, cannot assign a high probability to the event of receiving 100 percent of the promised services. She assigns 5 percent to indicate she cannot rule it out, since she might be missing specific

Table 4.1 Probability distribution for expected degree of service completion (x)

	Percentage of service completed					
Person	**100%**	**75%**	**50%**	**25%**	**0%**	**Total**
A	0.90	0.10	0	0	0	100%
B	0.05	0.10	0.15	0.50	0.20	100%
C	0	0	0	0.10	0.90	100%

inside information of the company, but obviously she finds it quite implausible. She is more bound to believe that, most likely, only 25 percent of the promised service will be delivered. C, being both a true pessimist and fully aware of the costs involved, will have an inverse probability distribution to that of A.

As we can see in Table 4.1, each of our advertisement readers has made up her own mind about the likelihood that these services could be delivered in the manner in which they are advertised. It should also be noted that the information given to them by the company remains the same. Hence it is circumstances such as personality traits and prior knowledge that lead them to assess the situation differently.

The expected value of a random variable, also called expectation, mean, or first moment, is the weighted average of all possible values this random variable can take. In the case of a discrete random variable, we use the probabilities of its occurrence as weights to compute this average, whereas in the case of a continuous random variable, we use the densities.

In our example, the probability assigned by each personality to each of the percentages in the table gives the probability for each value of the random variable "service completion" to occur as per each individual's opinion of what might happen in the future. If x is the random variable, then we denote the probability that x occurs as $\mathrm{P}(x)$. In this instance, x refers to a number of service completion stages, and $\mathrm{P}(x)$ refers to the probability that the specified completion rate is achieved. The sum of all probabilities assigned by each person across the different outcomes must add up to 100 percent. For the sake of simplicity, our example refers to a discrete distribution, but in reality the variable "percentage completion" should be closer to a continuous distribution. The reason for this is that if we were to make a list of all the variables that would describe 100 percent of the service, we could qualify each of these variables again using a range. For instance, we might think "politeness" is a relevant variable. Politeness could be defined in a range from 0 to 100 percent and the specific level of politeness perceived could receive a grade of 53.668 or 75.99 out of 100. This would be the case for each one of the variables that we list as part of "service completion."

When it exists, the "expected value" may be intuitively understood using the law of large numbers, for as the sample grows and more observations are used to assess the expected value, it will resemble more truly the "true value," which is the actual outcome. Informally though, it can be interpreted as the long-run average of the results of many independent repetitions of an experiment. Given that the expected value of a random variable is the average of all values it can take, the intuition is that the expected value is what one would expect to happen on average.

Let us assume that a random variable Y can take value y_1 with probability p_1, value y_2 with probability p_2, and so on, up to value y_k with probability p_k. Then the expected value, mean or expectation of this random variable Y, $\mathrm{E}[Y]$, is defined as the sum of each value y multiplied by its own probability p. Given that the sum of all probabilities p_i must add up to one: $p_1 + p_2 + \ldots + p_k = 1$, the expected value is a weighted average. If all outcomes y_i are equally likely (i.e. $p_1 = p_2 = \ldots = p_k$), then the weighted average is a simple average. Otherwise, when the outcomes y_i are not equally probable, the weighted average takes into account the fact that some outcomes are more likely than others. Either way, the intuition is that the expected value of Y is what you expect to happen on average.

One aspect to consider, though, is that the value may not be "expected" in the strict sense of the word. For example, its existence may be unlikely or even impossible, such as having 0.5 fish in a fish tank at time *t*. Also, an estimated average might not give an accurate representation of a true state.

According to the example above, each person's expectations would appear as described in Table 4.2.

It is apparent that the same public information displayed in this corporation's advertisement has been interpreted very differently by the three people. It is also true that each of them construes reality using their personalities as processors of information and their knowledge about costs (lack of specific facts as to costs is also an information set) to assess the most likely future.

To avoid extremes, although there is no reason why we should, let us walk through the mental path of person B who has come up with a few quick numbers in her head in order to assess reality, though the numbers presented in the advertisement do not add up. On average, the cost of gas is \$0.185 per mile, monthly depreciation for this type of luxury vehicle stays around \$800, and the driver's salary and associated taxes are no less than \$35,000 yearly. She thinks there is no way this business can be for real. The \$0.25 charged as per the advertisement includes gas and so that would only leave \$0.065 to attend to any other expense and to make a profit.

Given that she has no inside information on the company, she wonders whether this firm could be purchasing the lower end of these "luxury" vehicles under special conditions with a price reduction of, let us say, 25 percent, and whether the drivers are somewhat inexperienced and hired on a per-hour basis

Table 4.2 Probability distribution for expected degree of service completion

	Percentage of service completed					
Person	**100%**	**75%**	**50%**	**25%**	**0%**	**Expected value (μ)**
A	0.9	0.1	0	0	0	97.5%
B	0.05	0.1	0.15	0.5	0.2	32.5%
C	0	0	0	0.1	0.9	2.5%

at an approximate cost of $8 rather than a full-time salary, which would have doubled the bill. If that were the case, then depreciation would run at about $300 per month and gas would be $0.153 per mile, provided the lower-end vehicle also gets selected and which would reduce consumption. In this combined scenario, we would have $0.097 per mile to attend to expenditures other than gas which still would not cover even the cost of the chauffeur.

Most likely, if she has any time for it, B will try to imagine a situation in which this business scenario is plausible. That is how the average human mind works. B might try to pull into the hypothetical business plan of this service company additional sources of income such as advertising on the cars, as unlikely or insufficiently large as these might be. She may also consider new concepts, such as thinking it could be possible that the chauffeurs are transporting goods in addition to passengers. B will have a series of ideas that modify the likely original cash flow scenario she had once rejected in her head, just so she does not have to discard the advertisement as a scam. It seems to her that, should the advertisement not be backed up by real services, the company would be too exposed. However, she cannot explain how this business would work in real life. On the other hand, a little voice tells her that neither can she explain the Theory of Relativity.

That night B goes to bed and dreams of a solution to the problem: the entire car fleet was purchased at one of those police-seized property auctions for an average $1,500 per vehicle, and the drivers are students of a chain of chauffeur's academies who need to complete a certain number of practice hours to graduate, hence the reduced rates of $3/hr for their services. Furthermore, the company has entered into a two-year contract to keep 80 percent of its fleet busy transporting displaced football players and management that had been temporarily occupying a local sports facility. Thus, given her calculations in this modified scenario, the advertisement was referring to renting the excess capacity of a fleet, and this was just the tip of the iceberg as far as revenues were concerned. Such was B's curiosity that upon visiting the general manager to ask questions and volunteer proposals, she receives an offer to make the company national and then public. Obviously, her enthusiasm and imagination had made an impact, particularly after she contributed a number of ideas on how to grow the business's cash flow.

The next morning, B wakes up feeling she has a better understanding of how things could work. Charged by the experience, she talks to a number of friends who are ready to agree that her vision "makes sense." Furthermore, quite a few of them offer contributions to support her intuition by providing additional ideas on how to extract more income from this car fleet as well as stories about people they knew who had made fortunes running similar kinds of operations. That night, B dreams again of the company, this time of her role in taking it public. Apparently, this is what the market was waiting for: a good marketing campaign securing 100 percent subscription to the "A driver for all" corporation with slogans like "You too deserve a chauffeur."

After B takes over, the company grows every night, for a couple of weeks. Many professionals think that at the cost of $0.25 a mile, it makes sense to hire a chauffeur rather than drive their own cars. Reacting to the situation, and given the constant growth in demand, the company keeps purchasing additional assets and hiring new staff to attend to the expanding customer base.

In a brief period of time, the case of "A driver for all" becomes so exemplary, with its young, visionary management team, simple down-to-earth mission, catchy marketing phrases, and neat presentation, that the press begins to take an interest in telling its story. One of the nice things about how the press presented this business case is that it made the story very simple so any Jane or Joe could get the gist of it and be enticed to learn more. Actually, the national press delivered "A driver for all" news in a way that would remind viewers of soap opera trailers of upcoming episodes. Three months into this trend and the general public is familiar with the names and faces of most people in the top positions in the company. Some of the management take it a step further by introducing their children, pets, and significant others in magazine features. The titles of some of these features read as follows: "How to become a millionaire at 25," "Three steps to success," "Follow your dream and get it," "Five things brilliant people do every morning," "Some dos and don'ts: success versus failure," and "Take your company public in five months." In summary, these articles presented their view of the sure recipe for quick success which is embodied in making money.

The unfortunate side effect of getting so much attention was that copy-cat companies began to pop up on every corner. It seems B is not the only one who never understood how the business worked. The "would be" competition has decided on an entry strategy: that of cutting prices further to drive the first company out of business. The competitors' main idea is that "A driver for all" would be pushed out of the market and, once in control, a new pricing strategy could be considered for their own ventures. And so, with abundant financial resources and the help of a couple of "A driver for all" defectors, these businesses begin to appropriate certain customers, then certain sectors, and finally whole geographical areas. It is sooner rather than later that our dream company folds and its assets are sold off to pay lawyers, employees, taxes, suppliers, and debt holders of different types. In a few weeks, "A driver for all" has disappeared, leaving nothing behind, and never remembered by its customers. The only sign of its prior existence is a permanent confusion in B's head. She constantly asks herself: What happened?

4.2. Detecting bubbles

As stated in Chapter I, there is no agreed-upon definition of the word "bubble." Nonetheless, this word is generally used to refer to stock prices that are not justified by the assets' fundamentals. When a bubble bursts, high price volatility results as a consequence of excess supply.

The very general definition or, better said, description of the word *bubble* introduced in the first chapter referred to a discontinuity in the price of the bubbled asset. The implications are twofold: a large price change has to occur in a short period of time, and this change can only be recognized after it has already happened. Hence, this definition does not provide clues as to how to identify bubbles in advance. Of course, the observation of a strong inflationary process in the price of an asset might warrant attention. And an early diagnosis for the causes of this inflationary run can help determine whether there is an objective justification for it which is unrelated to variables such as excess liquidity. For instance, when the price of a mineral skyrockets, if this increase is justified by the fact that new technology allows its use in multiple industrial processes as a cheap substitute of other previously used materials, then the change in price will respond to a "real demand". In this case the increase in demand and price is expected to last. On the other hand, if the price increase were to respond to a fashion and excess liquidity available in the market, then we might be facing a bubble.

Given that the definition of "bubbles" does not by itself allow market participants to recognize these events before the fact, that is, before their implosion, to differentiate a likely bubble from a regular inflationary process one needs to understand how pricing processes are formed when the assets we are looking at are not dominated by a bubble component.

Bubbles can be of totally different natures, some forming and disappearing without causing any disturbance and going unnoticed by the market, while others drag down an entire financial system when they burst. Consequently, in this sense, one should have a measure of "importance" and not be concerned with the absolute meaning of the word, but rather with the relative significance of the specific instance. Hence, in general, one should not worry about bubbles, but focus on those which have a relevant impact on a financial system.

Bubbles do not come with a tag, and most often they are not easily distinguishable initially from a regular supply and demand or price growth process in a specific sector. Mathematical models are not able to "catch them" in advance either, for reasons we will analyze extensively in later chapters. Here we summarize the main problems inherent in testing for bubbles.

In general, tests for bubbles fall under one of two categories. In the first, one sets the null hypothesis that a bubble exists and attempts to reject it in favor of an unspecified alternative. In the second, one assumes the null hypothesis that a model is correct and attempts to reject it in favor of another that includes a bubble.

Most of the criticism of the first method points to the fact that these are "joint tests" of the "no-bubble" hypothesis.[1] In this respect, Fama (1991)[2] declared that market efficiency can only be tested jointly with some model of equilibrium, such as an asset-pricing model. The "joint hypothesis problem" proposes that when a model yields a return significantly different from the observed market return, one cannot conclude if the market is inefficient or that the model is incomplete. That is, whatever the finding one cannot know if the result derives from the assumptions made about the model (the definition and relation between fundamentals and prices), or the assumptions made about the time series properties of the fundamentals.[3, 4] In addition, the second methodology has also been questioned on the grounds that the same attributes can be associated with both fundamentals and bubbles.[5]

In response to the criticism against both categories, more sophisticated methodologies have been proposed in later studies. In addition, other groups of researchers have tackled the subject completely differently, such as proposing rules for identifying cyclical bubbles or creating theoretical models where bubbles can occur. We shall review these in Chapter 7, but in this chapter, our objective is to review the sources of "value" to distinguish among different types of price growth: those related to true fundamentals and those pertaining to a bubble formation process.

Contemporary economists aim to find a model that can produce a complete set of future market outcomes and probabilities. It is within this context that Karl R. Popper pointed out that any such approach is futile "to the extent to which [historical developments] may be influenced by the growth of our knowledge" (Popper, 1957, pp. xi–xii).[6] Furthermore, he also declared: "Quite apart from the fact that we do not know the future, the future is not objectively fixed. The future is open: objectively open."[7]

In line with these thoughts, it is our contention here that any attempt to create models which will "automatically detect future bubbles" is overall an incomplete effort, since our capacity to predict will be confined to the general characteristics of the events, and not apply to a particular event. For instance, one might be able to determine an environment where bubbles can grow, and one might also then be able to detect a given bubble. Nonetheless, one will not be able to assert that the existence of conditions where bubbles can form will always create bubbles or that once a bubble exists, it can be always be captured by a "model," given that situations evolve, including those which make bubbles possible.

Consequently, given that bubbles form when the intrinsic value of the underlying asset differs from its market price, a first key step in detecting bubbles is understanding asset valuation. A second aspect would be to assess how this "value" is traded in the market and the distortions it can suffer given expectations, market efficiency, contagion, and other aspects of the market place.

Models are simplifications of reality. Furthermore, models might be plain wrong or incomplete. Also, models themselves cannot assess the "quality" of the variables used to test them. So to start, it is good to keep an eye on the basic intuition of the bubble definition and ask ourselves a simple question: Is the price of the asset justified by its fundamentals or is it not? If we recall, Formula 1.1 stated that the price of the asset x_t should be similar to the value of the fundamentals F_t, for the bubble process not to dominate the asset price.

$$x_t = F_t + B_t \tag{4.1}$$

where

x_t is the price of the asset today
F_t is the part of the price that corresponds to the "fundamentals"
B_t is the part of the price that corresponds to the bubble (what we cannot justify according to the firm's fundamentals)

When $x_t = F_t$ there is no bubble component in the price of the asset.

According to this statement, one should get an assessment of the fundamental value of an asset in order to be able to determine whether its price is dominated by a bubble component or not.

Neophytes in the financial field have often been misled to believe that with the help of a good spreadsheet, optimal decision-making takes about as long as it does to input the necessary data into the computer program. However, a rude awakening is about to take place in the coming pages. It will soon become clear that asset valuation is quite a bit more complicated than that. Finance is an area of social science, and even though it relies heavily on mathematical models, to a large extent the value of the output of such models depends on the "quality" of the values inputted for each variable, the variables included, and the relationships among them.

Business valuation refers to the estimation of a wealth-generating process (positive or negative) that will occur in the future, resulting from a given business strategy should certain conditions hold. Some of these conditions are exogenous while others are endogenous to the model. Because our business environment is perceived to change constantly, for example, with new information reassessing inflation expectations, the overall scenarios we are trying to represent also change, and a lot of hard thinking has to go into determining which values should be assigned to the different variables, and which variables are relevant to the model for a specific "valuation". Furthermore, information is most often incomplete and uncertain, forcing the evaluator to make educated guesses. These guesses are most often parameterized within current conditions. It is very hard to think "out of the box," but given the "uncertain future" (Who will make up the management team in five years' time? What will the overall market conditions be like? Will the sector be as relevant or will currently unknown technology affect it?) it might also be impossible to make correct assessments.

An additional consideration in a bubble scenario is that the models and variables used to value an asset might be affected by the bubble itself. This is another reason why tests cannot "capture" bubbles. Reflexivity and feedback relations established through information updates will affect the fundamental valuation of the asset. That is, even assuming one is able to choose the right variables and propose a complete model, it is possible the specific values assigned to these variables – for instance, growth expectations and costs – are already affected by the bubble psychology and thus render the valuation tainted.

4.3. Fundamentals of valuation

A financial security represents a claim on future CF. That is the reason why its intrinsic or fundamental value is the PV of the CF the owner of the security expects to receive over time.

The problem now is that when talking about CF, one could be referring to two different things: (a) the selling of the financial security, that is, the shares in a market such as the New York Stock Exchange (NYSE) for a given price, or (b) the CF a company generates over time as a result of its operations, that is, the wealth creation process. The latter could translate at least partially into dividends, should the company distribute them. Otherwise, the generated wealth would be retained in the company and accumulated or used for new net positive PV projects. Hence, prices of traded stocks change so that after an adjustment to reflect expected dividends or earnings, time value of money, and cash flow risk differentials, they equal the market's best forecast of their future price to reflect all information about their intrinsic value.

For these reasons, the stock price at which the shares are traded in a stock exchange derives its value from the company's operations. However, having said that, the share price comes out of "an interpretation process" that has to take place in the market in order to assign a value to the expected future wealth creation of the given company, which is actually the factor that has value. And as history proves, at any specific point during this process a lot can be lost in translation, particularly when bubbles are in effect.

So what is the fundamental value of a stock? Some people will say it is the price you can get or pay for a share when trading in the market. This is clear and simple. However, as easy and convenient as it may sound, we have long enough been admonished not to confuse price with value, even if they sometimes coincide. For instance, because of the uncertainty created by the political situation at various moments during the liquidity crisis in 2008, some financial assets could not be traded and one could not get prices, or vice versa, that is, one could not get prices and therefore there was no trading. If we follow the logic that value is established by a sale action, then these assets would have had no value. However, we know they did and they do, even if temporarily there was just no market for them or there actually was a market but it could not clear.

When talking about the translation of business valuation into market prices, there are several distinct processes going on at the same time. The first is the business concern, that is, creating wealth through the sale of products and/or services. To assess its value, one can follow a formal valuation process such as discounting the expected CF to be generated by the firm's future operations. The second is the estimation of such a wealth-generating process, which depends to a large extent on who is doing the valuation. The result of the analysis of the firm's assets' capacity to create value in the future is reflected in the firm's share price, volatility, and trading volumes. In an efficient market, this could be seen as "the best guess or the market's best forecast of the future price." The third process is the independent activity of buyers, the demand side of the equation which has its own capabilities, needs, and constraints; for instance, the need to place the available liquidity on the parquet. The fourth is the business activity of intermediaries who need to make money out of fees and commissions and who are also constrained by other variables such as "fulfilling expectations of returns." These needs will result in actions which although "unrelated to fundamentals" will impact prices. In addition, fundamental analysis is dynamic, as updates in value come from innovations, and opportunities arise out of the need to manage portfolios. Consequently, the process of valuation of a company is continuous, rather than static.

In the "real world" and under average conditions, income growth from operations is limited because it would most often be the result of changes in inflation and demand. To estimate income growth through inflation, one would have to calculate its differential impact on corporate expenses and revenues. Increases in demand for any product or service would also be constrained by a number of factors, such as population growth, income, or average education level of the targeted population.

If we accept the premise that income growth is bounded by the variables mentioned, then other things being equal, the value of the company cannot increase beyond the impact of its growth due to these variables. That is, if our company sells just one book for children aged three to six in a region where women have 0.2 children per head, one could hardly justify expecting 100 percent growth once we have already penetrated the market, unless we start doing something else. By extension then, accepting that the share price fairly reflects the value of the company, stock prices cannot increase in a proportion that represents growth larger than the actual amount representing revenue increase. Nonetheless, there is plenty of evidence that stock prices are far more volatile than corporate income. This is because "other things" are going on in the markets and the interests of the various parties to the process are not completely aligned. Any process which does not create real fundamental value or one where the creation of real value is augmented significantly by unrealistic expectations can be identified as a potential seed for a bubble as expectations can justify almost any price.

As stated, the valuation of fundamentals refers to the analysis of the value-generating processes of a company: the creation of new CF and the risks associated with the same. Among others, Hamilton (1986)[8] and Tirole (1985)[9] proposed that the value of a financial asset is the PV of its future payoffs. With respect to shares, Shiller (1981)[10] suggested that one could use dividends or earnings to evaluate the PV of these future payoffs. He argued that earnings are relevant to the pricing of shares, but only as indicators of future dividends, since individuals are concerned with returns. His model implies that the capital gains component of returns is just a reflection of information on future dividends. In his opinion "earnings are statistics conceived by accountants which are supposed to provide an indicator of how well a company is doing and there is a great deal of latitude for the definition of earnings." Therefore, he recommends the use of dividends as proxies for fundamentals.

Others seem to have a different view. Dividends are established only by a few corporations and after careful consideration of multiple issues, including company maturity within its life cycle, expected future earnings, expected free cash flows (FCF), and smoothing. In the long run, whatever the dividend policy, it should bear relation to actual corporate earnings, since expected earnings are realized in time, and dividends can only be paid out of actual earnings. By extension, no matter how stock prices are determined, if there is no mispricing, they should, in the long run, be related to corporate actual earnings.

Market traders may not fully agree on the expected future CF of corporations, and they may have different needs, which is the reason why some agents buy while others sell. However, if over a sustained period of time stock prices grow at a rate significantly different from the growth rate of the operating CF, the moment will come when market-clearing prices no longer reflect the PV of discounted cash flows (DCF). Consequently, in order to ascertain whether there is a bubble, one can determine whether changes in share prices reflect changes in fundamental values as represented by company results. We can then justifiably compare price and earnings' growth rates to determine whether there is a significant relationship between the two variables. Of course, determining that mispricing exists does not in itself make a price more representative of value. Nonetheless, it is the foundation stone in understanding when a bubble formation process has begun. The size and path of the bubble can later be estimated or can be eliminated through coordinated action to redetermine value.

Under the efficient market hypothesis (EMH), the prices of all stocks equal the discounted value of their rationally expected dividend payment streams:

$$P_0 = \sum_{t=1}^{\infty} \frac{D_t}{(1+K_e)^t} \tag{4.2}$$

where

P_0 is the price of the financial asset today
D is the future stream of dividend payments generated by this asset
K is the discount factor which considers the risk of the asset
t is the time period under consideration

So let us illustrate, with a very simple example, how this functions and review the basic finance concepts related to asset valuation and pricing.

As argued in earlier paragraphs, the dividend stream here represents a cash flow generated by the firm and paid to shareholders. In this instance, we shall consider it correlated with the actual earnings or, better said, FCF of the company. We can assume, for the sake of an example, that out of earnings, the firm retains 80 percent and distributes 20 percent in common dividends. The dividends a company distributes provide information on the expected generation and stability of future CF, among other things. If the dividend stream were expected to change in any period, we would have $D_1 = D_0 + g$ or $D_1 = D_0\ (1 + g)$, where g is the growth rate in the dividend and D_1 is the dividend in the next period. In Formula 4.2, K is the discount factor – in this instance the cost of equity, which considers the risk of the asset or dividends. The subscript t represents the time period in which the payment occurs. Should we consider this to be perpetually growing payments, the formula would be modified to:

$$P_0 = \frac{D_1}{K_e - g} \tag{4.3}$$

These three variables – the time period, the size of the CF, and the risk and related cost of financial resources represented by the cost of equity – are the sources of value creation. Thus, when assessing value, these are the variables that need to be considered.

When valuing a business or any financial asset, we look into future prospects to create a number of plausible scenarios which could produce a stream of CF. These CF could happen at different future periods and thus need to be discounted or brought back to the present, which also entails different assumptions about risk. Thus, when discounting, we use assumptions related to the size, the timing, and the risk of the CF. There is no single correct answer to these, and this is where finance diverges from other sciences. Here it is all based upon "guesses" and assumptions. In theory at least, one can justify almost any PV. This is where the "crux" of valuation lies. That is why assessing the plausibility of the assumptions is the key to a good valuation process. There is no single correct answer, but there are an infinite number of wrong answers.

To sort this out, we need to have a strong intuitive understanding of what the fundamentals of an investment option might look like. Thus, we will review some of the basics, referring to the path between a value-generating idea reflected in a business plan, the financial statements used by the corporation to transmit information about the plan, and the market value of a firm.

The value of a company, like that of any investment, depends on the amount of money the investment is able to produce for its owners. Hence, if securities traded in stock markets derive their prices from the professional activities carried out by their respective firms, these prices should reflect the firm's fundamentals and there should be a relation between the firm's wealth-generating capacity and the prices at which its stock is traded on the exchanges. Formula 4.2 describes this fact: the current value of a given stock depends on the dividends paid by the company, the timings, and risks associated with these CF. Taking it a step further, dividends can only be paid out of the company's treasury. Thus, in the long run, there has to be a strong relation between these payments and the firm's ability to generate cash from its principal operations. Formula 4.4 refers, more generally, to this issue.

$$x_t = aE\left(x_{t+1} \middle| I_t\right) + z_t \tag{4.4}$$

where

x_t is the price of the asset today
a is a parameter between 0 and 1
$E(.)$ is the mathematical conditional expectation
I_t is the relevant information set
z_t are the fundamentals

If the transversality condition which prevents departures from the fundamentals holds, its forward solution defines the fundamental value, z_t, as expressed in equation 4.4. The price of the asset today, $x_{t,}$ should be similar to the fundamentals z_t for the bubble process so as not to dominate the asset price. If this condition fails, the general solution defines the current stock price as already indicated by

$$x_t = F_t + B_t \tag{4.1}$$

where F_t is the fundamental solution defined in equation 4.1 as the mean of the expected fundamentals conditional on the relevant information set, and B_t is the rational bubble component that is removed by the imposition of the transversality condition. B_t is an unstable, self-fulfilling set of expectations that violate long-term equilibrium conditions. Therefore, the bubble term may give rise to explosive solutions.

If the bubble term in this valuation is close to zero, then share prices will reflect fundamentals rather than a bubble component. These fundamentals can be defined in a number of ways. Looking at expected dividends is one approach. However, one could also look at the corporate future FCF, or earnings before interest and taxes (EBIT) among other proxies for fundamentals. Regardless of the chosen variables to proxy value, in order to assess the existence of a bubble, one needs to have a strong intuition about the fundamental value of the specific company that is under observation.

Assessing the value of companies is a subjective task. The reason for this is that in doing so, we need to look into future prospects, given that shareholders' claims refer to future wealth creation. Past wealth creation might be an indication of future capacity, but that is all. In valuing a company we are making a statement about the most probable result of the implementation of a strategy and business plan. Consequently, we need to make a considerable number of hypotheses about the behavior of the market, the company, and the interactions between both in order to estimate future results. As we cannot predict the future, our main concern should be to link "reasonable" hypotheses to the result of the assessment. It is then fair to highlight that value in this sense is a concept resulting from a specific business strategy responding to the question of how productive a company's assets will be, including management.

Business valuation is complicated by the fact that various assessment methods exist and none are capable of resolving all the problems that can come up in different types of businesses. Furthermore, the value of a company under one set of managers will most likely be different from the value of the same company when it is managed by a different group. There are also a series of qualitative factors and concepts that are truly difficult to grasp and represent; for instance, strategic factors, potential mergers or economies of scale, cost savings from sharing resources, and so on. These aspects appear constantly in the assessment process and complicate the procedure to an even greater degree.

When thinking about assessments, it is also important to keep in mind the distinction between value and price. This is essential, particularly when thinking about bubbles. The valuation process enables us to determine a justifiable price for an economic unit through the application of an assessment method which uses a series of hypotheses about the development of the economy and the firm's future financial position. Different evaluators will bring in different hypotheses and managerial capacities. Consequently, the value of a firm to different groups will vary. Price, on the other hand, is the quantity finally paid after the negotiation process between buyer and seller. We shall qualify this comment later on to differentiate the pricing process resulting from the evaluation of the productive ability of the firm's assets from any other, such as a pyramid scheme.

Our purpose here is to explain the objective logic of valuing business assets while highlighting the processes and limitations of each of the three methods most often used: accounting, liquidation, and DCF. Using these techniques, we shall underline the elements that define the value of an operating company.

With the help of a set of very simple financial statements, we are going to use different methods to assess the value of a company named Blue Jacket, Inc. Tables 4.3 and 4.4 present an initial income statement and balance sheet for this corporation.

The income statement shows the earnings and payments due within one period. The balance sheet includes, in the assets column, the total of goods and rights owned by the company, and in the liabilities column, the obligations toward suppliers, bankers, and shareholders at a given point in time.

Table 4.3 Blue Jacket, Inc. income statement ($ millions)

Sales	1,500
Sales costs	1,000
Gross margin	500
General expenditure	100
Depreciation	50
EBIT	350
Interest (10.67%)	37
Earnings before tax (EBT)	313
Taxes (35%)	109
Earnings after tax (EAT)	203

Table 4.4 Blue Jacket, Inc. balance sheet ($ millions)

Cash	60
Accounts receivable	255
Inventory	100
Current assets	415
Fixed assets	500
Total assets	915
Suppliers	100
Long-term debt	400
Capital and reserves	415
Total liabilities and owners' equity	915

4.3.1. Accounting value

The accounting or book value of a company is estimated by deducting the value of the liabilities from the value of the assets. In the case of Blue Jacket, Inc. this would be

$$\text{Accounting Value} = 915 - 500 = 415$$

This method is often not the best approach to valuation. There are two main reasons for this: first, most figures in the balance sheet reflect historical values, and second, the method does not tell us anything about the future earning capacity of the corporation. We address these issues by assessing value through the DCF method.

4.3.2. Discounted cash flow value

In order to value a company using a DCF method, we need to create a business plan assessing the size and timing of the CF that are expected to be generated by the firm in the future. We can put these figures into our various formulae to estimate their PV. The mathematical expression is:

$$PV = \frac{FCF_1}{(1+r)} + \frac{FCF_2}{(1+r)^2} + \ldots + \frac{FCF_t}{(1+r)^t} \tag{4.5}$$

where *FCF* represents the free cash flow expected in each period, *r* the financial cost or return that considers the risk of each cash flow, and *t* the time period in which each cash flow happens.

Given that the PV aggregates all the amounts the business will generate in the future, it is the maximum value of a business. The advantage of the DCF method resides in its capacity to take into consideration the various elements of value creation within the company. The key disadvantage lies in the number of hypotheses required to calculate it. Thus, to begin, the future behavior of business operations has to be assumed in order to establish how and when these operations would yield money. Hence, we also need to determine the corporation's life span and make predictions about inflation, interest rates, and business risk in order to choose the most adequate rate to discount the money yielded by the transactions the company carries out.

To sum up, we have a method that is conceptually powerful, but its true value will depend on our capacity to estimate the *FCF*, *t*, and *r* correctly. If we do not know anything about the company we need to assess, if we are not familiar with its context, its history, or if we are not able to gather information about the sector and its potential development, we will not be able to come up with a value that reflects a fair estimation of the company.

For the purpose of this example, we summarize our hypotheses on the evolution of sales, costs, investments, and depreciation for Blue Jacket, Inc. for the next five years, in Table 4.5.

Table 4.5 Blue Jacket, Inc. changes for five-year forecast ($ millions)

Year	**1**	**2**	**3**	**4**	**5**
Δ Sales	1.1	1.09	1.08	1.07	1.06
Sales cost (as %)	70	70	70	70	70
Δ General expenditure	1.05	1.05	1.05	1.05	1.05
Investments	70	70	70	70	70
Δ Working capital	1.1	1.09	1.08	1.07	1.06
Δ Depreciation: 10 years					

According to the hypotheses stated in Table 4.5, sales are increasing at a decreasing rate, and therefore further investments in productive capacity such as fixed assets and working capital will also be necessary. In this case, we invest $70 million annually to renew fixed assets that will be depreciated over ten years. In addition, we also increase working capital since, given the sales growth, we would have to carry a larger balance in inventory, accounts receivable and payable, and cash. The hypotheses to obtain our financial forecasts have been worked out in Table 4.6.

Given that our interest is to determine the expected FCF, we need to discount to determine the PV of the firm's future earning capacity. In the upper part of Table 4.6, we introduce some changes with respect to the conventional income statement. For instance, for now we ignore the interest corresponding to servicing the debt. The reason is that we want to know how much money the assets of the company yield, regardless of how the company is financed. This would later be accounted for in the discount rate. Also, in addition to the regular accounts of the income statement, we have added others which can help us transform earnings into FCFs. The general expression for FCF is: earnings after tax, +/– noncash adjustments (such as depreciation, amortization, or depletion), +/– investments in fixed assets, +/– investment in current assets. Once these adjustments are made, we will have the FCF – the amount of cash available for new investments or distributions to the company owners once all current needs have been considered.

Table 4.6 shows the projected evolution of the FCF over the next five years. Thereafter, we propose a number of hypotheses needed to estimate the residual value (RV) of the company after that first five-year period. If we are certain that operations cannot go beyond that time, we will have to liquidate: sell the assets and pay off the debts. If, on the other hand, operations carry on beyond the fifth year, we create, for the sake of this example, four additional scenarios for the period starting year six. These are summarized in Table 4.7.

Liquidating a business is an option either when operations are "finished for good" or the productive capacity of the firm's assets is less than their market value; that is, when the company is worth more dead than alive. With respect

Table 4.6 Blue Jacket, Inc. five-year forecast ($ millions)

Year		**1**	**2**	**3**	**4**	**5**
Sales		1,650	1,799	1,943	2,079	2,203
Sales cost		1,155	1,259	1,359	1,454	1,542
Gross margin		495	540	583	624	661
General expenditure		105	110	116	122	128
Depreciation		57	64	71	78	85
EBIT		333	365	396	424	448
Interest						
EBT		333	365	396	424	448
Taxes (35%)		117	128	139	148	157
EAT		216	237	257	276	291
+Depreciation		57	64	71	78	85
–Investment in working capital		32	31	30	29	26
–Investment in fixed assets		70	70	70	70	70
= FCF		172	200	228	255	280
Year	**0**	**1**	**2**	**3**	**4**	**5**
Cash	60	60	60	60	60	60
Other net current assets	255	287	318	348	376	403
Current assets	315	347	378	408	436	463
Fixed assets	500	513	519	518	510	495
Total assets	815	860	897	926	946	958

Table 4.7 Blue Jacket, Inc. scenarios to calculate the RV

	Growth	
Years	**0%**	**5%**
15	Scenario 1	Scenario 2
∞	Scenario 3	Scenario 4

to Blue Jacket, Inc. the gains/losses resulting from the sale of the assets considered in Table 4.2 are as follows: current assets will sell for half their book value and fixed assets for 50 percent over the book value. As the differences between book and market value would have tax consequences, we also need to consider these gains and losses. Table 4.8 summarizes the estimations.

4.3.3. Residual value

According to Table 4.7, the second overall scenario created for the example assumes that the company is not liquidated but continues to grow for a number of periods: either 15 additional years or forever (infinite future periods). The growth over these two scenarios will be either 0 or 5 percent. Thus, we use the

Table 4.8 Blue Jacket, Inc. liquidation value in the sixth year ($ millions)

	Book value	Market value	Taxes savings	Net value
Cash	60	60	–	60
Other net current assets	403	202	71	272
Fixed assets	495	743	–87	656
Total assets	958	1,004	–16	**988**

assumptions in Table 4.7 to depict the future CF of Blue Jacket, Inc. after the initial five years when the option is not to liquidate the firm.

The first scenario in Table 4.7 assumes the FCF will remain constant through a 15-year life span, that is, there will be zero growth in earnings. To calculate the RV of the FCF corresponding to these 15 years, we can use the PV formula:

$$RV = \left[\frac{FCF_6}{(1+r)} + \frac{FCF_7}{(1+r)^2} + \cdots + \frac{FCF_{20}}{(1+r)^{15}}\right]$$

Given that we have constant FCFs and the discount rate does not change, the expression can be reduced to:

$$RV = FCF_5\left[\frac{(1+r)^{15}-1}{r(1+r)^{15}}\right]$$

and assuming a 12 percent discount rate:

$$RV = 280x\left[\frac{(1+0.12)^{15}-1}{0.12(1+0.12)^{15}}\right] = 1{,}907$$

In our first scenario, the RV at the beginning of the sixth year would be $1,907 million.

Our second scenario is a 15-year annuity with a 5 percent growing CF:

$$RV = [FCF_5(1+g)]x\left[\frac{1-\frac{(1+g)^t}{(1+r)^t}}{r-g}\right]; \quad RV = [280(1+0.05)]x\left[\frac{1-\frac{(1+0.05)^{15}}{(1+0.12)^{15}}}{0.12-0.05}\right] = 2{,}605$$

The third scenario is a zero-growth perpetuity:

$$RV = \frac{FCF_5}{r}; \quad RV = \frac{280}{0.12} = 2{,}333$$

Finally, the last scenario is a 5 percent growing perpetuity:

$$RV = \frac{FCF_5(1+g)}{r-g}; \quad RV = \frac{280(1+0.05)}{0.12-0.05} = 4{,}200$$

To observe the impact of these different assumptions on the value of the corporation, we insert these figures in Table 4.9. As we can see, depending on the assumptions, the RV of the firm in the fifth year ranges from \$1,907 to \$4,200. That is, the same company can be valued as double or half; it all depends on the hypotheses.

In either of our two general scenarios, liquidation or business continuation, to estimate the overall value of Blue Jacket, Inc. we also need to calculate the PV of the FCF estimated for the first five years. To these, we then add the residual portion of any of the five options proposed.

$$PV_1 = \frac{172}{1.12} + \frac{200}{1.12^2} + \frac{228}{1.12^3} + \frac{255}{1.12^4} + \frac{280}{1.12^5} + \frac{988}{1.12^6} = 1{,}297$$

We have assumed that the liquidation value is recovered at the end of the sixth year. That is why the FCF is discounted by the denominator corresponding to that year.

$$PV_2 = \frac{172}{1.12} + \frac{200}{1.12^2} + \frac{228}{1.12^3} + \frac{255}{1.12^4} + \frac{280}{1.12^5} + \frac{1{,}907}{1.12^5} = 1{,}879$$

In this case, since the RV is calculated at the beginning of the sixth year (which is the same as the end of year five), the denominator applied to the FCF is that of year five. We repeat this last estimation for scenarios 2–4 and replace 1,907 with 2,605, 2,333, and 4,200 respectively.

$$PV_3 = \frac{172}{1.12} + \frac{200}{1.12^2} + \frac{228}{1.12^3} + \frac{255}{1.12^4} + \frac{280}{1.12^5} + \frac{2{,}605}{1.12^5} = 2{,}275$$

$$PV_4 = \frac{172}{1.12} + \frac{200}{1.12^2} + \frac{228}{1.12^3} + \frac{255}{1.12^4} + \frac{280}{1.12^5} + \frac{2{,}333}{1.12^5} = 2{,}120$$

$$PV_5 = \frac{172}{1.12} + \frac{200}{1.12^2} + \frac{228}{1.12^3} + \frac{255}{1.12^4} + \frac{280}{1.12^5} + \frac{4{,}200}{1.12^5} = 3{,}180$$

Now we need to deduct the existing debt from the value of the assets. Given that we are simplifying as much as possible, we will ignore the different approaches

Table 4.9 RV for different scenarios

	Growth	
Years	**0%**	**5%**
15	Scenario 1 \$1,907	Scenario 2 \$2,605
∞	Scenario 3 \$2,333	Scenario 4 \$4,200

we could opt for to carry this step forward and their effect on value. Here we will just deduct the current value of the debt in the balance sheet ($400 million) from the discounted value of the assets. In the upper portion of Table 4.10, we show the results for the liquidation in year six scenario; the remainder can be found at the bottom of the same table.

We have calculated the PV of the FCFs obtained from our operations under different RV assumptions and we have deducted the value of the debt to come to different estimations of value. Bearing this in mind, the summary of the range of values is presented in Table 4.11.

4.4. Methods used to assess the business value of corporations

Above, we provided a simplified example of how to value a company using three different methods, but there are many other ways to do so. Generally, business valuation considers the specific situation of the various financial aspects of a company based on the financial statements, ratios, and funding policy. Here we try to group and summarize the most common methods to value businesses.

Table 4.10 Blue Jacket, Inc. ($ millions)

Year		**1**	**2**	**3**	**4**	**5**	**RV**
EBIT		333	365	396	424	448	
Interest							
EBT		333	365	396	424	448	
Taxes		117	128	139	148	157	
EAT		216	237	257	276	291	
+Depreciation		57	64	71	78	85	
–Investment in working capital		32	31	30	29	26	
–Investment in fixed assets		70	70	70	70	70	
= FCF		**172**	**200**	**228**	**255**	**280**	**988**
PV at 12%	1.297						
–Debt	400						
Company's value	**897**	Liquidation year six scenario					
Other scenarios:							
PV_{s1} at 12%	1,879		PV_{s2} at 12%			2,275	
–Debt	400		–Debt			400	
Company's value	**1,479**		Company's value			**1,875**	
PV_{s3} at 12%	2,120		PV_{s4} at 12%			3,180	
–Debt	400		–Debt			400	
Company's value	**1,720**		Company's value			**2,780**	

Table 4.11 Final value for different scenarios without debt ($ millions)

Book value		**DCF**	
		Growth	
$415	**Years**	**0%**	**5%**
Liquidation		Scenario 1	Scenario 2
	15	**$1,479**	**$1,875**
$897		Scenario 3	Scenario 4
	∞	**$1,720**	**$2,780**

4.4.1. Methods based on asset value

The models that fall under this heading try to determine the value of the company by estimating the value of its assets. Here, we try to define these methods and succinctly review the ones that would most often be encountered by readers.

4.4.1.1. Book value and adjusted book value

The book value, accounting value, or net worth of a firm is estimated by subtracting the book value of the liabilities from that of the total assets of the firm. The result is the firm's equity. We provided a numerical example earlier in the chapter. Once changes are made to include or exclude certain adjustments in the accounting entries, and to the extent that these are more refined and comprehensive, the book value will come closer to the market or liquidation value of the firm. These adjustments may include assessing the value of fictitious or intangible assets such as patents and set-up expenses; the valuation of inventories; the assessment of the customers, adjusting for doubtful receivables; the valuation of real estate assets, as these may be far from the current ones; the obsolescence of machinery and plant and equipment, including comparisons of amortized percentages to see whether they differ from reality; and so on. The adjustments can be numerous and are all directed toward bringing the book value closer to the market value. Thus, a more realistic approximation to the value of a firm is already advanced with the concept of adjustments that can be done by checking every item of the balance sheet to restate them at their market price. The difference between the market value of assets and the market value of liabilities gives us the adjusted book value. In general, the main difficulty in this method lies in gathering the required information.

4.4.1.2. Liquidation value

This is the value that results from a hypothetical liquidation of the company in parts, in the event that its assets are sold and its debts canceled. As shown in the earlier example, the method can be summarized by subtracting the market value of the liabilities and contingent liabilities (staff redundancy costs, tax issues, etc.) from the market value of assets. Thus, we perform most of the adjustments (those mentioned in the previous section) to book value as, to the extent that progress on these adjustments takes place, the book value approaches the liquidation value. This applies to the liabilities as well which would be reflected in the price should we "sell" those liabilities at the time of valuation. The valuation should also be adjusted to consider any liability that may arise from the liquidation of the company, even when the objective of this valuation might not be the sale of the company.

4.4.2. Methods based on stock exchange ratios

These models use the income statement of the company to determine its value through its future profitability.

4.4.2.1. The P/E ratio

The P/E of a share determines the earnings multiple that is paid at the stock exchange per share. When it comes to valuing stocks, this ratio is one of the better known and most frequently used metrics. P/E stands for the ratio of a firm's share price to its per-share earnings. Thus, to estimate it, we just divide the current stock price of a company by its earnings per share (EPS):

$$\text{P/E ratio} = \frac{\text{Market value per share}^{11}}{\text{EPS}^{12}}$$

Most frequently, the P/E is calculated using EPS from the last four quarters, also known as the trailing P/E, although sometimes the leading or projected P/E is also used. This latter comes from the expected earnings from the next four quarters. Another variation is a mix, and uses the EPS of the past two quarters and the estimates of the next two quarters.

The P/E ratio fluctuates significantly through cycles and varies with the market, the sector, and among companies. Historically though, the US average has stayed in the range of 15–25. A P/E ratio of 25, for instance, is interpreted to mean investors are willing to pay $25 for every $1 of earnings generated by the company. For this reason it is called the "multiple" of a stock.

Although the EPS is most often calculated using historical data, the P/E is taken to reflect the level of market optimism for a company's growth, given the share prices in the numerator impound expectations. A higher P/E than average is taken as a positive sign of growth. Consequently, high growth companies have higher ratios on average. Thus, it is only appropriate to compare firms within the same sector: technological companies among themselves, utilities against utilities, and so on.

Most often, the P/E ratio does not convey much information; however, it can be useful to compare the P/E of one company with that of another in the same industry, to the market in general, or to the company's own historical P/E ratios. Normally, companies in the same sector have a similar P/E, that is, unless there are significant differences in future expectations which can then result in important P/E differences. Other things being equal, a company will be a cheaper buy relative to another one if it trades at a lower P/E. However, in reality the P/E is actually quite difficult to interpret as the information it conveys can vary in value with the situation. Additionally, the use of accounting figures complicates matters, particularly if one is attempting to make comparisons between stocks traded in countries subject to different accounting standards.

4.4.2.2. Dividend yield

Dividends are the portion of the net income (NI) used to pay shareholders. Most often, this is the single stream of funding stockholders receive when they retain ownership of the shares. The dividend yield is estimated by dividing the annual dividends per share by the price per share; so the ratio shows how much a company pays out in dividends each year relative to its share price. In the absence of any capital gains, the dividend yield is the return on investment for a stock.

4.4.2.3. Market capitalization

Rather than a valuation method per se, market capitalization is the result of multiplying the number of outstanding shares of a listed company by the price per share. It is the value of a company on the exchange, so it is sometimes referred to as "market value." A firm's market value conveys investors' perceptions of its prospects.

Market value fluctuates over time, depending on the business cycle and numerous other factors, such as the sector in which the firm operates, the amount of leverage, or the overall economic conditions. One way in which this metric is used is in comparing it with the book value or shareholders' equity of the same firm, or with other firms in the same sector. Most often, a stock would be considered undervalued if its market value is significantly below book value, which means the stock is trading at a deep discount to book value per share.

4.4.3. Methods based on discounted cash flows

DCF methods assess value in relation to the wealth a firm will generate in the future.

4.4.3.1. Dividend discount method

The dividend discount method is based on the idea that the value of a share of stock to a shareholder is the current value of the stream of future dividends (D) discounted at a rate (k) that considers the risk of such CF:

$$PV = \frac{D_1}{1+k_e} + \frac{D_2}{(1+k_e)^2} + \frac{D_3}{(1+k_e)^3} + \cdots + \frac{D_t}{(1+k_e)^t}$$

$$PV = \frac{D_1}{1+k_e} + \frac{D_2}{\left(1+k_e\right)^2} + \frac{D_3}{\left(1+k_e\right)^3} + \cdots + \frac{D_t}{\left(1+k_e\right)^t} \tag{4.6}$$

Formula 4.6 considers individual dividend payments. When there is an infinite number of future dividends to consider, some formulas simplify their estimation by proposing a constant dividend growth which may start at some point in the future, often after a few initial years of higher dividend growth. Constant dividend growth considers future dividend payments as resembling a perpetual income beginning with the first dividend in the "next" period (D_1), where g is a constant annual dividend growth rate.

$$D_1 = D_0 + g \quad \text{or} \quad D_1 = D_0\left(1+g\right) \tag{4.7}$$

Should we consider this to be perpetual growing payments, the formula would be modified thus:

$$P_0 = \frac{D_1}{k_e - g} \tag{4.8}$$

Finally, in addition to assigning a given value to the growth rate, this can also be adjusted by the formula: g = ROE × (1–payout), implying that future growth will depend on the return on equity (ROE) and the funds retained within the firm (1–payout) to remunerate at that rate.

The main objection to this method is that a company is not only worth what it distributes to shareholders, but also the returns it is able to extract from retained earnings. This method is reasonably more applicable to companies that have a high payout and is not applicable to those which do not make dividend distributions.

4.4.4. Method based on discounted expectations

The method based on DCF attempts to value the future earning power of a firm. This method is the one most widely used when valuing a company and the one which provides the best guarantee. We provided an example of this method above, with the valuation of Blue Jacket, Inc. A variation of this method considers equity cash flows (ECF), rather than FCF, and discounts by the cost of equity instead of the cost of capital.

4.4.5. Models based on shareholder value creation

The economic value added (EVA) focuses on the value-creating power of the management of the firm. To do so, EVA compares the profitability created by the company with the firm's cost of capital. To the extent that the former exceeds the latter, value is being created for the shareholder. If the EVA is positive, the company is getting higher returns on their investments than their cost of capital (that margin is precisely the creation of value); if, on the other hand, the EVA is negative, it is destroying value, even if the earnings and the CF retain a positive growth rate.

4.5. Components of fundamental analysis

Using financial ratios or any other valuation tool is the last part of the valuation process. Before getting to that point, one should understand that a firm is in an environment where multiple forces interact in very different ways, all of which impact the evolution of the business to some extent. Thus, to assess a likely evolution, one might look at a succession of analyses of the various contexts: international, national, macroeconomic, sector, institutional, and lastly, company focused. For instance, if we attempt to forecast the future of the automotive sector, one might want to start by looking at the macroeconomic national and international environment to identify the main variables that might affect the financial statements of the sector and their probable evolution.

The analysis of economic variables can be carried out in various ways, but it is always useful to distinguish between structural and conjunctural situations. Structural analysis refers to the basic variables of economic growth and the main components of supply and demand. Thus, to determine what brings greater growth at each point of time, some fundamental aspects would refer to the components of GDP, assessing the weight of the external sector by its behavior, and domestic demand (especially private consumption and investment). Moreover, the behavior of the public sector is also extremely important: both public consumption and public investment can define an investment strategy in the sectors in which the execution is envisioned.

The conjunctural analysis is related to the variables present in the short term. It focuses on aspects such as current inflation, keeping in mind the effect it might have on interest rates and the evolution of sales, wages, and general costs; central bank monetary policy, that is, the growth of money supply and domestic credit which can overheat the economy and impact inflation and interest rates; the evolution of the secondary bond market which is normally measured in terms of differentials relative to the bonds of the most advanced economies such as those of Germany or the USA; and exchange rates against major currencies like the yen, the euro, or the (US) dollar. Other variables of interest may include the balance of payments and its funding: trade and current account deficit, fiscal policy, public deficit, taxation, and income available for consumption and savings, as well as the promotion of certain activities and sectors, such as the housing market or the labor market (flexibility, evolution of unemployment, unemployment coverage, measures to promote employment, etc.).

The analysis must not only focus on the estimates and evolution of the stated variables, but also, in the context of today's world, on the international comparison with regard to variables such as cost structure, technology, market position, and barriers between markets which is necessary and essential to qualify and quantify the company's growth strategy. Likewise, the regulatory and institutional framework of the sector must be analyzed as well.

Another important point is the position of the sector with respect to the economic business cycle. Each sector has its greatest growth as a function of the overall economic situation. More importantly, in each economic period, there is a sector that will offer better prospects. Sector rotation is an investment strategy that shifts investment assets from one sector of the economy to another to take advantage of these cycles. This strategy involves the sale of securities related to a particular investment sector to purchase securities in another sector.

Finally, all the specific variables that affect the company under the macroeconomic and sector analysis combined with the strategic/competitive analysis (i.e. Porter analysis) are treated before getting into the financials of a valuation. In this context, the following aspects and their possible evolution should be studied: barriers to entry (economies of scale, goodwill, cost of entry into the sector, distribution networks); power of suppliers (an excessive concentration of orders with few suppliers can jeopardize the margins of production); power of customers (it is equally dangerous to compromise a high percentage of sales by concentrating on a small number of customers); competition (industry growth, fixed cost structure, overcapacity, product differentiation, concentration, diversity of competitors); and substitutes (evolution of the price of substitutes, quality, costs of switching by customers, propensity to change). Additionally, one could consider other variables which are harder to quantify, such as a change in the management team of a company, or political change in a country.

4.5.1. Sensitivity analysis

In our example, we used the company's growth forecast to extract FCF from expected accounting earnings. In addition, a sensitivity analysis should also be performed. This analysis would study the impact of changes on key variables that could potentially affect the income statement of the company in a significant manner. The significance of these variations would be scrutinized to establish a range in value and risk exposure, but it is equally important to establish actions in the event that the situation differs from that originally expected. This type of analysis is useful in instances when firms are particularly dependent on a few variables, or when these variables are extremely volatile.

4.5.2. Seasonal analysis

When the contribution of a quarterly income to the yearly total income of a company is comparatively large, it can be concluded that there is seasonality. Seasonality may affect the share price given the reported quarterly results.

4.5.3. Nonfinancial analysis

Other aspects to consider in fundamental analysis refer to the expected behavior of the firm itself with respect to its products, competition, management model, the medium- and long-term strategy, and so on. For instance, with respect to the products offered by the company, among other variables, one should analyze the technology, the possible evolution of the market share, the maturity cycle of the products, and the experience curve. A second important aspect is that of their comparison with the competition. This is a very useful tool in the analysis as it helps draw conclusions about the position of the company, the possible evolution of its costs and market share, the price/quality ratio of its offering, the degree of international development of the sector, and the degree of fragmentation of the sector relationship.

4.5.4. Real options

DCF implicitly assumes that firms hold real assets passively. It ignores the options found in real assets – options that sophisticated managers can use to take advantage of opportunities. The DCF method does not reflect or pick upon the whole value that management can create, as when developed for bonds and stocks; investors in these securities are necessarily passive in the sense that one investor can be replaced by another. If we consider the firm as an investor in real assets, the management can add value to those assets by responding to changing circumstances. The management has the opportunity to act because many investment opportunities have real options embedded in them. In a simple sense, DCF misses this extra value because it implicitly treats the firm as a passive investor. Thus, in valuing a business one should consider the value of real options as well.

The following are three examples of real options:

a. Companies often cite strategic value when taking on negative net present value (NPV) projects, given the option to make follow-on investments. A close look at the project's payoffs may reveal a call option or follow-on projects, in addition to the immediate project's CF. That is, today's investments can generate tomorrow's opportunities.
b. The option to abandon a project provides partial insurance against failure, thus, the option to abandon has value. This is a put option: the put's exercise price is the value of the project's assets if sold or shifted to a more valuable use.

c. A third case is when one has the option to wait (and learn) before investing. This is equivalent to owning a call option on the investment project. The call is exercised when the firm commits to the project, but often it is better to defer a positive NPV project in order to keep the call alive. Deferral is most attractive when uncertainty is great and immediate project CF, which are lost or postponed by waiting, are small.

Real options encountered in practice are usually very complex. For example, with respect to the option to abandon a project one may be able to bail oneself out of a project at any time rather than on a single occasion. Estimating the value of this option can be complicated as, for instance, the price at which you can bail out is likely to change over time and would rarely be known in advance. Once you have abandoned it, you may be able to reinstate the project if business improves. Handling these complexities requires a large number-crunching computer and modeling the problem requires informed judgment.

The option pricing theory can also be applied to both bond and stock valuation. Black and Scholes (1973)[13] noted that corporate securities may be viewed as options (or contingent claims) on the value of the firm. A contingent claim is a security whose payoffs depend on the prices of one or more other securities. Thus, a stock option is a contingent claim whose value depends on the price of the stock on which it is written. Contingent claims analysis is the general term for the techniques used to value options and other securities whose payoffs depend on the values of other assets or securities. This analysis is applied to the pricing of corporate securities (including debt and equity), and the evaluation of capital investment opportunities.

All the corporations' securities can be viewed as contingent claims or options whose values depend on those of the firm's assets. This recognition has led researchers to delve into the pricing of complex securities, such as convertible and callable bonds using contingent claims analysis. Yet another subset of this third strain relates to developing research on the use of contingent claims analysis to evaluate "real asset investment opportunities." Such analysis generally involves conceptualizing the determinants of the value of an investment opportunity in a productive asset utilizing a portfolio of options.

4.5.5. Result and evaluation of the analysis

4.5.5.1. Management model

When looking into the future, as we do when attempting to value a business, it is important to observe if the firm's management model inspires confidence and trust: Can their model adapt to new challenges and markets? Is it effective in handling change and setting the right corporate structure? Can they shape their corporate image and communication strategy in accordance with the times? Can they stay ahead of their sector? Are they capable of diversifying risks? Do they have the capacity and readiness to generate long-term efficient strategies that are cost effective and credible? And finally, can they successfully implement the plans proposed in these strategies?

4.5.5.2. Medium- and long-term strategy

A company has to have clear and coherent objectives and be able to concisely explain these to the market. Thus, an obvious undesirable trait is to have a high "turnover" in corporate strategies. This finding would not be evidence of the firm's ability to adapt but rather of poor management. This does not mean that once a strategy is set it cannot be changed or varied in the short term as the strategy should be able to adapt itself to the concrete market situation, products, and competitors. However, the objectives of that strategy should always be clear. Fundamental analysis will assess the firm's strategy and evaluate the degree of compliance achieved and its impact on the value of the corporation.

4.5.5.3. Final comments on valuation

We have estimated the value of Blue Jacket, Inc. using various valuation methods and a series of hypotheses. As a result, we have obtained a range of values which depend upon our stated expectations of future growth. At the lower end of this range is an accounting value of $415 million, and at the upper end, in the scenario of 5 percent perpetual FCFs, is that of $2,780 million. So time and growth are two of the fundamental hypotheses we have to be mindful of when valuing a business or when evaluating the validity of the assessments made by others. The third variable which can alter value significantly is the discount rate which refers to the cost of the financial resources. So if instead of discounting by 12 percent, we discount by 10 or 25 percent, our results would be completely different. It is only logical then that for every business environment and business strategy, we create expectations, and thus assumptions, that fit the reality. For instance, if we have a business which produces uniforms for the military and our contract elapses in five years with no hope of renewal, it would be a mistake to assume future CF similar to those of the past. If we have a food distribution business in an area with a negative population growth it would be a mistake to assume the growth of other regions. That is, our assumptions have to make "business sense" at the time of the valuation.

To understand if the price of a share deviates from that which can be justified by the firm's fundamentals, a clear understanding of the value-generating factors and valuing technology is deemed necessary. Nonetheless, as already said, the issue is complex because none of the various assessment methods resolve all the problems that could come up in different types of businesses and circumstances. Furthermore, when bubbles are present, even balance sheet valuation is affected: the asset values and amount of borrowing is affected, the growth expected in the income statement does not come through, and thus the retained earnings are not there, so the whole thing is out of sync.

In addition to all that has been mentioned, the value of a business will be affected by the position of the party making the valuation. Whether it is because of the participant's level of optimism, or aversion to risk, or any other factor, all factors will influence the hypotheses used and the assessment process.

Ultimately, it is also true that "things" have varying worth to different people. However, all in all, although assessing the value of a business is a difficult and inevitably subjective process, as each business owner will view its value differently (e.g. the owner of a single shoe store might value another shoe store differently from the owner of a chain of shoe stores). So the key factor to keep in mind is that the value of an ongoing concern depends on the implementation of a business strategy and any hypothesis used in the process ought to respond to plausible outcomes.

The analysis performed in the valuation of fundamentals assumes the stock exchange of a country reflects the development of the key macroeconomic variables of that nation and its economic environment. These variables (interest rates, growth, etc.) are the fundamentals of the economy. Correspondingly, the price of a share of stock of a corporation must reflect the fundamentals of a business (expected earnings, CF, market positioning, etc.). The fundamental analysis is a way of "reasoning" the evolution of the share prices on the stock markets. Investors decide what happens in the long term based on short-term trends that need not be justified. Thus, analysts try to anticipate the future behavior of the listed securities, so that investors can make or break their positions before it is too late.

Fundamental analysis, then, looks at the firm's likely development considering the corporation's business plan. If such analysis is done in a knowledgeable and conscientious manner considering true and likely probabilities, errors, at least hypothetically, will be limited to those intrinsic to the market, such as deviations due to the systematic risk inherent in the economy. Market inefficiencies are the basis for this approach, as the investor intends to profit from them. Thus, the fundamental analysis continually seeks to uncover misvaluations assuming that sooner or later the market will converge on the correct rating.

In fundamental valuation, it is the analysis per se which is the primary objective. Thus, all conclusions and tools used in this process are subordinated to this goal: it is from the firm's forecasted performance that a share price is established to make investment decisions. The foundation for these decisions is the expected future behavior of the firm and, thus, its shares. The idea is not only to conclude the value the share should reach in a given period of time, but also the corporation's medium-term objective prices. That is, it is not only important to determine whether a firm will present good growth over the coming years, or whether its financial structure will improve, but also, once this is known, the price at which it is interesting to buy or sell the security. Therefore, one of the essential tools is to develop projections or assessments of future CF derived from certain business plans under specific economic scenarios. That is, forecasts must be done by bringing together all available information, so that the behavior of the firm could be anticipated and an evaluation of the company could be performed.

In principle, the fundamental analysis is not focused exclusively on shares, but also considers derivatives, fixed rent, credit risk, and so on. Hence, fundamental analysis can be described as a systematic analysis of all the information available in the market which affects the said asset and is undertaken with the objective of making good investment decisions.

4.6. The fundamentals of economic value creation

Business practices derived from management procedures designed to increase profits, such as methodologies based on process re-engineering, total quality, teamwork, and so on, all too often fail to establish a clear relation to economic value creation. For instance, objectives are established to create value in specific quantities and set periods, and slogans and corporate images are designed to reinforce the executive team's commitment to this objective. Nonetheless, economic value creation is much more complex than merely stating wishes and setting some well-intentioned goals.[14]

So, what is economic value creation? Intuitively we could say that creating value results in an increase in the worth of a firm's equity. Hence, designed objectives and processes which result in the attainment of this goal should fit the bill. However, a few examples might help us analyze this matter in greater depth:

1. In 1989, the Nikkei index was close to 40,000 and its P/E near 70. Seven years later, the P/E had fallen to 25 and the stock market index to 15,000.[15] How did several billion US dollars vanish during this period? What was the logic behind it? Could it be a coincidence, or were external factors involved?
2. In December 1995, General Electric's (GE) total market capitalization stood at $113 billion and its P/E was 17, while that of the S&P 500 was 19. In this environment, Jack Welch, GE's managing director, believed the company was undervalued and decided to announce a repurchase program involving some $10 billion of shares.[16, 17] By July 1998, after repurchasing $5 billion dollars' worth of shares, it had increased to $310 billion.[18] Could the additional $200 billion be the result of a specific communication policy? Is it possible to manipulate information so that the market will value shares in a certain way? Is this value creation sustainable in time?
3. In January 1999, just like every year, the well-known financial analysts in Spain published a report containing the expected evolution of the main

shares quoted on the stock market. One of the reports, on a well-known Spanish business group, summed up its forecast for the new year as follows[19]: sector under the process of liberalization with a strong increase in competition; expected loss of market share: 20 percent; 15 percent reduction in prices; no dividends would be distributed in the next two years; and large investments in Latin America, which are of unknown quality at present. Still, the analysts' recommendation was to purchase these shares. How could this conclusion be correct?

4. Lastly, in February 1999, Dell announced its 1998 results, which included a 48 percent growth in sales and a net profit of 55 percent.[20] Despite this apparently good news, the company shares fell by about $6.[21]

These examples enlighten us with respect to the logic which underlies value creation. Indeed, it is believed that if a company creates value by making sound decisions, financial issues will work themselves out naturally, and markets will value the business appropriately. But, to what extent is this true?

Economic profit is created when the capital invested in a firm generates a real profitability above and beyond the cost of the resources it employs. Economic profit has three essential components: the cash flow generated by the business, the future expectations for the business, and the related business risks. To create real wealth it is necessary to take into consideration the three components of economic profit. However, acceptance of this approach implies considering that the main objective of a company is to maximize its value. But is that true?

Given our discussion in earlier chapters, we could say that together with creating value (which is a different discussion), a key and primordial company objective is to perpetuate itself as a living organism. However, we shall put that aside for the time being and instead analyze the fact that the notion of economic value as economic profit leads to numerous consequences with real-world impact.

4.6.1. Accounting profit and economic value are two different things

If value creation is obtained through the creation of economic profit, then accounting profit has a number of limitations, to list a few, accounting profit provides information about the past, not the future; it does not refer to CF, but to book profits and losses; and it does not include any type of adjustment for inherent risks. Furthermore, due to the use of different accounting standards when analyzing the same business situation, the results might differ. These differences may be explained using various accounting criteria to reflect a number of aspects such as business mergers, deferred taxes, pension funds, foreign currency exchange, book reserves, appreciation of assets, financial derivatives, leasing, depreciation expenses, and intangible assets.

In addition, another source of uncertainty when using book value to measure value creation is the possibility of the data having been manipulated. One of countless examples published in the press is of Sony. On November 17, 1994, Sony announced it was going to readjust $2,700 million as a result of undeclared losses by its subsidiary, Columbia Pictures.[22] These losses resulted from films which were failures, for instance *Geronimo* and *The Last Great Hero*, but that had been camouflaged by the box office successes of others such as *Philadelphia* and *Sleepless in Seattle*.[23]

Events of this kind are so common that in September 1998, Arthur Levitt, president of the Securities and Exchange Commission (SEC), complained of the excessive use of auditing "tricks" US companies indulged in. In reference to company rectifications announcing losses when they had previously declared profits, he said: "When I see these errors come to light, I can't help wondering whether professional auditors are sufficiently trained and supervised in order to ensure that audits are correct."[24]

Even though extreme opinions can be heard as a result of such occurrences, such as "published book profits are accounting fiction,"[25] it could be justifiably argued that, when correct, book profits are often the result of an agreement on accounting norms. In reality, markets react to accounting information published in the press and accounting profit is an indicator of economic profit.

4.6.2. Growth does not equate to economic value creation

There is plenty of evidence to show that sales growth does not guarantee value creation and may even, on the contrary, indicate a loss of wealth.[26] One such case was that of Compaq, which in 1998 increased its market share from 13.1 to 13.8 percent[27] while the value of its shares decreased by 50 percent over the ensuing months.[28] Obviously, the objective is not to sell, but to do so while creating economic profit.

4.6.3. Value creation requires a reference to the future and implies risk

Strictly speaking, economic value creation takes place in the future as a result of putting present-day strategies to work. However, historical information provides clues with respect to the management team's capacities and the results derived from the adopted policies. Financial data helps generate a decision-making system that creates economic value. For example, when a sales prevision is drawn up, this estimate will be the result of a series of reasonable assumptions and strategies which will be acted upon to ensure the achievement of objectives. In other words, actual sales cannot be forecasted, but the expected sales resulting from a series of actions can be approximated.

The first concern when forecasting is to evaluate the assumptions and check the coherence of the plan to expose contradictions and ensure that it leads to the creation of economic value. Examples of typical errors are to give excessive weight to the most recent tendency; or the inclination to maintaining the status quo, such as the propensity to keep things the way they are; or to think in terms of what is already familiar.

Given that creating value is associated with the future and the future is uncertain, we need to accept that economic profitability is tied to risk. There are different types and intensities of risk and these vary with business and external factors. However, two major types of risk are those which result from operative business conditions and those which result from business financing conditions (i.e. the structure of the liabilities).

4.6.4. Value creation is not an easy or short-term task

Good management performs actions which create value sustainable in time. A study published in 2002 shows that over the period of analysis, none of the top US companies in terms of shareholder value creation matched the firms with the biggest increase in sales, net profit growth, or growth in equity profitability.[29] Maximizing growth does not equate to maximizing corporate profitability or shareholder value. On the contrary, in a number of studies, companies with moderate growth in sales or earnings show the highest rates of return and value creation for their owners.[30] It is also important to understand that sustained value creation is not a short-term task, with regard to either design or results. Thus, it is necessary to uncover instances when management teams act according to their wish for short-term results rather than long-term vision.

4.6.5. The role of communication in value perception

Top executives frequently declare that the market undervalues their companies. If markets are moved by value creation expectations, it is important to analyze whether there is a difference between real created value, that is, possible value which may be created, and the real and possible value which the market perceives. One should analyze whether a source of disparity is a matter of communication capacity on the part of the firm, rather than its capacity to create value.

5
Bubbles and Technical Trading

5.1. Introduction to technical analysis

Technical analysis is a security analysis technique based on the assumption that stock prices move in trends which can be used to forecast the future direction of securities.[1, 2] So, contrary to the main valuation techniques reviewed in Chapter 4, technical analysts do not attempt to measure a security's intrinsic or fundamental value. Instead, they dig into the historical trading path of a financial asset to find patterns that might suggest future activity. This choice of method is substantiated by the technicians' credo that prices in financial markets reflect all relevant economic factors affecting companies; hence, it is sufficient to study share prices and volume information alone.

Price refers to any combination of the open, high, low, or close for a given security over a specific time period which, taken together, reflect forces of supply and demand. By looking at price development over a period of time, one can appreciate how supply and demand forces unfold. Higher prices reflect increased demand, when the number of buy orders exceeds the number of sell orders, whereas lower prices reflect increased supply, when the number of sell orders exceeds that of buy orders.

The time frame for price data can be intraday (1 minute, 5 minutes, 10 minutes, etc.), daily, weekly, or monthly; hence, a given price may last for a few hours or many years. The intraday high reflects the strength of the buyers, the demand side of the market, while the low reflects the availability of the suppliers or sellers. The close represents the final price at which the market ended business for that day.

As stated, what distinguishes technicians from their fundamental counterparts is their exclusive reliance on historical prices, open interest, and volume data to assess the future. These, though, are not the only sources of information available to technicians given that other data, such as indicators and sentiment analysis, are also considered. Different technical traders use different tools.

Some rely on chart patterns; others use technical indicators and oscillators or combinations thereof. Chartism requires familiarity with multiple patterns and has its own lingo. For example, outside of that community who would know what a *dead cat bounce* means?

The analysis of charts that show prices over time is crucial to the study of markets. Within these charts, technicians especially search for patterns, such as head and shoulders or double top reversal, and study indicators such as moving averages. Technicians also study charts to look for forms such as lines of support, resistance, channels, and more obscure formations like flags or balance days. Given the important role of chart analysis in technical analysis, technicians are also called chartists.

5.1.1. Fundamental versus technical analysis

As mentioned above, technical analysis holds that prices in financial markets already reflect the economic factors that influence companies; hence the study of share prices and volume information alone. Thus, unlike fundamental analysis which looks at the actual facts regarding the company of interest, technical analysis does not analyze the firm's nature. Rather, it looks at how the market reacts to it and bases its recommendations and actions solely on the charts of the said information. In other words, technical analysis attempts to understand the emotions at play in the market by studying the market itself, as opposed to its components. Furthermore, technical analysis does not result in absolute predictions. Instead, it assists investors in anticipating what is "likely" to happen to prices over time. The ultimate aim of technical analysts is to detect early price patterns, trends, and changes in buying if the shares are expected to increase in value, or selling, if shares are expected to decrease in value.

The basic difference between fundamental and technical analyses is that a technical analyst approaches a security from the charts, while a fundamental analyst starts with the financial statements. By looking at the financial statements among other data, a fundamental analyst tries to measure a company's intrinsic value. In this approach, if the price of a stock trades below its intrinsic value, it is a good investment. On the other hand, technical traders believe there is no reason to analyze a company's fundamentals because these are all accounted for in the stock's price. Technicians believe that all the information they need about a stock can be found in its charts.

With respect to the time horizon, as compared to technical analysis, fundamental analysis takes a relatively long-term approach to analyzing the market. While technical analysis can be used over a time frame of weeks, days, or even minutes, fundamental analysis often looks at data over a number of years. The different time frames result from the nature of the investing style: it can take a long time for a company's value to be reflected in the market, so when a fundamental analyst estimates intrinsic value, a gain is not realized until the stock's market price rises to its "correct" value.

Furthermore, financial statements are filed quarterly and changes in EPS do not emerge on a daily basis like price and volume information does. Also, "fundamentals" are the actual characteristics of a business: it takes time to create new products, marketing campaigns, supply chains, and so on. This is partly the reason why fundamental analysts use a long-term time frame.

Not only is technical analysis more short term in nature than fundamental analysis, but the goals of a purchase (or sale) of a stock are usually different for each approach: in general and again in very simplistic terms, technical analysis is used for a trade, whereas fundamental analysis is used to make an investment. Investors buy assets they believe will increase in value, while traders buy assets they believe they can sell to somebody else at a higher price. The line between a trade and an investment can be blurry, but it does characterize a difference between the two approaches.

Although technical analysis and fundamental analysis are regarded as representing opposing market philosophies, market participants often combine the two. For example, some fundamental analysts use technical analysis techniques to determine the best time to enter into an undervalued security. The hope is that by timing the entry into a security, the gains on the investment may be increased.

Also, some technical traders might look at fundamentals to add strength to a technical signal and many blend pure technical analysis with other methods. For instance, fusion analysis,[3] such as the Bollinger approach, mixes fundamental and technical analysis. Technical analysis also approaches a mixture of quantitative analysis and economics, as with the use of neural networks to identify intermarket relationships.[4]

Efforts to decipher the markets' future do not fall short of blending financial astrology and technical analysis either, as proposed in Chris Carolan's "Autumn Panics and Calendar Phenomenon,"[5] demonstrating how technical analysis and lunar cycles may be combined.[6] Of course, as will be explained later in the second volume of this work, much of the calendar-related phenomena (i.e. January and Monday effects) have been shown to relate to other reasons such as tax and accounting facts. Public news and company-related information, such as investor and newsletter polls and magazines, also cover sentiment indicators used by technical analysts.[7]

5.1.2. Brief history of technical analysis

In the early eighteenth century, a rice trader named Homma Munehisa began to develop the method which would later evolve into our current candlestick technique.[8, 9] By the end of the nineteenth century, cofounder of Dow Jones and editor of *The Wall Street Journal*, Charles Dow, also began to develop modern technical analysis methods. Looking at primary, secondary, and tertiary trends and rebounds or falls within them, Dow helped decipher the market "buy and sell signals." Nonetheless, rather than focusing on specific securities, Dow used the average of a portfolio of values. Three of those indicators have remained to this day: Dow Jones Industrials, Dow Jones Transportation, and Dow Jones Services. Of the many theorems put forth by Dow, three stand out: price discounts everything, price movements are not totally random, and "what" is more important than "why." These can be summarized as follows:

1. Congruent with the strong and semistrong forms of market efficiency, technical analysts posit that the current price of a financial security fully reflects all relevant information, including fundamental factors. Thus, rather than considering these factors separately, the belief is that a company's fundamentals, along with broader economic factors and market psychology, are all priced into the stock. Consequently, this only leaves the analysis of prices and their movements, a product of the supply and demand for a particular stock in the market. The information captured by the price is then used to interpret the market's view of a financial security and form expectations regarding its future behavior.
2. The basis of technical analysis is that price movements follow trends. Given that technical analysis can be applied to many different time frames, it is possible to spot both short-term and long-term trends. However, technicians also acknowledge that there are periods when prices do not trend, and their task is to identify them.
3. The price of a security results from the equilibrium between supply and demand. By focusing on forecasting the direction of the future price, technicians are not concerned with likely explanations for these movements. Simply put, if the price goes up, it is because there are more buyers than sellers, and vice versa. In their minds, market participants react in similar manner under equivalent market stimuli. This repetitive nature of price movements is attributed to market psychology.

Many other pioneers have helped develop this field further. This includes Ralph Nelson Elliott and William Delbert Gann, who proposed their respective techniques in the early twentieth century. Many more technical tools and theories have been developed and enhanced in recent decades. Adherents of different schools (e.g. candlestick charting, Dow Theory, and Elliott Wave Theory) may ignore the other approaches, yet most traders combine elements from several schools.

5.1.3. Criticism of technical analysis

Even though technical trading is a common practice, the approach used by chartists has always aroused suspicion among scholars. For instance, during the

1960s and 1970s, this method was widely discredited by academic mathematicians and others who argued the past does not imply future developments and that the evidence in support of technical analysis is, at best, inconclusive.[10]

Even today, many academics consider technical analysis a pseudoscience, alleging a lack of rigor and conceptual basis.[11] Critics argue that these "patterns" are random effects on which traders impose causation. Detractors of technical analysis include well-known fundamental analysts, such as Warren Buffett who has been credited with saying: "I realized technical analysis didn't work when I turned the charts upside down and didn't get a different answer."[12]

Regardless of the irony and criticisms, technical analysts hold that fundamental analysis cannot explain certain market movements.[13] Supporters claim that the system works, and this technique helps identify trading opportunities.

The argument that forecasting based on historical observations is unreliable is a powerful one. However, we can also think of this from a different angle. The forces of supply and demand driving prices depend on the decisions of investors. These traders have "memory," and their behavior may be influenced by past events. We can assume, therefore, that at least to some extent, the behavior of markets and price developments are influenced by this "collective memory." Furthermore, recalling our discussion on reflexivity and the self-fulfilling prophecy, we can assert that regardless of what could or not be true with respect to charting clairvoyance, once people agree on an upcoming event, they can make it happen. This assertion will be better understood by the end of our discussion on technical trading or chartism; however here we provide an example.

Let us assume a trader identifies a level of resistance that has been tested repeatedly. As the security approaches this level once again, the trader might decide to sell, given that in her view, it is very unlikely prices will move above the current level. The fact that she and all other technical traders interpret this information in the same manner has trading consequences and thus price and volume effects.

5.2. The basics of technical analysis

Technical analysis assumes that stock prices reflect the aggregate knowledge of all market participants. To buy or sell, these agents discount all relevant information and agree on a given price at each point in time. Thus, analysts then focus on the bottom line: What is the price? Where has it been? Where is it going? To answer these questions, technicians inspect graphs and charts, looking for patterns or indicators, such as mathematical transformations of price or volume data to determine if a security is trending and its price direction. Technicians inspect lines of support, resistance, channels, flags, and pennants, among many other factors. In addition, they study relationships like those between prices and volume, and correlations between changes in options (implied volatility) and put/call ratios with price. Given that no indicator considered in isolation can provide a complete picture of the market situation, the final diagnosis is the result of multiple analyses. Below, we summarize the most basic tools used by technical analysts and classify and explain the terminology.

5.2.1. The chart in technical analysis

Charts are graphical representations of price series. A chart pattern is a distinct formation on a stock chart that creates a trading signal, or a sign of future price movements. Chart pattern reading is based on the assumption that history repeats itself: certain patterns may be seen many times, and these patterns signal a certain high probability move in a stock, thus giving traders signals to buy or sell. In Figure 5.1, we provide an example of a *line chart* showing a stock's price movement over a one-year period, where each point on the graph represents the closing price for each day the stock is traded.

To help interpret information correctly, charts include relevant data. The time scale, which runs horizontally (x-axis), refers to the range of dates to which the chart refers. The shorter the time frame, the more detailed the chart. Each data point may represent the open, high, low, and close price of the period, depending on the chart used. The price of the security is shown on the y-axis. By looking at the graph, we observe a general upward trend and a stock value increase, with prices around $93 at the beginning of 1999 and reaching $105 by May 2000.

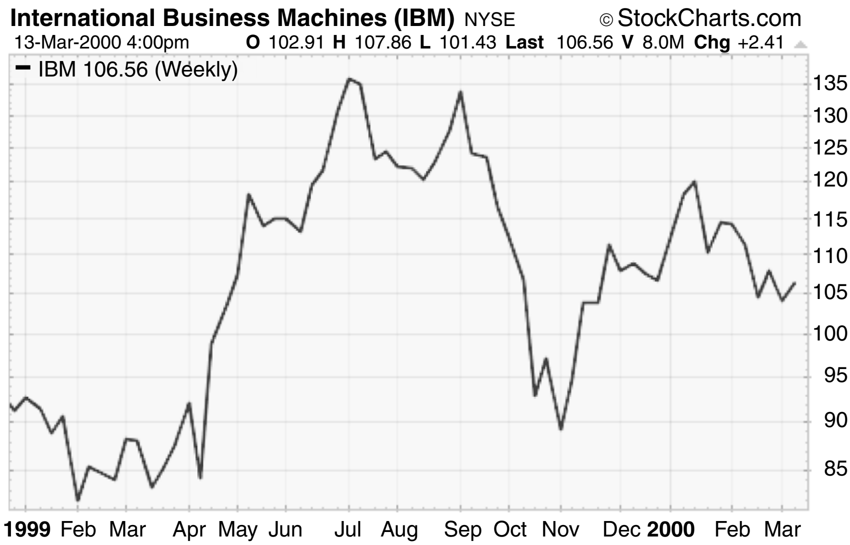

Figure 5.1 The line chart

Source: Chart courtesy of StockCharts.com

Most often one encounters charts with one of two different price scales: linear or logarithmic. If a price scale is built using a linear scale, the space between each price point (say 10, 15, 20 . . . as in our example in Figure 5.2) is separated by an equal amount. Thus, price changes such as 10 to 15 or 40 to 45 represent the same distance on the chart and the scale exhibits incremental values evenly on the *y*-axis. A $5 movement in price will look the same whether it goes from $10 to $15 or from $40 to $45. On the other hand, a semilog scale does not show incremental values in absolute numbers, but in percentage terms. So in a logarithmic environment, the distance between points is equal in terms of percentage changes. A move from $5 to $10 is a 100 percent gain and would appear to be a much larger change than a move from $35 to $40, which only represents a change of less than 15 percent. Figure 5.2 provides an example of each price scale.

There are several relevant considerations relative to the analysis of trend lines. The first is "validation," a rule used to confirm the existence of a suspected trend by taking into consideration a third point to validate a trend previously drawn with two other points. The second refers to the "distance" needed between data points to ensure the relationship among prices is relevant. In general, the space between the lows in an upward trend and the highs in a downward scenario will depend on the time frame under inspection and the price volatility. However, these points should not be too far apart, or too close to each other. If they were too distant, one could question whether there were any relationship between them at all.

Charts can be represented in different ways depending on the information that investors need to analyze. The four most common kinds are the line chart, the bar chart, the candlestick chart, and the point and figure chart (P&F). Figures 5.1 and 5.2 provide examples of *line charts* formed by connecting the shares' closing prices over a period of time.

The *bar chart* provides additional information over and above that of the line chart. In the lower part of Figure 5.4, we present an example where the vertical lines represent the high and low for the trading period, and the horizontal

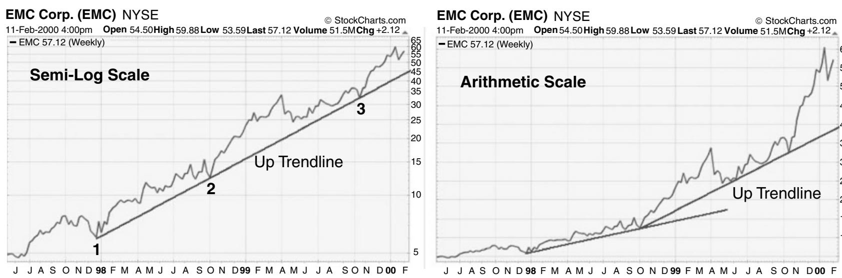

Figure 5.2 Logarithmic (log) price scale and linear (arithmetic) price scale
Source: Chart courtesy of StockCharts.com

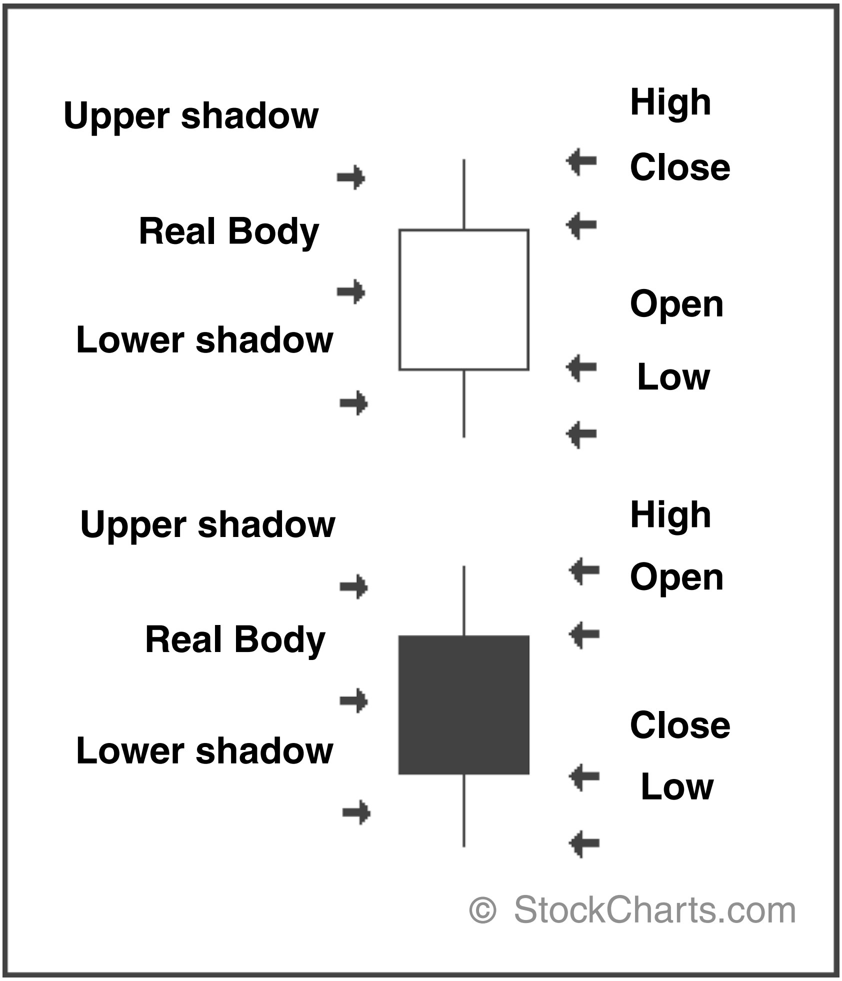

Figure 5.3 Candlestick formation
Source: Chart courtesy of StockCharts.com

dashes the close (on the right) and open (on the left). When the stock has gained in value, the horizontal dash will be shaded black; it will be red if the value has decreased over that period.

The *candlestick chart* resembles the bar chart except that differences between open and close prices are represented by a wide bar on the vertical line. For this chart, one needs the open, high, low, and close values for the period under consideration. The hollow or filled portion of the candlestick is called the "body." The "shadows" (also known as "wicks" and "tails") are the straight lines coming out of both ends of the body. Each body has two shadows. The

Figure 5.4 Candlestick chart versus bar chart
Source: Chart courtesy of StockCharts.com

upper end of the upper shadow marks the high, whereas the lower end of the lower shadow shows the low. The hollow candlestick shows that the closing price is higher than the opening price, with the bottom of the body representing the opening price and the upper, the closing price. On the other hand, a filled candlestick is drawn when the stock closes lower than its opening price, with the top of the body representing the opening price and the bottom of the body representing the closing price. A sketch of this scheme is shown in

Figure 5.3 and a comparison between the bar chart and the candlestick chart is provided in Figure 5.4.

The P&F removes noise and tries to neutralize the skewing effect of time on chart analysis. This is accomplished by reflecting price changes with no reference to time and volume. An example is provided in Figure 5.5. There are four things we can see on a P&F: support and resistance levels, and upward and downward trend lines. Each chart has a setting called the box size which defines the price range for each box. Each price must reach a value to warrant an X or an O, and the charts evolve as price changes occur. The upward price trends are shown by a series of Xs, while the downward trends are indicated by Os. Thus, no movement in price means no change in the P&F. Each chart has a second setting called the reversal amount that determines the amount a stock needs

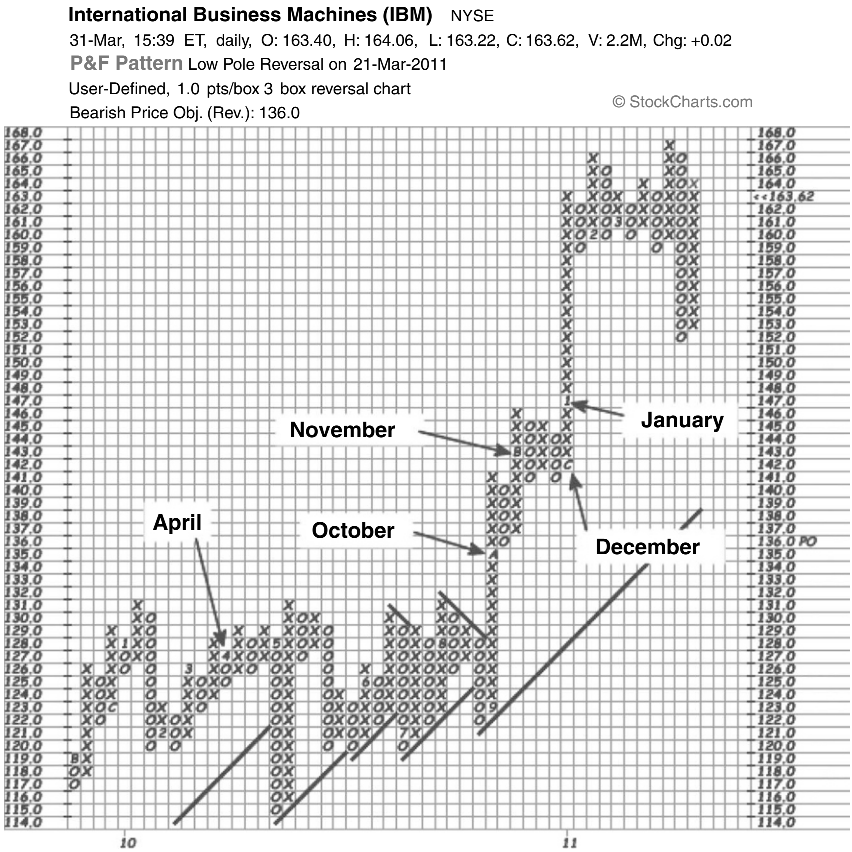

Figure 5.5 A point and figure chart

Source: Chart courtesy of StockCharts.com

to move in the opposite direction to warrant a column reversal. Whenever this reversal threshold is crossed, a new column is started and moves in the opposite direction. For instance, in a three-box reversal chart, column reversals are further filtered requiring a three-box minimum to reverse the current column. When the price trend has moved from one trend to another, it shifts to the right (signaling a trend change).

5.2.2. The trend

The *trend* line is a charting technique that draws a straight line connecting price points through time and then extends into the future. Understanding and identifying trends is crucial because they define the general direction followed by the price of a security. In addition, these lines are used to identify trend reversals. However, sometimes the trend is not so easily determined, especially when prices follow a series of highs and lows. Thus, in technical analysis, the movement of the highs and lows constitutes a trend in itself.

There are three types of trends: uptrends, downtrends, and sideways or horizontal trends. An uptrend line has a positive slope, showing that the net demand is increasing even as prices rise. The uptrend is considered solid so long as prices stay over this line. If there is a break below it, it would be a sign of weakening demand and indicate a likely change in trend. The reverse applies to a downtrend line. When there is little movement up or down, it is considered a sideways or horizontal trend. Figure 5.6 shows an example of an uptrend.

Figure 5.6 Upward trend

Source: Chart courtesy of StockCharts.com

Any trend direction can be identified as a long-term or primary major trend, an intermediate or secondary trend, or a short-term or tertiary trend. A major trend lasts longer than one year, an intermediate trend between one and three months, and a short-term trend less than a month. Figure 5.7 gives an idea of what these three trend lengths might look like. In this chart, we also see a trend line drawn at the lows of an upward trend, representing the support the stock has every time it moves from a high to a low. The line highlights the point at which agents might be able to anticipate when a stock's price will begin to move up again. As we can see, the short-term trend line has a steep slope, the medium-term trend line has a moderate slope, and the long-term trend line has a significantly gentler slope.

A *channel* is the addition of two parallel trend lines that highlight areas of support and resistance. The upper trend line connects the highs and the lower one the lows. In Figure 5.8 we provide an example of an ascending channel.

5.2.3. Support and resistance

The key concepts discussed next are those of support and resistance. *Support* is the price level at which demand is thought to be strong enough to prevent the price from declining further, and *resistance* is the price level at which selling is considered to be strong enough to prevent any further price increases. Thus, support is the price level through which a stock or market seldom falls, and resistance, the one over which the price does not rise.

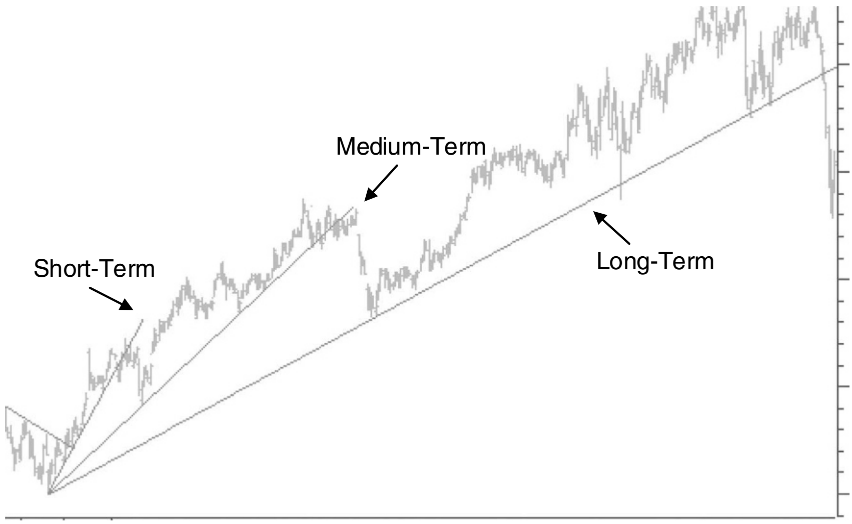

Figure 5.7 Long-, medium-, and short-term trends chart
Source: Chart courtesy of StockCharts.com

Figure 5.8 Chart with ascending channel
Source: Chart courtesy of StockCharts.com

From the point of view of market psychology and supply and demand, support and resistance are important because they indicate trading activity undertaken by those willing to buy (in the case of a support) or sell (in the case of resistance). Consequently, these levels are monitored since they help test and confirm trends and identify trend reversals. For example, at the support level, demand overcomes supply, thus preventing any further declines. On the other hand, as prices approach the resistance level, sellers become more interested in selling and buyers are less inclined to purchase. When these trend lines are broken, the supply and demand and the psychology behind the stock's movements are thought to have shifted, in which case, new levels of support and resistance will eventually be established.

Support and resistance points affect the way agents trade. For instance, traders may avoid placing orders close to or at those points, given that often prices do not reach these targets and the areas around them may be affected by great volatility. Bullish agents may place the orders a few points above the support level and bearish ones right below the same.

Round numbers such as 10, 35, 50, 100, and so on are one kind of support and resistance levels that have a major psychological component. These are important because they often represent major psychological turning points at which many traders make decisions to buy or sell. For instance, purchasers may buy a large number of shares once the price starts to fall toward a round number. This makes it more difficult for the price to fall below that level. On the other hand, sellers will also start to sell off shares as prices move toward another round peak number, also creating difficulties to move past this upper level. It is the increased buying and selling pressure at these levels that makes them important points of support and resistance. However, if either of these levels is broken, it is thought that supply and demand has shifted, its role reversed, with the support level becoming the resistance and vice versa. An example of support and resistance is provided in Figure 5.9, where the solid line at the top is the resistance level and that at the bottom is the level of support.

5.2.4. Gaps

In bar and candlestick charts, a gap is an empty space between consecutive trading periods. It occurs when the price differences between these periods are very large. The three main types of gaps are breakaway, runaway or continuation, and exhaustion. These form at the start, the middle, and the end of a trend respectively. A large gap may occur when, for instance, the trading in one period is $10 to $15 and the next trading period opens at $30. These gaps indicate something relevant has happened to the securities and are particularly significant when accompanied by an increase in volume.

5.2.5. Volume

The trading activity around a security is measured by volume, estimated as the number of shares or contracts that trade over a chosen period of time. Volume is as important as price because price changes are preceded by volume, and

Figure 5.9 Support and resistance chart
Source: Chart courtesy of StockCharts.com

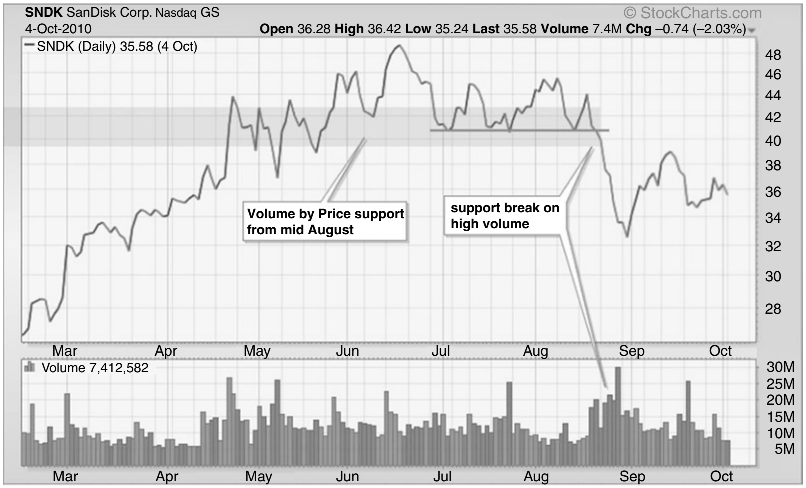

Figure 5.10 Volume chart
Source: Chart courtesy of StockCharts.com

volume is used to confirm trends and chart patterns. That is, if volume is starting to decrease in an uptrend, it is usually a sign that the upward run is about to end. Also, any price movement with high volume is seen as more relevant than a similar one with less volume. Thus, volume is closely monitored. We can see an example of volume in Figure 5.10.

On a chart, volume is usually represented as a histogram (vertical bars) below the price chart. These bars show the number of shares that have been traded during the period and illustrate trends just as prices do. Volume should move with the trend. Thus, if prices are moving in an upward trend, volume should increase and vice versa. When volume and price diverge, it means there is a contradiction between the two indicators, as, for example, if we see a clear upward trend on declining volume.

5.3. Chart patterns in technical analysis

Earlier in this section, we introduced the use of charts in technical analysis as a tool that helps identify trading opportunities. The presumption is that some price patterns are recurrent, given the technicians' belief in history repeating itself. In this part of the chapter, we will describe the most common chart patterns used in technical analysis.

5.3.1. Head and shoulders

Head and shoulders is a reversal chart pattern that signals the security is bound to move against the previous trend. Figure 5.11 shows one version of this pattern.

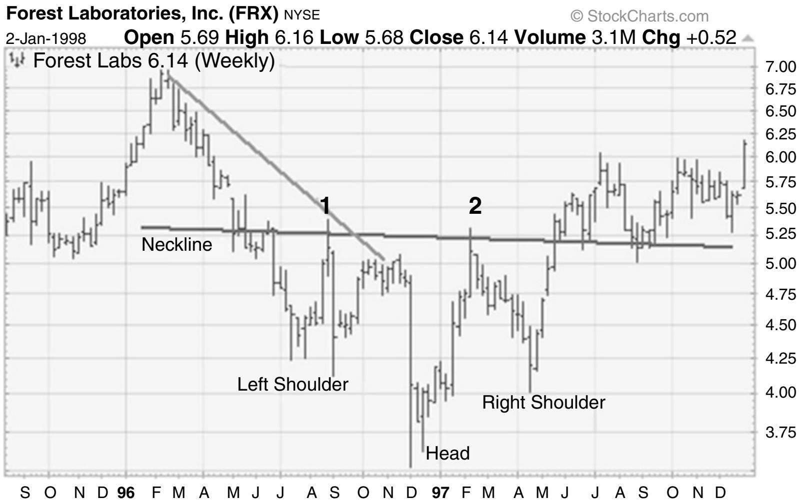

Figure 5.11 Head and shoulders chart

Source: Chart courtesy of StockCharts.com

Head and shoulders top is a chart pattern formed at the high of an upward movement and signals that the upward trend is about to end. Head and shoulders bottom, illustrated in Figure 5.11, is used to signal a reversal in a downtrend. Both these patterns are similar in that they are made up of four different parts: two shoulders, a neckline, and a head, each with a high and a low.

5.3.2. Cup and handle

A cup and handle chart is a bullish continuation pattern in which the upward trend has paused before continuing in an upward direction. This price pattern forms the shape of a cup, preceded by an upward trend and followed by the handle. An example can be seen in Figure 5.12. The handle is most often formed by a downward/sideways movement in prices. Once the price crosses the resistance lines in the handle, the upward trend continues. The time frame for this type of pattern can range from several months to over a year.

5.3.3. Double tops and bottoms

These two patterns are formed after a sustained trend, when a signal indicates the trend is about to reverse. The pattern occurs when support or resistance levels are tested twice but with no final breakthrough, often signaling intermediate and long-term trend reversals. In the case of the double top pattern in the left side of Figure 5.13, the price has twice attempted to get over a certain threshold

Figure 5.12 Cup and handle chart
Source: Chart courtesy of StockCharts.com

before the trend reverses with the price heading down. In the case of a double bottom, once the price tries unsuccessfully to break a support, it finally enters a new upward trend.

5.3.4. Triangles

Triangles respond to three types of formations: symmetrical, ascending, and descending, all lasting anything from two weeks to several months. The symmetrical triangle is a pattern of two trend lines converging toward each other. This pattern assumes the upward or downward breakthrough, which is a confirmation of a trend in that direction. When the upper trend line is flat and the bottom is upward sloping, this is an ascending triangle, a bullish pattern in which an upside breakthrough is expected. If the lower trend line is flat and the upper descending, it is a descending triangle, a bearish pattern in which a downside breakthrough is expected.

5.3.5. Flag, pennant, and wedge

Flag and pennant are short-term continuation patterns formed after a sharp price movement takes place and is followed by a sideways movement. The

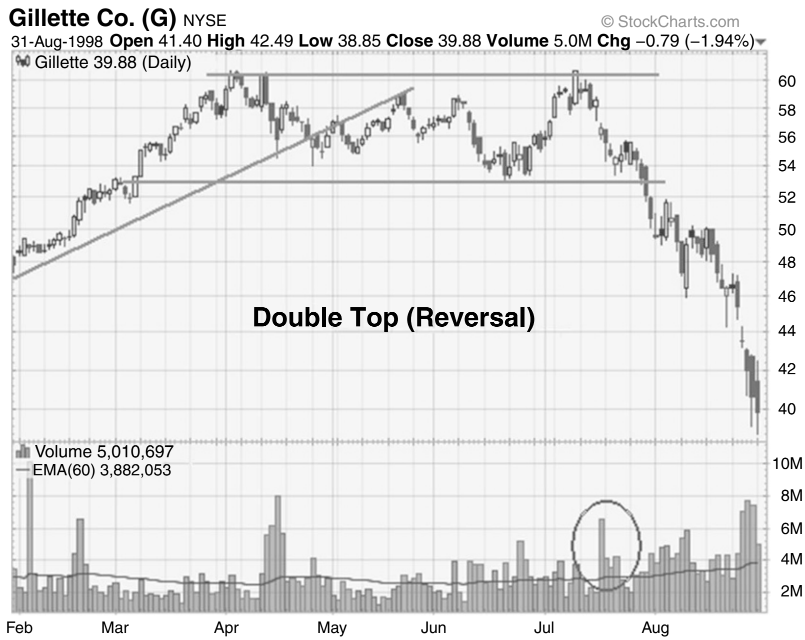

Figure 5.13 Double tops and bottoms chart
Source: Chart courtesy of StockCharts.com

patterns are completed after yet another sharp movement occurs in the direction of the initial trend. In the case of the flag, the middle section shows a channel pattern, whereas in a pennant we find converging trend lines. In either case, the trend is expected to continue once the price breaks the upper trend line.

The wedge chart pattern forms over a period of three to six months and signals either a continuation or a reversal. An example of a falling wedge is provided in Figure 5.14. If the price were to rise over the upper trend line, it would form a continuation pattern, but if it were to move below the lower trend line it would imply a reversal pattern.

5.3.6. Moving averages in technical analysis

When volatility is high and trends cannot be observed easily, technicians use moving averages. A moving average is a set of numbers, each of which is the average of the corresponding subset of an initial set of numbers. For instance, if one had a data set with 50 data points, the first value of the moving average might be the arithmetic mean of data points 1 through 10. The next value would be the average of data points 2 through 11, and so forth, until the final value, which would be the simple average of data points 41 through 50. The

Figure 5.14 Wedge chart

Source: Chart courtesy of StockCharts.com

purpose of using moving averages is to "smooth" the observations and ease the interpretation of the trend.

For instance, if one were to eliminate daily fluctuations by constructing a weekly moving average, the quarterly trend would become obvious. If we are interested in seeing the main trend a security follows over a decade, we could reduce daily observations to quarterly ones. The analyst will choose the length of the moving average and the point used in its estimation. That is, they will decide whether to use a 90-day moving average or a 60-day one or whether to use the closing price or the opening price, and which is more appropriate, the high or the low.

Moving averages can be estimated in several ways. In our previous example, the size of the subset being averaged was constant, as we always had ten digits ([1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10]/10). However, this is not a must. A *cumulative average* is a type of moving average where each value is the average of all previous data points in the full data set. So the size of the subset being averaged grows by one as each new value of the moving average is calculated. A moving average could also use a *weighted average*, perhaps to place more emphasis on the more recent data. The most common moving averages are simple, weighted, and exponential.

Figure 5.15 provides an example where the security is in an upward trend when the prices are above the moving average and in a downward trend when the prices are below it. Thus, the moving average helps us assess the trend reversal after the price crosses its line.

5.3.7. Indicators and oscillators[14] in technical analysis and strategy

Indicators are used to identify market conditions and help traders choose and apply their strategies. A given strategy will define how trends, volume, volatility, and momentum indicators are interpreted and used in order to make educated guesses about future market activity. Obviously, the adequate indicator would depend on the strategy, trading style, and risk tolerance the agent intends to follow. A trader who seeks long-term moves with large profits might focus on a trend-following strategy and thus utilize a trend-following indicator such as a moving average. A trader interested in small moves with frequent

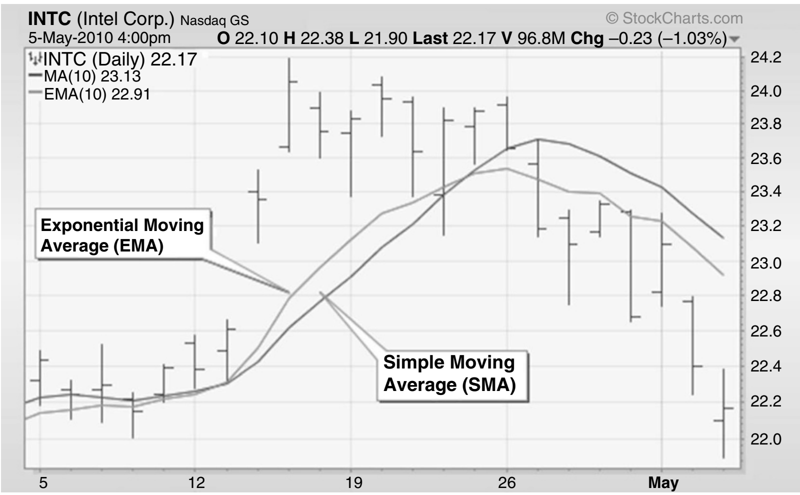

Figure 5.15 Upward and downward trends with moving average

Source: Chart courtesy of StockCharts.com

small gains might be more interested in a strategy based on volatility. Again, different types of indicators can be used for confirmation.

Indicators employ the price and volume of a security to measure money flow, trends, volatility, and momentum. They are utilized as secondary measures to confirm price movements and patterns. That is, indicators do not "stand alone" but should be used in conjunction with other aspects of technical analysis.

Furthermore, indicators can be leading or lagging behind change. In the first case, the indicator has a predictive quality as it precedes the price movement, but in the case of the latter, given that it follows the changes, it is used as a confirmation tool. Indicators assist in identifying trading opportunities by forming buy and sell signals through crossovers and divergence. Crossovers happen either when the price moves through the moving average, or when two different moving averages cross over each other. Divergence happens when the direction of the price trend and the direction of the indicator trend move in opposite directions, signaling that the direction of the price trend is weakening.

In order to adapt the indicator to the circumstances, traders enter the input necessary to define the key variables, such as the length of the historical time series to be used in their estimations. This can be done with indicators in the public domain, in readymade statistical packages or statistical packages developed by financial investors.

In a chart, the technical indicators are squiggly lines often found above, below, and on top of the price information. Indicators that use the same scale as prices are typically plotted on top of the price bars and are therefore referred to as "overlays." Table 5.1 presents the four main types of technical indicators, along with a few representative examples.

To get an idea of how these indicators work, we succinctly and selectively explain some of them:

Accumulation/distribution line, a volume indicator: The accumulation/distribution line measures money flows in a security by using the buy-to-sell ratio. This is done by comparing the price movement to the volume moved within a

Table 5.1 The four basic categories of technical indicators

Type	Examples
Volume	Accumulation/distribution line, On-balance volume (OBV)
Volatility	Bollinger bands, Average true range, Standard deviation
Trend	Aroon, Aroon oscillator
Momentum/strength	Average directional index, Moving average convergence, RSI, Stochastic oscillator

period. For example, a security has an accumulation/distribution line that trends upward if there is more buying than selling.

Bollinger bands, a volatility indicator: Bollinger bands are based on the standard deviation of prices to estimate the upper and lower limits of "normal" price movements.

Aroon, a trend indicator: The Aroon is used to measure the direction and magnitude of a trend and to predict the start of a new one. This indicator is comprised of two lines: an "Aroon up" and an "Aroon down." The first measures the amount of time that has elapsed since the highest (lowest) price was reached during the period chosen by the analyst. The trend is up when the Aroon up line is above the Aroon down, and vice versa. For instance, a 15-day Aroon up measures the number of days since a 15-day high. A 15-day Aroon down measures the number of days since a 15-day low. In this sense, the Aroon indicators are quite different from typical momentum oscillators, which focus on price relative to time. Aroon is unique because it focuses on time relative to price.

Relative strength index (RSI), a momentum indicator: The RSI is used to signal instances when securities have been overbought or oversold by comparing the average price change of advancing periods to that of declining periods.

So, what is a strategy? In technical analysis the established strategy describes the conditions under which the trader will act. Using the indicators listed above, as well as many others, this strategy will define what trades will be established, how they are going to be managed, and when they will be closed. Thus, traders use multiple indicators to implement a strategy. For instance, *trade filters* will help identify the setup conditions, that is, a price that has closed above the 100-day moving average, while *triggers* will detect the right timing for a specific action, such as when the price reaches one tick above the bar that breached the 50-day moving average. Also, redundancy in indicators is frequent, as the first will provide an indication that the second can confirm as a believable signal. A moving average strategy, for example, might employ the use of a momentum indicator for confirmation that the trading signal is valid, whereas opposing signals might indicate that the signal is less reliable and that the trade should be avoided. Each indicator and combination thereof requires research to determine the most suitable application with respect to the trader's style and risk tolerance.

5.4. Chart analysis

Technical analysis can be as complex or as simple as you want it to be. Here, we describe a simplified version of an instance where we are interested in buying stocks, with the focus being on spotting bullish situations.

The first step would be to identify the overall trend. This can be accomplished through trend lines, moving averages, or peak/trough analysis. We would then locate areas of congestion and previous highs above the current price and mark the resistance levels. A break above resistance would be considered bullish.

Momentum is usually measured with an oscillator. If the oscillator is above its exponential moving average or positive, then momentum will be considered bullish, or at least improving. For stocks and indices whose volume figures are available, an indicator that uses volume is utilized to measure buying or selling pressure. In assessing relative strength, the price relative is a line formed by dividing the security by a benchmark. For stocks in the USA, this is often the share price divided by the S&P 500. The plotting of this line over a period of time will tell us if the stock is outperforming (rising) or underperforming (falling) the major index. The final step is to synthesize the above analysis to ascertain the following: strength of the current trend, maturity or stage of current trend, reward-to-risk ratio of a new position, and the potential entry levels for new long position.

5.5. Fibonacci

One of the greatest European mathematicians of the Middle Ages, the Italian Leonardo Pisano Bigollo was credited with discovering the unique properties of the Fibonacci sequence. He introduced this to the West. The Fibonacci sequence is built with terms which are the sum of the two preceding ones (0, 1, 1, 2, 3, 5, 8, 13, etc.), extending to infinity. This sequence is interesting for its many unique mathematical properties. Among them, the *golden ratio* (also called the golden mean, the golden section, phi, and the divine proportion) is probably the better known, given it seems to describe a variety of natural relationships and proportions in nature. The golden ratio can be summarized thus: the quotient of the adjacent terms approximates to 1.618. This number is important because many things in nature have dimensional properties that adhere to the ratio of 1.618, from atoms to large celestial bodies. For instance, the golden ratio is often encountered when taking the ratios of distances in simple geometric figures such as the pentagon, the pentagram, and the dodecahedron, and has connections with continued fractions and the Euclidean algorithm for computing the greatest common divisor of two integers. Its proponents assure that if you split the number of female bees by the number of male bees in any given hive, you will get 1.618. They also claim sunflowers, which have opposing spirals of seeds, have a 1.618 ratio between the diameters of each rotation. This ratio can be seen in many other relationships between different components throughout nature, that it seems unavoidable.

Curiously enough, nature is not the only one to rely on this innate proportion to maintain balance. Financial markets are also believed to conform to it.

Here, we take a look at some technical analysis tools that have been developed to take advantage of the ratio, and examine some ways in which it can be applied to finance.

5.5.1. Fibonacci in technical analysis

The Fibonacci sequence is constructed by terms which are the sum of the two preceding terms. Thus, the following would describe a part of the series: 0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, 610, and so on. The sequence extends to infinity and contains many unique mathematical properties, some of which we review below:

a. After 0 and 1, each number is the sum of the two prior numbers (1 + 2 = 3, 2 + 3 = 5, 5 + 8 = 13, 8 + 13 = 21, etc.).
b. A number divided by the previous number approximates 1.618 (21/13 = 1.6153, 34/21 = 1.6190, 55/34 = 1.6176, 89/55 = 1.6181). The approximation nears 1.6180 as the numbers increase.
c. A number divided by the next highest number approximates 0.6180 (13/21 = 0.6190, 21/34 = 0.6176, 34/55 = 0.6181, 55/89 = 0.6179, etc.). The approximation nears 0.6180 as the numbers increase. This is the basis for the 61.8 percent retracement we will discuss later.
d. A number divided by one that is two places higher approximates 0.3820 (13/34 = 0.382, 21/55 = 0.3818, 34/89 = 0.3820, 55/144 = 0.3819, etc.). The approximation nears 0.3820 as the numbers increase. This is the basis for the 38.2 percent retracement commented on below. Also, note that 1 – 0.618 = 0.382.
e. A number divided by one that is three places higher approximates 0.2360 (13/55 = 0.2363, 21/89 = 0.2359, 34/144 = 0.2361, 55/233 = 0.2361, etc.). The approximation nears 0.2360 as the numbers increase. This is the basis for the 23.6 percent retracement.
f. 1.618 refers to the Golden Ratio or Golden Mean, also called Phi. Both Phi and its inverse 0.618 can be found throughout nature.

When used in technical analysis, the golden ratio is typically translated into three percentages – 38.2, 50, and 61.8. However, more multiples can be used when needed, such as 23.6, 161.8, and 423 percent. There are four primary methods of applying the Fibonacci sequence to finance: retracements, arcs, fans, and time zones.

5.5.1.1. Fibonacci retracements

Fibonacci retracements are ratios found in the sequence that are used to identify potential reversal levels. They employ horizontal lines to alert investors of a potential trend reversal, and indicate areas of support or resistance. Retracements are based on the prior move. A bounce is expected to retrace a portion of the prior decline, while a correction is expected to retrace a portion of the prior advance. Once a pullback begins, chartists can identify specific Fibonacci retracement levels for monitoring. As the correction approaches these retracements, chartists should become more alert to a potential bullish reversal.

The way retracements are calculated is by first locating the high and low of the chart. Five lines are then drawn: the first at 100 percent (the high on the chart), the second at 61.8 percent, the third at 50 percent, the fourth at 38.2 percent, and the last at 0 percent (the low on the chart). The most popular Fibonacci retracements are 61.8 and 38.2 percent, often rounded to 38 and 62 percent. The other common retracements are 23.6 and 50 percent. After an advance or decline, chartists apply Fibonacci ratios to define retracement levels and forecast the extent of a correction or pullback or to forecast the length of a countertrend bounce. These retracements can be combined with other indicators and price patterns to create an overall strategy. The chart in Figure 5.16 provides an example of retracements.

5.5.1.2. Fibonacci arcs

Fibonacci arcs help anticipate the support and resistance levels and areas of ranging on a chart. These are found by determining the high and low of a chart and with a compass-like movement, drawing three curved lines at 38.2, 50, and 61.8 percent from the desired point.

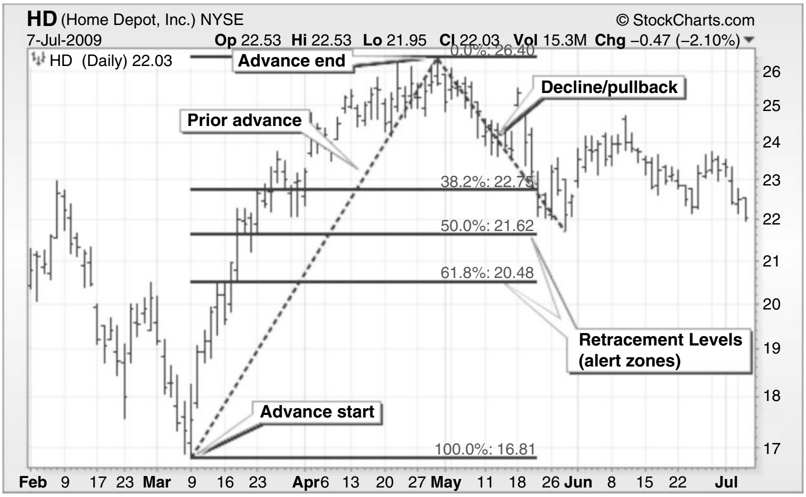

Figure 5.16 Fibonacci retracement chart
Source: Chart courtesy of StockCharts.com

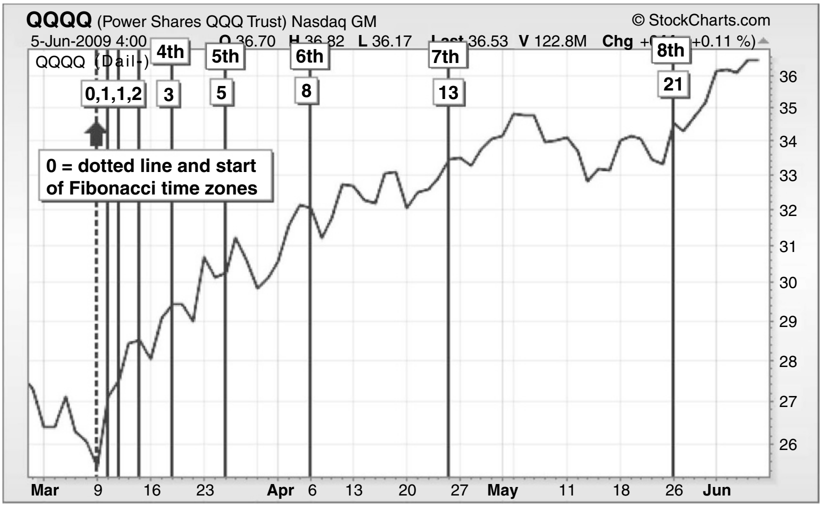

Figure 5.17 Fibonacci time zones
Source: Chart courtesy of StockCharts.com

5.5.1.3. Fibonacci fans

Fibonacci fans are composed of diagonal lines that indicate areas of support and resistance. After the high and low of the chart is located, an invisible vertical line is drawn though the rightmost point. This invisible line is then divided into 38.2, 50, and 61.8 percent, and lines are drawn from the leftmost point through each of these points.

5.5.1.4. Fibonacci time zones

Fibonacci time zones are a series of vertical lines drawn to indicate areas in which major price movement may be expected. They are composed by dividing a chart into segments with vertical lines spaced in increments that conform to the Fibonacci sequence (1, 1, 2, 3, 5, 8, 13, etc.). An example is provided in Figure 5.17.

Traders use combinations of Fibonacci methods. For example, a trader may observe the intersecting points, using a combination of the Fibonacci arcs and resistances. She might use Fibonacci studies along with other forms of technical analysis such as candlesticks, momentum oscillators, moving averages, or Elliott Waves to predict the extent of the retracements after different waves.

5.6. Market predictions based on wave patterns, Elliott Waves

In the late nineteenth century, Ralph Nelson Elliott suggested that financial market cycles and forecast trends could be analyzed by identifying peaks in investor psychology, price movements, and other factors. After scrutinizing 80 years of market data, Elliott stated that, although somewhat chaotically, markets traded in repetitive cycles. His proposal states that market prices unravel in identifiable patterns,[15] with upward and downward swings caused by crowd psychology movements between optimism and pessimism. These trading patterns, named *waves,* were then used to make stock market predictions.

In Elliott's model, market prices alternate between an impulsive, or *motive* phase, and a corrective phase on all time scales of trend. Impulse waves move with the trend or larger degree wave, while corrective waves move against it. When the larger degree wave is up, advancing waves are impulsive and declining waves are corrective. When the larger degree wave is down, impulse waves are down and corrective waves are up.

A basic impulse advance forms a five-wave sequence. That is, there are five waves in its pattern. These are five lower-degree waves, alternating between motive and corrective character. We can see an example in the chart in Figure 5.18. Clearly, the entire wave, the bigger trend, is up. Waves 1, 3, and 5 are impulses, and 2 and 4 are smaller retraces of 1 and 3. These are corrective waves because they move against this bigger trend.

A basic corrective wave goes against the main trend and subdivides into three smaller degree waves, typically a, b, and c, starting with a five-wave countertrend impulse, a retrace, and another impulse. The chart in Figure 5.19 shows an abc corrective sequence where waves a and c are impulse waves, given that they happen in the direction of the larger degree wave. Wave b is a corrective

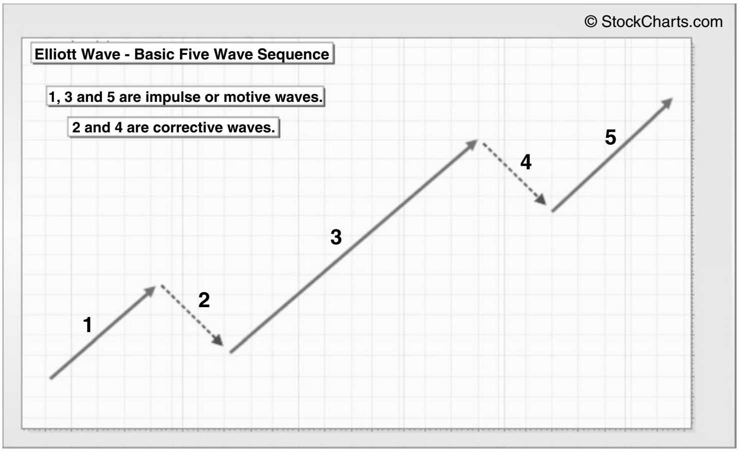

Figure 5.18 Elliott Waves, basic sequence
Source: Chart courtesy of StockCharts.com

wave as it moves against the larger degree wave. The larger degree wave is shown by the fact that the entire trend is downward.

The Elliott Wave sequence sums up a total of eight waves: a basic five-wave impulse sequence and a three-wave corrective sequence. The complete sequence is therefore split into two distinct phases: the impulse phase and the corrective phase. In Figure 5.20, we see the complete cycle: 1-2-3-4-5 is the Bigger

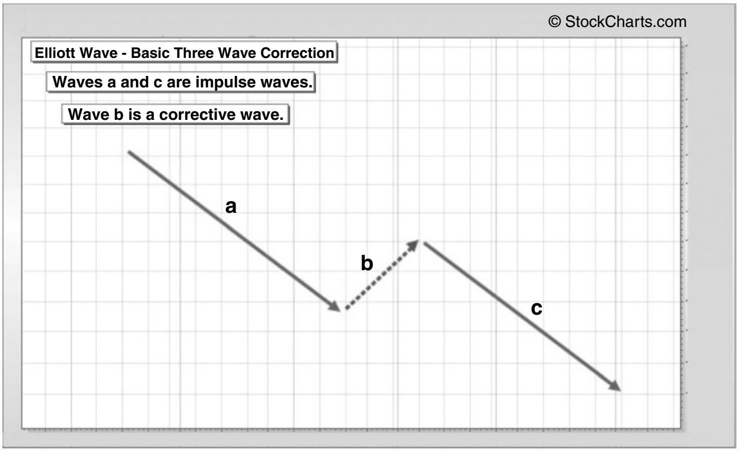

Figure 5.19 Elliott Waves, basic sequence correction
Source: Chart courtesy of StockCharts.com

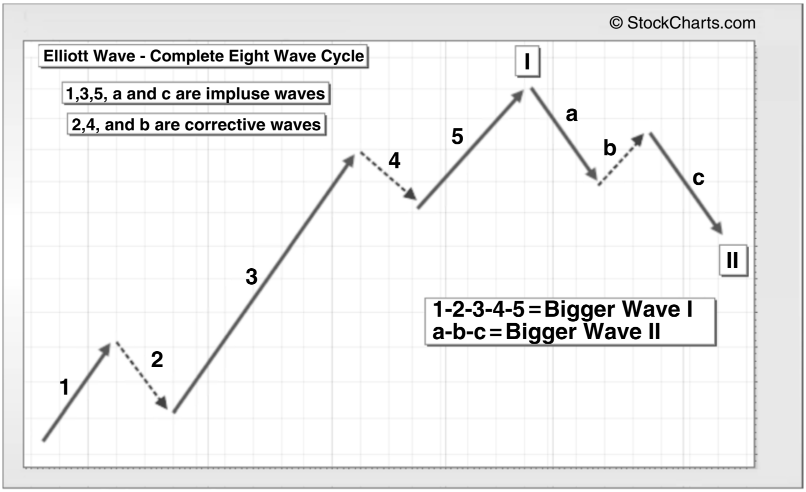

Figure 5.20 Elliott Waves, complete eight-wave cycle
Source: Chart courtesy of StockCharts.com

Wave I, whereas the abc corrective phase, representing a correction of the larger impulse phase, is the Bigger Wave II. Overall, 1, 3, 5, a, and c are impulse waves while 2, 4, and b are corrective waves.

These eight-wave charts show two larger degree waves (I and II) as well as the lesser degree waves within them. Waves 1-2-3-4-5 are one lesser degree than Wave I. By extension, Wave I is one larger degree than Waves 1-2-3-4-5. Waves abc are one lesser degree than Wave II.

After the initial five waves up and three waves down, the sequence begins again and the self-similar fractal geometry begins to unfold according to the five- and three-wave structure, which it underlies one degree higher. The completed motive pattern includes 89 waves, followed by a completed corrective pattern of 55 waves.[16]

5.6.1. Wave levels

Elliotticians, Elliott Wave analysts, hold that each individual wave reflects the psychology of the moment,[17] and interpreting this is the key to the application of the Elliott Wave Principle (EWP). As shown in the earlier charts, each degree of a pattern receives a name, and the symbols for each wave indicate both function and degree – numbers for motive waves, letters for corrective waves. The Elliott Wave Theory brands waves from largest to smallest. Degrees are relative because they are defined by form, not by absolute size or duration. That is, waves of the same degree (level) may be of very different size and/or duration.[18] We summarize them in Table 5.2.[19]

Table 5.2 Elliott Wave classification

Wave classification	**Degrees (duration)**	**Degree labels**
Grand Supercycle	Multicentury	((I)) ((II)) ((III)) ((IV)) ((V)) ((a)) ((b)), ((c))
Supercycle	Multidecade (about 40–70 years)	(I) (II) (III) (IV) (V) (a) (b) (c)
Cycle	One to several years (or even several decades under an Elliott Extension)	I II III IV V a b c
Primary	A few months to a couple of years	((1)) ((2)) ((3)) ((4)) ((5)) ((A)) ((B)) ((C))
Intermediate	Weeks to months	(1) (2) (3) (4) (5) (A) (B) (C)
Minor	Weeks	1 2 3 4 5 A B C
Minute	Days	((i)) ((ii)) ((iii)) ((iv)) ((v)) ((a)) ((b)) ((c))
Minuette	Hours	(i) (ii) (iii) (iv) (v) (a) (b) (c)
Subminuette	Minutes	i ii iii iv v a b c

Source: Robert R. Prechter, Jr. (2003) 'A New Perspective on the Quantitative Self-Affinity of Elliott Waves', *Journal of Technical Analysis*, Summer–Fall, 25–32.

The labeling convention used here identifies the degree or level of the wave, which represents the size of the underlying trend. The uppercase Roman numerals represent the large degree waves, the simple numbers represent the medium degree waves, and the lowercase Roman numerals represent the small degree waves. The trends start with the largest degree (Grand Supercycle) and work their way down to waves of lesser degree. For example, the Cycle wave is one larger degree than the Primary wave. Conversely, the Primary wave is one lesser degree than the Cycle wave. Wave 1 of (1) would indicate that Wave 1 is part of a larger degree Wave (1). Wave 1 is a lesser degree than Wave (1).

In reality, most chartists only use between one and three wave degrees on their charts. It can grow quite complicated, trying to apply nine wave degrees or even just using the labeling convention in the order provided. Chartists using one to three wave degrees can simply label the highest degree waves with uppercase Roman numerals (I, II, III, IV, V, a, b, c), the middle degree waves with numbers (1, 2, 3, 4, 5, A, B, C), and the lowest degree waves with lowercase Roman numerals (i, ii, iii, iv, v, a, b, c). This provides three distinct groups for labeling various waves.

To use the theory in everyday trading, the trader determines the main wave, or Supercycle, goes long, and then sells or shorts the position as the pattern runs out of steam and a reversal becomes imminent.

5.6.2. Fractal nature

The basic five-wave impulse sequence and three-wave corrective sequence have been explained. After the initial five waves up and three waves down, the sequence begins again and the self-similar fractal geometry begins to unfold according to the five- and three-wave structure which it underlies one degree higher. An Elliott Wave is fractal in that the structures described by Elliott are self-similar patterns appearing at every degree of trend.[20] This means that wave structure for the Grand Super Cycle is the same as for the Minuette. No matter how big or small the wave degree, impulse waves take on a five-wave sequence and corrective waves take on a three-wave sequence. Any impulse wave subdivides into five smaller waves. Any corrective wave subdivides into three smaller waves. The charts in Figures 5.21 and 5.22 show the fractal nature of the Elliott Wave in action.

5.6.3. Rules and guidelines

To interpret Elliott Waves, one needs to understand three basic "nonnegotiable" rules and a few guidelines that help chartists improve their counting. The rules apply to a five-wave impulse sequence. These are as follows:

a. Rule 1: Wave 2 never retraces more than 100 percent of Wave 1.
b. Rule 2: Wave 3 can never be the shortest of the three impulse waves, namely Waves 1, 3, and 5.

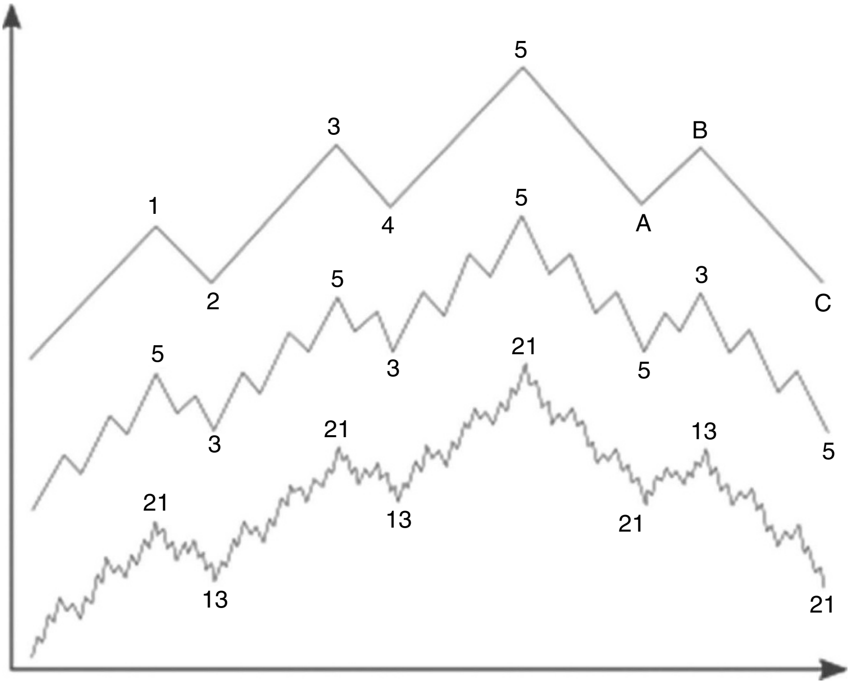

Figure 5.21 The fractal nature of Elliott Waves

Sources: From R. N. Elliott's essay, "The Basis of the Wave Principle," October 1940, in Wikipedia (from Wikimedia Commons, the free media repository in https://commons.wikimedia.org/wiki/File:Elliott_wave.svg)

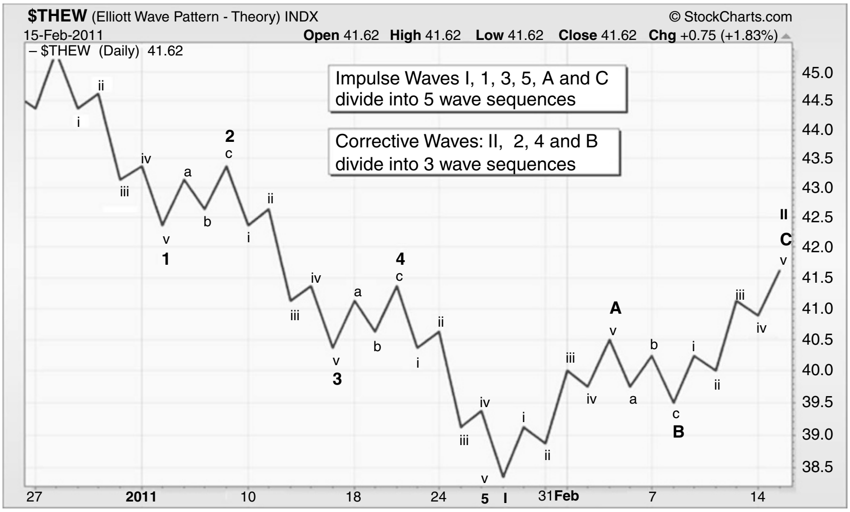

Figure 5.22 Elliott Wave patterns

Source: Chart courtesy of StockCharts.com

c. Rule 3: Wave 4 can never overlap with Wave 1, except in the rare event of a diagonal triangle.

An example can be seen in Figure 5.23.

While the rules are a "must," guidelines are bendable and subject to interpretation. The three main guidelines are as follows:

a. Guideline 1: When Wave 3 is the longest impulse wave, Wave 5 will approximately equal Wave 1.
b. Guideline 2: The forms for Wave 2 and Wave 4 will alternate. If Wave 2 is a sharp correction, Wave 4 will be a flat correction. If Wave 2 is flat, Wave 4 will be sharp.
c. Guideline 3: After a five-wave impulse advance, corrections (abc) usually end in the area of prior Wave 4 low.[21]

We do not want to go into further detail, as our purpose is merely to provide an overall idea of the technique. In Figure 5.24, we can visualize the three main guidelines.

5.6.4. Criticisms

Elliott published his theory of market behavior in a number of books and articles. More recently, other popular works have been written that trumpet this theory's

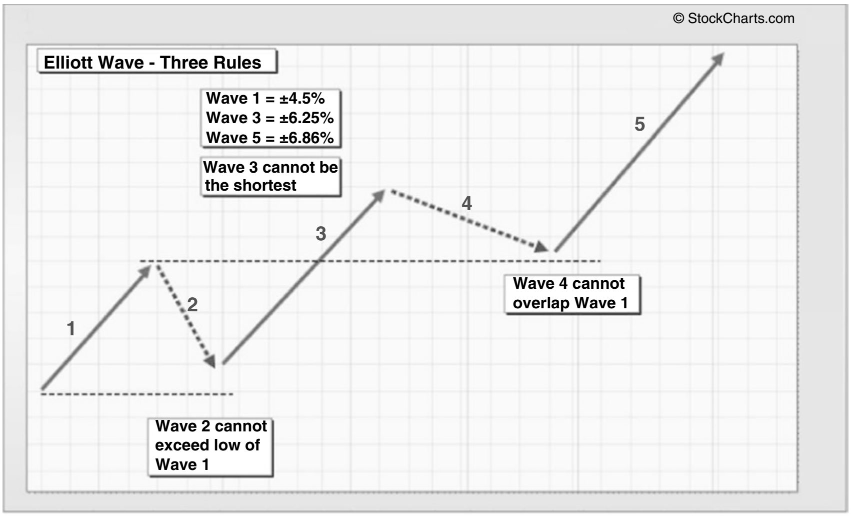

Figure 5.23 The three rules of Elliott Waves
Source: Chart courtesy of StockCharts.com

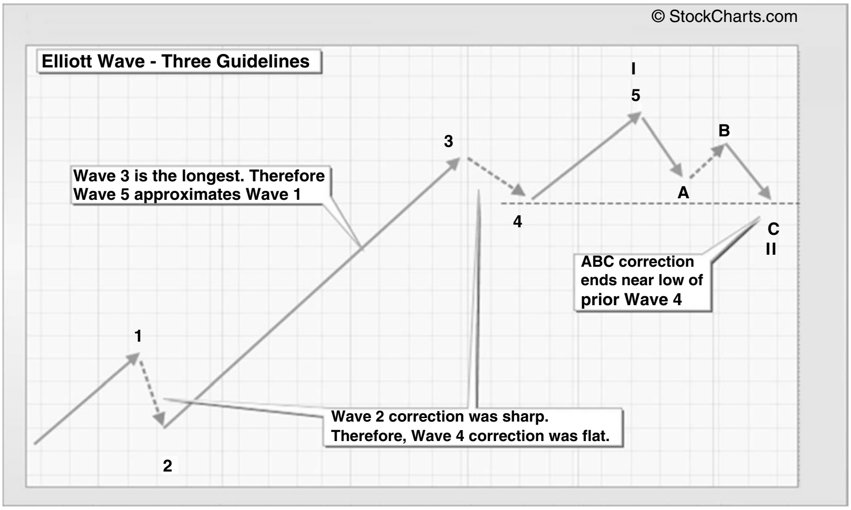

Figure 5.24 The three main guidelines of Elliott Waves
Source: Chart courtesy of StockCharts.com

worth and contributions to the field. Still, even though many market technicians use the EWP as a component of their trade tools to forecast opportunities, the empirical validity of the EWP remains the subject of much debate.[22, 23, 24]

On the one hand, its supporters state that it provides a probability framework on when to enter or get out of a particular market.[25, 26] For instance, renowned physicist Didier Sornette is said to be intrigued by the relationship between the log-periodic structures and the "Elliott Waves,"[27, 28] while others, such as Glenn Neely,[29] have elaborated on Elliott's concepts to develop their own forecasting methods.

However, the assessment of the theory remains mixed.[30] Mandelbrot has questioned whether Elliott Waves can actually predict financial markets.[31] Critics also warn that although eloquently argued, the wave principle is too vague to be useful. For instance, in Aronson's view,[32] the method's loosely defined rules and ability to postulate a large number of nested waves of varying magnitude with no consistent precision on when waves start or finish is the main reason EWP fits any segment of market history. In addition, Elliott Wave forecasts are prone to subjective revision. Furthermore, critics also argue that markets have evolved, but the theory has not. Given that today, markets function differently from those when the principles were formulated, and that these changes have affected the wave patterns, the theory is thought to be outdated and skewed in its forecasting accuracy.

The dispute also involves the relationship between the principle and the Fibonacci ratios. For instance, Batchelor and Ramyar[33] concluded that the idea that prices retrace to a Fibonacci ratio or round fraction of the previous trend lacks any scientific rationale. Furthermore, they found no significant difference between the frequencies with which price and time ratios occur in cycles in the Dow Jones Industrial Average (DJIA) and frequencies which we expect would occur at random in such a time series.[34] These findings were later challenged by Robert Prechter,[35] who insisted the data show "Fibonacci ratios do occur more often in the stock market than would be expected in a random environment."

5.7. Does technical analysis help identify trading opportunities?

Earlier in the chapter, we discussed some of the criticisms facing technical trading. After having presented a summary of the tools most commonly utilized by these traders, we now aim to assess the impact of the contention made earlier that this technique helps identify trading opportunities. Given the chartists' premise that price patterns exist and are relevant in forecasting, we now wish to discuss the relation between the statements listed below and the possible consequences deriving from the use of these charting techniques, especially in the context of bubbles and contagion:

a. Prices are random.
b. Technical rules result in coordinated "buy and sell orders" at specific junctures with such "signals" being instantaneously shared by a large number of traders.
c. The implications of statement (b) in terms of its impact on reflexivity or the self-fulfilling prophecy and contagion. That is because collective concerted trading actions are expected when signals are pre-announced by charting methods.

Does technical trading have a disproportionate impact on market actions that have no justification from the point of view of fundamental variables, actions that are triggered by historical data and reinforced by reflectivity and contagion mechanisms? If so, what can explain these actions?

5.7.1. Prices are random

In his 1900 doctoral thesis, "The Theory of Speculation," Louis Bachelier set forth his conclusion that "there is no useful information contained in the historical price movements of securities."[36] It took 60 years until Nobel-winning economist Paul Samuelson uncovered this study and, together with Eugene Fama, expanded on Bachelier's findings. Samuelson's work resulted in the Random Walk Theory,[37] which asserts that stock prices continuously react to new information and consequently move in random and unpredictable ways. His proposals can be summarized as follows: (a) market prices are the best estimates of value; (b) price changes follow random patterns; and (c) future news and stock prices are unpredictable.

Concurrently with Samuelson's work, Eugene Fama also investigated the random fluctuations of market prices. His findings led to the conceptual framework of the "Efficient Market Hypothesis." Market efficiency ensures that prices resulting from agreements between buyers and sellers are the best estimate of fair market values. Market efficiency does not equate to perfect pricing, or to the nonexistence of mispriced securities. Rather, these precepts assert that because prices reflect all publicly known information at any point in time, mispriced securities cannot be singled out ex ante. Given that news is unpredictable and random, so are these securities' price movements. Consequently, stock or time pickers pondering whether a share will go up or down in the short term, or the best time to get in or out of the market, is in essence making up their minds by tossing a coin.

In his paper "The Behavior of Stock Market Prices,"[38] Fama further examined the assumption that access to huge amounts of information by big brokerage firms provided managers and analysts with a competitive advantage which resulted in higher performance. However, contrary to expectations, Fama's findings showed that no such advantage existed. Furthermore, later work led Fama to conclude that trends in capital markets cannot be identified in advance. In a study published with Kenneth French,[39] the authors defended that 96 percent of historical returns in diversified stock portfolios could be explained by exposure to market, size, and value risk factors.

Many studies have focused on whether stock pickers have a superior ability to choose winner assets. For instance, a paper published by Terrance Odean,[40] considering the activity of 10,000 discount brokerage accounts, concluded that investors commonly overestimated the profit potential of their stock trades. In a second paper written with Brad Barber,[41] the authors analyzed 66,465 individual trading accounts pertaining to the years 1991–1996. This time, their results showed that during the period under consideration, active investors earned annual returns of 11.4 percent versus the 17.9 percent achieved by market.

Other studies have focused on the picking ability of specific "gurus." For example, a paper by Barras, Scaillet, and Wermers[42] looked at the period between 1975 and 2006 to analyze the performance of 2,076 mutual fund managers. Once more, the findings showed that 99.4 percent of these managers exhibited no genuine stock picking abilities, while the remaining 0.6 percent had winnings "statistically indistinguishable from zero."

Time pickers or market timers claim to predict future market movements. In forecasting the direction of the market, their objective is to get in before the market goes up and to get out before it goes down. These market timing abilities have also earned the attention of researchers and numerous studies have been published to examine such declared competencies.

One interesting example is that of the CXO Advisory Group, a firm which tracks public forecasts of self-proclaimed market timing gurus. In 2005, the group wished to answer the question, "Can equity market experts, whether self-proclaimed or endorsed by others (such as publications), reliably provide stock market timing guidance?" The group proceeded to investigate this issue by collecting 6,582 forecasts for the US stock market.[43] These were public forecasts offered by 68 experts, bulls and bears, using technical, fundamental, and sentiment indicators. The gathered forecasts included those in archives, and thus the final sample covered the period 1998–2012.

According to the CXO aggregate grading results, the terminal accuracy, that is, the percentage of times that market timing gurus got it right, was 46.9 percent. If we average by guru rather than across all forecasts, the terminal accuracy was then 47.4 percent. The chart in Figure 5.25 tracks the inception-to-date accuracy of all 6,582 graded forecasts in the sample.

The studies cited above suggest no evidence to support the idea that following a given trading strategy can provide a means to consistently outperform the market. The knowledge possessed by one analyst, from balance sheets to marketing material to trading prices and volumes, is already built into the stock price and shared with the millions of other traders who have the same information. What none of them possess is the knowledge of what would move the stock in the future. Such events are unpredictable and impossible to forecast.

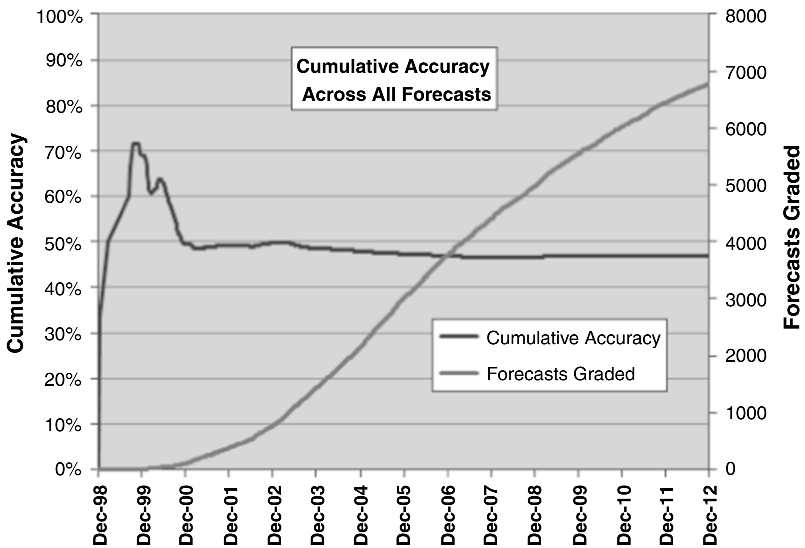

Figure 5.25 Market timing gurus' terminal accuracy

Source: Courtesy of the CXO Advisory Group, LLC (http://www.cxoadvisory.com)

So why would anyone believe that the analysis of historical data can provide a trading advantage? First, we need to consider that investment managers, such as stock mutual fund managers, are compensated and retained in part due to their performance relative to that of their peers. Taking a conservative or contrarian position may result in performance unfavorable to peers. This could make customers go elsewhere and could affect the investment manager's own employment or compensation. The typical short-term focus of US equity markets and elsewhere exacerbates the risk for investment managers who do not participate in commonly shared strategies, particularly during the building phase of a bubble. In attempting to maximize returns for clients and maintain their employment, they may rationally participate in a bubble they believe to be forming, as the risks of not doing so outweigh the benefits.[44]

Also, if we review our box of psychological biases presented in Chapter 3, a heuristic known as *representativeness*[45] explains why people often attempt to predict future uncertain events by considering a short history of data. Technical analysis *is* representativeness. Traders focus on partial information to decipher the broader picture this history is representative of. The listed psychological biases propose some explanations for why a large number of people trust technical analyses:

Communal reinforcement: Communal reinforcement happens when a belief results from the claim a community makes repeatedly, disregarding the existence of any empirical evidence that could back the claim.

Confirmation bias: Confirmation bias is a type of selective thinking which happens when one highlights and searches for information that confirms one's pre-existing beliefs, while concurrently ignoring any evidence that might contradict the same.

Extrapolation: Investors search their memories for "equivalencies." They think they see something they have experienced in the past and make decisions on market movements, assuming the perceived trend will repeat itself. When agents extrapolate, they project historical data into the future on the same basis; if prices have risen at a certain rate in the past, they will continue to rise at that rate forever. The argument is that investors tend to extrapolate past extraordinary returns on investment of certain assets into the future, causing them to overbid on those risky assets in order to continue to capture those same rates of return.

Hindsight bias: Looking back, people think that, had they been more attentive, past events could have been predicted. However, innovations move the markets and past events, as much as we want to rationalize them, could not have been predicted in advance.

Overconfidence: People think they are smarter than the average and can outperform the market.

Selective thinking: Selective thinking results from focusing on favorable evidence in order to justify one's convictions.

Self-attribution bias: People take personal credit for gains and blame others – or outside factors – for losses. Thus, success is linked to personal skill while losses are due to bad luck.

Self-deception: Self-deception occurs when we mislead ourselves into accepting what we know to be false by disregarding evidence that supports the opposing position.

However, regardless of the reasons for following technical rules in trading and whether the analysts' assessments are right or wrong, this methodology results in coordinated "buy and sell orders" at specific time junctures, with such "signals" being instantaneously shared across a large number of traders. Can this trading strategy impact the market?

Of course, there has been a great deal of theoretical and empirical exploration of the proposition that irrational investor errors cause market misvaluation of assets. This includes some examination of whether there is contagion in biases across different investor groups, or from analysts to investors. Thus, herding is another explanation that can be used in behavioral finance to explain some of the effects of following technical trading rules. This process speaks to the fact that investors tend to buy or sell in the direction of the market trend. This is helped by technical analysts who detect those trends and establish their trading strategies accordingly.

The tendency of institutional investors to chase return trends has received considerable attention in financial economics. Also known as positive feedback trading or momentum investing, trend-based strategies call for buying (selling) financial assets with high (low) recent returns. Grinblatt, Titman, and Wermers (1995[46]) provide the first empirical evidence that mutual funds, especially those that are growth-oriented, tend to buy, on average, recent winners. These findings were confirmed from larger samples of institutional investors. The evidence is especially strong for actively managed mutual funds.[47, 48]

Despite the large body of work on the subject and the proposed psychological explanations, very little is known about why institutional investors respond to past returns when formulating their investment strategies. Financial theoretical studies typically assume this type of trading behavior rather than provide a rationale for the same.[49, 50] However, it is tempting to link the evidence on trend-chasing to return predictability; perhaps institutional investors also implement mechanical strategies designed to exploit the momentum anomaly, as documented in some research.[51]

Nonetheless, while some institutional investors indeed specialize in momentum, they are trend-chasers even in markets that do not exhibit the same. Given this observation, Alti, Kaniel, and Yoeli (2012[52]) propose and test an explanation for trend-chasing at the stock level which does not rely on arguments based on the existence of market anomalies.

The main premise of their thesis is that investors face uncertainty regarding the *precision* of their private information. Consequently they await confirmation from subsequent good news to revise their estimates and establish new stock positions. While such news impact stock prices, at the same time, it increases investors' estimates of the precision of their information. With low information quality, the latter effect dominates and drives investors to purchase. These actions can generate trading behavior that resembles trend-chasing. However, investors do not trade stocks based on past returns per se; rather, it is the news that drives stock returns which also triggers trade by affecting investors' confidence in the validity of their initial analyses. Their findings confirm that initiations of stocks with relatively low information quality tend to follow high abnormal return and point to a strong link between information quality and the tendency to chase return trends. A second finding is that mutual funds base their investment decisions in part on their past return experiences with individual stocks.

There is a great deal of theoretical literature on informed trading dynamics. Some of these works present multiperiod models in which some informed investors trade early and others trade late.[53, 54, 55] In these, similar to the confirmation effect in Alti et al.'s model, price movements resulting from early trades help late investors better interpret their own signals and affect their trading behavior.

Other papers generate trading behavior that resembles trend-chasing for various reasons. For instance, Wang (1993, 1994[56]) develops dynamic asset-pricing models that feature risk-averse and asymmetrically informed investors. In these models, uninformed investors infer informed investors' past signals through stock price realizations and respond by rebalancing their portfolios. For instance, a high realized return reveals that the uninformed investors had underestimated the expected dividend of the stock and hence underinvested in it, causing them to buy. In this way, uninformed investors resemble trend-chasers.

Other works have also generated return-sensitive trading patterns due to risk aversion by informed investors.[57, 58] These investors partially reverse their initial positions to reduce their risk exposure once prices more fully reflect their private information, either after the arrival of additional informed traders or the arrival of public information. Thus, informed investors act in contrarian ways (e.g. selling the security after a price run-up), whereas uninformed investors resemble trend-chasers by taking the other side of these trades.

The idea that investors gradually learn about their ability to analyze a stock has been tackled in a number of studies, some of which emphasize the asset-pricing implications of behavioral biases.[59, 60] For instance, some of them analyze the impact of biased self-attribution on informed investors' trading dynamics. One such study by Daniel et al. (1998[61]) models private information as being long lived, in which case biased self-attribution causes trend-chasing by informed investors (e.g. buying more of the stock after a confirming signal increases the stock price). In contrast, private information is short-lived according to Gervais and Odean (2001[62]), which rules out trend-chasing by construction.

5.7.2. Trend-chasing in creating bubbles

We have mentioned the generalized belief that bubbles can result from irrationally valuing assets based solely upon their returns in the recent past without resorting to a rigorous analysis of their underlying "fundamentals." Experimental and mathematical economics has tried to shed light on these issues through modeling and experiments. For instance, a study by Caginalp et al.[63, 64] modeled trading, considering that an asset's supply and demand depended on its valuation as well as other factors such as price trend. In their analysis, the authors reached the conclusion that the bubble would be larger if there was an initial undervaluation. To start with, "value-based" traders would buy the undervalued asset, thus creating an uptrend, which would then attract the "momentum" traders and a bubble would be created. Later experiments confirmed the importance of trend-based investing in creating bubbles.[65, 66, 67]

Asset flow equations have been used to study the formation of bubbles where it was shown that a stable equilibrium could become unstable with the influx of additional cash or the change to a shorter time scale on the part of momentum investors.[68] Thus, a stable equilibrium could be pushed into an unstable one, leading to a trajectory in price that exhibits a large "change" from either the initial stable point or the final stable point.

Stability and changes in equilibrium have also been analyzed in the literature. Chiarella et al. (2002)[69] develop a discrete time model of asset price dynamics, based on the interaction of fundamentalists and chartists. The former hold an estimate of the fundamental value of the asset, and their demand is an increasing function of the difference between that estimate and the current price. The latter base their trading decisions on an analysis of past price trends, their demand being a function of the expected return differential with an alternative asset. The analysis in this paper shows how the behavior of the two types of traders affects local and global dynamics of key parameters: the "strength" of fundamentalist and chartist demand and the "speed of adjustment" of chartists' expectations. For instance, for sufficiently low values of strength of chartist demand, the equilibrium is stable for a wide range of values of the fundamentalist parameter "strength of demand." While when "strength" of chartist demand is sufficiently high, that is, chartist demand is relatively strong, the ability of the fundamentalists' demand to stabilize the system is restricted to a fairly narrow range of the parameter.

Furthermore, focusing on the global behavior observed in this case, the authors show that when the equilibrium is locally stable, other dynamic phenomena arise, such as chaotic transients before the convergence on the stable equilibrium or the coexistence of attractors. This happens for sufficiently high values of "speed of adjustment" of chartists' expectations and "strength" of fundamentalist demand. By increasing the parameter "speed of adjustment," the equilibrium becomes unstable via a Neimark–Hopf bifurcation. When this bifurcation occurs at sufficiently high values of the parameter "strength" of fundamentalist demand, an increase of "speed of adjustment" leads to oscillatory behavior which may become chaotic around the equilibrium. As reaction speed of adjustment of chartists' expectations approaches one, the dominating state is an attracting cycle (although coexisting with a repellor) with a wide basin of attraction. A stochastic version of the model demonstrates how the interaction between the deterministic nonlinear dynamic phenomena and a simple external noise process can generate the typical market patterns of volatility clustering, fat tails, peaked return distributions, and skewedness.

So, the question remains: Is it really possible to benefit from chasing the trends? The existence of limit cycle and convergence show that there could be trends that are worth chasing, while the possible existence of chaotic motion under the same trading rules shows that there may be trends that are hardly predictable. The experiments have shown that transitions between dynamic regimes take place when the balance changes between aggressiveness and reactivity of the chartist traders and the speed of the fundamentalist traders. This suggests that as the trader's "mood" changes, the price dynamics might be switching between those regimes, thus allowing all traders to benefit at certain stages and lose at others. Some additional important findings of this project showed that traders need to modify their behavior in order for dynamic regimes to exist in networked markets, and that for certain networks, it was very hard to trigger chaotic motion or limit cycle motion. The conclusion is that a stable market structure can exist, which can prevent speculators from destabilizing the market.

There are countless papers that investigate different aspects of agent models. The evolution of bubble processes have been considered to be related to the model structures of heterogeneous agent model (HAM), which can be described as time-varying agents trading weights. For instance, Brock and Hommes (1997, 1998[70]) introduced a HAM with fundamentalist and chartist agents. Fundamentalists believe that prices will revert to the asset's intrinsic value and therefore base expectations on the deviation of the market price from the fundamental value. Consequently, they will invest in assets that are undervalued and sell those that are overvalued. Chartists, on the other hand, speculate on the persistence of deviations from the fundamental value. They extrapolate information from previous prices, expecting trends to continue in the same direction. They will buy (sell) when past prices increase (decrease). These premises summarize two distinct trading behaviors. Fundamentalist behavior is assumed to have a stabilizing effect on market prices, bringing the asset price closer to its intrinsic value, while chartists tend to have a destabilizing effect and drive market prices away from the intrinsic value of the asset. Thus, the more chartists or trend-followers there are, destabilizing the stock market, the more likely events will turn into bubbles, particularly when the stock price is dominated by them. Thus, a motivation to use HAMs is that they are applicable to portraying the dynamic behavior of these agents, as well can be utilized to directly examine whether the bubbles that have occurred are positively associated with chartist weights.

Intrigued by the power of HAMs, Ma and Yin[71] analyzed the dynamic relationship between chartist weights and volatility clustering, as well as market bubbles in international exchanges. The authors covered the period 1996–2011 with a sample of 50 international stock exchanges. Their findings include the following: (a) HAM is significant in most markets, supporting the existence of dynamic agent weight changes. Agents may adopt different trading strategies, from momentum to contrarian, in terms of different return volatility. (b) Volatility clustering is significantly positively enhanced by chartist weights. (c) Market bubbles are associated with higher chartist weights.

Many works have tackled the issue of price dynamics displaying short-term momentum and long-term mean reversion in markets with positive feedback traders, fundamentalists, and rational agents.[72] Interactions between rational investors and noise traders following positive feedback strategies – buy when prices rise, sell when prices fall – can reproduce short-term return momentum and mean reversion patterns of asset returns.[73] Rational traders can destabilize the market by initially driving prices beyond fundamentals and then selling out at even higher prices to the feedback traders.[74] These findings also apply to the foreign exchange (FX) market, with trend-chasers and investors trading on mean reversion to fundamentals and show that it can generate prolonged periods of overvaluation.[75] Although rational agents have the ability to, at least potentially, correct the mispricing bubbles created by noise traders, the bubbles may persist when rational agents are unsure about the exact timing of their occurrence. In this setting, it can also be optimal for rational agents to jump the bandwagon and follow the strategy of positive feedback traders.

Short-run momentum and long-run mean reversion in markets have also been explained through models of news watchers and momentum traders.[76] Models with fundamentalists and positive feedback traders are capable of generating equity market returns with heavy tails, excess kurtosis, and volatility clustering.[77] A crucial ingredient of these models is the presence of a core of nonrational positive feedback traders and chartists who expect past price changes to continue in the future. Also, consistent with positive feedback trading, stock portfolio adjustments of individual investors reflect past market movements.[78] Momentum trading by institutional investors has been documented as well.[79] Other papers show evidence of trend-chasing behavior in commercial banks' investments in real estate and among professional forecasters of the commercial real estate market.[80]

Equivalent findings extend to the housing market. A 2011 study by Roy Kouwenberg and Remco C. J. Zwinkels[81] estimates a behavioral agent-based model with boundedly rational investors for the US housing market. Their model uses quarterly time-series data on prices and rents for the aggregate stock of owner-occupied housing for the period 1960–2009. In this work, there are two groups of investors: fundamentalists, who expect the house price to revert to its fundamental value based on rents and chartists, who extrapolate past price trends. Investors are allowed to switch between groups, conditional on recent performance. The estimation results show that fundamentalists and chartists are usually present in the market in roughly equal proportions. From 1992 until 2005, however, the proportion of chartists in the market was substantially above the long-term average, such that the house price level climbed far above its fundamental value. In an out-of-sample assessment, the model predicts the decline of the housing market from 2006 onwards. Finally, the estimated model generates boom–bust price cycles endogenously. In this paper, the authors try to improve forecasts for housing market prices by estimating a behavioral HAM with positive feedback traders. The added value of this type of model is that the proportion of market participants applying positive feedback trading rules versus fundamental based rules is time-varying. As such, it not only can explain the underreaction–overreaction anomaly, it is also capable of generating bubble periods.

Momentum strategies are used by technical traders who buy stocks that are rising in value and sell (or short-sell) stocks that are dropping. In effect, the momentum trader desires to buy shares during periods of excess demand and sell during periods of excess supply. Thus, to accurately measure the costs of implementing a momentum strategy, one has to be aware not only of the fact that such trades desire great immediacy in typically less liquid (more expensive) stocks, but also that momentum traders wish to trade on the side of the market where there is a reduced supply of liquidity (e.g. buying when there is a relative increase in the supply of buyers and a relative decrease in the supply of sellers). In such environments, the trading associated with momentum strategies exerts pressure on prices (price impact) that will certainly be greater than the price impact implicit in unconditional costs.

A study by Donald B. Keim[82] documents the costs of implementing actual momentum strategies. This study questions whether it makes sense to follow the trend or act upon other technical analysis strategies once we come to the "real world" and assess the real costs associated with the same. The author examines the trade behavior, and the costs of those trades, for three distinct investor styles: momentum, fundamental/value, and diversified/index for 33 institutional investment managers executing trades in the USA and 36 other equity markets worldwide, covering more than 1.6 million trades worth $1.1 trillion.

The results show that momentum traders do indeed condition their trades on prior price movements, that is, buys are more likely to be made in rising markets and sells are more likely in falling markets. This is in contrast to the index/diversified managers in the sample whose buys and sells are unrelated to recent prior price movements, and for the value managers whose buys (sells) are more likely to follow recent price declines (price increases). The findings show clear evidence of trend-chasing by the momentum traders.

With respect to the costs of implementing the strategies, Keim examines the trade costs of each strategy conditional on past market returns to assess how the market environment in which a trade is made affects the cost of execution and the profitability of the strategy. Conditioning on past market returns is important because momentum traders are buying (selling) when the stock price is rising (falling) and the market for the stock has excess buyers (sellers). Under such conditions, their trade costs (e.g. for purchases of stocks on a rising trajectory) are expected to be higher than the unconditional average due to a combination of (a) increased demand for liquidity on the momentum trader's side of the market due to the existence of other like-minded traders; and/or (b) reduced supply of liquidity due to fewer sellers/owners of recently appreciated stocks who don't wish to realize their capital gains.

As it turns out, market environment is an important determinant of price impact and costs for trades that are made conditional on prior market returns are significantly greater than for unconditional costs, especially for momentum traders. For example, according to this study, one-way average price impacts for momentum traders in the USA are 1.21 percent when buying stocks in a rising market and 1.37 percent when selling in a falling market. Adding opportunity costs, commissions, and other explicit costs of transacting inflates these costs to 1.82 percent for buys in rising markets and 1.96 percent for sells in falling markets. The trade costs reported set a very high hurdle rate for the profits implied by the simulated strategies.

5.7.3. The self-fulfilling prophecy and bubbles

The self-fulfilling prophecy asserts that making predictions often leads component actors to behave in ways that make the "prophecy" come true. In economics, a development of this theory has been popularized by George Soros under the generic name of "reflexivity." Interestingly, a consequence of reflexivity is that a statement that would have been false otherwise is made true by its own announcement. Several propositions can be derived from this idea, but the one most relevant to our context is that, in situations that have thinking actors, the actors' point of view is always partial and hence distorted. These distorted views can influence the situation to which they relate, because false views lead to inappropriate actions.

Technical trading rules can trigger reflexive actions and feedback loops. For instance, we can say that reflexivity would be in place when rising/falling prices signal to buy/sell and hence affect further price increases. Feedback loops work to move prices both up and down and work to amplify disequilibrium. One example was posed earlier when presenting a situation in which the forces of supply and demand driving prices depend on the decisions of investors who "collectively" agree on trading rules, that is, identifying resistance or support levels, and acting together in the same direction in response to such observations when given a shared expectation. The fact that all technical traders interpret this information in the same manner has trading consequences, and thus price and volume effects.

Is there a broad market reflexivity cycle in which technical trading rules are substantial actors? The causes of bubbles remain disputed by those who are convinced that asset prices often deviate strongly from intrinsic values. Many explanations have been suggested, and research has recently shown that bubbles may appear even without uncertainty,[83] speculation,[84] or bounded rationality.[85] In such cases, bubbles may be argued to be rational, when investors are at every point fully compensated by higher returns for the possibility that the bubble might collapse. It has also been suggested that bubbles might ultimately be caused by processes of price coordination[86] or emerging social norms.[87]

6
Contagion

6.1. Background on crises and contagion

Despite the vast amount of literature written on the propagation of shocks, there is no formal definition of "contagion." For instance, the World Bank refers to three different definitions: the broad definition, the restrictive definition, and the very restrictive definition, which can be respectively explained as follows:[1]

> Contagion is the cross-country transmission of shocks or the general cross-country spillover effects. Contagion can take place both during "good" times and "bad" times. Then, contagion does not need to be related to crises. However, contagion has been emphasized during crisis times;
>
> Contagion is the transmission of shocks to other countries or the cross-country correlation, beyond any fundamental link among the countries and beyond common shocks. This definition is usually referred to as excess co-movement, commonly explained by herding behavior;
>
> Contagion occurs when cross-country correlations increase during "crisis times" relative to correlations during "tranquil times".

Other sources also constrict the definition of this term to crisis environments, specifically alluding to the change of co-movements or shifts in cross-market linkages. For instance, Dornbusch, Park, and Claessens (2000, p. 3)[2] state that "contagion is a significant increase in cross-market linkages after a shock to an individual country (or group of countries), as measured by the degree to which asset prices or financial flows move together across markets relative to this co-movement in tranquil times," whereas Forbes and Rigobon (2002)[3] use the term *shift-contagion*, focusing on a change in the strength of market interconnections. Finally, contagion is also used to point specifically to the "transmission" factor itself, thus making use of a more general definition of it as a highly transmittable disease.[4]

An example of contagion repeatedly studied in financial literature was the effect of the 1997 flotation of the Thai baht. This triggered financial turmoil across East Asia. As a result, the countries affected hardest, Indonesia, Korea, Malaysia, and the Philippines, saw a devaluation of their currencies approaching 75 percent by December of that year. A second example was the reaction of a number of markets to Russia's 1998 defaulting on its sovereign bonds. The impact of the event was felt in places as distant as Hong Kong, Brazil, or Mexico, as well as in many other emerging markets and the riskier segments of developed markets. In addition to the decline in the value of the currencies, the economic impact on these countries included declines in equity prices, spikes in the cost of capital, shortage in the availability of international capital, and reductions in economic output.

But how do these impacts happen, and how do we measure them? What constitutes contagion and what does not? There are multiple transmission mechanisms and hence methods through which they can be analyzed. For instance, the literature speaks about correlations, even though correlations found among markets do not imply contagion. The literature also talks about shift-correlation, which entails a change in prior correlation. Contagion might be present if this is the consequence of an event. In addition to correlation or shift-correlation in asset returns, research also looks at other measurements such as changes in the probability of speculative attacks and the transmission of shocks or volatility.

Given that the term "contagion" has multiple meanings in the literature, we should note that here we use it to allude to the instance when an event is immediately followed by effects which impact other markets. Thus, we refer to a situation whose impact is transmitted from one place to the other and that generates a change locatable in time and space. This impact has to be immediate rather than gradual and have consequences that can be observed in hours or days. Instances in which there is no immediate reaction, even when some might eventually emerge in a gradual manner, are defined as *spillovers*. Common external shocks, such as changes in international interest rates or oil prices, are also not automatically termed "contagions." Even though the effects of shocks may be transmitted as a disease, these effects do not need to be bad per se. Consequently, it is adequate to think of contagion as a dynamic evolution propelled by events in one "market" which later affects other markets. The definition of contagion should suggest consideration of financial as well as macroeconomic fundamentals. Their impact can be thought of in terms of contagion effects.

These statements seem to imply that when one is talking about a financial crisis, one should be able to identify a geographical origin, say country A, and an event, say excess liquidity. The originators, receivers, and channels of transmission need to be identified, evaluated, and mapped. This is a considerable challenge. For instance, when referring to the 2008 crisis, most people understand the financial turmoil started in the US financial system (country A), within its subprime market (problem in country A). It was the collapse of this market whose ripple effects spread over the entire financial system and later, overseas. Country B (or countries B, C, D, and so on, the countries suffering from contagion effects) would then need to be identified and the mechanisms through which the crisis was transmitted to it (or to each of these countries) determined.

Three key elements have been proposed to distinguish contagion from other instances, such as spillover effects. The first is a reversal in international capital inflows. Contagion usually follows once an increase in international capital inflows helps "inflate" the bubble. The "sudden stop" problem, a drastic reversal of capital flows, helps deflate the bubble and, more often than not, worsens the ensuing crises.[5] The second is an "unexpected" announcement that sets off chain reactions. Since the announcement is unanticipated, investors are not able to adjust their portfolios in anticipation of it.[6] Finally, the third is a leveraged common creditor such as commercial banks, hedge funds, bondholders, or any other creditor who helps spread the contagion.[7]

Intimately related to the concept of contagion is that of "systemic risk," which also has multiple definitions. Here, we assume these meanings to say that it is the risk that a shock will materialize in such imbalances that it will spread, impairing the functioning of financial systems and reducing economic growth. This risk represents "general systemic vulnerabilities" rather than individual types of risks that affect single institutions. It is the transmission of disturbances between interconnected elements of the system and the aggregation of their effects which may ultimately have a negative impact on the real economy. For instance, credit risk, liquidity risk, and so on, when considered separately before a crisis can be directly related to a given institution(s); however, the interactions between them and their aggregation leads to systemic risk. In addition, as systemic risk spreads across entities, markets, and countries, the loss of confidence becomes a basic feature of its evolution.

Definitions of systemic risk most often focus on its various aspects. For instance, the definition might be related to the correlated exposures of financial institutions, information asymmetry, feedback effects, asset bubbles, contagion and negative externalities, or other factors that follow from systemic risk.[8] The general lack of consensus on what constitutes systemic risk and the complex nature of this phenomenon results in the need for various measures and principles to evaluate the same.[9] Here, it is worth highlighting Dow's (2000)[10] observation that in the most prevalent types of systemic risk, moral hazard plays a key role in disrupting the motives of financial institutions. However, regardless of how it materializes, systemic risk causes a loss in confidence and increased uncertainty about the functioning of the financial system and its parts. The concept of systemic risk lies in the contagion effect and negative impact on the real economy.[11] The bottom line is that, regardless of the precise definition, the emergence of systemic risk is a key cause of financial instability.

6.2. Channels of propagation: the empirical evidence

The 2008 financial crisis has yet again shown the need to improve our understanding of the channels of propagation and the nature of systemic risk in the financial sector. Contagion is one key propagation mechanism of systemic risk and crises.[12] In this section, we cover a partial review of the literature which has analyzed different financial mechanisms of contagion for some of the more well-known bubbles.

6.2.1 Financial linkages

Financial linkages as key means of contagion have been analyzed by a growing strand of literature. Overall, six main sources of systematic risk can be identified: common exposure to asset price bubbles, mispricing of assets, fiscal deficits and sovereign default, currency mismatches in the banking system, maturity mismatches, and liquidity provision.[13] In the banking sector, contagion might result from common asset exposure,[14] domino effects through the payments system or interbank markets due to counterparty risk,[15] or price declines and resulting margin requirements.[16] Also, as reviewed in the first two chapters of this book, common creditors and capital flows are examples of financial linkages.

Within the banking sector, much of the empirical literature has focused on the existing evidence of contagion via direct linkages between banks, such as through the mutual claims financial institutions have on each other. In fact, it appears that higher interbank exposure to a failed bank and weak fundamentals can generate large deposit withdrawals.[17] In this respect, recent events in the banking sector in Europe suggest that changes in expectations and coordination failure of depositors may not only be a source of individual bank runs but also an important channel of systemic risk. For instance, just in July 2012 (as a continuing result of the 2008 crisis), runs at Spanish banks caused outflows of $74 billion, equivalent to 7 percent of the Spanish GDP.[18] More recently, in Greece, bank deposits shrank by $30 billion between January 2015 and May 2015, once again leftover effects of the 2008 crisis.[19]

However, the analysis of the contagion of deposit withdrawals across banks has proven to be laborious and intricate. One reason is that it is not easy to disentangle contagion as a cause of correlated deposit withdrawals across banks from other potential explanations such as correlated liquidity shocks across households, correlated performance shocks across banks due to macroeconomic events, or common exposure to asset shocks.[20]

Still, a key insight from some of this research is that the possibility for contagion depends on the precise structure of the interbank market. That is, the effects of the same shocks depend upon the structures in a lending banking system. For instance, a "complete" structure of claims in which every bank has symmetric exposures to all other banks is much more stable than an "incomplete" structure where banks are only linked to one neighbor.[21] Also, "disconnected" structures where banks exclusively retain bilateral symmetric exposures are more prone to contagion than "complete" structures, but they avert spreading contagion to other banks. Finally, it is apparent that it is the values of the model's parameters which determine the possibility for contagion in a system with money-center banks where the institutions on the periphery are linked to banks at the center but not to each other.[22] The findings that determine which characteristics make a banking system more prone to contagion are insightful on their own; however, they do not shed light on the specific features of the system which are decisive in making the interbank market more prone to contagion.[23]

It is also difficult to extend findings from theoretical models to the more complicated network systems of the real world. For this reason, using data on actual exposures to test for the possibility of contagion in more complex structures, researchers have turned to simulation and the use of network analysis to identify patterns associated with contagion. An interesting conclusion thus derived is the negative and nonlinear relation found between contagion and capital and between contagion and the level of interbank lending to other assets. It seems that an initial increase in interbank lending (from a low level to a higher one) will not affect contagion as losses are absorbed by capital. However, when interbank lending exceeds a certain threshold, contagion rises swiftly. An increasing number of connections in the interbank network have two effects. On the one hand, adding links increases the channels through which contagion may occur. On the other hand, any further increase heightens resilience as losses are spread across a larger number of counterparts. The net effect varies with the degree of connectivity and the amount of capital in the system.[24]

Some examples of possible channels of contagion on the liability side of the banking system are bank runs through multiple equilibriums and fear of further withdrawals,[25] the common pool of liquidity,[26, 27] information about asset quality,[28, 29] the rebalancing of portfolios,[30] the fear of direct effects,[31] and the strategic behavior of potential lenders.[32] On the asset side we can observe both direct and indirect effects. Among the former are interbank lending,[33] the payment system,[34] security settlement system,[35] FX settlement system, derivative exposures, and equity cross-holdings.[36] Among the latter are asset prices.[37]

One aspect worth mentioning, particularly considering its role in the 2008 crisis, is the actions of arbitrageurs. In critical instances, when prices are significantly distorted and the participation of arbitrageurs most vital, the latter may leave the market. The reason numerous fully invested arbitrageurs liquidate their positions include the fear that further adverse price movements might later cause drastic outflows of funds.[38] This adds to the list of transmission mechanisms of instability.

The role of liquidity is similarly applicable.[39] Any leveraged investor will need to liquidate asset holdings to face margin calls. These liquidations will involve the better-valued assets, those whose prices have not diminished. However, in following this strategy the investor causes the prices of additional assets to fall, spreading the crisis to other assets and markets.

Commercial banks can worsen a crisis by calling loans and drying up credit lines in their home country and anywhere else they may have exposure.[40] These actions might respond to the need to rebalance the overall risk of the bank's asset portfolio, recapitalize, and bring the ratios back to their targets, following initial losses. However, by inducing a sudden stop in capital flows in the form of bank lending, commercial banks help spread crises.[41]

There are numerous examples of mutual funds serving as mechanisms in the spreading of crises. For instance, in the case of the Mexican crisis in 1994, US-based mutual funds had a key role in spreading shocks throughout Latin America. This happened when they sold assets from one country when prices fell in another.[42]

We should also consider the role of asymmetric information in these kinds of events.[43] Considering the premise that the long-term values of financial assets are determined by macroeconomic risk factors shared across countries and country-specific factors, contagion occurs when "informed" investors rebalance their portfolio's exposure to the shared macroeconomic risk factors in other countries to respond to private information on a country-specific factor. Given that "uninformed" investors are not aware of the reason for the change in demand, they respond as if the motive were related to information relative to their own country. The result is a shock that generates contagion across the countries' assets and markets. Thus, contagion can occur when the two countries share at least one underlying macroeconomic risk with a third country, through which portfolio rebalancing can take place. This scenario is more applicable to liquid markets and countries with a larger share of internationally traded financial assets rather than smaller, illiquid markets which are likely to be underrepresented in international portfolios.

Somewhat related is the "wake-up call hypothesis."[44] This hypothesis states that once investors realize the weaknesses prevalent in the country with the crisis, they abandon other markets which might share these characteristics.

6.2.2 Common creditors and capital flow linkages

In their role as common creditors, banks are a key force in the transmission of crises. The major crises of the last 30 years attest to this fact. For instance, in the 1980s, it was the US banks that were doing the bulk of the lending to Latin America, whereas in the 1990s, it was the European and Japanese banks lending to the transition economies in Asia and the Spanish banks lending to those in Latin America.

Collectively, the Asian crisis countries (with the exception of the Philippines) made up 65 percent of the Japanese banks' emerging market loan portfolio in the crisis years. The latter's lending to Asia had increased at a rate of 25 percent between 1994 and 1997, reaching $124 billion by the start of the crisis. In addition, the international bank lending to the Asian crisis countries had grown at a pace of about $40 billion year after year. However, following the flotation of the Thai baht on July 2, 1997, credit lines were cut and bank outflows of about $47 billion began to ensue.[45]

The banks were not alone. In addition, the Thai crisis also triggered equity outflows through mutual funds that were deeply invested in these Asian economies. The countries with the most liquid financial markets in the region, Hong Kong, Singapore, and Taiwan, were the worst affected by the withdrawals. For these three markets, the abnormal withdrawals relative to the mean flow during the whole period was around 10 percent.[46]

Similar events were recorded for other periods and economies. For example, three years before the Russian crisis in the mid-1990s, international bank lending to the region grew at 14 percent annually. This time, it was the German institutions championing the exposure, with about 20 percent of their emerging market portfolio concentrated in the area. When the crisis hit, capital turned around and the bank flows which had been reaching the region at an average of $28 billion per year turned into outflows of $14 billion the year after. This downturn helps explain why other transition economies were affected by the Russian crisis.

The sudden halting of important capital inflows have been found to play a key role in the wake of a crisis. This inflow of capital can have numerous sources such as banks, funds or other financial institutions, or bondholders. One consideration is that on average, debt contracts may have short maturities. Thus, sooner rather than later, investors and financial institutions are faced with the choice of rolling over their debts or opting out. Consequently, when investors and financial institutions are highly leveraged and deeply exposed to the crisis country, they are also ready to back out on short notice.

It is important to realize that the effects of these flows differ greatly according to the specific situation. For instance, low volumes of international capital flows generally do not instigate major international impacts when financial crises occur. This is because investors and institutions have no need, or at least a less urgent need, to adjust their portfolios given the lower levels of exposure when the shock occurs. Also, when the shock is anticipated, portfolios are adjusted in advance, before the event takes place. In this respect, for instance, anecdotal evidence shows that for all the crises experienced in Latin America during the 1980s (hyperinflations and other defaults in Bolivia in 1985 and Peru, Argentina, and Brazil in 1990), the international repercussions were inconsequential.

6.2.3 Trade linkages

Trade linkages have also been proposed as an "obvious" channel for the propagation of crises across borders, including those that involve competition in a common third market. However, even though this seems intuitively logical and many empirical studies have attempted to find support for it, they have been largely unsuccessful. As a result, many studies conclude that trade links, whether directly via bilateral trade or through a third party, are a more "plausible" explanation than a real one. That is, contrary to what one might expect, trade links do not seem to be a channel of transmission, except in a few exceptional circumstances. This conclusion is twofold: first, contagion is transmitted among countries with little or no trade. Second, contagion might not happen among countries with important trade linkages. Thus, financial linkages are significant transmission mechanisms explaining the propagation of shocks, but that is not the case for trade linkages.[47, 48, 49, 50]

6.2.4 The surprise element

In many cases the start of a crisis and the level of contagion are related to an element of "surprise." This is even more the case when the countries potentially affected share a common lender who is unguarded to the initial shock, as when the lender is surprised by events and has no time to rebalance portfolios and reduce exposure to the affected country. This is in contrast to those instances where the crisis is anticipated; for instance via earlier downgrades of credit ratings and widening interest spreads, agents can limit losses by reducing exposure and hedging their positions.

6.3. Social learning, information cascades, and herding

6.3.1 Social learning

In an earlier section, we briefly reviewed some contagion literature from the point of view of the proposed transmission mechanisms among markets. However, other aspects of contagion are key to the analysis of bubble formation and the spreading of crises. These are related to social learning processes and the way this is achieved in markets and in society at large. In general, models replicating these behaviors use simple information structures and choices to study the completeness of information revelation and the rate at which it occurs.

There are many common examples of how the behavior of others influences our own.[51] For instance, before leaving home on an autumn day, we might glance out the window to assess the temperature outside. If the pedestrians we see are wearing heavy jackets, we might infer it is a cold day and don our own coats before exiting.

Imitative behavior influences almost every human activity, including investment and financial transactions. For example, if well-known financiers such as George Soros and Warren Buffett make a certain investment and this becomes public knowledge, other investors will make efforts to acquire the same financial asset. Ultimately, the collective imitative behavior of a large group of market agents acting in the same direction as a "herd" can affect the price of the asset. In the case of this example, the herding behavior can be explained with the likely assumption that Soros and Buffet possess "superior" information the other players do not have. However, other motives can also be listed to explain this imitative behavior. For instance, it could be that an agent is rewarded for "copying" the actions of others through compensation schemes that penalize deviation from the crowd, or that this herding behavior results from the agent being more comfortable when conforming to the opinion of the majority.

In most social learning models, agents base their decisions, at least partially, on the experience of others. However, social learning is an important aspect of the process of technology adoption as well, when "technology" is understood in a broad sense. Agents may learn by observing the actions of their neighbors and the payoffs resulting from these choices. For example, we might see how our neighbors choose a school for their children or what kind of vehicle they purchase. We might also see how satisfied they are with these choices, whether these decisions are reconsidered at subsequent stages or the various neighbors are homogeneous in their choices.

Social learning results from observing the actions of others, including instances in which information asymmetry prevails. For example, in *information cascades,* individuals attempt to profit from the information of others. It is the underlying assumption that "others" know better which often drives individuals to ignore their own knowledge and follow "the crowd."

In his 1993 paper, Alan Kirman[52] summarizes some aspects of economic models that can be comparable to the example of the ant contagion process he uses in this work. The first aspect is that it involves agents making different choices. These agents can be grouped according to "how" each particular choice is made, for instance, given the agent's taste, expectations, and so on.

The second aspect is that agents "recruit" others in their specific choices. This recruitment can be the result of one agent convincing others of his superior information; or it can be the outcome of some externality by which the second agent, having observed the action of the first, follows in the latter's choices; or it may come about from a general externality of the Keynesian "beauty contest type"[53] or be the outcome of technological spillover.

The third aspect is that the process is intrinsically dynamic. Consequently, once a given "regime" has been followed for a period of time, there will be a change. These changes can occur as the result of an evolutionary process in which convergence is achieved after the success of one strategy leads to the self-conversion of the players, or in a number of other ways.

In markets, the switching behavior of a given population has been related to processes that involve contagion, mimicking, or herding.[54] Such influence may be rational, but market players are often said to be irrational and their actions the result of convergent reactions, as "herds," or some contagious emotional response to events. However, as we know, many variables can have an impact on these events. Furthermore, a fully rational market may react to information in ways the researcher cannot observe. It could also be that, in retrospect, analyst forecasts and market prices are wrong and corporations react en masse to changing fundamental conditions.

6.3.2 Information cascades and herding

Let us imagine the following situation. An airplane lands at a new airport where information boards are not yet displayed. Consequently, all the passengers have to somehow find their way to the area where baggage is delivered. As you walk down the corridor, you come to a crossway where it is only possible to go left or right. Everyone has a "judgment" or "opinion" which we classify as a "private imperfect signal." Let everyone have a private signal "left," with probability 2/3 of the true best choice being to go to the left, or "right" if that be the correct way to proceed. So, everyone's signal is equally good, but the signal not perfect.

Now, assume that you are the third person in line and the two people ahead of you went left. In this case, even if your private signal/intuition tells you to go "right," it is optimal to go "left." The reason for this is that the first person had an "l" signal, because she went left. The second saw the first go "left" and would have figured out that the first individual's signal was "left." If the second person's private signal was "left," she would have walked left too. If her signal was "right," she would have been aware of one right and one left signal, and she might have walked either way.

Having seen both people preceding you walk "left," you know that the first had a "left" signal and the second had a better than even chance of having had a "left" signal as well. The actions of your predecessors have given you more than one "left" signal. Hence, even if your private intuition points to a "right" signal, rationally you should go "left."

Everyone after you will be aware that your choice was unrelated to your private information. However, the remainder of the crowd will be in the same situation and, for them too, it will be optimal to make the same decision and go left. A major consequence of *information cascades* is that a thousand rational individuals will choose "left" just because the first two walked in that direction, even if the true best option was to go "right."

In conclusion, cascades predict the occurrence of massive social imitation which, occasionally, will lead the "herd" to select the wrong option. Given that everyone knows choices in a cascade are made on the basis of very little information, these are "fragile" in that just a little bit of new public information can make a big difference. Thus, an *information cascade* is a situation in which, based on observation of the actions of others, every subsequent agent makes the same choice, independent of her own private signal. Everyone is acting rationally on an individual level, and, even if collectively the participants have strong information in favor of the correct action, each and every agent may choose the wrong one. Furthermore, given everyone is aware of the scarcity of the information with which choices are made, a hint of public information or an unusual signal can change the options of a long-standing informational cascade. That is, even if the thousand individuals before have chosen to go left, seeing an airport helper carrying a wheelchair to the right might induce the next thousand passengers to take the opposite path. Hence fragility is an integral component of the information cascade theory.

There are numerous theoretical applications of information cascades. For instance, some explore the impact of social structure (multidimensional links) on informational externalities, the impact of the bandwagon effect and individual preferences in prices and the conditions which prevent or encourage cascades such as information revelation, market conditions and sectors, and learning, networks in cascade processes.[55]

Obviously, there is much literature published on information cascades[56] with a mounting ton of empirical evidence on the existence of herd or "crowd" behavior in speculative markets.[57] Though various authors take different approaches to the subject, there are two key ingredients of an *information cascade* model: sequential decisions with subsequent agents observing the decisions (not information!) of prior agents and a limited action space, such as an adopt/reject decision. In addition, there are some technical assumptions. Among them are private information that is bounded and imperfectly informative, homogeneity of agents, and a constant cost of adoption. The relaxation of these generally decreases the speed at which cascades form.[58] A last assumption, exogenous ordering, can also be relaxed, but it often leads to an increase in the speed with which cascades emerge.[59]

Information cascade herding happens when individuals disregard their own information to copy the actions of others. In these scenarios, the existing aggregate information becomes so overwhelming that a single piece of private information is not strong enough to reverse the decision of the crowd. Thus, one makes the choice to mimic others and ignore one's own knowledge. This scenario holds for an individual and anyone acting after the initial person. It is the domino-like effect that is then referred to as a "cascade."

Notwithstanding what has been established, the imitative behavior is sometimes due to specific reasons and thus is branded in concrete ways. For instance, if the imitative behavior is due to positive reputational externalities resulting from acting as part of a group, this is then called *reputational herding*.[60, 61, 62]

Researchers are investigating the incentives investment advisors are offered when deciding whether to herd and, in particular, whether economic conditions and agents' individual characteristics affect the likelihood of herding. From the perspective of the managers who are concerned about their reputations in the workforce, herding might be rational.[63]

Particularly among professional investors, the less information you have, the more important your incentive to follow the consensus. Herding among investment newsletters, for instance, is found to decrease with the precision of private information.[64]

Another variation is known as *investigative herding* and occurs when an analyst investigates information she thinks others might examine later. In this scenario, the agent would prefer to keep the information private. However, she can only profit from it if others follow in her footsteps and trade in the same market. Thus, she will be left holding the asset unless the information is shared and others push the price in the direction anticipated, and thus cause the desired price change.[65, 66, 67]

Finally, *empirical herding* refers to observations of "herding" without reference to a specific model or explanation. There is indeed evidence of herding and clustering among pension funds, mutual funds, and institutional investors when a disproportionate share of investors engage in buying or selling the same stock. These works suggest that clustering can result from momentum-following, also known as "positive feedback investment," that is, buying past winners or repeating the predominant buy or sell patterns from the previous period.

6.4. Theories of contagion

Research on bubbles has tried to answer the question of whether security prices are justified by the assets' fundamentals. The realization that, occasionally, prices of financial assets deviate from their fundamental values has been explained by speculative behavior that disappears after a certain period of time. Noise trader models and psychological or behavioral models are being used to explain these changes in asset prices[68, 69, 70, 71] as well as the transmission of their impact through the financial system.

Contagion encompasses a series of mechanisms through which financial crises are transmitted across markets. These mechanisms can be grouped under two headings: (a) investor behavior that gives rise to the possibility of herding and fads; and (b) economic linkages through trade or finance, mainly cross-border capital flows and common creditors.

It is extremely difficult to establish well-defined categories to group the research on contagion. First, the sheer size of the existing literature is gargantuan, but second and most important, the study of this phenomenon has been approached from various angles, using different models, assumptions, and data; therefore, it can be found under multiple headings. Here we review some key works to offer an informative overview of the most important aspects the work encompasses, even though these categories are neither exhaustive nor mutually exclusive.

6.4.1 Herding and cascades

Standard models of informed speculation assume traders have long-term horizons. Nevertheless, if traders have short-term horizons, they may herd on the same information and, when no better information is available, they may rely on the actions of other market participants they believe are better informed.[72, 73] In addition, a speculator is more willing to buy or sell if she sees most traders buying and selling. That is, trading in itself is a signal that others follow to trade.

Herding has been theoretically linked to many economic activities,[74] such as investment recommendations,[75, 76] price behavior of IPOs,[77] fads and customs,[78] earnings' forecasts,[79] corporate conservatism,[80] economic conditions and agents' individual characteristics, and delegated portfolio management.[81] The first works on this subject are by Banerjee (1992),[82] Welch (1992),[83] and Bikhchandani, Hirshleifer, and Welch (1992)[84] who show that *information cascades* and/or "herding" will eventually occur with some degree of certainty.

Herd behavior is imitative behavior that results in the contagion of opinion and actions. The main explanation for it is that the actions of the "other" might be influenced by better information, at least in the mind of the observer. This assumption may be false but, nevertheless, may lead to self-reinforcing fluctuations. If learning is sequential, then a *cascade* occurs.

Under herd behavior, bubbles are explained as the consequence of infection among traders which results in clearing prices that deviate from fundamental values. The higher the returns, the more willing the speculators are to follow the crowd, since many follow positive feedback strategies. This could partially explain excessive stock market volatility during these periods. The group thus magnifies price shocks.[85] In this context, bubbles are thought of as a temporary phenomenon, which leads to price fluctuations around fundamental values and eventually moves far from these.

There are two main reasons why it might be rational for speculators to trade over short horizons. First, some traders may not want to invest all their wealth in long-term investments, allowing them to take advantage of any good opportunities that might come along.[86] Second, some speculators (e.g. money managers) may need to prove their skills to justify a high salary and the authority to manage an important portfolio.[87]

There are at least three other reasons for this phenomenon. First, institutional investors may share an aversion to stocks with certain characteristics. For example, they may try to avoid stocks with lower liquidity that would be difficult to sell or prefer stocks that are less risky although they offer lower returns.[88] Second, given that all managers receive the same information and analyze the same indicators, they may trade in the same direction if they reach the same conclusions on what investment strategies to follow.[89, 90] And third, managers may ignore their private information because of the risk inherent in acting differently from other managers.[91]

In the General Theory, Keynes (1936)[92] states his doubts about the will of long-term investors to fight against market trends and ensure efficient investment. The reason Keynes provides for this is that if investors act according to their own information and beliefs, the resulting atypical behavior may hurt their careers:

> It is the long-term investor, he who most promotes the public interest, who will in practice come in for most criticism, wherever investment funds are managed by committees or boards or banks [. . .] Worldly wisdom teaches that it is better for reputation to fail conventionally than to succeed unconventionally.

Keynes proposes managers "follow the herd" because they worry about how others will assess their professional performance. If real, this kind of herd behavior may have important consequences. One example is the explanation provided for the bull market before the crash of October 1987:

> The consensus among professional money managers was that price levels were too high – the market was more likely to go down than up. However, few money managers were eager to sell their equity holdings. If the market did continue to go up, they were afraid of being perceived as lone fools for missing out on the ride. On the other hand, if the market declined they would take comfort in numbers, how bad would they look if everybody else had suffered the same fate?[93]

Following this line of thought, work has been done to model instances where the analysts with higher reputations herd to protect their status and salary.[94] In this line, it has also been shown that recommendations by investment bankers may result from a conflict of interest that lies in the compensation structure for equity research analysts as well as in the business practices imposed by the firm itself.[95]

Managers observe other agents' activities and may trade in the same manner if they feel these other managers are better informed, or if they see most acting in a similar fashion.[96] In support of this theory, Graham (1999)[97] developed a model in which an analyst is likely to herd if his private information is inconsistent with strong public information. In accordance with this hypothesis, his results indicated that a newsletter analyst is prone to herd on Value Line's recommendations.

Somewhat related is the evidence that a few funds follow other leader funds in their purchase patterns. This finding has been supported by the results that institutional investor trading patterns contribute to serial correlation in daily stock returns.[98] It has also been shown that stocks having the largest trade imbalances among investment companies (dollar purchases exceeding dollar sales) usually follow prolonged periods of positive abnormal stock returns.[99]

The informational price theory (IPT) explains how prices reflect information on the expectations from future earnings. In the market, some agents purchase information relevant to their trading while others derive it from the new price levels that result from trading by "informed agents." Because of this observed imitative conduct, it has been said that capital gains are most likely determined by the behavior of other agents.[100] This dual role of prices – affecting demand both through the budget constraint and through expectations – leads to different price elasticities from traditional models in which they play only the first role.[101]

In this respect, one other consideration is that hedging strategies create additional supply as prices fall. Uninformed investors may not be able to differentiate hedging activities from trades resulting from new information, consequently reducing their expectations. Unobserved supply shocks affect prices to a much larger extent than observed ones because investors would be unwilling to absorb the available additional supply until prices have substantially declined.

For example, after the stock market crash on October 19, 1987, economists and financial observers tried to find the reasons for the sudden change in prices. Since the negative trends identified had existed for months, the Brady Commission focused on internal market causes such as "price insensitive" strategies like portfolio insurance. The Brady Report suggested that institutional sellers drove prices down. Some analysts responded to this assessment by pointing out that though portfolio insurers sold about $6 billion worth of stock and index futures, portfolio insurance strategies were followed by less than 3 percent of stock market funds, representing no more than 15 percent of the total volume traded on October 19. Although the impact on prices would depend upon the elasticity of the demand for stocks, some studies suggest that the amount of selling could not, on its own, explain the price drop.[102, 103, 104]

One of the problems derived from the IPT approach is that since investors are rewarded for buying and releasing information on the actions of the market, observers cannot distinguish when actions result from information on "fundamentals" or from "other agents' actions."[105] When informed agents trade against uninformed agents, the uninformed can be led to believe that the asset is underpriced. By purchasing the asset at a premium, the uninformed agent's behavior can result in a bubble, since an increase in price tends to provoke a further increase, resulting in self-fulfilling bubbles. When the bubble bursts, the informed agent will be better off, but only at the expense of the uninformed one. Thus, a bubble can result from completely rational payoff-maximizing behavior on the part of all market participants.[106]

The IPT proposes that the theoretical asset market model shows excessive volatility relative to fundamentals if news is of high quality relative to prior information, or news arrives infrequently.[107] Some of the observers of the 1997 Asian crisis suggested that since the information available at the time in Thailand was incomplete and of poor quality, market participants could not evaluate the true state of the economy and therefore joined the stampede. It has also been noted that prices for commercial property in the Bangkok area fell by approximately 30 percent at the end of 1997, while the decrease in the value of the equity of real estate companies was much larger. The general opinion is that a lack of real estate transactions kept the prices up.[108]

A seminal paper by DeLong, Shleifer, Summers, and Waldman (1990)[109] demonstrated that shorter horizons on the part of "smart money" traders allow the behavior of noise traders to have a greater impact on asset prices. They argued that

> if sophisticated investors' horizons are long, arbitrage becomes less risky and prices approach fundamental values. (p. 713)

An inefficiency created by short horizon speculation is that traders might focus on poor quality data or on variables that have no relation to fundamentals. Speculators' demands depend on the information they observe. In forming these demands, they take as a given the number of speculators who are informed, the trading strategies of these speculators, and the pricing strategy of the market maker. Additionally, liquidity traders have inelastic demands of the asset. Two main classes of models in which short trading horizons can lead to inefficiencies are noise trader models and behavioral models.

6.4.2 Noise trader models

The buy and sell actions of investors move the prices of assets in financial markets. Hence, any deviation from a random walk in the price trajectory of a financial asset must be traced back to the behavior of investors. Accordingly, the mechanisms that may lead to positive feedbacks on prices, one of the main strategies followed by technical traders, are of particular interest as this results in the continuation of a trend. That is, as the price of a financial asset moves, traders bet on the persistence of the direction of the change so that a large cumulative move in the same direction ensues. Negative feedback, the opposite of positive feedback, tends to regulate growth toward equilibrium. However, in contrast, positive feedback results from it, in that the higher the price or return in the recent past, the higher the price growth in the future. Thus, when uncorrected, positive feedbacks can produce runaways so far from equilibrium that it can lead to the unleashing of other effects and the generation of crashes.

In a dynamic theory of asset price bubbles, positive feedback is of key relevance.[110] This is because speculative trends, the trend betting, ends up dominating fundamental beliefs that drive prices away from their fundamental values. This type of growth results in systems becoming increasingly susceptible to exogenous shocks, which eventually precipitate a crash. There are countless mechanisms in the stock market and the behavior of investors that may lead to such positive feedbacks. "Herd" or "crowd" behavior based on imitation processes is one of them. For instance, one can imagine a simple model in which investors choose trading actions based on the interactions and information they glean from others. It is the repetition of these interactions that consolidates into a cooperative phenomenon in which the market can suddenly "solidify" a global opinion.

Noise trading models assume less than fully rational traders. The agents are called naïve traders, noise traders, or chartists and make investment decisions based upon the study of past trends.[111] Positive feedback investors are those who buy securities when prices rise and sell when prices fall. Among other reasons, positive feedback results from extrapolating expectations about prices, trend-chasing, stop-loss orders (which prompt selling in response to declines in prices), and the liquidation of the positions of investors who cannot meet margin calls.

DeLong et al. (1990)[112] examine "positive feedback" traders who extrapolate past price trends and drive the asset price away from its fundamental values. In this work, the authors propose that if rational speculators trade on good news, they recognize that the initial price increase will stimulate buying by positive feedback traders. In anticipation of these purchases, informed rational speculators buy more stock, driving prices higher than fundamentals might advise. The next day, positive feedback traders buy in response to the current day's price increase, thus keeping prices above fundamentals. This will happen even as rational speculators are selling and, consequently, bring prices closer to fundamentals.

In an extension of their earlier work, DeLong et al.[113] develop a subsequent model with rational sophisticated traders and "noise traders." In this setting, the proportion of noise traders increases when they are making higher returns than sophisticated traders. This idea has been interpreted as "recruitment" to a successful strategy. The key assumption is that speculators think the bubble will last at least until they complete their trade.

Concurrently with these ideas, works by Blanchard and Watson (1982)[114] and Leach (1991)[115] suggest that when agents trade in a bubble environment, they only need to determine whether the last trader will execute his trade, since the agents want to make sure they can sell before others buy. Following this line of thought, a further development was presented by Wang's (1993)[116] study. This model proposes that noise results in an increase in prices due to the larger compensation component presented for greater risk.

Gennotte and Leland (1990)[117] define a crash as a discontinuity in the relationship between the underlying environment and stock prices. They state that a crash may begin with a small shock, such as a change in information, which translates into a large impact on prices and volume that is traded in the market. This could develop from investors hedging against losses and using portfolio insurance and stop-loss strategies. Their model reflects the intuition that when there are relatively few informed investors, markets may be much less liquid and more prone to being shaken if an unexpected supply shock occurs. According to the authors, the reason for this (backed by observation) is that supply shocks lead to lower prices, which in turn prompt uninformed investors to revise their expectations downward. Hedging induces greater excess supply and further decreases in prices. Furthermore, stop-loss strategies worsen the situation by generating a cascade effect of early exit, triggering an avalanche of further exit.[118]

Additional work on the effect that asymmetric information has in the formation and implosion of bubbles[119] demonstrates a bias toward myopic investment and concludes that insider traders could not realize profits if mandatory disclosure of information existed.[120]

To test the existence of contagion, one may look at variables such as volume, money flows, and volatility, since noise trader models propose a causal relation between each of these variables and stock returns. This is because these relationships are consistent with two assumptions made in the models: the trading strategies pursued by noise traders cause stock prices to move, and noise traders use positive feedback trading strategies.[121] Also, herd behavior can explain some of the incremental stock market variability; if many people follow the same trading rules, price shocks will be magnified.[122]

Many studies have tested for causality between stock price and volume.[123, 124, 125] This line of research tries to clarify whether trading volume helps predict prices, since volume represents information. Large positive price changes result in capital gains and provoke further transactions. The Granger Causality Test and the Baek and Brock Test have been used to analyze this relationship.[126] Evidence for the positive causal bidirectional relationship is, however, not supportive of the EMH.[127, 128]

Furthermore, there is a group of papers that investigate empirical clustering. These papers suggest that clustering can result from momentum-following (positive feedback investment such as buying past winners or repeating the predominant buy or sell pattern from the previous period). This should not be confused with herding behavior, but can still result in prices deviating from fundamentals.[129, 130, 131, 132]

6.4.3 Behavioral models

Behavioral models try to "make known" or "disclose" the process of contagion. In many cases, this process is approached using probabilities to analyze the dynamics of systems (like the market) where the units that constitute the system interact.[133] In our case, the units of the system are the speculators who trade in the market, and it is through their interactions that mutual infection of attitudes and opinions result.[134, 135, 136, 137]

Over time, many researchers have used behavioral models of financial markets.[138] After Kindleberger (1989)[139] explained the importance of psychological factors in historical financial crises, several authors have built models of stock market dynamics that explicitly include contagion of opinion and behavior. Thus, the findings of these models offer a behavioral explanation of the trading mechanism.

Lux's paper (1995),[140] in which the author explicitly models the psychological factors that influence the behavior of traders, the "unsophisticated traders, who do not possess information about fundamentals" is a very good example of such a behavioral explanation. In this work, Lux describes the formation of the traders' expectations and how these depend to a great extent on the behavior of others. Accordingly, the author uses a process of "mutual mimetic contagion" among speculators that is consistent with Topol's (1991)[141] theory of mimetic contagion, where market traders gather information about fundamentals from the bid and ask prices of others who, on the other hand, may be as uninformed as they are.

Following the same line of thought, Kirman (1993)[142] points to three issues observed in experiments with ants, which apply to behavior in financial markets. First, there is the fact that the morphological constraints and environment under which the species operate changes over time. Second, survival does not imply optimality. Third, even though a certain behavior may be locally optimal, the current situation may not be globally optimal. The type of interaction construed in the ants model does not depend on specifying any proximity between the agents (those who are more likely to meet). Hence, this approach can be extended to explicitly model the communication network and to study "change of opinion" within this framework. An important feature of Kirman's 1993 model is the suggestion of perpetual change in the process. This reflects the belief that financial markets are characterized by changes in the collective mood of investors, which result in periods where assets are undervalued or overvalued.

Seminal work was also presented by Shiller and Pound (1986),[143] who provide evidence consistent with the existence of herd behavior in the stock market. The authors survey institutional investors to determine the factors that go into their decision to buy a particular stock. Their findings show that the purchase of stocks that show price increases is motivated by the opinion of others (other investment professionals, newsletters, etc.). Nevertheless, when selecting stocks with more stable prices, research into fundamentals is comparatively more important. These findings suggest that it is likely that money managers invest in stocks even when their fundamentals advise otherwise. Using a discussion of classical epidemic models, the authors make the point that direct "interpersonal communication among peers seems to produce the kind of attention and reassurance that leads to changes in behavior."

Banerjee (1992)[144] also explores the idea that speculators base their decisions on other peoples' strategies and disregard their own information. In this study, the author starts by examining the behavior of consumers who, when choosing between two restaurants, select the one with more people in it. The reason more people enter this restaurant is that consumers assume the fuller restaurant offers better food, primarily because more people opt for it. This points to a retro-feedback mechanism. Hence, the author examines the implications of decisions that are influenced by what others are doing, particularly the fact that the decisions of others may reflect potentially important information that is not present in the public domain. In the context of sequential decision-making, people paying attention to what others are doing end up doing what everyone else is doing (that is, herding behavior), even when one's own private information encourages a different course of action. Once more, the herd externality is of the positive feedback type: if we join the crowd, we induce others to do the same. Hence, the signals perceived by the first few decision-makers, albeit random and not necessarily correct, determine where the first crowd forms, and from then on, where people will line up. This characteristic of the model also captures, to some extent, the phenomena of "excess volatility" in asset markets, or the frequent and unpredictable changes in fashions suggesting that "information externalities" is the most widespread explanation of herding.[145]

In conclusion, the behavioral view of bubbles finds support in experimental studies. These studies set up artificial markets with finitely lived assets and observe that price bubbles arise frequently. The presence of bubbles is often attributed to the lack of common knowledge of rationality among traders. Traders expect bubbles to arise because they believe that other traders may be irrational. Consequently, optimistic media stories and analyst reports may help create bubbles not because investors believe these views but because the optimistic stories may indicate the existence of other investors who do, thus destroying the common knowledge of rationality.

To identify instances in which bubbles can be born, behavioral models base their assumptions on empirical facts and evidence documented in psychology while attempting to explain other known return anomalies. Overall, behavioral models capable of generating bubbles can be grouped into four categories. While the first group of models (referring to *differences of opinion and short-sale constraints*) is only capable of creating positive bubbles, the rest can also engender negative ones.

6.4.3.1 Differences of opinion and short-sale constraints

The first class of models considers a scenario of disagreements between investors and short-sale constraints. A price bubble appears when optimistic investors fail to take into account the pessimistic investors' view who, short-sale constrained, cannot sell and consequently do not impound their opinions into prices. The resulting equilibrium is then above the fundamentals. The bubble disappears when investor opinions converge to a common valuation, or when the short-sale constraint is relaxed. Works such as Miller's (1977)[146] utilize this setting. Miller presents a simple static model for an overvaluation generated by disagreement and short-sale constraints. A number of dynamic models have been explored, following Miller's intuition.[147, 148] In the dynamic setting, the price bubble grows over the one under the static scenario, due to the fact that agents pay a premium over their private valuations in the hope of reselling later at higher prices.

Some works have analyzed the correlation between disagreement and ownership volume represented by the number of funds holding a stock. In these cases, disagreement about the value of a stock is smaller, as a larger cross-section of funds owns it. The results indicate that stocks with a smaller breadth of mutual fund ownership tend to be overpriced as, over time, they underperform in comparison to similar stocks. In addition, it is also found that stocks with higher levels of analyst disagreement are overvalued from the start as these underperform compared to those where disagreement is smaller.[149, 150]

6.4.3.2 Feedback trading

Feedback trading mechanisms where investors trade based on recent price movements have been proposed as a possible explanation of the price momentum phenomenon.[151, 152, 153] Under these models, an asset experiences an initial price increase in response to good news. Thereafter, it is feedback traders who push the price up, reacting to the observation of other feedback traders' buys. This retro-feedback mechanism continues so that prices exceed fundamentals until the rate of new capital inflow slows down and the growth rate of the price diminishes as well. Eventually the supply of new capital is exhausted, which causes the bubble to implode. Just as in a pyramid or Ponzi scheme, those feedback traders who get in early profit at the expense of those who come in later.

The model proposed by Hong and Stein (1999)[154] includes two groups of boundedly rational traders: news watchers and momentum traders. While news watchers observe private signals about the asset fundamentals, they do not trade on historical price changes as momentum traders do, ignoring information related to fundamentals. In this model, price changes occur gradually as the assumption is that innovations diffuse slowly through the news watchers. Given that news watchers are unable to assess where they are within the new cycle, prices overshoot the fundamentals and the latecomers end up losing money.

DeLong et al. (1990)[155] introduce a group of rational speculators in a similar setting. Contrary to expectations, these do not ensure that prices reflect the fundamentals but, while attempting to take advantage of feedback traders' irrationality, further destabilize them. The model shows how rational traders do not always trade against the anticipated mispricing. On the contrary, they align themselves with the mispricing, buying more of the asset today in order to resell it at inflated prices in the future, further contributing to bubble formation and maintenance.

The behavioral view of bubbles also finds support in experimental studies. These studies set up artificial markets with finitely lived assets and observe that price bubbles arise frequently. In these models, the presence of bubbles is mainly attributed to the lack of common knowledge of rationality among traders. Thus, news media has been found to play a key role in this process, reinforcing the behavior of feedback traders by reporting more frequently about assets with higher levels of investor attention.[156] In support of this finding, other works have also shown that financial journalists dedicate a comparatively disproportionate amount of coverage to Internet stocks which presented positive news during the Internet-bubble growth period. The news stories that were generally positive turned negative after the bubble's collapse.[157]

6.4.3.3 Biased self-attribution

Biased self-attribution is a cognitive bias according to which people selectively recognize signals that confirm their beliefs while dismissing others as noise. Daniel, Hirshleifer, and Subrahmanyam (1998)[158] introduce a model in which biased self-attribution results in bubbles. In this model, traders observe a noisy private signal which they use to form their initial valuations before a noisy public signal is revealed. This second signal is pure noise and should be ignored. Nonetheless, because of the self-attribution bias, traders use this signal to revise their valuation. Further public signals contradicting the initial private valuation are ignored, resulting in unchanged prices, while those confirming initial beliefs cause a price revision that inflates the bubble. However, once the aggregation of contradictory public signals grows large enough, the traders' confidence in their own private signal diminishes, deflating the bubble. Asem and Tian (2010)[159] present more recent empirical evidence in support of this approach.

6.4.3.4 Representativeness heuristic and conservatism bias

Representativeness heuristic results in overreaction because traders focus excessively on attention-grabbing news, while conservatism bias has the opposite effect given the traders delay, a response to reconsider their models in the face of relevant but not so "enticing" news. Both biases result in the formation of bubbles.

Barberis, Shleifer, and Vishny (1998)[160] use these biases to create a model in which traders' "imagined trends" cause them to underreact to relevant information such as current earnings realizations. That is, given that traders assume earnings follow a trending or mean-reverting process, they assess future earnings innovations as either being of the same sign as past innovations or by expecting past innovations to be reversed. The salient signal that makes traders choose the trending model is a run of earnings innovations of the same sign, which they take as a pattern. Thus, they ignore the small chance that any corporation can continue to grow for a long period of time. Pricing errors will remain prolonged until the aggregation of innovations makes investors reconsider and switch from the trending to the mean-reverting model when prices revert to fundamentals. As an example, it has been shown that closed-end country funds overreact to salient news but underreact to everyday news related to fundamentals that affect their portfolios.[161]

6.4.3.5 Bubbles in experimental settings

The behavioral view of bubbles finds support in experimental studies. A large number of experimental studies have consistently replicated the phenomenon. The experiments tested the assumptions behind theoretical models, specifically whether lack of experience, lack of common knowledge of rationality, and short-sale constraints play a role in the emergence of bubbles.

Many of these studies have built on the experimental design put forth by Smith, Suchanek, and Williams (1988).[162] In the experiment, the market consists of traders who are endowed with the asset and cash. They are free to trade the asset over the course of the experiment, which consists of 15 (or 30) trading periods, each lasting a maximum of 240 seconds. At the end of each trading period, the asset pays an uncertain dividend derived from a known probability distribution. All information is common knowledge. The traders who wish to buy or sell one unit of the asset can type their bid or ask for the prices on the computer screen, but only the highest bid and the lowest sell offers are displayed to the whole market. To accept an offer, the trader confirms the trade, which is recorded at the accepted price. Thereafter, the new highest bid and lowest offer prices are displayed once again. The traders' cash endowments are adjusted at all times by the accumulated capital gains and losses generated from trading, as well as by the accumulated dividends. Traders can continue to purchase the asset units so long as they possess sufficient cash to cover the purchase price. Short-selling is prohibited. At the close of the market, a trader's endowment is equal to the sum of the capital gains and losses from trading and the dividends earned.

The experiment revealed several interesting things. One was that price bubbles appeared frequently (a bubble was observed in 14 out of 22 experiments), its frequency greater the less experienced the subjects were. Furthermore, the mean price in the first trading period was always below the expected value of future dividends, consistent with traders exhibiting risk aversion. The initially low price could have created an expectation of future capital gain, possibly giving rise to a bubble. The collapse of the bubble was preceded by a decline in the number of "buy" relative to "sell" offers made by the traders and was accompanied by a lower trading volume than the bubble's rise.

A particularly insightful outcome of this experiment was that it proved unnecessary to provide traders with divergent dividend expectations in order to induce trade. Even in the presence of common knowledge and common priors, trades were motivated by the differing price expectations that arose because of the uncertainty regarding the actions of other traders and diversity in risk attitudes.

More recent work with a similar setup tried to assess whether traders' experience helps prevent bubbles from appearing in experimental markets.[163] In the context of one of these experiments, a subset of traders had experience as they had previously participated in other rounds of the game. The authors found that mixing experienced and inexperienced traders, even when the proportion of the experienced traders was only one-third of the whole, eliminated or substantially reduced the incidence of bubbles. Given that in real markets, the proportion of experienced traders is greater and their experience more substantial, the authors questioned whether bubbles can be realistically attributed to the prevalence of inexperienced traders at all.

In some follow-up experimental studies, the authors relaxed the short-sale constraints with varying consequences. In one such study, the results showed that short-selling made experimental markets more efficient and moved trading prices closer to the fundamentals.[164] However, a second paper published the opposite finding: short-selling did not make markets more efficient. Still, the experimental markets retained many of the properties associated with positive or negative asset bubbles: high transaction volume, large swings in price relative to the fundamentals, and sustained trading at prices different from the fundamentals.[165]

6.4.4 Recent generation of studies

Following on some of the mentioned works, a more recent generation of studies has centered on the impact that incentives, market frictions, and nonstandard preferences have on the creation and maintenance of bubbles. These three broad areas refer to the perverse incentives enjoyed by key market figures responsible for the spread of correct information to the market during bubble episodes; the preference for riding bubbles in states where asymmetry of rewards results from risk taking in scenarios of limited liability; and the herding of traders due to nonstandard preferences and incentive structures.

6.4.4.1 Reputational herding

In their 1990 paper, Scharfstein and Stein[166] propose a model where, according to some incentive structures, managers herd in order to maximize their "labor reputation." This model differentiates between "smart" and "dumb" managers based on their ability to understand the value of an investment and the signals they emit. The signals of smart managers are correlated with the truth, and those of the dumb with noise. In addition, good returns are also a function of random "luck". Unsurprisingly, the authors show that, given the market updates, its assessments of skill are partly based on whether a manager acts like the others and that managers will disregard private information to follow the actions of others. This behavior is inefficient from the societal perspective, but rational for managers concerned solely with their reputations.

Other works have shown similar results. For instance, Hong et al. (2000)[167] also report reputational incentives for analyst herding and show that organizational pressure to conform is particularly costly to the less experienced analysts. Evidence of "conformity" also affects stock buy/sell recommendations. This is because findings of revisions in proposals by a security analyst are positively related to revisions in the buy/sell recommendations of the next two analysts, which are also correlated with the prevailing consensus forecasts and consistent with the agency effect of reputational herding.[168]

Shiller's (2002)[169] literature review of behavioral finance also focuses on mechanisms influencing decision-making. Among others, the author proposes that given the constraints in time and resources that investment managers face, deriving information from the actions of other investment professionals is common. De facto interpretation of peers' actions as a result of their owning relevant private information explains herding in investment decisions. Relatedly, Lamont and Frazzini (2008)[170] find that mutual funds are compelled to invest in "high-sentiment" stocks and industries by investors who allocate money dynamically. One example proposed by the authors relates to what they call the "dumb money" phenomenon and presents an example of investors' preference for high-tech stocks during the dot-com bubble. This preference helped sustain the bubble.

With respect to herding as a key mechanism of bubble propagation in 2008, DeMarzo, Kaniel, and Kremer[171] present a rational general equilibrium model where relative wealth or the preference for "keeping up with the Joneses" is necessary to sustain bubbles. This is a finite-horizon overlapping generations' model in which the agents' main concern is consumption. In order not to fall too far behind their peers' wealth during the bubble's upside, if dependence is strong, agents prefer to participate as long as others do it too. Competition over future investment opportunities makes agents' utilities dependent on the wealth of their cohort and induces relative wealth concerns. Even though the bubble is likely to burst and lead to important losses, agents' relative wealth concerns make them afraid of trading against the crowd.

Just as in cascades, reputational herding occurs when an agent chooses to ignore her private information and mimics the actions of another agent who has acted previously. However, reputational herding models have an additional layer of mimicking, resulting from positive reputational benefits one enjoys as a member of a group. In this setting, interesting evidence has shed light on the fact that a forecaster's age is a significant variable that determines her willingness to follow the crowd. As a forecaster ages, evaluators develop tighter prior beliefs about her ability, and hence her incentive to herd with the group diminishes. On the other hand, the incentive for a second-mover to disregard her own information and mimic the leader increases with her initial reputation, as she is concerned about protecting her current status and income. That is, a second-mover incentive to mimic the leader "decreases with her own ability but increases with her initial reputation, the strength of public information that is consistent with the leader's actions, and the level of correlation across informative signals."[172]

6.4.4.2 Limited liability

One key element of bubble propagation during the 2008 crisis was "limited liability." This referred to the fact that most agents would benefit from a rising bubble, but not suffer a symmetrical loss upon its implosion. This effect had also been observed during prior crises. Thus, in their 1993 paper, Allen and Gorton[173] analyze "limited liability" as a likely incentive for money managers to choose to ride the bubble. In their model, there are also skilled and unskilled managers; the first group identifies and earns returns from undervalued investments, while the second does not have such skills. Hence, the unskilled managers invest in bubbles, thinking they can always get out and sell before the implosion. The limited liability structure bounds the downside, incentivizing this kind of behavior. Subsequently, in their 2000 model, Allen and Gale[174] also claim limited liability is key in bubble formation. The essence of their argument resides in this because borrowers get to keep the upside of their investments but are protected on the downside. This payoff structure generates a preference for risk among borrowers who then choose to ride the bubble. The borrowers' preference for risky assets starts the bubble, the size of which is a function of the riskiness of the asset.

6.4.4.3 Perverse incentives

With respect to the aftermath of 2008, three examples of misaligned incentives and incentive asymmetry were repeatedly cited as contributors to bubble formation and maintenance: rating agencies' inability and/or unwillingness to lower ratings for instruments and countries/institutions; auditors' reporting practices geared to hide the true financial condition of some corporations; and equity analysts advising customers to act in ways that went against their best interests. Two reasons behind these actions were the higher trading volume and investment profits gained by the culprits during the bubble upside, while the downside risk was being bounded by the limited liability of shareholders and management and the implicit and explicit government guarantees.

7
Bubbles

7.1. Introduction

Every major financial market crash has provoked an avalanche of studies trying to prove or disprove the existence of a bubble that might help explain the crisis. The discussion that ensues typically reflects on whether, as prices increased, investors realized the assets were overvalued. As it is often made clear that this should have been the case, explanations for the participation of agents in the formation of the bubble also begin to mount. During these inflationary events, speculators may trade because they think they will get out in time or because the expected high returns rewards them sufficiently even in the event of a crash.[1, 2, 3, 4] However, regardless of the motives for trading, when a bubble bursts, there is great discontinuity in market-clearing prices, and high price volatility follows as a result of excess supply.

Even though much has been written about bubbles, there is still no exact definition for this word. In general, and in this text, a bubble is formed when asset prices are no longer justified by their fundamentals. More formally, the role of the bubble is clear in the following pricing model built according to the EMT. Here, the price is determined by a competitive market and reflects all relevant information.[5]

$$x_t = aE\left(x_{t+1} \middle| I_t\right) + z_t \tag{7.1}$$

where

x_t is the price of the asset today
a is a parameter between 0 and 1
$E(\cdot)$ is the mathematical conditional expectation
I_t is the relevant information set
z_t are the fundamentals

If the transversality condition preventing departures from the fundamentals holds, its forward solution defines the fundamental value, z_t, as

$$F_t = \sum_{\tau=0}^{\infty} a^{\tau} E\left(z_{t+\tau} | I_t\right) \tag{7.2}$$

where

$E(z_{t+t} | I_t)$ is the mean of the expected fundamentals conditional on the relevant information set, and

I_t is a stochastic process since the realization of this information set depends on the variables not included in it.

If this condition fails, the general solution defines the current stock price, x_t, as

$$x_t = F_t + B_t \tag{7.3}$$

where F_t is the fundamental solution defined above, and B_t is the rational bubble component that is removed by the imposition of the transversality condition. B_t is an unstable, self-fulfilling set of expectations that violate long-run equilibrium conditions. Therefore, the bubble term may give rise to explosive solutions.

$$E\left(B_{t+1} | I_t\right) = a^{-1} B_t \tag{7.4}$$

If 7.2 is deterministic, as time goes to infinity, B will also approach infinity.

$$B_{t+1} = a^{-(t+1)} B_0 \tag{7.5}$$

If B_0 is nonzero, it will give rise to an ever-increasing divergence between the price, x, and its fundamental solution, even though it is usually assumed that it will not go negative. When 7.2 is stochastic, at each point in time, there is a probability that the bubble will burst.

Werner (1997)[6] shows that asset prices can be decomposed into a fundamental value and a pricing bubble where the fundamental value obeys the present value rule. The bubble may grow at a negative rate during some periods, such as when it bursts, or in those times when the discount rate r is negative. However, the bubble component may not be negative, given that a growing bubble will eventually make the asset price negative, and anticipating such an occurrence will suffocate the growth of the bubble at an early stage.

Relating these observations to valuation, we can define a positive bubble as one that occurs when an asset's trading price (P_t) exceeds the discounted value of its expected future CF:

$$P_t > E_t\left[\sum_{\tau=t+1}^{\infty} \frac{CF_{\tau}}{(1+r)^{\tau-t}}\right] \tag{7.6}$$

where r is the appropriate discount rate, required rate of return, or cost of capital used here interchangeably.

Formula 7.6 refers to Tirole's[7] (1985) model in which the market fundamentals' value of the asset is the discounted present value of its future payoffs in his example, proxied by the expected dividend payments, D_t. Tirole's proposal was that if the asset's price in the market is above that which can be justified by its fundamentals, then there is a bubble. The difference between the price of the asset and its fundamental value, $F_t - x_t$, is the part of the price that corresponds to the bubble, B_t, which cannot be justified on the basis of firm's fundamentals. Thus, when $x_t = F_t$ there is no bubble component in the price of the asset, and the size of the bubble at any given time may be assessed by using $F_t - x_t = B_t$.

Under the EMH, the prices of all stocks (P_t) equal the discounted value of their rationally expected dividend payment streams (D_t), the proxy for fundamentals:

$$P_0 = \sum_{t=1}^{\infty} \frac{D_t}{(1+k_e)^t} \tag{7.7}$$

where

P_0 is the price of the financial asset today

D_t is the future stream of dividend payments generated by this asset, $D_1 = D_0 + g$ or $D_1 = D_0\,(1+g)$, where D_1 is the dividend in the next period and g the growth rate of the dividend payment

k is the cost of financial resources, the risk-adjusted discount factor used to bring to present value the future CF

t is the time period under consideration during which the payment occurs

Should we consider the stream of CF in 7.7 a perpetual growing stream of payments, then

$$P_0 = \frac{D_1}{K_E - g} \tag{7.8}$$

This formula, the Gordon model used in the valuation chapter, can be adapted to different types of CF and not necessarily the dividend. Thus, when valuing the firm according to its fundamentals, we replace D with the actual FCF, ECF, or any other CF of the company deemed appropriate for the case, accordingly adjusting the discount rate as well. The time, the size of the CF, and the risk-adjusted cost of the financial resources represented here by the cost of equity k, or later more generally by r, are the sources of value creation. Thus, these are the variables we should concern ourselves with when assessing value.

The value of a company, such as that of any investment, depends on the wealth the said investment is able to bring its owners. Hence, if the securities traded in stock markets derive their prices from the professional activities carried out by their respective firms, these prices should reflect the firm's fundamentals. That is, there should be a relation between the firm's wealth generating capacity and the prices at which its shares are traded in the exchanges. Our formula (7.7) describes this fact: the current value of a given stock depends on the dividends to be distributed by the company and the timings, the costs, and the risks associated with these CF. Taking these ideas a step further, it can be assumed that dividends can only be paid out of the company's treasury. Hence, over the long run, there must be a strong relation between these payments and the firm's ability to generate cash from its regular operations. Our formula (7.1) refers more generally to this issue. The implication is that the price today (x_t) depends upon the fundamentals (z_t) and the expectation of the price tomorrow $E(\cdot)$, considering a relevant information set (I_t).

If the transversality condition, which prevents departures from the fundamentals, holds, its forward solution defines the fundamental value, z_t, as expressed directly in equation (7.2) and indirectly to formula (7.9) below. The price of the asset today, x_t, should be very similar to the firm's value according to its fundamentals z_t in order for the bubble process to not dominate the asset price.

If this condition fails, the general solution defines the current stock price as already indicated in (7.3), where F_t is the fundamental solution defined in equation (7.2) as the mean of the expected fundamentals conditional on the relevant information set, and B_t is the rational bubble component that is removed by the imposition of the transversality condition.

Given that, over the long run the dividend payments can only come out of the treasury of the firm, the "back up" for our dividend stream is the stream of future *FCF* the firm will generate during each period *t*; those we discount by the cost of the financial resources (*r*) to estimate the company's *PV* is

$$PV = \frac{FCF_1}{(1+r)} + \frac{FCF_2}{(1+r)^2} + \cdots + \frac{FCF_t}{(1+r)^t} \tag{7.9}$$

Thus, we have a conceptually powerful valuation method we can use to estimate the fundamentals' worth, even though the precision of our results will depend on our capacity to correctly project the *FCF*, *t* and *r*, variables which can house the bubble component. If we do not know anything about the company we are to assess, if we are not familiar with its context or its history or if we are not able to gather information on the management team and the sector in question (including its potential development), we will not be able to come up with a value that reflects a fair estimation of the company. In these circumstances, or in addition to the said data, we might gather further information about the company from the prices at which the stock is traded in the financial markets.

Value investors and short sellers use different trading strategies: the first specialize in finding undervalued assets and investing in them, whereas the latter attempt to locate overvalued assets to sell them short. Given that maintaining a short position carries important costs, such as a potentially unlimited loss, the risk that the borrowed asset is called back in advance, and a fee paid on the margin account, trading against overvaluation is more expensive than trading against undervaluation. Consequently, the former are more common than the latter.

Mispricing may result from chartists and feedback traders trading according to historical price information rather than fundamentals and uninformed or naïve investors deriving their information from current market values and volume information, hence perpetuating trends irrespective of the fundamental or fair values underlying the financial asset. Also, institutional restrictions may contribute to an unjustified price level. For example, when it comes to a downward correction, many institutions are forced to sell once a firm's market capitalization falls below the institution's investible universe. This selling pressure not only causes a further decline in price, but it is also likely to reduce the stock's liquidity. All other things being equal, a less liquid asset is also a less attractive asset, putting further downward pressure on estimations and thus forcing additional price drops. Feedback traders betting on trends will further magnify these effects.

Overall, there is abundant historical evidence of the fact that market prices cannot be consistently explained through present-value models. This realization has encouraged the academic community to research bubbles as a tool to model price deviations from present value relations. The earlier literature was dominated by models in which all agents were assumed to be rational, and a bubble could still exist. In more recent studies, though perfect rationality is not assumed, the relationship between bubble episodes and high trading volume, under which conditions bubbles burst and why arbitrage forces fail to ensure prices, reflect fundamentals at all times.

The behavioral view of bubbles, covered in the previous chapter, finds support in experimental studies. Over the years, academic study has expanded to explore the effects of perverse incentives and bounded rationality. The new generation of rational models identifies the incentive to herd and the limited liability compensation structure as pervasive problems that encourage professional money managers to invest in bubbles. Another problem that contributes to bubbles is that information intermediaries are often not incentivized to report findings that do not conform to their expectations.

At present, and once more, the relevance of research on bubbles has been brought to the forefront of the academic community, particularly in light of the US real estate bubble (2008) and the even more recent Chinese debacle. Thus, the question of why bubbles are so prevalent and how they can put entire financial systems at risk is once again a matter of concern for academics and policymakers.

7.2. An overview of bubbles

In general, tests for bubbles fall under one of two categories. The first category includes tests in which a bubble is hypothesized to exist while the researcher hopes to reject it in favor of an alternative that is not specified. Given these are joint tests of the no bubble hypothesis, their main problem relates to the assumptions made about the model (the definition and relation between fundamentals and prices) or the assumptions made about the time series properties of the fundamentals.[8, 9] The tests in the second category are hypothesized to be correct but are then rejected in favor of another one which includes a bubble. This methodology has also been questioned because the same attributes can be associated with both fundamentals and bubbles.[10] In response to the critiques, later studies have proposed more sophisticated methodologies. A third group of researchers has tackled the subject in a completely different manner, proposing rules for identifying cyclical bubbles or creating theoretical models and experimental conditions where bubbles can occur.

There are mainly two academic schools of thought on bubbles. There are those who believe that bubbles exist, even if, due to methodology constraints, it is difficult to unequivocally prove their existence. Then there are those who believe bubbles are an impossibility, even if, over time, their existence has not been thoroughly disproved. This divergence in opinion results from a disagreement on the definition of fundamentals and the assumptions, power, and appropriateness of the models and econometric tools utilized to test for bubbles.

Those who believe financial asset prices sometimes deviate from their fundamental values have suggested several reasons to explain why bubbles are created and maintained. For example, it has been proposed that irrespective of fundamentals, self-fulfilling expectations can and do drive prices.[11] Another theory is that as long as another investor is willing to purchase the same asset at a higher price, fundamentals may be irrelevant to some speculators.[12] Lastly, it has also been said that short-term investment horizon decisions can be influenced by considerations unrelated to fundamentals, such as liquidity needs, accounting issues, and so on.[13]

Regardless of the arguments proffered above, there are those within the financial community who do not believe bubbles can be formed and explain every financial crisis in terms of fundamentals. They hold that there are strong theoretical arguments to support the theory that prices do not diverge from present value levels.[14] Also, there are multiple econometric difficulties in testing whether stock prices are more or less explosive than dividends (as a proxy for fundamentals).[15] And, furthermore, those who don't believe that bubbles can be formed say prices can be justified by fundamentals not observed by the researcher as no one can assume to know all the fundamentals.[16]

Bubbles may present different characteristics. Some are similar to *pyramid* and *Ponzi* schemes where people benefit so long as new individuals are added to the chain of buyers and crash as soon as this is no longer possible. *Price level bubbles* are described as arbitrary, explosive, time-dependent elements that appear in the price solutions of dynamic models.[17] Stationary *sunspots* are an equilibrium during which people demand an asset for speculative purposes.[18] *Intrinsic bubbles* are very stable and persistent rational bubbles, which are deterministic functions of fundamentals that rest on aggregate dividends.[19] *Stochastic bubbles* start and crash continuously. *Deterministic bubbles* go on forever (as in cases of hyperinflation), have an explosive component, and cannot be reduced to stationarity by differencing. *Contagious bubbles* happen when stock price increases in one country or market provoke stock price increases in others.[20] *Rational bubbles* occur when an asset's price is influenced by its own expectation. This allows the price to depart from the assets' fundamentals, either permanently or temporarily, without assuming irrational investors. Rational bubbles give evidence to the belief that other variables, in addition to market fundamentals, affect share prices. In this case, arbitrage prevents any special trade opportunities and easy profit is not, therefore, characteristic of this type of bubble.

A number of studies have classified bubbles in order to identify the group that prevails under different market conditions. For example, a group of researchers attributes the occurrence of *speculative bubbles* to low interest rates. Sakakibara (1986)[21] provides an instance of this view, establishing it in a general equilibrium framework with rational agents. The author shows that a stationary sunspot equilibrium exists if and only if the interest rate falls below a certain level. In a 1989 paper, Gilles[22] shows that a specific kind of consumer patience or some kinds of technological pathologies are enough to ensure the existence of bubbles. Grossman and Yanagawa (1993)[23] agree with Gilles (1989)[24] in that bubbles are more likely to occur when households are patient and when investments in accumulative assets are very productive. It has also been proposed that intrinsic bubbles provide a more convincing explanation of the deviations from present value than do rational bubbles. This is because they can maintain deviations from fundamental values that seem stable over long periods of time.[25]

The new generation of rational models investigates how incentives, market frictions, and nonstandard preferences may play a role in creating and sustaining bubbles. The broad areas of these studies are nonstandard preferences and incentive structures that cause agents to display herding in their investment decisions; limited liability, which induces a preference for riding bubbles; and perverse incentives of key market players, whose role is to disseminate correct information to market participants during bubble episodes. Mainly related to contagion and herding, these have already been reviewed in Chapter 6. In this chapter, we introduce different kinds of bubbles and discuss some of the classical literature on them. Given the volume of the literature, we leave other texts and a more focused analysis of the models for the second volume of this work.

7.2.1. Rational and near-rational growing bubbles

To see how rational growing bubbles can occur, we consider an asset that is in *fixed supply and has an infinite life*. The agent's motives for trading are not relevant here; we just assume that all traders are risk neutral, have *the same discount rate "r" per period, and an identical information set "I" in period "t."* In addition, the asset distributes a dividend D, in period t. The first-order condition in the agents' maximization of intertemporal utility is

$$P_t = \frac{E(D_{t+1}|I_t) + E(P_{t+1}|I_t)}{1+r} \tag{7.10}$$

Equation (7.10) brings us back to our earlier discussion by implying that the current price of the asset, P_t, results from CF expected to be earned by the owner of the asset: the discounted sum of the expected dividend payment, D_{t+1}, and the next period's resale price $E(P_{t+1}|I_t)$. The expectation is conditioned by the information set I_t. A bubble in prices satisfies this condition if the bubble in t equals the expected discounted bubble in $t+1$. Equilibrium is then satisfied by bubbles B_t such that

$$B_t = \frac{E(B_{t+1})}{1+r} \tag{7.11}$$

or

$$B_{t+1} = (1+r)B_t + z_t \tag{7.11*}$$

where z_t is a random term with a zero mean and no autocorrelation. Because expectations are rational, there are two solution components to the difference equation in expectations: an arbitrary bubble, B_t, and the discounted expected value:

$$P_t = \sum_{i=1}^{\infty} \frac{E(D_{t+i}|I_t)}{(1+r)^i} + B_t \tag{7.12}$$

Thus, the rational price must equal an intrinsic fundamental value, the discounted dividend stream, and a bubble term, whether this last is worth zero or otherwise.

These kinds of bubbles must grow at a rate r, as they need to provide some return to sustain their own growth, ergo the participation of traders. One "complication" is that the intrinsic value depends on both the information set I_t and the discount rate (or expected return) r. When agents form different beliefs based on the same information or when they have different information sets and/or when discount rates vary over time, an agent's estimated intrinsic value is one of an infinite number of solutions.[26] Consequently, it is difficult to disentangle observed prices from unobserved intrinsic value in practice. This is particularly the case when we consider the role of "expectations" in the formula. However, it is useful to assess deviations from a "strict" estimation of intrinsic value if we keep in mind the basics of fundamentals' valuation reviewed in Chapter 4.

7.2.1.1. Models

In a perfectly rational world, where *complete information is shared*, the classical literature concludes that a bubble can exist in the price of an *infinitely lived* asset *if the bubble's growth rate is equal to the discount rate*. The idea is as follows. Assume the price of an asset, P_t, includes the fundamental component (DCF), P_t^F, and the bubble component, B_t. Then, $P_t = P_t^F + B_t$. The market price of this infinitely lived asset then equals the sum of the DCF and the present value of the future bubble component:

$$P_t = E_t\left[\sum_{\tau=t+1}^{\infty} \frac{CF_\tau}{(1+r)^{\tau-t}}\right] + \lim_{T\to\infty} E_t\left[\frac{B_T}{(1+r)^{T-t}}\right] \tag{7.13}$$

If the bubble grows at a constant rate r_B, so $B_T = B_t(1+r_B)^{T-t}$, and this rate is smaller than the discount rate $r_B < r$, the present value of the bubble is zero. Hence, there is no bubble. If the bubble's growth rate is larger than the discount rate $r_B > r$, the present value of the bubble is infinite and there can be no bubble either. If the expected growth rate of the bubble is the same as that of the return ($r_B = r$), only then can the bubble component of the price exist without bursting.

However, having said that, the idea is that rational bubbles must grow in "expected terms" per period at a rate r. The implication is that this growth rate per period does not need to be exact. Let us say that we expect the bubble to burst each period with probability p, so long as the bubble in the next period, conditional on *not* bursting, is $B_{t+1} = (1+r)B_t/(1-p)$ and that the expected growth rate is still r. The expected life of this kind of bubble is $1/p$ periods, and the probability that it lives n periods is $(1-p)^n$. As n grows, the probability converges to zero, meaning bubbles can exist in a rational trader world regardless of whether they know the bubble will burst at some point. The bubble's growth is consistent with rational expectations (RE) and the EMH because prices reflect the discounted price of its future and, given that current prices are adjusted for the expectations, traders cannot realize excess profits.[27]

The condition return $r_B = r$ constricts the set of possible bubbles. For instance, we can argue that bubbles cannot exist if the asset's price has an upper bound. This would be the case of an asset with a close substitute with which consumers can replace the bubble asset once the price becomes too high. There are a number of reasons why bubbles can be ruled out, some of which are discussed below. However, the possibility of rational growing bubbles does not depend on the restrictive assumptions used to derive (7.13).

In the same scenario (of rationality with perfect information), a second conclusion is that a bubble cannot exist in the case of a *finitely lived* asset. To illustrate this idea, think of *assets with a finite lifetime* of T periods, like bonds. In period T, no rational trader will pay more than the terminal value of the asset when it is liquidated for its fundamental value. As agents are informed of the fact that the bubble will burst at T, or, equivalently, that the life of the asset is limited to T, no trader with RE will pay more than the discounted terminal value in period $T - 1$. Then, it would also implode at $T - 1$. The same idea implies the bubble cannot exist at $T - 2$, $T - 3, \cdots$, all the way down to the present time. No rational bubble can start because all traders can anticipate with certainty its end.

An equivalent argument can be made for an instance where there are known wealth constraints (including borrowing ability). The bubble has a finite date in this case because it must stop growing at some point, and so the argument made above applies.[28]

Bubbles can exist within the *finitely lived asset* scenario when there is no common knowledge and a short-sell constraint bind;[29] one such instance is when speculators are *heterogeneously* informed about T and cannot short-sell the asset. This situation applies when the agent knows the asset is overvalued, but she does not know that others know as well. The agent's private information gained from market prices is only revealed in the final period, the outcome of a complex information structure which results in various states. The hope that she can hold to the asset and sell it later at an overvalued price is justified by her "expectation" that others are uninformed.[30]

In addition to the fact that *heterogeneous* information among agents and *restrictions on short selling* do allow for the existence of bubbles, if discount rates differ across agents or time periods, appropriate generalizations of the solution (3) hold. Also, when traders are risk averse, and when they must grow *faster* than $1 + r$ at a premium, bubbles can still exist.[31]

Getting out of the *finitely lived asset* scenario and moving into an *infinitely lived asset with a bubble component*, then, as T goes to infinity, the bubble portion of the price and the price-to-cash-flow ratio also approach infinity. This is a special case of rational bubbles in which the bubble is a function of the fundamentals rather than time. Froot and Obstfeld (1991)[32] called these *intrinsic bubbles* because they are deterministic functions of the fundamentals. Under this specification, the bubble relative to the fundamental must explode but can even disappear when the cash flow falls to zero.

Methodological and economic reasons may rule out rational bubbles. Furthermore, to be able to grow, price bubbles need certain "environmental" conditions: assets must be durable, supply has to be limited, exchange markets have to be effective, and there has to be a social mechanism for coordinating the common belief that a bubble exists and will continue to grow.[33] Thus, when these conditions are not present, in general (but just in general) we may be able to rule out the appearance of a bubble.

In addition to the instances already listed above, such as when money supply is limited and the life of the asset known and finite, we can also rule out rational growing bubbles when *the number of traders is finite, they have rational expectations and rational trading strategies*. The argument is as follows: given all traders will eventually retire, there is a date at which each agent will leave the market to spend her gains on consumption. Near retirement, the incentive to buy the asset is negative chiefly because it has an opportunity cost. For instance, each trader knows she has a chance of dying during each period, and the probability of this event increases with time. Traders with higher probabilities have higher discount rates than market discount rates which determine the bubble's growth rate. Given that buying the asset has negative discounted expected value for these traders, they will quit and retire. Since traders will eventually depart after they have sold the asset at bubble prices, the remaining traders play a negative-sum game. Anticipating this, nobody will initially buy at a bubble price.

Although the arguments above rule out rational bubbles under many conditions, the empirical evidence suggests that bubbles do occur. Hence, it may be useful to speculate on "near-rational" theories of bubbles.[34] We summarize a few of these ideas, as some of these proposals enter the realm of "contagion" and the theories discussed in the previous chapter.

Obviously there can't be any rational bubbles when there are no RE, such as in "adaptative expectations" and "near-rational bubbles." For example, optimism has been called near rational if it results from a "preference" or from "psychological conditions."[35] In this context, it is the instance when a finite number of traders expect to earn profit from participating in the bubble even after the initial winning agents leave behind a negative-sum game. From the perspective of the theory, each trader is equal under RE and thus should expect an equal portion of the negative-sum pie. Thus, expectations of earning a positive profit are not rational, even though people tend to be unrealistic about their own trading abilities and choose to harbor optimistic expectations.[36] Thus, optimism among a finite number of traders can result in a near-rational bubble if each of the traders left after the departure of the winner expects her profits to be positive, even though the aggregate amount of profits are negative.

Near-rational bubbles might also happen when the limits of the market's scope are not clear.[37] This can be seen in the example discussed earlier, in which each agent knows there is a bubble which should eventually burst but is uncertain about the extent of the knowledge of other traders. So we can say that each agent i has a guess T_i about the opinion of the other traders with respect to the likely life of the bubble. Under common knowledge, all the T_is should be zero. However, in the limited common knowledge context, the trader with the largest value of T_i will be stuck with the asset when the bubble bursts.[38]

7.2.1.2. Sunspots

Sunspots are *extrinsic* random variables upon which agents coordinate their actions, a formal representation of apparent *irrational* behavior. In these markets, traders respond to *information unrelated to fundamentals*. The volatility is incited by the speculators' actions and reactions to others and their predictions of other people's responses. These responses result in self-fulfilling optimal actions. That is, these *extrinsic random variables* do not affect economic fundamentals directly but have an effect on equilibrium outcomes because they influence expectations.

Sunspot models are complete rational-expectations and general-equilibrium models that offer an explanation for excess volatility. In terms of economic modeling, sunspots are realizations of *extrinsic uncertainty*, given uncertainty that does not impact the fundamentals of the economy. In proper sunspot equilibrium, the allocation of resources depends to a significant extent on sunspots. In this case, we say that *sunspots matter*; otherwise, sunspots do not matter.[39]

The original work on sunspots is derived from Jevons (1884),[40] who tried to explain the business cycle by its relation to sunspot activity. If sunspot activity affects economic fundamentals (weather patterns that affect crops, etc.), this is an instance of *intrinsic* uncertainty. However, if the sunspots do not significantly affect fundamentals, but have a substantial impact on the economy, then they matter beyond their effects on the fundamentals.

In modern economics, the term is no longer related to the solar phenomena. Instead, it is used to describe random variables that have no impact on preferences, allocations, or production technology but affect equilibrium such that the market outcome or allocation of resources varies in a manner unrelated to economic fundamentals. In other words, the outcome depends on an *extrinsic* random variable that matters only because people think it matters.

Thus, *market uncertainty*, the uncertainty generated by the economy, can be created by the economy itself through shocks to fundamentals (e.g. technologies), or it can be adopted from outside the economy as a result of coordinating the plans of individual agents. Uncertainty about economic fundamentals, which provides an explanation for volatility in the same, is *intrinsic uncertainty*. The market uncertainty which is not transmitted through fundamentals can then be driven by *extrinsic uncertainty*: that is, market psychology, self-fulfilling prophecies, and "animal spirits" are collectively known as sunspots.[41] Basically, sunspot theories provide a basis for the idea that business cycles might be caused by arbitrary, self-fulfilling beliefs.[42]

The sunspot equilibrium framework supplies a basis for RE modeling of excess volatility resulting from sources other than randomness in economic fundamentals. It aims to prove the existence of equilibria which result from *extrinsic* uncertainty. Proper sunspot equilibria can exist in a number of economic situations, including asymmetric information, externalities in consumption or production, imperfect competition, incomplete markets, and restrictions on market participation. For instance, the 2008 financial meltdown can be viewed as being partly sunspot driven.

Heterogeneity of beliefs is not the only source of sunspot equilibria, and any departure from the finite, perfect-market competitive equilibrium economy is *in itself* a separate source of sunspot equilibria allocations. These violations can occur when there are natural restrictions on participation in the securities markets, where there is market incompleteness, and when the overlapping generations model is not a finite model. The overlapping generations model is a relatively easy place to find relevant sunspots.

Markets may never be fully able to rule out sunspots, as it is almost impossible to bring in enough contingent claims that span all random events that could conceivably be important.[43] However, even if the contingent claims markets are complete, sunspots can still matter if agents have different beliefs regarding their prior probability.[44] For instance, pronouncements by renowned financial gurus could be sunspots. Speculators might recognize that these announcements carry no relevant information. However, they might still expect the announcements to affect prices in a manner related to self-fulfillment of beliefs.

Bank runs can also be explained by sunspots. Banks attempt to smooth consumption between depositors who can afford to wait and those who cannot. People "run" on a bank when they expect others to run. The problem can begin when those regarded as "patient" panic.[45] Even though sunspot-driven bank runs are possible, they are typically not mere randomizations over certainty equilibria. The reason for this is that if the probability of a run is small, the optimal banking contract will tolerate it. However, if the probability is large, it will not.[46] In the equilibrium, depending upon the realization of the sunspot signals, a full bank run, a partial bank run, or no bank run will occur.

Sunspots arise naturally in dynamic economies in which expectations play a key role. They matter when markets are incomplete, when participation in markets is restricted, when the horizon is infinite, and when preferences and/or technologies are nonconvex. Hence sunspot volatility is here to stay. Complete immunization of an economic system might not be feasible, or even desirable. This is because, though possible at times, government-designed policies that immunize an economy from sunspots can come with high costs and allow substantial inefficiencies into the system.[47, 48, 49]

7.2.2. Fads

"Fads" are social fashions in financial asset markets. Fads can be grouped according to where the "faddishness" is located within the stock price discounted dividend model.

In the first group, prices may fluctuate because the utility (such as what you get from artwork) from holding the financial asset might change over time.

In the second, these fluctuations may be explained by changes in beliefs in the asset's future intrinsic value. In the last group, the fluctuations may be due to fads in expected returns, a function of *r* which replaces the discount rate *r*. For instance, the impact in *r* could be dominated by variations in perceptions of risk, differences caused by cognitive errors or social forces.[50] Examples of belief fads include

1. the Tulip Bubble: investing in a product or service because it is "fashionable";
2. the South Sea Bubble[51] which was driven by unrealistic beliefs about the business prospects of English companies in the Spanish colonies; or prices of some high-technology stocks which skyrocketed as investors translated scientific success into economic success.[52]

A fad is a deviation between market prices and intrinsic value, F_t, that slowly reverts to its mean of zero:[53]

$$P_t = \sum_{i=1}^{\infty} \frac{E(D_{t+i})}{(1+r)^i} + F_t \tag{7.14}$$

with

$$F_{t+1} = CF_t + e_t \tag{7.15}$$

where C calibrates the speed of decay of the fad, and e_t is a zero-mean independent error term. When C equals zero, the fad disappears instantly; when it equals $1 + r$, the fad is a rational growing bubble. If it is less than one, the fad is not rational. In this situation, (7.14) does not satisfy the equilibrium condition: the expected return on the faddish part of the price is less than *r;* hence the agents would sell the asset, making the fad disappear. On the other hand, if C approaches one, the fad may decay so slowly that investors cannot profit from betting on its disappearance.

Distinctions between types of fads may be useful to help take adequate measures against their occurrence. For instance, a type of fad may persist if there is some incentive for agents to coordinate actions. An example of this is professional investors who coordinate their actions to minimize the variance between their performance and that of their colleagues. These traders will not go against the crowd because they are evaluated relative to the returns made by other professional investors.[54] In this context, investment is a coordination game and participating is utility maximizing. The fad carries on, since acting differently would put the individual at risk.

A large proportion of the work that provides evidence for fads derives from "price variance tests," tests of excess volatility in asset prices which use the "variance inequality" or "variance bounds" concept for the analysis of prices.

The essence of the concept is as follows. When a forecasted price and the realized price are the same, the forecasting error is zero. If there is an error, it should be unrelated to the price, since otherwise, it would constitute evidence of forecasting errors. Researchers thus attempt to decipher whether asset prices are too variable to be rational forecasts of future intrinsic values.[55]

Overall, the results of most studies comparing the variance of stock prices to that of a perfect foresight price (or ex post intrinsic value) constructed from historically realized dividends show that actual stock prices are significantly more variable than perfect foresight price. However, even though excessive volatility is acknowledged by a large proportion of researchers, the tests have serious methodological flaws which have been addressed in the literature. For instance, it has been claimed that the additional volatility of stock prices can be explained by risk aversion,[56] that sample size is relevant as it (this factor) might introduce a small-sample bias in estimating variances,[57] or the nonstationarity of the time series of dividends or prices.[58] Lastly, the large price fluctuation could be a response to the expectation of an extremely rare event that might affect dividends severely.[59]

More recent studies have addressed the methodological concerns listed here.[60] In general, the findings show large violations of the discounted dividend model, albeit much smaller than what was earlier predicted. However, objections to the newer methodological techniques also persist.[61] Two such criticisms are that it is difficult to assess the statistical significance of the variance in equality violations, and that the excess volatility found could be due to changes in discount rates.[62] Overall, two conclusions emerge: stock prices fluctuate more than perfect foresight prices, and prices might appear to fluctuate excessively, thanks to the small-sample bias, nonstationarity issues in the data, or changes in discount rates.[63]

Both rational growing bubbles and irrational fads can exist; however, they are distinct phenomena, even when stochastic bubbles that do not live long enough to grow large might appear to be very similar to fads.[64] Since growing bubbles are rational and fads are not, one may ask whether apparent excess stock price volatility is due to rational bubbles.[65] However, this seems unlikely[66] for the following reasons: First, a price bubble will affect both the actual price and the perfect foresight price. Hence, a bubble will not make actual prices appear more volatile than the perfect foresight price. Also, the dividend-to-price ratio should show a trend, as the bubble model predicts this ratio will fall as prices grow away from dividends.[67] Finally, rational bubbles cannot be negative, but there is evidence suggesting fads of asset underpricing do take place.[68] Hence, the excess volatility in stock prices might be explained by fads rather than rational bubbles. However, given that the causes underlying the growth of bubbles and irrational fads are different, we might need additional evidence to prove their differences.

7.2.3. Information bubbles

The information set available to the traders is of paramount importance in the setting and updating of the prices of financial assets. Prices reveal information; naïve traders can learn from smart or informed traders if they can correctly interpret the information provided within prices. This statement is true to a certain extent, as, given the amount of "noise" in the market data (like liquidity-motivated trading, positive feedback trading strategies, etc.), the tandem information and prices is not linked clearly enough to reveal information perfectly.[69] If prices do not reveal all information, or if agents have different information or different models of their economic world, prices may deviate from their intrinsic values and an "information bubble" may be formed.[70] Such deviations may or may not be rational. Information bubbles are typically temporary and usually small in size.

Another way to analyze this problem involves considering the process of information release, such as when agents observe prices based on trades which are actually consummated.[71] Traders have an incentive to gather such information because early trades are understood to be profitable. However, since early trades do not reflect the aggregation of information and show that noise and changing beliefs can cause prices to depart from true value temporarily at any point in time, these prices might not reflect perfect information.[72] Furthermore, an information bubble can also arise when no agent has any information, but all of them ignore that each of their peers is in the same situation. Noisy trading might lead people to read erroneous information in prices, leading to a greater disequilibrium. The ensuing positive feedback trading strategies might cause prices to increase and hence lead to the strengthening of belief.[73] These information bubbles have been called *mirages*, because people think they see information which is not actually there.

Information bubbles are difficult to detect because researchers cannot assess the quality of information available to all traders at every point.[74] However, the fact that asset prices are much more volatile during trading hours than nontrading hours has been accepted as evidence of the same.[75] The ratio of per-hour trading-hour volatility to per-hour weekend volatility, for instance, is about 70:l. This seems too large a difference to be explained purely on the basis of the amount of news published during the different schedules.

In 1968, administrative reasons led to the US exchanges being closed on Wednesdays. An interesting experiment was developed to take advantage of these circumstances. The premise was that the same amount of news was generated on all Wednesdays, no matter whether the exchanges were open or not. However, weeks when the market was closed on Wednesday were found to have about 83 percent as much variance as those during which the markets were open, leading to the conclusion that the most important determinant of volatility is the number of days the exchange is functional. This hypothesis was supported with evidence from other markets as well.[76] In addition, findings related to the volatility within different trading hours of the day, such as lunch hour versus other times, helped shape the idea that trading-hour volatility is high because trading is "self-generating." That is, traders have to occupy their hours with something, and observing trades leads to further trades, creating "mini-speculative bubbles."[77] These seemed to occur at hours during which traders can observe each other trade and suggests that these are information bubbles.[78]

According to some sources, one-fourth of the variance of price changes in the market is due to noisy trading.[79] Others estimate that about 4 to 12 percent of price volatility is due to self-generating trading.[80] Overall, noise can account for important empirical anomalies such as excess volatility, large return differences between equities and bonds, and the impact of dividend and capital structure policy on prices.[81] Noise trading pushes prices away from fundamental values in a number of ways. Furthermore, when traders use fundamentally different models, the interaction between classes of agents such as fundamentalists and chartists can produce information bubbles.

7.2.4. Bubbles in experimental settings

A large number of experimental studies have consistently replicated the bubble phenomenon. The experiments test the assumptions behind theoretical models, such as whether lack of experience, lack of common knowledge, and short-sale constraints play a role in the emergence of bubbles. One advantage to this setup is that mistaking an incorrect specification of intrinsic value for a bubble can be avoided. The reason for this is that in an experiment, intrinsic values are controlled. Nonetheless, a key problem remains in generalizing from experiments to natural settings.

Many experiments are modeled on the work of Smith, Suchanek, and Williams (1988).[82] In their pioneering work, subjects traded an asset which lived for 15 periods and paid a random dividend to whoever held the asset at the end of each period. Since the asset expired with a liquidation value, assuming risk neutrality, its intrinsic value in period T was its expected liquidation value plus the stream of $15 - T$ periods of expected dividends. This setting was repeated a number of times and the finding in most instances was that prices spiraled well above the intrinsic value in temporary bubbles that crashed in the last few periods. A bubble was observed in 14 out of 22 instances.

Also, as individuals became more experienced in the "game," crashes occurred earlier. Other findings included the fact that price bubbles appeared more frequently around less experienced subjects. Lastly, consistent with risk aversion, it was noted that the average price during the first trading period was way below the expected value of future dividends. The initially low price might have helped create expectations of future capital gains, possibly giving rise to a bubble. Finally, the collapse of the bubble was preceded by a reduction in the number of "buys" relative to "sell" offers made by traders and was accompanied by a lower trading volume than the bubble's rise. However, since it was public knowledge that the asset had a finite life, these were not rational bubbles, and they were attributed to a "greater fool" type uncertainty about other traders, and optimism.

An interesting outcome of this experiment was that it highlighted the fact that agents do not need to face divergent dividend expectations in order to induce trade. Even with common knowledge and priors, trades were motivated by differing expectations with respect to the actions of others and diversity in risk attitudes.

In a similar setup to that of Smith, Suchanek, and Williams (1988), Dufwenberg, Lindqvist, and Moore (2005)[83] tested whether trading experience affected bubble formation in experimental markets. In this work, a subset of the traders selected had already participated in several rounds of the game. When experienced and inexperienced traders were mixed, the results showed a substantial reduction in bubble formation, even if the proportion of experienced traders was just 30 percent. This led the authors to conclude that there is no relation between bubble incidence and trading experience.

Contradictory findings are also prevalent in experimental market tests. For instance, allowing short-selling has been found to have a positive impact on experimental markets, making them more efficient and bringing trading prices closer to the fundamentals,[84] while in other studies, the opposite result was obtained.[85] Still, these experimental markets showed high transaction volume, large swings in price relative to the fundamentals, and sustained trading at prices different from the fundamentals, that is, many of the characteristics associated with bubbles.

7.2.5. Related accounting literature

Since our proxies for fundamentals are to a certain extent accounting variables, it is also appropriate to briefly mention some of the accounting literature models and findings that relate to the relationship between accounting earnings and stock prices.

In general, accounting literature has investigated the relationship between corporate revenues and financial asset prices in one of two ways. The first approach is to use event study methodology to investigate the impact of earnings announcements on stock prices. The second is to create different trading strategies to determine whether financial information is impounded in share prices.

Overall, it has been shown that the average explanatory power of single or multiple factor market models is minimal. That is, most researchers report small coefficients and R^2s close to zero. For example, Lev (1989)[86] suggests that the R^2s in earnings-returns regressions are "too low" to be economically important.

This finding remains true even when one takes into consideration the effect of news on returns.[87] In contrast, it has also been found that the explanatory power of these models for specific firms can be very large.

Several authors have proposed different reasons to explain the above findings. One explanation is that prices may respond to information that becomes public throughout the quarter. It would then be possible that current events will not be reflected in the accounting earnings of the current period.[88]

Another explanation provided for the lack of apparent forecasting accuracy is analyst over-optimism. On average, the annual forecasting error (the difference between the expected and actual earnings growth), has been found to be around 7 percent.[89, 90]

In understanding the findings reported by different authors, one should be aware that the results obtained from different studies may not be directly comparable. For example, one should not directly compare the R^2 of the models where level data was used (i.e. prices) with those models where returns were used (the percentage changes in prices). The latter will always be smaller thanks to scaling differences. Also, one cannot compare the results across models in which data of different interval periods was used (quarterly earnings versus annual earnings etc.). The reason for this is that the longer the period, the higher the R^2.

There is a general consensus that expected returns are notoriously difficult to predict for many reasons, including modeling and econometric problems. Therefore, reaching general conclusions regarding the predictability of returns of different models is not a straightforward endeavor, and consequently, the literature provides ample evidence of conflicting findings. That remains the case whether we look at the contagion literature, the bubble literature partially covered in the next section, or the accounting literature as it relates to price predictability.

7.3. A partial history of the classical literature on bubbles

As stated in Section 7.2, in general, there are two main categories of tests for bubbles. In the first category, one sets the null hypothesis that a bubble exists and attempts to reject it in favor of an unspecified alternative. Most critiques of this method point out that these are joint tests of the no-bubble hypothesis, that is, the assumptions made about the model (the definition and relation between fundamentals and prices) or the assumptions made about the time series properties of the fundamentals.[91, 92] Within the second category, one sets the null hypothesis that a model is correct and attempts to reject it in favor of another that includes a bubble. This methodology has also been questioned because the same attributes can be associated with both fundamentals and bubbles.[93]

In response to the critiques against both categories of tests, later studies have proposed more sophisticated methodologies. In addition, a third group of researchers has tackled the subject in a completely different manner, proposing rules for identifying cyclical bubbles or creating theoretical models where bubbles can occur.

The following is a summary of some of the most relevant classical research published on the subject. Unfortunately, given the extent of the literature, this is merely a small representation.

In 1985, Tirole[94] designed a model in which the market fundamentals' value of the asset is the discounted present value of future payoffs from holding an asset. Therefore, if the asset's price is above market fundamentals, a bubble can be said to exist. This is a general equilibrium model that assumes an infinite succession of overlapping generations of asset holders having finite investment horizons. The author showed that as long as the growth rate of the economy is equal to or greater than the asset's required rate of return, bubbles cannot be ruled out.

An interesting paper by O'Connell and Zeldes (1988)[95] presents the general conditions that make Ponzi games possible. The authors propose that the existence of a Ponzi game depends upon the conditions of the lender-country economy, rather than that of the holder's. As an example, they suggest the case of the Third World debt situation, where the opportunity to roll over the debt perpetually depends on the health of the lenders' economic conditions rather than the strength of the borrowers' economy. The authors show that any monetary equilibrium can be replicated by a Ponzi game equilibrium with finitely lived debt. Nevertheless, it is also shown that some monetary equilibriums cannot be replicated by a Ponzi game equilibrium with positive coupon consols. These are consolidated annuities in the form of perpetual government bonds, which are redeemable at the option of the government. The holder does not know with any certainty when or if these will be ever redeemed, so the principal invested "seems to have been lost" in a way. However, consolidated annuities offer a yield in the form of coupons that provide periodic interest payments.

O'Connell and Zeldes's (1988) findings are similar to those of Tirole's (1985). Although their model is different in that it has a growing number of asset holders with infinite planning horizons, nevertheless, in both, the bubble grows exponentially at a rate equal to the required rate of return. The reason for this is that speculators would not hold an overvalued asset unless they expect it to be even more highly valued in the next period.

Concurrently, Diba (1988)[96] proposes that bubbles can arise in some dynamic linear rational expectations (DLRE) models. The author is critical of these models because, in his view, they imply that the existence of a bubble depends on the condition that the bubble be born on the first day of trading. In his opinion, the problem lies in determining whether the bubble exists from the beginning of trading, or whether the tests used are capable of detecting its presence.

Diba and Grossman (1987),[97] (1988)[98] argue for bounding the start of rational bubbles; otherwise, they argue, the bubble would grow permanently, increasing the difference between the share's price and the value of its fundamentals. One interesting result of this is that once population growth is allowed, the transversality condition does not rule out the existence of rational bubbles. Nevertheless, one can always eliminate the existence of bubbles by assuming that, at some point in the future, the wealth of new traders will be insufficient to buy the asset on which a bubble has grown. On the other hand, this could be taken as an argument in support of the bursting of bubbles and as explanatory of crisis events.

A practical challenge in RE modeling, and particularly in dynamic expectations modeling, is that a model can have multiple equilibria. Thus, multiple formal solutions often materialize, even if these are fully specified general equilibrium models. Fundamental solutions, bubbles, and sunspots have been used to describe various RE equilibria. With multiple equilibria, the predictive power of the RE hypothesis on its own, is weak; hence, additional criteria for finding "reasonable" solutions are needed. Examples of selection criteria proposed in the literature include stationarity of equilibria, minimum variance solutions, minimal state variable solutions, and expectational stability of equilibria.

It is within this framework of RE that Evans and Honkapohja (1992)[99] develop a model with multiple equilibrium solutions. Their hypothesis is that disequilibrium stability analysis, with respect to the formation of expectations, is a valid criterion for choosing among the elements present within the set of RE equilibria. They analyze the expectational stability (E-stability) of the different solutions of a linear RE model in which the endogenous variable depends on its own lagged value as well as expectations (formed in the past) of its current and future values. The results show that the continuum of bubble solutions cannot be strongly expectational-stable. Nevertheless, for certain parameter values, a particular solution that would usually be classified as a bubble solution may be strongly expectational-stable.

A year later, Allen, Morris, and Postlewaite (1993)[100] introduced a finite period, general equilibrium model of an economy with asymmetric information. The authors provide several samples of bubbles that occur when certain conditions are met. For a *strong* bubble to arise, one in which the price is higher than the dividend with probability one, they list three necessary conditions. These are (a) every trader has private information on the period and state in which the bubble is born, (b) every trader is restricted from short selling at some future time, and (c) agents' trades are not known.

According to Taylor (1977),[101] irrelevant random variables have been shown to cancel the solution to DLRE models (those in which multiple equilibria are equivalent to rational bubbles). The implication of this finding is that bubbles can augment the variance of asset prices. Thus, because of their contribution to the price variance, bubbles can cause prices to respond to seemingly irrelevant news as well as display disproportionate reactions to relevant news.

Partial equilibrium analysis of DLRE models suggests that stochastic bubbles are more frequent than deterministic bubbles. In this respect, Shiller (1978)[102] proposes that no finite set of initial or terminal conditions can exclude the existence of a stochastic bubble term in the solution to a DLRE model. Therefore, finding bubbles in the solutions of DLRE models does not necessarily imply that prices will conflict with fundamentals. It could also be that the price of an asset coincides with market fundamentals for a while, at least until the birth of a speculative bubble. In addition to these findings, Blanchard (1979)[103] shows that an existing stochastic bubble can burst at a random date and, in later models, he indicates that stochastic bubbles can burst and restart repeatedly.[104, 105]

The stochastic bubbles arising in DLRE models cannot be easily located in the stochastic general equilibrium model because the perfect-foresight general equilibrium propositions of Tirole (1985)[106] and O'Connell and Zeldes (1988)[107] cannot be directly generalized to an RE's framework. The problem is that the differences in the wealth between the old and new agents introduce nonlinearity into the models, while the partial equilibrium analysis of DLRE models assumes linearity. Diba (1988)[108] uses a stochastic version of the O'Connell–Zeldes (1988) model and gets around this problem by canceling out the wealth differences.

In 1981, Shiller[109] developed another test to determine whether prices reflect market fundamentals. This test is dubbed "variance bound" because it compares the variability in the price of the security with that in the fundamental value of the asset. The test assumes that the fundamental price of a stock today will depend on all expected future dividends adjusted by a discount rate. Consequently, today's price is a predictor of future returns. If the market price is consistent with market fundamentals, the share price should equal market fundamentals and the volatility of the estimated CF (the fundamentals' price) will not be greater than that of actual CF (the returns). Using dividends and prices from the S&P 500, Shiller shows that stock prices are much more volatile than the present value of discounted dividends and, therefore, cannot account for the variability of stock prices.

Though interesting, Shiller's (1981) findings have been questioned for the following reasons. First, investors discount future CF in different ways. But in his paper, it is assumed that the interest rate is constant. Second, one can assume that the market value of corporate shares equals the sum of all future dividends, if appropriately discounted. Nevertheless, one cannot determine people's expectations. Third, it has also been suggested that dividends are less volatile because firms smooth out these payments.

Furthermore, the methodology proposed by Shiller has been questioned on the grounds that the findings from "variance bound" tests may be due to a misspecification of the economic model.[110] Therefore, the test will not help determine whether there are bubbles in asset prices. Furthermore, volatility tests have been proposed to be inferior to Euler-equation tests as tests of discount-rate models.[111]

Over time, several other authors have presented results that help gain a better understanding of the bubble idiosyncrasy. Ikeda and Shibata (1992),[112] for example, examine a type of bubble that rests on market fundamentals. Their model of stock prices is in continuous time with dividends growing stochastically. In these bubbles, in contrast to Shiller's (1981) findings, stock prices can be less volatile than fundamental prices. The bubbles themselves show different patterns, and their correlation with market fundamentals changes over time. Also, the authors introduce crash risks (affecting the stochastic stability of the bubbles) to allow the bubbles to crash and exhibit different stochastic switching processes.

Security prices are believed to have a bubble component if market fundamentals are growing at a slower rate than the price of the corresponding asset. So, in this respect, Diba and Grossman (1988)[113] propose tests based on the stationarity properties of stock prices and dividends. The general idea behind them is that one can assess whether changes in security prices are consistent with those in market fundamentals by determining whether the growth rates are comparable. The tests performed by Diba and Grossman (1988) and others, and Campbell and Shiller (1987),[114] do not reject the hypothesis that prices conform to fundamentals, but rather suggest that real stock prices and real dividends are cointegrated. In their opinion, the deviations of stock prices from market fundamentals can be attributed to variables that possess a stationary mean.

Co-integration analysis seeks to answer the question of whether there exists a stationary, linear combination of two nonstationary variables. The intuitive idea behind such a relation is that although each of the two variables follows a random walk (i.e. they separately never converge to a traditional equilibrium) they are tied together and can never drift far apart from each other. This in turn means that knowledge of the level of one of these variables provides information on the value of the other, although past values of each have no explanatory power about their corresponding present level. If, in the long run, such a relationship between prices and dividends exists, we ought to expect these series to be cointegrated and, therefore, confirm the long-run predictability of asset prices defended in much of the literature.

The problem with Diba and Grossman's (1988)[115] findings is that, given the sample sizes, the stationarity tests may not have sufficient power to detect any possible rational bubble. In theory, one could generate a bubble with an explosive mean whose fluctuations in a finite sample could follow the behavior of a stationary bubble. Therefore, the key is to differentiate between bubbles that burst, with explosive conditional mean, and bubbles that do not burst, with a stationary unconditional mean of zero. Charemza and Deadman (1995)[116] propose that unless rational and intrinsic bubbles are constrained from bursting, these can exist and not be captured by the unit root tests.

Within the same general idea of examining the relationship between two time series, Allen and Gorton (1988)[117] work with a finite horizon model in which bubbles can occur. In order to examine the growth in market fundamentals and security prices, they evaluate trends in the data. If the trends are found in stock prices and dividends, the series are "differenced" (last year's price is subtracted from the current price). The differenced data are tested and, if trends are found, both series are differenced again until the transformed data has no trends. If the series of market prices is differenced more times than that of market fundamentals, it can then be assumed that a bubble is present in market prices. If, for example, trends are found for both prices and dividends, but disappear once they are differenced, this would mean that prices in that period have not grown consistently faster than dividends. In this scenario, stock prices would not include a bubble component.

Another version of this idea was put forth by Han (1996),[118] who uses Canonical Cointegrating Regression to investigate the long-run relationship between prices and dividends. The present value model implies that the levels, or the log levels, of stock prices and dividends are cointegrated when there are no rational bubbles. The author examines both the deterministic and the stochastic components of prices and dividends in order to determine the validity of the present value model. Nevertheless, his findings show that neither the levels nor the log levels of stock prices and dividends are cointegrated. Thus, he concludes that rational bubbles may exist in the deterministic component of the stock price.

Looking into the cofounding effects of information from fundamentals versus that from trading, Friedman and Aoki (1992)[119] propose that, even if the asset's price is not a sufficient statistic for aggregate information, traders respond to prices. This results in additional volatility. Consequently, the theoretical asset market model shows that prices are more volatile than fundamental values. This is particularly true when news is infrequent or of high quality relative to prior information. Nevertheless, in the authors' opinion, prices may still not be biased away from their fundamental values. The results of the tests performed by Campbell and Shiller (1987),[120] and Diba and Grossman (1988)[121] using this method do not seem to contradict the hypothesis that prices conform to market fundamentals. Nevertheless, the ability of these tests to detect the existence of bubbles has been questioned. For instance, Evans (1991)[122] built a model with periodically collapsing bubbles not detectable by using standard tests to determine whether stock prices are more explosive or less stationary than dividends.

To assess the existence of rational bubbles, West (1987)[123] proposes yet another approach. His idea is to use the constant expected return model to test the null hypothesis of no bubbles. The process is to perform a battery of specifications tests of a model's one period intertemporal equilibrium condition and then to test the cross-equation restrictions dictated by the model's market fundamentals solution on the projections of price and dividend on lagged variables. If the tests do not reject the intertemporal equilibrium condition but reject the specification of market fundamentals, one concludes that bubbles are present. The thought is that, in contrast to the misspecification of market fundamentals, a bubble would not bring about the rejection of the one-period intertemporal equilibrium condition. In this way, one can get around the issue of differentiating bubbles from unobserved fundamentals.

West (1987) applies his test to Shiller's (1981) annual S&P 500 data and strongly rejects the specification of the market fundamentals equation. He performs further tests of linearized models allowing for time variation in expected returns and finds that rational bubbles are present in real stock prices. Obviously, though, his conclusions rest on the power of these tests. West (1987) assumes the dividend forecasting equations are stationary in either the levels of real dividends or their first differences. Nevertheless, Flood, Hodrick, and Kaplan (1986)[124] argue that these tests may in fact have low power to detect misspecifications of the one-period intertemporal equilibrium condition.

It is apparent that all the tests used to determine the existence of rational bubbles in stock prices have been subject to criticism of some kind. The stationarity tests that are not very sensitive to a fundamental's misspecification may not detect some rational bubbles; and West's (1987)[125] methodology may require a better model of time variation in expected returns. Although the following two papers propose quite different approaches to testing for the existence of bubbles, they are still subject to the same problem of joint tests of the no-bubble hypothesis.

Renshaw (1990)[126] argues that consensus forecasts are more accurate and have a better track record than most forecasting systems based on parsimonious models. His proposal is then to (a) compare the dividend yield, price-earnings ratios, and money supply to some historical levels; and (b) to observe two times a consecutive-year double-digit growth of the annual financial returns for the S&P index.

Looking at this from a different angle, Chirinko and Schaller (1996)[127] propose that a bubble in prices could provide firms with a cheap source of financing and might therefore influence investment spending. They evaluate these issues using Q and Euler equations to be estimated in a simultaneous equations model where their fundamentals are expressed as the expected present value of future CF. The idea is that these equations reflect different information about the stock market and investment decision-making. Using US data for 1911–1987, their results suggest that bubbles exist, but that real investment decisions are based on fundamentals.

The second category of tests examines returns on empirical attributes of bubbles such as autocorrelation, skewness, and kurtosis.[128] These attributes result from the runs of positive abnormal returns and crashes. Hence, for example, there is autocorrelation of positive returns. Skewness is also present because of the bubble innovations. If the bubble continues, its growth rate is positive and small, whereas if the bubble bursts, there is a rare but large negative change. Kurtosis results from the mixture of low variance return distributions when the bubble is small and the higher variances that occur as the bubble grows. This means that if a time series of observations is scrutinized, one would see rather small values but, occasionally, very large absolute values. The assumption is that information lumping can cause price changes to have a huge variance.

Within the category of bubble tests that examine the properties of the data, McQueen and Thorley (1994)[129] test an implication from the rational speculative bubbles model where investors are aware of the bubble but believe prices will continue to increase for a while. In this scenario, one can expect a long run followed by a crash. The new implication suggests the probability that a run of positive abnormal returns ends should decline with the length of the run. The authors use duration dependence tests and adapt them for use with discrete stock runs data. They deal with nonlinearity by allowing the parameters to change with the persistence and sign of the run. In support of the authors' insight, duration dependence is an attribute unique to bubbles, whereas autocorrelation, skewness, or kurtosis can result from time-varying risk premiums, asymmetric news of fundamentals, and the arrival of information in batches. The results are consistent with the existence of bubbles, since there is evidence of duration dependence in monthly real stock returns.

Using equivalent tests to examine the empirical attributes of returns, Blanchard and Watson (1982)[130] also fail to reject the no-bubble hypothesis on the gold market using runs tests (autocorrelation). Nevertheless, they do find support for it in their tail test (kurtosis). In addition, Evans (1986)[131] finds support in the FX markets using a median test (skewness). The findings of these two papers are questioned mainly due to the fact, already stated, that autocorrelation, skewness, and kurtosis could also be associated with fundamentals.

Many authors have contributed their findings and thoughts to supporting the proposition that it is difficult to find irrefutable evidence for the existence of bubbles. Several authors state that one of the main problems is that market fundamentals are not directly observable.[132] It has also been argued that if the market price of an item partially depends on the expected future price change (self-fulfilling elements in expectations), the market can start a price bubble when the equilibrium price is set arbitrarily.[133]

This conveys some of the problems we might encounter when providing evidence for the existence of bubbles. In addition, within the first category of tests, failing to reject the presence of a bubble cannot be strictly interpreted as proving the bubbles' existence. In reference to the second category of tests, the rejection of a structural model cannot be solely attributed to the presence of the bubble. It may simply be that the model is mis-specified. Therefore, the findings under this method will be contested once again.

7.4. Frequently asked questions about bubbles

7.4.1. How are bubbles started?

There are several theories and no overall agreement on to how to answer this question. Nonetheless, a general conclusion is that the birth of bubbles is made possible by certain underlying macroeconomic environmental conditions, such as excess credit. Also, the hypotheses offered include the behavior of "rational" and "not rational" or "quasi rational" market agents. In addition, models consider misaligned incentives and nonstandard preferences. Taken as a whole, these theories encompass a number of economic conditions and market imperfections, specific behaviors, and interconnectivities between the micro and macro sectors across assets and boundaries.

As far as asset prices are concerned, the bubble life cycle follows a number of distinct phases that can be grouped around three key events. The first phase starts with an increase in money supply of some sort, whether it is the result of financial liberalization, a decision by the central bank to increase lending, a sudden boost in foreign CF, or some other equivalent event. The resulting expansion in credit is accompanied by an increase in the prices of assets, which continues as the bubble inflates. During the second phase, this rate of increase levels off, and as CF diminish, the bubble bursts and asset prices collapse, often in a short period of time. The last phase is characterized by the default of firms and agents who borrowed in order to buy the assets at their inflated prices. A banking and/or FX crisis may follow this wave, and the difficulties associated with these cause further and long-lasting problems in the real sector of the economy(ies).

Ignoring some of the macro conditions discussed in the first three chapters, we can group explanations on the origins of speculative asset price bubbles into two main categories. The first refers to divergence of opinion among investors and short-sales constraints, and the second proposes that once a bubble begins, it is difficult for smart money to eliminate the mispricing due, for example, to limits in arbitrage.[134] Thus, in conclusion, investors disagree on asset values. There are several reasons for this, and here we list a few.

In the context of the traditional models discussed earlier in this chapter, Diba and Grossman (1987)[135] rationalized the reason why bubbles are present from the start of the asset's trading and cannot be formed subsequently. On the other hand, the behavioral models presented thereafter argue bubbles may result from overreaction to innovations in fundamentals.

In the latter context, the origins of many bubbles can be explained by changes in the valuation of financial assets. One suggestion is that bubbles can be related to technological innovations resulting in productivity gains that were, originally, difficult to evaluate. This idea was revisited by Bernanke, Gertler, and Gilchrist (1999)[136] who added the effect of the "financial accelerator" resulting from the cheaper cost of capital that higher value firms can benefit from when, in a context of credit market frictions, they can offer better collateral guarantees. The interrelation between technological innovation and cheaper credit can retro-feedback to stimulate further investment with the consequent impact of expected CF and, thus, current valuation.

A second proposal relating bubbles to the birth of new technologies is presented by Hong, Scheinkman, and Xiong (2008)[137] and is posited within the behavioral context of agent disagreement and short-sale constraints. The bubble, according to them, is born out of a signaling equilibrium: when some investors fail to take into account the incentives of the advisors within the group of "technical-savvies" to inflate their assessments, a pricing bubble arises. Its size is maximized when there is a mix of "smart" and "naïve" investors in the economy.

A third idea refers to the "money illusion," when investors think of money in nominal rather than real terms. This is related to some of the psychological biases discussed in the third chapter. For example, the nominal purchasing price of a house can serve as an anchor for a reference price, even when the real price can be derived. The effect of anchoring, a form of framing effect, is described by Brunnermeier and Julliard (2008)[138] in the context of housing frenzies. In this work, the authors argue that a reduction in inflation can fuel housing prices if agents suffer from money illusion. The reason is that, when making renting-versus-owning decisions, many investors fail to recognize that if mortgage payments are fixed, in time, inflation lowers their real cost. Meanwhile, annual hikes in rental expenditures keep up with inflation. When inflation expectations are high, the agent would choose renting over buying, thus pushing down prices, while the opposite would happen when inflation expectations fall. The authors' findings show that changes in price-to-rent ratios are affected by the nominal rather than real interest rate.

Overall, to explain bubbles and contagion, behavioral models assume at least some traders are "boundedly rational." The behavior of these agents has been grouped into four categories:[139]

a. Differences of opinion and short-sale constraints: In this group, we have optimistic investors and pessimistic short-sale-constrained investors whose personalities and situations impact their actions and, thus, future outcomes.
b. Feedback trading: Feedback trading behavior results in a trading strategy based on recent price movements. When an asset's price increases, feedback traders push the price even further by purchasing it. This action attracts additional attention from other feedback traders, who, through their dealings, keep pushing prices upward to a point where they exceed fundamentals.
c. Biased self-attribution: This bias refers to a cognitive process by which people recognize signals that confirm their prior beliefs while disregarding others that contradict their earlier opinions.[140]

d. Representativeness heuristic and conservatism bias: These cognitive biases represent departures from optimal Bayesian information processing. Under representativeness heuristic, investors overreact to salient news by putting too much weight (high probability) on such signals relative to their base probabilities. The opposite happens under conservatism bias, when investors underreact to less attention-grabbing signals assigning low probability weights. Both biases can lead to the formation of price bubbles.

However, it has also been proposed that the "financial accelerator" can amplify positive shocks to the fundamentals.[141] This statement responds to a financial theory which asserts that a small change in financial markets can result in a large change in economic conditions and create a feedback loop. Thus, the term "financial accelerator" is used for the economic shocks amplification and propagation mechanism, which aims to explain how small economic shocks can have large and persistent effects on the aggregate economic activity due to market imperfections.

In addition to the financial accelerator, several channels have been proposed that might spread financial crises. Some models emphasize investor behavior that results in herding and fads, influencing the behavior of capital flows and financial markets and exacerbating booms and busts. Other models stress economic linkages through trade or finance. Thus financial linkages – cross-border capital flows and common creditors – and investor behavior figure most prominently in theoretical explanations of contagion.

Lastly, the vulnerabilities of the system need to be met by adequate regulation, supervision, and intervention. Thus, defective leadership resulting in excess liquidity due to governmental and private actions; lack of supervisory effectiveness from governmental agencies over private corporations, within private corporations, and from agencies into both government and private activity; unethical attitudes and actions which affect governance and regulatory and supervisory activities; and a general lack of transparency either due to the necessary information not being made public or being tainted. All these are serious problems that have lasting effects on markets and bubble formation.

7.4.2. Why do bubbles implode?

Within the context of the models discussed, the implosion of bubbles can be explained in a number of ways. For instance, a bubble may burst when uncertainty about the value of an asset (e.g. new technology) disappears and when the impediments to short-selling are gone. In addition, bubbles can burst when the feedback trading loop is broken. This can happen once the supply of new feedback traders is exhausted, when the inflow of new capital is cut off, or when the costs of trading rise. A bubble can also deflate when positive sentiment is reversed or after a strong negative signal.[142] Thus, when shared expectations of future price increases can no longer be sustained, busts may happen.

In order to keep growing, bubbles need capital inflows. Once the inflow of new capital slows down, prices begin to flatten out, the mood starts to change, and the bubble begins to deflate. Of course, it is difficult to predict when this reversal of sentiment will occur. If it is linked to the slowing of the bubble's growth, it will run parallel to the collapse of new investors, as would be the case with bubbles that result from a sudden expansion of credit. These will also deflate when credit tightens. In addition to a halt in foreign capital, governments may also tighten credit or take other measures to burst a bubble. For example, they may increase capital gains or transactions taxes. In all these instances, the said actions will increase costs and reduce feedback trading. Thus, governments may use regulation to help deflate bubbles. Lastly, a bubble will burst if arbitrageurs attack it by selling short a sufficient amount of the overvalued asset.

Bubbles on finitely lived assets have to implode before or on the asset's final trading day. This was the case of the Chinese warrants bubble; in this context, due dates were public knowledge. Hence the expectation was that the bubble would implode at any point up to that date, and that was exactly what happened. But rather than deflating at the last possible moment, the bubble deflated gradually through consecutive and accelerating price adjustments that largely happened over the last six trading days before the expiration of the option.[143]

7.4.3. What are the consequences?

The following is a partial list of the consequences of bubble incidents:

1. Bubbles create "fictitious" wealth and destroy real wealth. Even if these are inflationary occurrences, the final implosion of significant bubbles brings about a deflationary period during which wealth vanishes from the overall economy.
2. Bubbles reallocate wealth from some sectors of society to others. Smart money will initiate the cycle by ripping most profits in capital revaluations and fees, while unsophisticated investors and latecomers will be left holding most of the undervalued assets.
3. Bubbles transfer wealth from the future to the present. A share of future demand is brought to the present, stealing expectations as well as future consumption. This shift does not create sustainable wealth; instead, the resources that should have been available to make future demand possible are consumed in the present since bubbles are fueled by debt.
4. Bubbles steal liquidity from the system and are fueled by easy credit. Given that liquidity goes into financing the asset bubble expansion, these resources are no longer available to do the same in other sectors of the economy. Upon the bursting of the bubble, as assets are devalued and balance sheets require adjustments, credit simply disappears.

5. Bubbles induce misallocation of resources. Given the spectacular returns produced during the initial phases of bubble growth, suppliers misallocate their resources and, when unable to cope with demand during the initial phase, they overestimate future demand. The ensuing excess capacity leads to waste of capital (particularly worrisome in capital-intensive sectors) and eventually destroys know-how, production, and distribution capabilities.
6. Bubbles result in a post-implosion balance-sheet-induced recession. After the bursting of a bubble, the affected assets will be re-valued at their new lower level, while the value of the liabilities will not change. Thus, from that time on, the corporations' main objective is to reduce debt rather than to maximize revenue. At the aggregate level, during the deleverage process, no one will borrow money independently of the interest rate. Consequently, money entering the local banking system cannot revert to the local economy. Furthermore, if both the government and the private sector deleverage at the same time, the economy is weakened.

7.4.4. Should the government intervene?

It was back in the 1920s when the American journalist Edwin Lefèvre wrote thus:

> Nowhere does history indulge in repetitions so often or so uniformly as in Wall Street. When you read contemporary accounts of booms or panics, the one thing that strikes you most forcibly is how little either stock speculation or stock speculators today differ from yesterday. The game does not change and neither does human nature.[144]

Since those words were written, many booms and crashes have affected asset markets around the world. At the time of writing, we wonder whether another crisis is building in China. As of now, it is too early to present a diagnosis of the latter case; however, the observed financial frictions of the 2008 crisis have once more raised questions on financial stability and what can be done to ensure it. Issues such as the resiliency of a financial system, the extent to which risk is exogenous, the quantification of systemic risk, and the role of financial innovation in the destabilization of the system need to be re-assessed. Given the seriousness of the effects of the implosion of bubbles, responsibilities need to be assigned for control of their development and remedies that can be offered upon their implosion.

The phenomenon of risk shifting shows that when agents borrow to invest in pre-existing assets, there can be an increase in prices beyond fundamentals. Risk shifting results from the limited capacity of lenders to assess the risk of their borrowers' investments. Due to risk shifting, borrowers bid up prices on the fixed-supply asset, exacerbating the bubble[145] and the ensuing crisis. Credit expansion interacts with risk shifting by encouraging investors to fund risky investments on the current date. Thus, credit expansion has a contemporaneous effect on asset prices. In addition, the anticipation of future credit expansion can increase current prices and thus further increase the likelihood of an eventual crisis.

A general assumption is that arbitrageurs will eliminate mispricing. However, there is plenty of evidence[146, 147, 148] to the contrary and several reasons for it. The first is that arbitrageurs have to close or scale down their bets to meet margin calls for short positions if the bubble continues to grow.[149] A second problem relates to the possibility that fundamentals may change and the assets may lose value. Eliminating mispricing in these circumstances can be costly, given that the risk of potential mispricing is high and the costs of arbitrage are as well.

Informational asymmetries can also play an important role in mispricing. When the former are large, trading costs will increase as a function of the size of trade. The potential impact on the arbitrageurs' profits will dissuade them from the trade[150] and many arbitrageurs will be required to burst a bubble. A forceful synchronized effort will be necessary, particularly when the impact of each is relatively small.

Findings in the literature have shown that[151] upon becoming aware of the bubble, arbitrageurs optimally choose not to short-sell the overvalued asset, but rather ride the bubble for a period of time. Thus, under certain conditions, arbitrageurs may choose to participate in the mispricing, ultimately increasing their profits at the expense of other investors. Furthermore, if they wait long enough, the speculative attack may be deferred until the bubble bursts for exogenous reasons.

The relationship between the banking system and financial crises is also a matter of concern.[152] Many deficiencies were observed during the 2008 crisis. Here, it is sufficient to say that, though to different degrees, banks in general showed significant failings in corporate governance, particularly with respect to risk and liquidity management. How the banking system deals with risk sharing during crises and how defaulting and the resultant disruption in the financial sector spills over into the real economy and the distortions this creates – need to be considered.[153] Default costs to be borne by banks ex post and the impact of the same on the functioning of the financial system (as well as new levels of risk) are also a matter of concern.

Recent studies have dug into the role that explicit or implicit government guarantees play in leading to risk-shifting behavior and high asset prices in the transmission of crises. Also, uncertainty about the future course of credit creation in the economy and its interaction with the agency problem in intermediation has been proven crucial to determining the extent of asset price bubbles and tracking ensuing developments. Such uncertainty is often the result of government policies and should be considered when policy revision is demanded.

The systemic risk built up during the expansion period materializes in a crisis after the deflation of the bubble. Temporary shocks can have persistent effects on economic activity as they impact the net worth of levered agents. Financial frictions lead to the amplification of shocks through leverage and prices. Externalities include spillovers and contagion. For instance, direct contractual impact could be a domino effect due to interconnectedness. Indirect price effect, such as fire sales externalities, credit crunch, and liquidity spirals, could also take place. The amplification through prices works through adverse feedback loops, as declining net worth of levered agents leads to a drop in prices of assets concentrated in their hands, further lowering these agents' net worth and the interconnections of the various sectors and agencies within the system.

Works characterizing the financial system dynamics have highlighted a number of important considerations.[154] First, the system's reaction to shocks is highly nonlinear. The implication is that equivalent shocks that take place at various times will impact the system differently. Second, the system's reaction to shocks is asymmetric, with the various sectors carrying the weight of the losses in different proportions and at various times. Third, the endogenous risk generated by the system dominates volatility dynamics and affects experts' precautionary motives. When changes in asset prices are not driven by fundamentals, but rather by the constraints of market participants, incentives to hold cash increase. Fourth, after moving through a high volatility period, the system can get stuck in a recession, with the resulting low growth and misallocation of resources. Finally, lower exogenous risk can lead to more extreme volatility. Independently of the exogenous risk, systems enter volatile regimes every so often as low risk environments are conducive to greater buildup of systemic risk. This happens because low fundamental risk leads to higher equilibrium leverage.

Financial innovation allows experts to hedge their idiosyncratic risk. Still, this can be self-defeating since it could lead to higher systemic risk. That is, the risk does not disappear but instead, spreads to other parts of the system. Several examples of this were in evidence during the 2008 crisis. For instance, securitization of home loans into mortgage-backed securities allows originating institutions to unload some of the risks onto others who purchase these products. Institutions also shared risks through mechanisms such as credit-default swaps, the integration of commercial and investment banks, and various intermediation chains.[155] However, hedging idiosyncratic risks results in an increased capacity to take on more leverage, making the system less stable.

Balance-sheet recession can have long-term consequences. Trauma resulting from the crises will prevent borrowing, even once the balance sheets are repaired. Fiscal stimulus will be required to overcome this trend and get the economy back on its "normal track."

It has been repeatedly stated that there is a need to change the way macroeconomic policy is conducted after a large bubble implodes.[156] The reason is that periods of economic development during which profit maximization is the overall goal of corporations are not comparable to scenarios where the opposite (debt reduction) is the goal. In the latter scenario, low interest rates have no impact, as companies need to restore the health of their balance sheets before they embark on any further capital investments. Given that all firms will be trying to minimize the value of their liabilities at the same time, the danger of economic contraction is present and measures need to be taken in order to prevent a fall into depression.

Once a bubble has erupted, there are certain important measures that must be taken to ensure the health of the economy. However, these fall under the "remedies" section. But can bubbles be prevented? Should governments intervene to prevent bubble growth? The analyses of bubble formation implies that many policy, governance, and supervisory decisions are contributors to these processes; hence, it seems that changes in the way these are carried out should help put a stop to bubbles right at the start.

Even though the post-bubble period brings a lot of soul searching, there are profound disagreements on whether governments and central banks should intervene at all. Some of the reasons for this come out of a "laissez-faire" mentality, but others are related to the unintended consequences of changes in regulatory, fiscal, and policy environments.

In proposing potential paths of action, one needs to be alert toward counterintuitive concerns since policies to recapitalize the financial sector can lead to moral hazard in normal times. For example, capital requirements can occasionally harm welfare or impede growth. Consequently, when considering policies, it is important to understand how they affect the entire equilibrium at various times. Thus, the proposed regulation should consider the system as a whole, rather than the impact on one sector alone.[157] It should be adapted to the times, so that counter-cyclical regulation might be considered in order to encourage some behaviors in good times and relax constraints in downturns.

Both in this and the previous chapter we have reviewed the key academic literature on asset bubbles. The behavioral models explain bubbles by assuming limited investor rationality and the potentially destabilizing role of rational agents. A newer generation of rational models attributes the existence of bubbles to various incentive problems faced by key economic agents. All these models provide valuable insights into the recent and prior bubbles, all of which should be studied to enhance system protection mechanisms. However, these works stop short of recommending that governments should actively intervene to halt the formation of bubbles.

In some sectors of society, the existence of bubbles is still questioned. Even when recognized, some doubt their negative consequences and consequently argue in favor of no intervention. The arguments used most often to defend the "no intervention" posture include the following:[158]

a. Bubbles are difficult to identify.
b. Bubbles might not damage the economy and may in fact play a positive role.
c. Deflating prices might be counterproductive overall.
d. Piercing a bubble will hurt those holding the bubble asset.
e. A bubble may have arisen as a result of some market inefficiency.
f. Central banks do not have the capacity to deflate bubbles without inducing collateral damage.

Of course, the answer to these arguments depends upon what model of economic growth one has in mind and what one thinks of the responsibilities of the public sector. Also, one should consider whether one concurs with the fact that the gains made by those in the "sophisticated" sector should happen at the expense of those in the "naïve" sector. That is, after a bubble implodes, a small number of individuals end up wealthier, while the community at large ends up poorer. One might argue that "this is the way it should be."

Furthermore, bubbles may be prevented with the simultaneous use of a number of mechanisms in coordination with various agencies. One such mechanism is correcting inefficiencies, but yet again, such inefficiencies must be identified. At the heart of this dilemma lie moral hazards and many conflicts of interests which impede such easy identification and resolution.

Notes

Preface

1 *A Random Walk Down Wall Street,* Burton Gordon Malkiel, W. W. Norton & Company, 1973.
2 *A Short History of Financial Euphoria*, John Kenneth Galbraith, Penguin Books, 1994.
3 *The General Theory of Employment, Interest and Money*, John Maynard Keynes, Palgrave Macmillan, 1936.

1. Introduction to Bubbles and Contagion

1 A. B. Krueger (2013), Chairman, Council of Economic Advisers, 'Land of Hope and Dreams: Rock and Roll, Economics and Rebuilding the Middle Class', June 12, Cleveland, OH. https://www.whitehouse.gov/sites/default/files/docs/hope_and_dreams_-_final.pdf
2 US Department of the Treasury, April 2012, The Financial Crisis Response in Charts, p. 3. http://www.treasury.gov/resource-center/data-chart-center/Documents/20120413_FinancialCrisisResponse.pdf. Retrieved 1 November 2013.
3 Financial Crisis Inquiry Commission (FCIC) formed by the USA Congress, 2011, The Financial Crisis Inquiry Report, p. xv, http://www.gpo.gov/fdsys/pkg/GPO-FCIC/pdf/GPO-FCIC.pdf. Retrieved 1 November 2013.
4 http://elpais.com/elpais/2015/08/02/media/1438529480_493891.html. Retrieved 18 August 2015. https://www.sepe.es/contenidos/que_es_el_sepe/estadisticas/datos_avance/pdf/empleo/evolparoseries.pdf. Retrieved 18 August 2015. http://www.sepe.es/contenidos/que_es_el_sepe/estadisticas/datos_avance/datos/datos_2015/AV_SISPE_1505.pdf. Retrieved 18 August 2015.
5 Wikipedia, Euro zone Crisis: European Sovereign Debt Crisis Spain, http://en.wikipedia.org/wiki/European_sovereign-debt_crisis#Spain. Retrieved 1 November 2013.
6 Wikipedia, Eurozone Crisis: European Sovereign Debt Crisis and Bond Credit Rating, http://en.wikipedia.org/wiki/European_sovereign-debt_crisis and http://en.wikipedia.org/wiki/Bond_credit_rating. Retrieved 20 August 2013.
7 British Broadcasting Corporation, 2010, Moody's downgrades Portugal debt, http://www.bbc.co.uk/news/10610673
8 G. Tremlett (2011) 'Spain Changes Constitution to Cap Budget Deficit', *The Guardian*, August 26, http://www.theguardian.com/business/2011/aug/26/spain-constitutional-cap-deficit
9 Ycharts (2013) 'Spain 10 Year Government Bond Yield', http://ycharts.com/indicators/spain_10_year_government_bond_yield
10 http://www.theguardian.com/business/2010/apr/23/greece-activates-eu-imf-loans
11 C. Ziotis and N. Weeks (2010) 'Greek Bailout Talks Could Take Three Weeks as Bond Repayment Looms in May', *Bloomberg LP*, April 20, http://www.bloomberg.com/news/2010-04-20/greek-talks-on-61-billion-aid-may-take-three-weeks-as-bond-payment-looms.html
12 J. Ewing and J. Healy (2010) 'Cuts to Debt Rating Stir Anxiety in Europe', *The New York Times*, April 27, http://www.nytimes.com/2010/04/28/business/global/28drachma.html?_r=1&

13 http://www.reuters.com/article/2014/04/23/us-greece-deficit-idUSBREA3M0G620140423. Retrieved 15 September 2014.
14 L. Phillips (2011) 'ECB Man to Rule Greece for 15 Weeks', *EUobserver*, http://euobserver.com/economic/114250. Retrieved 29 December 2011.
15 H. Smith (2012) 'IMF Official Admits: Austerity is Harming Greece', *The Guardian* (Athens), http://www.theguardian.com/business/2012/feb/01/imf-austerity-harming-greeve. Retrieved 1 February 2012.
16 http://uk.reuters.com/article/2013/04/11/uk-greece-unemployment-idUKBRE93A08220130411
17 http://www.bbc.com/mundo/noticias/2015/07/150712_grecia_escenarios_crisis_euro_az. Retrieved 18 August 2015.
18 Dan Bilefsky (2010) 'Three Reported Killed in Greek Protests', *The New York Times*, Retrieved 5 May 2010. http://www.nytimes.com/2010/05/06/world/europe/06greece.html?src=me&_r=0
19 http://sdw.ecb.europa.eu/browse.do?node=bbn3146
20 http://www.ecb.europa.eu/stats/money/long/html/index.en.html
21 Financial Crisis Inquiry Commission (FCIC) formed by the USA Congress, 2011, The Financial Crisis Inquiry Report, p. 16, http://www.gpo.gov/fdsys/pkg/GPO-FCIC/pdf/GPO-FCIC.pdf. Retrieved 1 November 2013.
22 Ibid., p. 18.
23 J. Tirole (1985) 'Asset Bubbles and Overlapping Generations', *Econometrica*, 53(6), 1499–1528.
24 R. P. Flood and P. M. Garber (1980) 'Market Fundamentals vs. Price Level Bubbles: The First Tests', *Journal of Political Economy*, 88(4), 745–770.
25 J. D. Hamilton and C. H. Whiteman (1985) 'The Observable Implications of Self-Fulfilling Expectations', *Journal of Monetary Economics*, 16(3), 353–373.
26 C. Mackay (1841) *Memoirs of Extraordinary Popular Delusions and the Madness of Crowds* (London: Richard Bentley) archived from the original on 31 March 2008, retrieved 15 August 2008.
27 A. Goldgar (2007) *Tulipmania: Money, Honor, and Knowledge in the Dutch Golden Age* (Chicago: University of Chicago Press).
28 P. M. Garber (2000) *Famous First Bubbles: The Fundamentals of Early Manias* (Cambridge: MIT Press), p. 81.
29 http://www.bloomberg.com/news/2013-04-05/how-the-south-sea-bubble-created-u-k-s-modern-monarchy.html
30 G. S. Shea (2007) 'Understanding Financial Derivatives during the South Sea Bubble: The Case of the South Sea Subscription Shares', School of Economics and Finance University of St. Andrews, www.econ.upf.edu/docs/seminars/shea.pdf
31 C. Mackay (1841) Memoirs of Extraordinary Popular Delusions and the Madness of Crowds (London: Richard Bentley), archived from the original on 5 February 2008, p. 35.
32 G. Campbell and J. Turner (2010) '"The Greatest Bubble in History": Stock Prices during the British Railway Mania', Queen's University Belfast, 21820, March, http://mpra.ub.uni-muenchen.de/21820/
33 A. Kraay and J. Ventura (2005) 'The Dot-Com Bubble the Bush Deficits, and the US Current Account', *Working Paper*, 11543, http://www.nber.org/papers/w11543
34 http://www.homeinsight.com/details.asp?url_id=23
35 J. J. Murphy (2013) *Trading with Intermarket Analysis: A Visual Approach to Beating the Financial Markets Using Exchange-Traded Funds*, Enhanced edn (John Wiley & Sons). Hoboken, NJ.

36 E. R. Porras (2001) 'Burbujas y contagio en la nueva economía', *Información Comercial Española, ICE: Revista de Economía*, 791, 115–130.
37 D. Sornette (2009) *Why Stock Markets Crash: Critical Events in Complex Financial Systems* (Princeton, NJ: Princeton University Press).
38 Financial Crisis Inquiry Commission (FCIC), p. 84
39 Ibid., p. 85.
40 Ibid.
41 Ibid.
42 Chairman B. S. Bernanke, *Monetary Policy and the Housing Bubble*, at the Annual Meeting of the American Economic Association, Atlanta, GA, 3 January 2010, p. 16 (www.federalreserve.gov/newsevents/speech/bernanke20100103a.htm).
43 Financial Crisis Inquiry Commission (FCIC) formed by the USA Congress, 2011, The Financial Crisis Inquiry Report, pp. 86–87, http://www.gpo.gov/fdsys/pkg/GPO-FCIC/pdf/GPO-FCIC.pdf
44 Recent Changes in US Family Finances: Evidence from the 2001 and 2004 Survey of Consumer Finances, http://www.federalreserve.gov/econresdata/scf/scf_2004.htm
45 An MBS is an asset-backed security secured by a mortgage or a collection of mortgages. The mortgages are sold to a financial institution that bundles the loans together ("securitizes") into a security that can be sold to investors.
46 A CDO is a type of structured asset-backed security (ABS). Originally developed for the corporate debt markets, over time CDOs evolved to encompass the mortgage and MBS markets.
47 In general, Alt-A mortgages are obtained by borrowers with less than full documentation, lower credit scores, higher loan-to-values, and more investment properties.
48 Financial Crisis Inquiry Commission (FCIC) formed by the USA Congress, 2011, The Financial Crisis Inquiry Report, Note 2, p. 214, http://www.gpo.gov/fdsys/pkg/GPO-FCIC/pdf/GPO-FCIC.pdf (2. Inside Mortgage Finance, *The 2009 Mortgage Market Statistical Annual*, 2, *The Secondary Market* (Bethesda, MD: Inside Mortgage Finance, 2009), p. 13, "Non-Agency MBS Issuance by Type").
49 Ibid., Note 4, p. 214 (4. *2009 Mortgage Market Statistical Annual*, 2:373, "Commercial MBS Issuance").
50 With $639 billion in assets and $619 billion in debt, Lehman Brothers, the fourth-largest US investment bank with 25,000 employees worldwide, filed for bankruptcy on September 15, 2008. Lehman's bankruptcy filing, the largest in history, contributed significantly to the vanishing of $10 trillion in market capitalization from global equity markets in the ensuing month. Given the firm's status and the impact that its demise had in the markets, the US government's decision to let the firm fail has been severely criticized. One aspect that puzzled investors was the fact that other institutions, such as Bear Stearns, did get the needed support.

Source: Case Study: The Collapse of Lehman Brothers http://www.investopedia.com/articles/economics/09/lehman-brothers-collapse.asp#ixzz3rSNfOHCN
51 Ibid., p. 389.
52 Ibid., p. 390.
53 C. Kindleberger and R. Aliber (2005) *Manias, Panics and Crashes: A History of Financial Crises*, 5th edn (Hoboken, NJ: John Wiley & Sons), p. 104.
54 Financial Crisis Inquiry Commission (FCIC) formed by the USA Congress, 2011, The Financial Crisis Inquiry Report, p. xvi, http://www.gpo.gov/fdsys/pkg/GPO-FCIC/pdf/GPO-FCIC.pdf

55 Ibid., pp. xvii–xxv.
56 Ibid., p. 451.
57 A. Greenspan (1996) 'The Challenge of Central Banking in a Democratic Society', Annual Dinner and Francis Boyer Lecture of The American Enterprise Institute for Public Policy Research, Washington, DC, 5 December. http://www.federalreserve.gov/boarddocs/speeches/1996/19961205.htm
58 A. Greenspan (2010) written testimony for the 'Financial Crisis Inquiry Commission (FCIC)', Subprime Lending and Securitization and Government-Sponsored Enterprises (GSEs), day 1, session 1: The Federal Reserve, 7 April, p. 9. http://www.gpo.gov/fdsys/pkg/GPO-FCIC/pdf/GPO-FCIC.pdf
59 M. K. Brunnermeier (2008) *Bubbles, the New Palgrave Dictionary of Economics*, in Steven N. Durlauf and L. E. Blume, eds (Basingstoke: Palgrave Macmillan).
60 F. Allen, S. Morris, and A. Postlewaite (1993) 'Finite Bubbles with Short Sale Constraints and Asymmetric Information', *Journal of Economic Theory*, 61, 206–229.
61 A. Scherbina and B. Schlusche (2011) 'Asset Price Bubbles: A Survey, Draft', December, http://ssrn.com/abstract=1793420
62 K. A. Froot and M. Obstfeld (1991) 'Intrinsic Bubbles: The Case of Stock Prices', *American Economic Review*, 81, 1189–1214.
63 G. L. Kaminsky, C. M. Reinhart and C. A. Végh (2003) 'The Unholy Trinity of Financial Contagion', *Journal of Economic Perspectives*, 17(4), 51–74.
64 K. Daniel, D. Hirshleifer and A. Subrahmanyam (1998) 'Investor Psychology and Security Market Under- and Overreactions', *Journal of Finance*, 53, 1839–1885.
65 E. Asem and G. Y. Tian (2010) 'Market Dynamics and Momentum Returns', *Journal of Financial and Quantitative Analysis*, 45, 1549–1562.
66 N. Barberis, A. Shleifer and R. Vishny (1998) 'A Model of Investor Sentiment', *Journal of Financial Economics*, 94, 307–343.
67 B. S. Bernanke, M. Gertler and S. Gilchrist (1999) 'The Financial Accelerator in a Quantitative Business Cycle Framework', in J. B. Taylor and M. Woodford, eds, Handbook of Macroeconomics (Elsevier). Chapter 21, pp 1343–1390, *Handbook of Macroeconomics*, Volume 1, Edited by J. B. Taylor and M. Woodford © 1999 Elsevier Science B.V. All rights reserved. Also published as Bernanke, B., Gertler, M. and S. Gilchrist (1998), "The financial accelerator in a quantitative business cycle framework", NBER Working Paper Series No. 6455.
68 S. Bikhchandani, D. Hirshleifer and I. Welch (1992) 'A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades', *Journal of Political Economy*, 100(5), 992–1026.
69 G. Calvo and E. Mendoza (2000) 'Rational Contagion and the Globalization of Securities Markets', *Journal of International Economic*, 51(1), 79–113.
70 A. V. Banerjee (1992) 'A Simple Model of Herd Behavior', *Quarterly Journal of Economics*, 107(3), 797–817.
71 R. Nurkse (1944) *International Currency Experience: Lessons of the Interwar Period* (Geneva: League of Nations).
72 S. Gerlach and F. Smets (1996) 'Contagious Speculative Attacks', *CEPR Discussion Paper*, 1055. (London: Center for Economic Policy Research)
73 A. Lahir and C. A. Végh (2003) 'Delaying the Inevitable: Optimal Interest Rate Defense and BOP Crises', *Journal of Political Economy*, 1, 1–62.
74 A. Shleifer and R. W. Vishny (1997) 'The Limits of Arbitrage', *Journal of Finance*, 52(1), 35–55.

75 G. Calvo (1998) 'Capital Market Contagion and Recession: An Explanation of the Russian Virus', Mimeograph. (College Park, MD: University of Maryland)
76 G. Kaminsky and C. Reinhart (2000) 'On Crises, Contagion, and Confusion', *Journal of International Economics*, 51(1), 145–168.
77 R. Basu (1998) 'Contagion Crises: The Investor's Logic', Mimeograph, University of California, Los Angeles.
78 M. Goldstein (1989) *The Asian Financial Crisis* (Washington, DC: Institute for International Economics).
79 G. Calvo and E. Mendoza (2000) 'Rational Contagion and the Globalization of Securities Market', *Journal of International*, 51(1), 79–113.
80 L. Kodres and M. Pritsker (2002) 'A Rational Expectations Model of Financial Contagion', *Journal of Finance*, 57(2), 769–800.
81 This section follows closely from J. P. Rodrigue (2009), http://people.hofstra.edu/jean-paul_rodrigue/jpr_blogs.html
82 H. P. Minsky (1992) 'The Financial Instability Hypothesis', *Working Paper*, 74 http://www.levyinstitute.org/pubs/wp74.pdf
83 H. P. Minsky (1986) 'Stabilizing an Unstable Economy', http://digamo.free.fr/minsky86.pdf
84 C. P. Kindleberger (1978) *Manias, Panics, and Crashes: A History of Financial Crises* (New York: Basic Books).
85 Mackay, Charles (1841), *Memoirs of Extraordinary Popular Delusions and the Madness of Crowds* (London: Richard Bentley), archived from the original on 31 March, 2008, retrieved 15 September, 2008.
86 Schama, Simon (1987), *The Embarrassment of Riches: An Interpretation of Dutch Culture in the Golden Age* (New York: Alfred Knopf), ISBN 0-394-51075-5.
87 Krelage, E. H. (1942), *Bloemenspeculatie in Nederland* (Amsterdam: P.N. van Kampen & Zoon).
88 Garber, Peter M. (2000), *Famous First Bubbles: The Fundamentals of Early Manias* (Cambridge: MIT Press), ISBN 0-262-07204-1

2 Macro "Players" in Bubble Formation and Contagion Processes

1 Financial Crisis Inquiry Commission (FCIC) formed by the USA Congress, 2011, The Financial Crisis Inquiry Report, p. 134, http://www.gpo.gov/fdsys/pkg/GPO-FCIC/pdf/GPO-FCIC.pdf
2 S. Meier and C. Vits (2009) 'Weber Says ECB Has Used Room to Cut Interest Rates (Update3)', *Bloomberg*, June 23, http://www.bloomberg.com/apps/news?pid=newsarchive&sid=a5S5Boes29lo, visited 23 July 2013.
3 Declarations by former Fed governor, F. Mishkin, Federal Reserve Board. Financial Crisis Inquiry Commission (FCIC) formed by the USA Congress, 2011, The Financial Crisis Inquiry Report, p. 133, http://www.gpo.gov/fdsys/pkg/GPO-FCIC/pdf/GPO-FCIC.pdf, also available at http://fcic.law.stanford.edu/interviews/view/187
4 P. Krugman, interview by FCIC, October 6, 2010. Financial Crisis Inquiry Commission (FCIC) formed by the USA Congress, 2011, The Financial Crisis Inquiry Report, http://www.gpo.gov/fdsys/pkg/GPO-FCIC/pdf/GPO-FCIC.pdf, p. 104, also available at http://fcic.law.stanford.edu/interviews/view/226

5 Financial Crisis Inquiry Commission (FCIC) formed by the USA Congress, 2011, The Financial Crisis Inquiry Report, p. 420, http://www.gpo.gov/fdsys/pkg/GPO-FCIC/pdf/GPO-FCIC.pdf, originally from Ben S. Bernanke, "Monetary Policy and the Housing Bubble," Speech at the Annual Meeting of the American Economic Association, Atlanta, Georgia, 3 January 2010, www.federalreserve.gov/newsevents/speech/bernanke20100103a.htm

6 John Maynard Keynes popularized the paradox of thrift in the 'Treatise on Money,' New York, Harcourt, Brace and company (1930). The central point of the work is the distinction between investment and saving: inflation will result when investment exceeds saving, whereas recession will result if saving exceeds investment. Contrary to the idea that thrift is required in hard times, the recommendation is that during an economic depression spending should be encouraged and saving discouraged. In Keynes's words: 'the engine which drives enterprise is not thrift, but profit'.

7 H. P. Minsky (1992) 'The Financial Instability Hypothesis', Levy Economics Institute, *The Economics Working Paper Archive*, http://www.levyinstitute.org/pubs/wp74.pdf

8 M. Draghi, F. Giavazzi and R. C. Merton (2003) 'Transparency, Risk Management and International Financial Fragility', *National Bureau of Economic Research*, Working Paper 9806 Cambridge, MA, June 2003, http://www.nber.org/papers/w9806

9 The following pages refer to a number of conferences and articles on the topic of microfinancial risks presented by Merton et al. For the full content please visit the original sources cited here: *At the fifth annual CFA Institute European Investment Conference on 19th October 2012 in Prague, R. C. Merton gave a presentation on analyzing and managing macrofinancial risk*. On a New Approach for Analyzing and Managing Macrofinancial Risks, R. C. Merton Center for Financial Studies Seminar, Goethe University, 21 January 2013; Draghi, Giavazzi and Merton, 'Transparency, Risk Management and International Financial Fragility'.

10 R. Merton, M. Billio, M. Getmansky, D. Gray, A. W. Lo and L. Pelizzon (2013) 'On a New Approach for Analyzing and Managing Macrofinancial Risks', *Financial Analysts Journal*, 69(2), CFA Institute.

11 Similarly, a household has operating assets plus a house (a physical building) as total assets, mortgage plus other loans as debt, and net worth as equity. The building itself is, however, the main collateral backing up the mortgage.

12 A bond bought at a price lower than its face value, with the face value repaid at the time of maturity.

13 A put or put option is a contract between two parties to exchange an asset (the underlying), at a specified price (the strike), by a pre-determined date (the expiry or maturity). One party, the buyer of the put, has the right, but not an obligation, to re-sell the asset at the strike price by the future date, while the other party, the seller of the put, has the obligation to re-purchase the asset at the strike price if the buyer exercises the option. The value of an option depends upon its time to maturity, the spot price of the underlying asset, the strike price, the risk-free annual rate, and the volatility of returns of the underlying asset.

14 Merton et al. 'On a New Approach for Analyzing and Managing Macrofinancial Risks'.

15 The delta, also called the hedge ratio, measures the change in the value of an option as the value of the underlying asset changes. The hedge ratio increases as the value of the underlying asset falls. See J. C. Hull (2000) *Options Futures and Derivatives* (Upper Saddle River, New Jersey: Prentice Hall), pp. 310–319.

16 R. C. Merton (1974) 'On the Pricing of Corporate Debt: The Risk Structure of Interest Rates', *Journal of Finance*, 29(2), 449–470.

17 R. C. Merton (1977) 'An Analytic Derivation of the Cost of Deposit Insurance and Loan Guarantees', *Journal of Banking & Finance*, 1(1), 3–11.
18 M. Draghi, F. Giavazzi and R. C. Merton (2004) 'Transparency, Risk Management and International Financial Fragility', *Geneva Reports on the World Economy, Geneva: International Center for Monetary and Banking Studies*, 4.
19 D. F. Gray and S. Malone (2008) *Macrofinancial Risk Analysis* (Hoboken, NJ: John Wiley & Sons).
20 D. F. Gray, R. C. Merton and Z. Bodie (2006) 'A New Framework for Analyzing and Managing Macrofinancial Risks of an Economy', *Harvard Business School*, Working Paper 07-026, http://www.hbs.edu/faculty/Publication%20Files/07-026.pdf
21 Merton et al. 'On a New Approach for Analyzing and Managing Macrofinancial Risks'.
22 M. Billio, M. Getmansky, D. Gray, A. Lo, R. C. Merton and L. Pelizzon. 2013. 'Sovereign, Bank, and Insurance Credit Spreads: Connectedness and System Networks', *MIT Working Paper*. For more information on this type of analysis, see M. Billio, M. Getmansky, A. Lo and L. Pelizzon (2012) 'Econometric Measures of Connectedness and Systemic Risk in the Finance and Insurance Sectors', *Journal of Financial Economics*, 104(3), 535–559.
23 Billio et al., 'Sovereign, Bank, and Insurance Credit Spreads'. For more information on this type of analysis, see Billio et al., 'Econometric Measures of Connectedness and Systemic Risk'.
24 There are various ways to look at interconnection and feedback; see D. F. Gray, M. Gross, J. Paredes and M. Sydow. 2013. 'Modeling Banking, Sovereign, and Macro Risk in Contingent Claims Analysis (CCA) Global VAR', https://www.imf.org/external/pubs/ft/wp/2013/wp13218.pdf
25 B. S. Bernanke, M. Gertler and S. Gilchrist (1999) 'The Financial Accelerator in a Quantitative Business Cycle Framework, 1st edn, in J. B. Taylor and M. Woodford, eds, *Handbook of Macroeconomics* (Elsevier).
26 B. S. Bernanke, M. Gertler and S. Gilchrist (1996) 'The Financial Accelerator and the Flight to Quality', *Review of Economics and Statistics*, 78(1), 1–15.
27 Bernanke, Gertler and Gilchrist, 'The Financial Accelerator'.
28 B. S. Bernanke and M. Gertler (1990) 'Financial Fragility and Economic Performance', *The Quarterly Journal of Economics*, 105(1), 87–114.
29 A. Korinek (2011) 'Systemic Risk-Taking: Amplification Effects, Externalities, and Regulatory Responses', *ECB Working Paper 1345*. SSRN 1847483.
30 J. M. Keynes (1936) *The General Theory of Employment, Interest, and Money* (New York: Harcourt Brace).
31 J. A. Schumpeter (1934) *Theory of Economic Development* (Cambridge, MA: Harvard University Press).
32 H. P. Minsky (1975) *John Maynard Keynes* (New York: Columbia University Press).
33 H. P. Minsky (1986) *Stabilizing an Unstable Economy* (New Haven, CT: Yale University Press).
34 Minsky, 'The Financial Instability Hypothesis'.
35 M. K. Brunnermeier, G. Gorton and A. Krishnamurthy (2011) 'Risk Topography', *NBER Macroeconomics Annual*, 26(1), 149–176, http://scholar.princeton.edu/sites/default/files/risk_topography_0.pdf
36 F. H. Knight (1921) *Risk, Uncertainty, and Profit* (Boston, MA: Hart, Schaffner & Marx; Houghton Mifflin Company).
37 R. J. Caballero (2009) 'Sudden Financial Arrest', *Tenth Jacques Polak Annual Research Conference International Monetary Fund*, 5–6, https://www.imf.org/external/np/res/seminars/2009/arc/pdf/caballero1.pdf

38 B. S. Bernanke and M. S. Gertler (1989) 'Agency Costs, Net Worth, and Business Fluctuations', *American Economic Review*, 79(1), 14–31; Bernanke, Gertler and Gilchrist, 'The Financial Accelerator'; C. T. Carlstrom and T. S. Fuerst (1997) 'Agency Costs, Net Worth, and Business Fluctuations: A Computable General Equilibrium Analysis', *American Economic Review*, 87(5), 893–910; D. Gale and M. Hellwig (1985) 'Incentive-Compatible Debt Contracts: The One-Period Problem', *Review of Economic Studies*, 52(4), 647–663.
39 G. Gorton and G. Pennacchi (1990) 'Financial Intermediaries and Liquidity Creation', *Journal of Finance*, 45(1), 49–71; T. V. Dang, G. Gorton and B. Holmström (2010) 'Ignorance and the Optimality of Debt for Liquidity Provision', *Working Paper*, Yale University.
40 J. Geanakoplos (2003) '"Liquidity, Defaults, and Crashes," in Advances in Economics and Econometrics: Theory and Applications', Eighth World Congress, Volume 2, *Econometric Society Monographs*, Cambridge University Press.
41 M. K. Brunnermeier and L. H. Pedersen (2005) 'Predatory Trading', *Journal of Finance*, 60(4), 1825–1863.
42 V. V. Acharya, D. Gal and T. Yorulmazer (2011) 'Rollover Risk and Market Freezes', *Journal of Finance*, 66(4), 1177–1209.
43 For example, under Basel II the risk weight on sovereign debt rated AA- and higher is zero. Basel III is similar to Basel II in this respect.
44 C. M. Reinhart and K. S. Rogoff (2009) *This Time is Different: Eight Centuries of Financial Folly* (Princeton, NJ: Princeton University Press).
45 M. K. Brunnermeier, L. Garicano, P. R. Lane, M. Pagano, R. Reis, T. Santos, D. Thesmar, S. van Nieuwerburgh and D. Vayanos (2011) 'European Safe Bonds: ESBies', Euro-nomics.com
46 C. M. Reinhart and K. S. Rogoff (2011) 'From Financial Crash to Debt Crisis', *American Economic Review*, 101(5), 1676–1706.
47 J. R. Barth, A. P. Prabhavivadhana and G. Yun (2012) 'The Eurozone Financial Crisis: Role of Interdependencies between Bank and Sovereign Risk', *Journal of Financial Economic Policy*, 4(1), 76–97.

3. Contributors to the Bubble Formation and Contagion Process

1 W. Charles (2008) *Adverse Selection, The New Palgrave Dictionary of Economics*, 2nd edn (New York: Palgrave Macmillan).
2 J. E. Stiglitz (1998) 'The Private Uses of Public Interests: Incentives and Institutions', *Journal of Economic Perspectives*, 12(2), 3–22.
3 J. O. Ledyard (2008) *Market Failure, The New Palgrave Dictionary of Economics*, 2nd edn (New York: Palgrave Macmillan).
4 P. Krugman and R. Wells (2006) *Economic*, 1st edn (New York: Worth Publishers).
5 I. Palacios-Huerta (2003) 'Time-Inconsistent Preferences in Adam Smith and David Hume', *History of Political Economy*, 35(2), 241–268.
6 J. J. Laffont (2008) *Externalities, The New Palgrave Dictionary of Economics*, 2nd edn (New York: Palgrave Macmillan).
7 J. E. Stiglitz (1989) 'Markets, Market Failures, and Development', *American Economic Review*, 79(2), 197–203.
8 K. J. Arrow (1969) 'The Organization of Economic Activity: Issues Pertinent to the Choice of Market versus Non-market Allocations', *Analysis and Evaluation of Public*

Expenditures: The PPP System, Washington, DC, Joint Economic Committee of Congress, 1–16.

9 H. Gravelle and R. Rees (2004) *Microeconomic* (Essex, England: Prentice Hall, Financial Times), pp. 314–346.

10 D. Weimer and A. R. Vining (2004) *Policy Analysis: Concepts and Practice* (Englewood Cliffs, NJ: Prentice Hall).

11 See M. C. Jensen and W. H. Meckling (1976) 'Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure', *Journal of Financial Economics*, 3(4), 305–360.

12 See L. D. Brown and M. S. Rozeff (1978) 'The Superiority of Analyst Forecasts as Measures of Expectations: Evidence from Earnings', *Journal of Finance*, 33(1), 1–16; D. Givoly (1982) 'Financial Analysts' Forecasts of Earnings: A Better Surrogate for Market Expectations', *Journal of Accounting and Economics*, 4(2), 85–108.

13 See D. Givoly and J. Lakonishok (1979) 'The Information Content of Financial Analysts Forecasts of Earnings: Some Evidence on Semi-Strong Efficiency', *Journal of Accounting and Economic*, 2, 165–186; T. Lys and S. Sohn (1990) 'The Association Between Revisions of Financial Analysts' Earnings Forecasts and Security Price Changes', *Journal of Accounting and Economics*, 13, 341–364; J. Francis and L. Soffer (1997) 'The Relative Informativeness of Analysts' Stock Recommendations and Earnings Forecast Revisions', *Journal of Accounting Research*, 35(2), 193–212.

14 See L. Brown, G. Foster and E. Noreen (1985) 'Security Analyst Multi-year Earnings Forecasts and the Capital Market', *Studies in Accounting Research*, #21 (Sarasota, FL: American Accounting Association).

15 See H. Lin and M. McNichols (1998) 'Underwriting Relationships, Analysts' Earnings Forecasts and Investment Recommendations', *Journal of Accounting and Economic*, 25(1), 101–128; P. Dechow, A. Hutton and R. Sloan (2000) 'The Relation Between Analysts' Forecasts of Long-Term Earnings Growth and Stock Price Performance Following Equity Offerings', *Contemporary Accounting Research*, 17(1), 1–32.

16 See P. Healy and K. Palepu (1993) 'The Effect of Firms' Financial Disclosure Strategies on Stock Prices', *Accounting Horizons*, 7, 1–11; P. Healy and K. Palepu (1995) 'The Challenges of Investor Communications: The Case of CUC International, Inc.', *Journal of Financial Economics*, 38, 111–141.

17 J. Warner, R. Watts and K. Wruck (1988) 'Stock Prices and Top Management Changes', *Journal of Financial Economics*, 20, 461–493; M. Weisbach (1988) 'Outside Directors and CEO Turnover', *Journal of Financial Economics*, 20, 431–461.

18 C. Noe (1999) 'Voluntary Disclosures and Insider Transactions', *Journal of Accounting and Economics*, 27, 305–327.

19 D. Aboody and R. Kasznik (2000) 'CEO Stock Options Awards and the Timing of Corporate Voluntary Disclosures', *Journal of Accounting and Economics*, 29, 73–100.

20 G. Waymire (1984) 'Additional Evidence on the Information Content of Management Earnings Forecasts', *Journal of Accounting Research*, 22, 703–719; B. Ajinkya and M. Gift (1984) 'Corporate Managers' Earnings Forecasts and Symmetrical Adjustments of Market Expectations', *Journal of Accounting Research*, 22, 425–444.

21 G. Pownall and G. Waymire (1989) 'Voluntary Disclosure Credibility and Securities Prices: Evidence from Management Earnings Forecasts', *Journal of Accounting Research*, 27, 227–246.

22 D. Diamond and R. Verrecchia (1991) 'Disclosure, Liquidity, and the Cost of Capital', *The Journal of Finance*, 66, 1325–1355; O. Kim and R. Verrecchia (1994) 'Market Liquidity and Volume Around Earnings Announcements', *Journal of Accounting and Economics*, 17, 41–68.

23 P. Healy, A. Hutton and K. Palepu (1999) 'Stock Performance and Intermediation Changes Surrounding Sustained Increases in Disclosure', *Contemporary Accounting Research*, 16, 485–520; D. Gelb and P. Zarowin (2000) 'Corporate Disclosure Policy and the Informativeness of Stock Prices', *Working Paper*, New York University, New York, NY. http://papers.ssrn.com/paper.taf?abstract_id=235009
24 M. Welker (1995) 'Disclosure Policy, Information Asymmetry and Liquidity in Equity Markets', *Contemporary Accounting Research*, 11(2), 801–828.
25 C. B. Barry and S. J. Brown (1984) 'Differential Information and the Small Firm Effect', *Journal of Financial Economics*, 13(2), 283–295; C. B. Barry and S. J. Brown (1986) 'Limited Information as a Source of Risk', *The Journal of Portfolio Management*, 12, 66–72.
26 C. A. Botosan (1997) 'Disclosure Level and the Cost of Equity Capital', *The Accounting Review*, 72(3), 323–350.
27 M. Lang and R. Lundholm (1993) 'Cross-Sectional Determinants of Analysts Ratings of Corporate Disclosures', *Journal of Accounting Research*, 31, 246–271.
28 G. A. Akerlof (1970) 'The Market for "Lemons": Quality Uncertainty and the Market Mechanism', *The Quarterly Journal of Economics*, 84(3), 488–500.
29 M. Spence (1973) 'Job Market Signaling', *Quarterly Journal of Economics*, 87(3), 355–374.
30 W. I. Thomas and D. S. Thomas (1928) *The Child in America: Behavior Problems and Programs* (New York: Alfred A. Knopf), pp. 571–572; E. Grunberg and F. Modigliani (1954) 'The Predictability of Social Events', *Journal of Political Economy*, 62(6), 465–478; H. Simon (1954) 'Bandwagon and Underdog Effects of Election Predictions', *Public Opinion Quarterly*, 18, 245–253; G. Soros (1987) *The Alchemy of Finance: Reading the Mind of the Market* (Simon & Schuster, 1988) (paperback: Wiley, 2003), pp. 27–45; G. Soros (2008) *The New Paradigm for Financial Markets*, 1st edn (London: Public Affairs TM).
31 R. K. Merton (1948) 'The Self-Fulfilling Prophecy', *The Antioch Review*, 8(2), 193–210.
32 Grunberg and Modigliani, 'The Predictability of Social Events'.
33 Simon, 'Bandwagon and Underdog Effects of Election Predictions'.
34 R. E. Lucas (1976) 'Econometric Policy Evaluation: A Critique', in K. Brunner and A. Meltzer, eds, *The Phillips Curve and Labor Markets*, Carnegie-Rochester Conference Series on Public Policy 1 (New York: American Elsevier), pp. 19–46.
35 Soros, *The Alchemy of Finance*, pp. 27–45.
36 Soros, *The New Paradigm for Financial Markets*.
37 O. J. Flanagan (1981) 'Psychology, Progress, and the Problem of Reflexivity: A Study in the Epistemological Foundations of Psychology', *Journal of the History of the Behavioral Sciences*, 17, 375–386.
38 S. Sandri (2009) *Reflexivity in Economics; An Experimental Examination on the Self-Referentiality of Economic Theories* (Heidelberg: Physica Verlag).
39 D. W. Hands (1994) 'Blurred Boundaries: Recent Changes in the Relationship Between Economics and the Philosophy of Science', *Studies in the History and Philosophy of Science*, 25, 751–772; D. W. Hands (1995) 'Social Epistemology Meets the Invisible Hand: Kitcher on the Advancement of Science', *Dialogue*, 34, 605–621; D. W. Hands (1998) 'Reflexivity', in J. Davis, D. Hands and U. Mäki, eds, *The Handbook of Economic Methodology* (Cheltenham: Elgar), pp. 413–416; D. W. Hands (2001) *Reflection without Rules: Economic Methodology and Contemporary Science Theory* (Cambridge: Cambridge University Press); U. Mäki (1999) 'Science as a Free Market: Reflexivity Test in an Economics of Economics', *Perspectives on Science*, 7, 486–509; J. Davis and M. Klaes (2003) 'Reflexivity: Curse or Cure?', *Journal of Economic Methodology*, 10(3), 329–352; Grunberg and Modigliani, 'The Predictability of Social Events'; E. A. Devletoglou

(1961) 'Correct Public Prediction and the Stability of Equilibrium', *Journal of Political Economy*, 69, 142–161; M. C. Kemp (1962) 'Economic Forecasting When the Subject of the Forecast Is Influenced By the Forecast', *American Economic Review*, 52, 492–496; M. Galatin (1976) 'Optimal Forecasting in Models with Uncertainty When the Outcome Is Influenced By the Forecast', *The Economic Journal*, 86, 278–295; E. Grunberg (1986) 'Predictability and Reflexivity', *American Journal of Economics and Sociology*, 45, 475–488.

40 Davis and Klaes, 'Reflexivity: Curse or Cure?'.

41 Grunberg, 'Predictability and Reflexivity'.
Grunberg and Modigliani, 'The Predictability of Social Events'.
Devletoglou, 'Correct Public Prediction and the Stability of Equilibrium'.
Kemp, 'Economic Forecasting'.
Galatin, 'Optimal Forecasting in Models'.

42 Soros, *The Alchemy of Finance*, pp. 27–45.

43 Soros, *The New Paradigm for Financial Markets*.

44 R. Cross and D. Strachan (1997) 'On George Soros and Economic Analysis', *Kyklos*, 50, 561–574.

45 J. Calandro (2004) 'Reflexivity, Business Cycles and the New Economy', *The Quarterly Journal of Austrian Economics*, 7(3), 45–69.

46 Sandri, *Reflexivity in Economics*.

47 S. Joshi, J. Parker and M. Bedau (2002) 'Financial Markets Can Be at Sub-Optimal Equilibria', *Computational Economics*, 19, 5–23.

48 A. Subrahmanyam and S. Titman (2001) 'Feedback from Stock Prices to Cash Flows', *Journal of Finance*, 56, 2389–2413.

49 Y. Luo (2005) 'Do Insiders Learn from Outsiders? Evidence from Mergers and Acquisitions', *Journal of Finance*, 60, 1951–1972.

50 Q. Chen, I. Goldstein and W. Jiang (2008) 'Price Informativeness and Investment Sensitivity to Stock Price', *Review of Financial Studies*, 20, 619–650.

51 T. Bakke and T. Whited (2006) 'Which Firms Follow the Market? An Analysis of Corporate Investment Decisions', *Working Paper*, University of Wisconsin. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=891570

52 C. Polk and P. Sapienza (2009) 'The Stock Market and Corporate Investment: A Test of Catering Theory', *Review of Financial Studies*, 22, 187–217.

53 P. Fulghieri and D. Lukin (2001) 'Information Production, Dilution Costs, and Optimal Security Design', *Journal of Financial Economics*, 61, 3–42.

54 E. Ozdenoren and K. Yuan (2008) 'Feedback Effects and Asset Prices', *Journal of Finance*, 63, 1939–1975.

55 G. M. Angeletos, G. Lorenzoni and A. Pavan (2008) 'Wall Street and Silicon Valley: A Delicate Interaction', *Working Paper*, MIT and NBER. http://faculty.wcas.northwestern.edu/~apa522/WSSV-new.pdf

56 S. Puddu (2009) 'Real Sector and Banking System: First and Second Round Effects. A Non-linear VAR Approach', *Working Paper*, presented at the MAFIN09 conference, Reykjavik. http://smye2009.org/file/081_Puddu.pdf

57 E. Oikarinen (2009) 'Interaction between Housing Prices and Household Borrowing: The Finnish Case', *Journal of Banking and Finance*, 33, 747–756.

58 B. Bernanke and M. Gertler (1989) 'Agency Costs, Collateral and Business Fluctuations', *American Economic Review*, 79, 14–31.

59 B. Bernanke, M. Gertler and S. Gilchrist (1998) 'The Financial Accelerator in a Quantitative Business Cycle', *NBER Working Paper*, 6455. http://ideas.repec.org/p/nbr/nberwo/6455.html

60 M. Uribe and V. Yue (2005) 'Country Spreads and Emerging Countries: Who Drives Whom?', *Journal of International Economics*, 69, 6–36.

61 For a general review of the housing prices and credit relationship, see C. Goodhart and B. Hofmann (2008) *House Prices and the Macroeconomy: Implications for Banking and Price Stability* (Oxford: Oxford University Press).

For a set of 20 industrialized countries, see B. Hoffman (2003) 'Bank Lending and Property Prices: Some International Evidence', *The Hong Kong Institute for Monetary Research, Working Paper* 22.

T. H. Koh, R. S. Mariano, A. Pavlov, S. Y. Phang, A. H. Tan and S. M. Wachter (2005) 'Bank Lending and Real Estate in Asia: Market Optimism and Asset Bubbles', *Journal of Asian Economics*, 15, 1103–1118.

62 N. Barberis, A. Shleifer and R. Vishny (1998) 'A Model of Investor Sentiment', *Journal of Financial Economics*, 49, 307–343; K. Daniel, D. Hirshleifer and A. Subrahmanyam (1998) 'Investor Psychology and Security Market Under- and Overreactions', *Journal of Finance*, 53, 1839–1884.

63 N. Jegadeesh and S. Titman (1993) 'Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency', *The Journal of Finance*, 48(1), 65–91; N. Jegadeesh and S. Titman (2001) 'Profitability of Momentum Strategies: An Evaluation of Alternative Explanations', *The Journal of Finance*, LVI, 2.

64 L. Festinger (1985, first published 1957) *A Theory of Cognitive Dissonance* (Stanford, CA: Stanford University Press, retrieved 6 March 2013, reissued by Stanford University Press 1962).

65 D. G. Spencer and S. Myers (2006) *Social Psychology*, 3rd Canadian edn (Toronto: McGraw-Hill Ryerson).

66 Festinger, *A Theory of Cognitive Dissonance*.

67 N. R. Carlson and H. C. Donald (2010) *Psychology: The Science of Behaviour*, 4th Canadian edn (Toronto: Pearson).

68 D. Sornette (2003) *Why Stock Markets Crash Critical Events in Complex Financial Systems* (Princeton, NJ: Princeton University Press).

69 C. J. Keller and M. W. Stocker (2008) 'Eyeing Executive Compensation', *The National Law Journal*, http://knowledgenetwork.labaton.com/Eying-Executive-Compensation.cfm

70 Ibid.

71 Ibid.

72 T. M. Tse and R. Merle (2008) 'The Bonuses Keep Coming', *Washington Post*, January 29. L. Story (2008) 'On Wall Street, Bonuses, Not Profits, Were Real', *The New York Times*, December 18.

73 Ibid.

74 M. Faulkender, D. Kadyrzhanova, N. Prabhala and L. Senbet (2010) 'Executive Compensation: An Overview of Research on Corporate Practices and Proposed Reforms', *Journal of Applied Corporate Finance*, 22(1), 107–118.

75 Jensen and Meckling, 'Theory of the Firm'.

76 Ibid.

77 See, for example, R. A. Haugen and L. W. Senbet (1981) 'Resolving Agency Problems of External Capital through Options', *Journal of Finance*, 36, 629–647.

78 M. Jensen and K. J. Murphy (1990), 'Performance Pay and Top-Management Incentives', *Journal of Political Economy*, 98(2), 225–264.

79 The relation between pay and performance is not uniform across firms and across industries. In particular, pay was relatively less aligned with performance among large firms and at regulated utilities.

80 B. J. Hall and J. B. Liebman (1998) 'Are CEOs Really Paid Like Bureaucrats?', *Quarterly Journal of Economics*, 113, 653–691.
81 J. Core, G. Wayne and D. Larcker (2003) 'Executive Equity Compensation and Incentives: A Survey', *Economic Policy Review*, 9, 27–50.
82 B. J. Hall and K. J. Murphy (Summer 2003) 'The Trouble with Stock Options', *Journal of Economic Perspectives*, 17(3), 49–70.
83 K. J. Murphy (1999) 'Executive Compensation', in O. Ashenfelter and D. Card, eds, *Handbook of Labor Economics*, 3b (Elsevier Science North Holland), pp. 2485–2563.
84 R. Fahlenbrach and R. M. Stulz (2009) 'Bank CEO Incentives and the Credit Crisis', *Working Paper* 1009-13, Dice Center, Ohio State University. http://ssrn.com/abstract=1439859
85 Ibid.
86 M. C. Jensen and K. J. Murphy (1990) 'CEO Incentives: It's Not How Much You Pay, But How', *Harvard Business Review*, 3, 138–153.
87 USA Committee for Economic Development (2008) Built to Last: Focusing Corporations on Long-Term Performance, http://www.ced.org/reports/single/built-to-last-focusing-corporations-on-long-term-performance
88 Ibid.
89 US Commodity Futures Trading Commission (2012) CFTC Glossary: A Guide to the Language of the Futures Industry, http://www.cftc.gov/consumerprotection/educationcenter/cftcglossary_s, retrieved 28 August 2012.
90 US Commodity Futures Trading Commission (2008) Staff Report on Commodity Swap Dealers and Index Traders with Commission Recommendations, http://www.cftc.gov/ucm/groups/public/@newsroom/documents/file/cftcstaffreportonswapdealers09.pdf, retrieved 27 August 2012.
91 'In a setting in which speculation is not possible, bubbles and crashes are observed. The results suggest that the departures from fundamental values are not caused by the lack of common knowledge of rationality leading to speculation, but rather by behavior that itself exhibits elements of irrationality'. V. Lei, C. Noussair and C. R. Plott (2001) 'Nonspeculative Bubbles in Experimental Asset Markets: Lack of Common Knowledge of Rationality vs. Actual Irrationality', *Econometrica*, 69(4), 831–859.
92 US Commodity Futures Trading Commission (2012) 'Speculative Limit', http://www.cftc.gov/IndustryOversight/MarketSurveillance/SpeculativeLimits/index.htm#P6_864, retrieved 21 August 2012.
93 US Commodity Futures Trading Commission (2012) 'CFTC Approves Notice of Proposed Rulemaking Regarding Regulations on Aggregation for Position Limits for Futures and Swaps', http://www.cftc.gov/PressRoom/PressReleases/pr6260-12, retrieved 21 August 2012.
94 D. Cho and B. Appelbaum (2010) Obama's 'Volcker Rule' Shifts Power Away from Geithner, *The Washington Post*, retrieved 13 February 2010.
95 G. S. Becker (1976) *The Economic Approach to Human Behavior* (Chicago: The University of Chicago Press).
96 H. A. Simon (1955) 'A Behavioral Model of Rational Choice', *Quarterly Journal of Economics*, 69, 99–118.

H. A. Simon (1979) 'Rational Decision-Making in Business Organizations', *American Economic Review*, 69, 493–513.
97 G. Gigerenze and R. Selten (2002) *Bounded Rationality: The Adaptive Toolbox* (Cambridge, MA and London: MIT Press).
98 Answers Corporation, Bounded Rationality: Definition, http://www.answers.com/topic/bounded-rationality, retrieved 12 April 2009.

99 A. Tversky and D. Kahneman (1971) 'Belief in the Law of Small Numbers', *Psychological Bulletin*, 76, 105–110.

100 K. E. Stanovich and R. F. West (2000) 'Individual Differences in Reasoning: Implications for the Rationality Debate', *Behavioral and Brain Sciences*, 23, 645–665.

101 D. Kahneman and A. Tversky (1973) 'On the Psychology of Prediction', *Psychological Review*, 80, 237–251.

102 H. E. Pashler (1998) *The Psychology of Attention* (Cambridge, MA: MIT Press).

103 S. A. Sloman (2002) 'Two Systems of Reasoning', in T. Gilovich, D. Griffin and D. Kahneman, eds, *Heuristics and Biases* (New York: Cambridge University Press), pp. 379–396; K. E. Stanovich (1999) *Who is Rational? Studies of Individual Differences in Reasoning* (Mahwah, NJ: Erlbaum); D. Kahneman and S. Frederick (2002) 'Representativeness Revisited: Attribute Substitution in Intuitive Judgement', in T. Gilovich, D. Griffin and D. Kahneman, eds, *Heuristics and Biases: The Psychology of Intuitive Judgment* (New York: Cambridge University Press), pp. 49–81; S. Sloman (1996) 'The Empirical Case for Two Systems of Reasoning', *Psychological Bulletin*, 119(1), 3–22; K. E. Stanovich and R. F. West (2002) 'Individual Differences in Reasoning: Implications for the Rationality Debate?', in T. Gilovich, D. W. Griffin, D. Kahneman, eds, *Heuristics and Biases: The Psychology of Intuitive Judgment* (New York: Cambridge University Press), pp. 421–440.

104 D. Kahneman (2003) 'Perspective on Judgment and Choice, Mapping Bounded Rationality', *American Psychologist*, 58(9), 697–720, http://www2.econ.iastate.edu/tesfatsi/JudgementAndChoice.MappingBoundedRationality.DKahneman2003.pdf

105 E. T. Higgins (1996) 'Knowledge Activation: Accessibility, Applicability, and Salience', in E. T. Higgins and A. Kruglanski, eds, *Social Psychology: Handbook of Basic Principles* (New York: Guilford Press), pp. 133–168.

106 For a review, see S. Fiske (1998) 'Stereotyping, Prejudice, and Discrimination', in D. T. Gilbert, S. T. Fiske and G. Lindzey, eds, *The Handbook of Social Psychology*, 4th edn (New York: McGraw-Hill), pp. 357–411.

107 G. Klein (1998) *Sources of Power: How People Make Decisions* (Cambridge: MIT Press).

108 A. Tversky and D. Kahneman (1986) 'Rational Choice and the Framing of Decisions', *Journal of Business*, 59, S251–S278.

109 K. J. Arrow (1982) 'Risk Perception in Psychology and Economics', *Economic Inquiry*, 20, 1–9.

110 H. A. Simon and J. R. Hayes (1976) 'Understanding Process: Problem Isomorphs', *Cognitive Psychology*, 8, 165–190.

111 D. Kahneman and A. Tversky (1979) 'Prospect Theory: An Analysis of Decision Under Risk', *Econometrica*, 47(2), 263–292, http://www.hss.caltech.edu/~camerer/Ec101/ProspectTheory.pdf

112 D. Kahneman and D. Lovallo (1993) 'Timid Choices and Bold Forecasts: A Cognitive Perspective on Risk Taking', *Management Science*, 39(1), 17.

113 R. H. Thaler (1985) 'Mental Accounting and Consumer Choice', *Marketing Science*, 4, 199–214.

114 R. H. Thaler (1999) 'Mental Accounting Matters', *Journal of Behavioral Decision Making*, 12(3), 183–206.

115 D. Kahneman and I. Ritov (1994) 'Determinants of Stated Willingness to Pay for Public Goods: A Study in the Headline Method', *Journal of Risk and Uncertainty*, 9, 5–38.

D. Kahneman (2000) 'A Psychological Point of View: Violations of Rational Rules as a Diagnostic of Mental Processes (Commentary on Stanovich and West)', *Behavioral and Brain Sciences*, 23, 681–683.

116 D. Kahneman, P. P. Wakker and R. Sarin (1997) 'Back to Bentham? Explorations of Experienced Utility', *Quarterly Journal of Economics*, 112, 375–405.

117 A. Tversky and D. Kahneman (1974) 'Judgment under Uncertainty: Heuristics and Biases', *Science*, 185, 1124–1131.

118 T. Gilovich, D. Griffin and D. Kahneman (2002) *Heuristics and Biases* (New York: Cambridge University Press).

119 D. Kahneman, P. Slovic and A. Tversky (1982) *Judgment under Uncertainty: Heuristics and Biases* (New York: Cambridge University Press).

120 J. Coval, J. Jurek and E. Stafford (2009) 'The Economics of Structured Finance', *Journal of Economic Perspectives*, 23(1), 3–25.

121 N. Gennaioli, A. Shleifer and R. W. Vishny (2011) 'A Model of Shadow Banking', *Economics Working Papers,* Department of Economics and Business, Universitat Pompeu Fabra. Working Paper 17115. Available at: http://www.nber.org/papers/w17115

http://economics.ucdavis.edu/faculty/kdsalyer/LECTURES/Ecn235a/PresentationPapersF11/Shleifer_Vishny_ShadowBanking.pdf

122 P. Bolton, J. Scheinkman and W. Xiong (2006) 'Executive Compensation and Short-Termist Behaviour in Speculative Markets', *The Review of Economic Studies*, 73(256), 577.

123 Tversky and Kahneman, 'Rational Choice and the Framing of Decisions'.

124 E. Shafir (1993) 'Choosing versus Rejecting: Why Some Options are Both Better and Worse Than Others', *Memory & Cognition*, 21, 546–556.

125 Kahneman and Tversky, 'Prospect Theory'; A. Tversky and D. Kahneman (1981) 'The Framing of Decisions and the Psychology of Choice', *Science*, 211, 453–458; Tversky and Kahneman, 'Rational Choice and the Framing of Decisions'; D. Kahneman and A. Tversky (2000) *Choices, Values and Frames* (New York: Cambridge University Press).

126 J. von Neumann and O. Morgenstern (1947) *Theory of Games and Economics Behavior*, 2nd edn (Princeton, NJ: Princeton University Press).

127 Ibid.

128 M. Allais (1953) 'Fondements d'une théorie positive des choix comportant un risque et critique des postulats et axioms de l'école americaine', *Econométrie*, 40, 257–332.

129 Kahneman and Tversky, 'Prospect Theory'; Tversky and Kahneman, 'The Framing of Decisions and the Psychology of Choice'.

130 Kahneman and Tversky, 'Prospect Theory', 47, 263–292.

131 S. Plous (1993) *The Psychology of Judgement and Decision Making* (New York: McGraw-Hill), p. 97.

132 Ibid., p. 98.

133 A. Tversky and D. Kahneman (1992) 'Advances in Prospect Theory: Cumulative Representation of Uncertainty', *Journal of Risk and Uncertainty*, 5, 297–323.

134 Ibid.

135 R. Hastie and R. M. Dawes (2001) *Rational Choice in an Uncertain World: The Psychology of Judgment and Decision Making* (Thousand Oaks, CA: Sage Publications), pp. 216, 294.

136 Kahneman, 'Perspective on Judgment and Choice, Mapping Bounded Rationality'.

137 D. Kahneman, J. L. Knetsch and R. H. Thaler (1991) 'The Endowment Effect, Loss Aversion, and Status Quo Bias', *Journal of Economic Perspectives*, 5, 193–206; Tversky and Kahneman, 'Advances in Prospect Theory'.

138 R. O. Swalm (1966) 'Utility Theory: Insights into Risk Taking', *Harvard Business Review*, 44, 123–136; Tversky and Kahneman, 'Advances in Prospect Theory'.

139 W. Samuelson and R. Zeckhauser (1988) 'Status-Quo Bias in Decision Making', *Journal of Risk and Uncertainty*, 1, 7–59.

140 R. Mehra and E. C. Prescott (1985) 'The Equity Premium: A Puzzle', *Journal of Monetary Economics*, 15, 145–161.

141 S. Benartzi and R. Thaler (1995) 'Myopic Loss Aversion and the Equity Premium Puzzle', *Quarterly Journal of Economics*, 110(1), 73–92.

142 T. Odean (1998) 'Are Investors Reluctant to Realize Their Losses?', *The Journal of Finance*, 53(5), 1775–1798.

143 K. Erta, S. Hunt, Z. Iscenko and W. Brambley (2013) 'Applying Behavioral Economics at the Financial Conduct Authority', *Financial Conduct Authority Occasional Paper* (London), 1.

144 S. DellaVigna (2009) 'Psychology and Economics: Evidence from the Field', *Journal of Economic Literature*, 47(2), 315–372, http:www.aeaweb.org/articles.php?doi=10.1257/jel.47.2.315; http://eml.berkeley.edu/~sdellavi/wp/01-DellaVigna-4721.pdf, retrieved 9 September 2014.

145 N. Barberis and R. Thaler (2003) 'A Survey of Behavioral Finance', in G. M. Constantinides, M. Harris and R. Stulz, eds, *Handbook of Economics of Finance* (Amsterdam: Elsevier Science), 1063–1085; DellaVigna, 'Psychology and Economics'.

146 K. Arrow (1958) 'Utilities, Attitudes, Choices: A Review Note', *Econometrica*, 26(1), 1–23.

147 The *Europe Economics* review for DG Sanco (2008) claims that while sometimes we could question the validity of consumers' preferences and whether his choices are really in his interest, we should mostly respect them as reflecting 'genuine' preferences.

148 See D. Laibson (1997), T. O'Donaghue and M. Rabin (1999, 2006), B. Madrian and D. Shea (2001), DellaVigna and Malmendier (2004), R. H. Thaler and S. Benartzi (2004) and P. Heidhues and B. Köszegi and (2010); D. Laibson (1997) 'Golden Eggs and Hyperbolic Discounting', *Quarterly Journal of Economics*, 112, 443–477; T. O'Donoghue and M. Rabin (1999) 'Doing It Now or Later', *American Economic Review*, 89(1), 103–124; T. O'Donaghue and M. Rabin (2006) 'Optimal Sin Taxes', *Journal of Public Economics*, 90(10–11), 1825–1849; B. Madrian and D. Shea (2001) 'The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior', *Quarterly Journal of Economics*, 116, 1149–1187; S. DellaVigna and U. Malmendier (2004) 'Contract Design and Self-Control: Theory and Evidence', *Quarterly Journal of Economics*, 119(2), 353–402; R. H. Thaler and S. Benartzi (2004) 'Save More Tomorrow: Using Behavioral Economics to Increase Employee Saving', *Journal of Political Economy*, 112, S164–S187; P. Heidhues and B. Kőszegi, 'Exploiting Naivete about Self-Control in the Credit Market', *American Economic Review*, 100(5), 2279–2303.

149 S. Frederick, G. Loewenstein and T. O'Donoghue (2002) 'Time Discounting and Time Preference: A Critical Review', *Journal of Economic Literature*, 40, 351–401.

150 Laibson, 'Golden Eggs and Hyperbolic Discounting'; O'Donoghue and Rabin, 'Doing It Now or Later'; O'Donaghue and Rabin, 'Optimal Sin Taxes'; Madrian and Shea, 'The Power of Suggestion'; DellaVigna and Malmendier, 'Contract Design and Self-Control'; Thaler and Benartzi, 'Save More Tomorrow'; Heidhues and Kőszegi, 'Exploiting Naivete about Self-Control in the Credit Market'.

151 Consumers tend to regret the choice after the fact, recognizing that it was not in their best interest. There is a large range of academic literature that documents and seeks to explain these inconsistencies in credit present biased card borrowing – see,

for example, L. M. Ausubel (1999), P. Heidhues and B. Kőszegi and (2010), H. Shui and L. M. Ausubel (2005) and many others.

L. M. Ausubel (1999) 'Adverse Selection in the Credit Card Market', *Working Paper*, Department of Economics, University of Maryland; Heidhues and Kőszegi, 'Exploiting Naivete about Self-Control in the Credit Market'; H. Shui and L. M. Ausubel (2005) 'Time Inconsistency in the Credit Card Market', *Working Paper*, Department of Economics, University of Maryland.

Given the opportunity, consumers often welcome an opportunity to pre-commit to making decisions on a consistent forward-looking basis. For example, an experiment by N. Ashraf et al. (2006) found that consumers with more time-inconsistent preferences were significantly more willing to take up a commitment savings product and considerably increased their saving rates as a result.

N. Ashraf, D. Karlan and W. Yin (2006) 'Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines', *Quarterly Journal of Economics*, 121(2), 635–672.

152 Thaler and Benartzi, 'Save More Tomorrow'.

153 Kahneman and Tversky, 'Judgment under Uncertainty'; Kahneman and Tversky, 'Prospect Theory'; D. Kahneman and A. Tversky (1991) 'Loss Aversion in Riskless Choice: A Reference-Dependent Model', *Quarterly Journal of Economics*, 106(4), 1039–1061; Odean, 'Are Investors Reluctant to Realize Their Losses?'; D. Genesove and C. Mayer (2001) 'Loss Aversion and Seller Behavior: Evidence from the Housing Market', *Quarterly Journal of Economics*, 116(4), 1233–1260; J. A. List (2004) 'Neoclassical Theory versus Prospect Theory: Evidence from the Marketplace', *Econometrica*, 72(2), 615–625; P. Heidhues and B. Kőszegi (2008) 'Competition and Price Variation When Consumers are Loss Averse', *American Economic Review*, 98(4), 1245–1268.

154 S. Gächter, H. Orzen, E. Renner and C. Starmer (2009) 'Are Experimental Economists Prone to Framing Effects? A Natural Field Experiment', *Journal of Economic Behavior & Organization*, 70, 443–446.

155 G. Loomes and R. Sugden (1982) 'Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty', *The Economic Journal*, 92, 805–824; M. Huysentruyt and D. Read (2010) 'How Do People Value Extended Warranties? Evidence from Two Field Surveys', *Journal of Risk and Uncertainty*, 40, 197–218; T. Baker and P. Siegelman (2013) 'Protecting Consumers from Add-On Insurance Products: New Lessons for Insurance Regulation from Behavioral Economics', *U of Penn, Inst for Law & Econ Research Paper*, 13(1). Available at SSRN: http://ssrn.com/abstract=2199569 or http://dx.doi.org/10.2139/ssrn.2199569

156 U. Malmendier and G. Tate (2005) 'CEO Overconfidence and Corporate Investment', *Journal of Finance*, 60(6), 2661–2700; U. Malmendier and G. Tate (2008) 'Who Makes Acquisitions? CEO Overconfidence and the Market's Reaction', *Journal of Financial Economics*, 89(1), 20–43; A. Sandroni and F. Squintani (2008) 'Overconfidence, Insurance, and Paternalism', *American Economic Review*, 97(5), 1994–2004; M. Grubb (2009) 'Selling to Overconfident Consumers', *American Economic Review*, 99(5), 1770–1807.

157 D. A. Moore and P. J. Healy (2008) 'The Trouble with Overconfidence', *Psychological Review*, 115(2), 502–517.

158 M. Rabin (2002) 'Inference by Believers in the Law of Small Numbers', *Quarterly Journal of Economics*, 117(3), 775–816.

Kahneman and Tversky, 'Judgment under Uncertainty'.

Tversky and Kahneman, 'Belief in the Law of Small Numbers'.

159 G. Loewenstein, T. O'Donoghue and M. Rabin (2003) 'Projection Bias in Predicting Future Utility', *Quarterly Journal of Economics*, 118(4), 1209–1248; M. Conlin, T. O'Donoghue and T. J. Vogelsang (2008) 'Projection Bias in Catalog Orders', *American Economic Review*, 97(4), 1217–1249.

160 See:

Thaler, 'Mental Accounting and Consumer Choice'; Thaler, 'Mental Accounting Matters',; D. Prelec and G. Loewenstein (1998) 'The Red and the Black: Mental Accounting of Savings and Debt', *Marketing Science*, 17(1), 4–28; D. Read, G. Loewenstein and M. Rabin (1999) 'Choice Bracketing', *Journal of Risk and Uncertainty*, 19, 171–197.

161 Kahneman and Tversky, 'The Framing of Decisions and the Psychology of Choice'; B. M. Barber and T. Odean (2008) 'All that Glitters: The Effect of Attention and News on the Buying Behaviour of Individual and Institutional Investors', *Review of Financial Studies*, 21(2), 785–818; N. Chater, S. Huck and R. Inderst (2010) 'Consumer Decision-Making in Retail Investment Services', *Report to the European Commission (SANCO)*; Office of Fair Trading (2010b) 'The Impact of Price Frames on Consumer Decision Making'.

Also see:

X. Gabaix and D. Laibson (2006) and S. Huck, J. Zhou and C. Duke (2011) for key firm responses.

X. Gabaix and D. Laibson (2006) 'Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets', *Quarterly Journal of Economics*, 121(2), 505–540; S. Huck, J. Zhou and C. Duke (2011) 'Consumer Behavioral Biases in Competition: A Survey', *A Report for the Office of Fair Trading*, OFT1324.

162 I. P. Levin, S. L. Schneider and G. J. Gaeth (1998) 'All Frames are Not Created Equal: A Typology and Critical Analysis of Framing Effects', *Organizational Behavior and Human Decision Processes*, 76, 149–188.

163 D. Maheswaran, D. M. Mackie and S. Chaiken (1992) 'Brand Name as a Heuristic Cue: The Effects of Task Importance and Expectancy Confirmation on Consumer Judgments', *Journal of Consumer Psychology*, 1, 317–336.

164 See:

Kahneman and Tversky, 'Judgment under Uncertainty'.

S. S. Iyengar and M. R. Lepper (2000) 'When Choice is Demotivating: Can One Desire Too Much of a Good Thing?', *Journal of Personality and Social Psychology*, 79(6), 995–1006.

S. Benartzi and R. H. Thaler (2001) 'Naive Diversification Strategies in Defined Contribution Savings Plans', *American Economic Review*, 91(1), 79–98.

N. Stewart (2009) 'The Cost of Anchoring on Credit Card Minimum Payments', *Psychological Science*, 20, 39–41.

165 D. Kahneman (2003) 'Maps of Bounded Rationality: Psychology for Behavioral Economics', *The American Economic Review*, 93(5), 1449–1475. doi:10.1257/000282803322655392.

166 See Cialdini (2006), Schultz et al. (2008) and Sunstein (2011b) for a discussion of social norms as 'nudges'.

R. B. Cialdini (2006) *Influence: The Psychology of Persuasion* (New York: Harper Business); Competition Commission (2009) *Market Investigation into Payment Protection Insurance*. Final Report; P. W. Schultz, J. M. Nolan, R. B. Cialdini, N. J. Goldstein and V. Griskevicius (2008) 'The Constructive, Destructive, and Reconstructive Power of Social Norms', *Psychological Science*, 18(5), 429–434; C. R. Sunstein (2011b) 'Empirically Informed Regulation', *University of Chicago Law Review*, 78(4), 1349–1429.

167 A. Eliot, P. Dolan, I. Vlaev, C. Adriaenssens and R. Metcalfe (2010) 'Transforming Financial Behaviour: Developing Interventions That Build Financial Capability', *A Report for the Consumer Financial Education Body.*
168 D. Ariely (2008) *Predictably Irrational* (New York: Harper Collins).
169 C. Duhigg (2012) *The Power of Habit: Why We Do What We Do in Life and Business* (New York: Random House).
170 G. Ellison (2006) *Bounded Rationality in Industrial Organization: In Richard Blundell, Advances in Economics and Econometrics: Theory and Applications*, W. Newey and T. Persson, eds. (Ninth Word Congress, Cambridge University Press); R. Spiegler (2011) *Bounded Rationality and Industrial Organization* (New York: Oxford University Press); Huck, Zhou and Duke, 'Consumer Behavioral Biases in Competition'; Oxera (2013) Behavioral Economics and Its Impact on Competition Policy – A Practical Assessment with Illustrative Examples from Financial Services. *Prepared for the Netherlands Authority for Consumers and Markets. https://www.acm.nl/nl/download/bijlage/?id=11161*
171 Gabaix and Laibson, 'Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets'.
172 M. Armstrong and S. Huck (2010) 'Behavioral Economics as Applied to Firms: A Primer', *A paper prepared for the Office of Fair Trading.*
173 Chater, Huck and Inderst, 'Consumer Decision-Making in Retail Investment Services'.
174 M. Bertrand, D. Karlan, S. Mullainathan, E. Shafir and J. Zinman (2010) 'What's Advertising Content Worth? Evidence from a Consumer Credit Marketing Field Experiment', *Quarterly Journal of Economics*, 125(1), 263–306;

Chater, Huck and Inderst, 'Consumer Decision-Making in Retail Investment Services'.

175 B. R. Handel (2011) 'Adverse Selection and Switching Costs in Health Insurance Markets: When Nudging Hurts', *Working Paper 17459*, Manuscript, UC Berkeley.
176 Chater, Huck and Inderst, 'Consumer Decision-Making in Retail Investment Services'.
177 J. Y. Campbell, H. E. Jackson, B. C. Madrian and P. Tufano (2011a) 'Consumer Financial Protection', *The Journal of Economic Perspectives*, 25(1), 91–113.
178 See S. DellaVigna and U. Malmendier (2006) for a good, early example of academic research that builds evidence for a behavioral bias in a market (present bias and projection bias, in choosing a gym contract and then using the service).

S. DellaVigna and U. Malmendier (2006) 'Paying Not to Go to the Gym', *American Economic Review*, 96(3), 694–719.

179 See B. Kőszegi and M. Rabin (2008) on the concept of true preferences and B. D. Bernheim and A. Rangel (2009) for welfare analysis when true preferences do not exist.

B. Kőszegi and M. Rabin (2008) 'Choices, Situations, and Happiness', *Journal of Public Economics*, 92, 1821–1832.

B. D. Bernheim and A. Rangel (2009) 'Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics', *The Quarterly Journal of Economics*, 124(1), 51–104.

180 D. Hirshleifer (2008) 'Psychological Bias as a Driver of Financial Regulation', *European Financial Management*, 14(5), 856–874.
181 See Bar-Gill (2012) and Campbell et al. (2011a) on the use of information disclosures, Sunstein (2010, 2011a, 2011b) on applying behaviorally-informed disclosure to policy and regulation and Federal Trade Commission (2008) on mortgage disclosure.

O. Bar-Gill (2012) *Seduction by Contract: Law, Economics and Psychology in Consumer Markets* (Oxford: Oxford University Press); Campbell, Jackson, Madrian and Tufano, 'Consumer Financial Protection'; C. R. Sunstein (2010) 'Disclosure and Simplification as Regulatory Tools', *Memorandum for the Heads of Executive Departments and Agencies*.

C. R. Sunstein (2011a) 'Informing Consumers through Smart Disclosure', *Memorandum for the Heads of Executive Departments and Agencies*.

Sunstein, 'Empirically Informed Regulation'; Federal Trade Commission (2008) 'Improving Consumer Mortgage Disclosures: An Empirical Assessment of Current and Prototype Disclosure Forms', *Bureau of Economics Staff Report*, Federal Trade Commission, Washington, DC.

182 M. Bertrand and A. Morse (2011) 'Information Disclosure, Cognitive Biases and Payday Borrowing', *Journal of Finance*, 66(6), 1865–1893.

183 See D. de Meza et al. (2008) and A. Eliot et al. (2010) for discussions of effectiveness of provision of education and advice to consumers.

D. de Meza, B. Irlenbusch and D. Reyniers (2008) 'Financial Capability: A Behavioral Economics Perspective', *FSA Consumer Research Paper*, 69; Eliot, Dolan, Vlaev, Adriaenssens and Metcalfe, 'Transforming Financial Behaviour'.

184 See Cabinet Office (2012) for a discussion of the presentation of information and citations.

Cabinet Office (2012) *Applying Behavioral Insights to Reduce Fraud, Error and Debt. Cabinet Office and Institute for Government (2010).* MINDSPACE: Influencing Behaviour through Public Policy (London: The Stationary Office).

185 See R. H. Thaler and C. R. Sunstein (2008) or C. R. Sunstein (2011b) on the use of defaults as nudges and J. J. Choi et al. (2005) for an early academic application.

R. H. Thaler and C. R. Sunstein (2008) *Nudge: Improving Decisions about Health, Wealth, and Happiness* (New Haven & London: Yale University Press); Sunstein, 'Empirically Informed Regulation'; J. J. Choi, D. Laibson, B. C. Madrian and A. Metrick (2005) 'Optimal Defaults and Active Decisions', *NBER Working Paper* 11074, National Bureau of Economic Research, Cambridge, MA.

186 See Beshears et al. (2008) for a discussion of these options.

J. Beshears, J. Choi, D. Laibson and B. Madrian (2008) 'How Are Preferences Revealed?', *Journal of Public Economics*, 92(8–9), 1787–1794.

187 See D. de Meza and D. J. Reyniers (2012), for example, for a model where banning hidden charges can make consumers worse off.

D. de Meza and D. J. Reyniers (2012) 'Every Shroud Has a Silver Lining: The Visible Benefits of Hidden Surcharges', *Economics Letters*, 116(2), 151–153.

4. Bubbles versus the Valuation of Fundamentals

1 E. F. Fama (1970) 'Efficient Capital Markets: A Review of Theory and Empirical Work', *The Journal of Finance*, 25(2), 383–417.

2 E. F. Fama (1991) 'Efficient Capital Markets: II', *The Journal of Finance*, 46(5), 1575–1617, http://links.jstor.org/sici?sici=0022-082%28199112%2946%3A5%3C1575%3AECMI%3E2.0.CO%3B2-L

3 R. P. Flood and P. M. Garber (1980) 'Gold Monetization and Gold Discipline', *NBER Working Paper*, 0544, National Bureau of Economic Research, Inc. Published as: R. P. Flood and P. M. Garber (1984) 'Gold Monetization and Gold Discipline', *Journal of Political Economy*, 92(1), 90–107.

4 J. D. Hamilton and C. H. Whiteman (1985) 'The Observable Implications of Self-fulfilling Expectations', *Journal of Monetary Economics*, 16(3), 353–373.
5 G. McQueen and S. Thorley (1994) 'Bubbles, Stock Returns, and Duration Dependence', *The Journal of Financial and Quantitative Analysis*, 29(3), 379–401.
6 K. R. Popper (1957) *The Poverty of Historicism* (London and New York: Routledge), pp. xi–xii.
7 K. R. Popper (1990) *A World of Propensities* (Bristol: Thoemmes Antiquarian Books).
8 J. D. Hamilton (1986) 'On Testing for Self-fulfilling Speculative Price Bubbles', *International Economic Review*, 27(3), 545–552.
9 J. Tirole (1985) 'Asset Bubbles and Overlapping Generations', *Econometrica*, 53(6), 1499–1528.
10 R. J. Shiller (1981) 'Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?', *American Economic Review*, 71, 421–436.
11 The price of the shares in the market.
12 EPS estimated as [(NI – dividends on preferred stock)/average outstanding shares] is the key variable in determining a share's price as it represents the portion of a company's profit allocated to each outstanding share of common stock.
13 F. Black and M. Scholes (1973) 'The Pricing of Options and Corporate Liabilities', *Journal of Political Economy*, 81(3), 637–654.
14 F. J. Lopez Lubian (March 2000) *Dirección de Negocios y Creación de Valor*, Una guía práctica, published by CISS/PRAXIS, Based on Chapter 1.
15 'Two Decades Later: Nikkei and Lessons from the Fall', http://www.aei.org/publication/two-decades-later-nikkei-and-lessons-from-the-fall/, visited 21 August 2015.
16 http://finance.yahoo.com/q/hp?s=GE&a=11&b=1&c=1995&d=07&e=1&f=1998&g=d, visited 21 August 2015.
17 GE (1997) Annual report, https://bib.kuleuven.be/files/ebib/jaarverslagen/GE_1997.pdf, visited 21 August 2015.
18 http://www.sec.gov/Archives/edgar/data/40545/0000040545-96-000026.txt, visited 21 August 2015.
19 F. J. López Durán (2002) *Valoración de Empresas*, Conference at Instituto de Empresa, Madrid, Spain, February 2.
20 Dell Computer Corporation Annual report, http://i.dell.com/sites/doccontent/corporate/financials/en/Documents/ar_fy1998-YIR.pdf, visited 21 August 2015.
21 Dell Inc. Historical closing prices, http://www.dell.com/learn/us/en/vn/corporate~secure~en/documents~dell-closing-costs.pdf, visited 21 August 2015.
22 http://www.icmrindia.org/casestudies/catalogue/Business%20Strategy2/Sony-Columbia%20Pictures%20Cross%20Border%20Acquisition.htm, visited 21 August 2015.
23 http://www.independent.co.uk/news/business/welcome-to-tinseltown-mr-morita-when-sony-bought-columbia-pictures-its-hardheaded-japanese-realists-found-themselves-in-a-fantasy-world-papered-with-money-fuelled-by-rumour-and-riven-by-bloodletting-after-spectacular-setbacks-including-the-last-action-hero-and-with-its-chief-sidelined-by-a-stroke-phil-reeves-reports-from-los-angeles-on-sonys-bid-to-show-that-it-isnt-lost-in-the-dream-factory-1369085.html, visited 21 August 2015.
24 https://www.sec.gov/news/speech/speecharchive/1998/spch220.txt
25 E. MacDonald (1999) 'Analysts Increasingly Look to Cash Flow over Earnings', *The Wall Street Journal*, http://www.wsj.com/articles/SB922926887562477208, visited 21 August 2015.

26 C. A. Ramezani, L. Soenen and A. Jung (2002) 'Growth, Corporate Profitability, and Value Creation', *Financial Analysts Journal*, http://online.sfsu.edu/alanjung/FAJPaper.pdf, visited 21 August 2015.
27 https://en.wikipedia.org/wiki/Market_share_of_personal_computer_vendors, visited 21 August 2015.
28 https://en.wikipedia.org/wiki/Compaq, visited 21 August 2015.
29 Ramezani, Soenen and Jung, 'Growth, Corporate Profitability, and Value Creation'.
30 M. J. Mauboussin (2011) 'What Shareholder Value is Really About', *HBR Online Forum: The CEO's Role in Fixing the System*, https://hbr.org/2011/10/ceos-must-understand-what-crea, 23 August 2015.

5. Bubbles and Technical Trading

1 C. D. Kirkpatrick and J. Dahlquist (2006) *Technical Analysis: The Complete Resource for Financial Market Technicians* (Upper Saddle River, NJ: FT Press), p. 3.
2 J. J. Murphy (1999) *Technical Analysis of the Financial Markets* (New York: New York Institute of Finance), pp. 1–5, 24–31.
3 C. H. Park and S. H. Irwin (2007) 'What Do We Know About the Profitability of Technical Analysis?', *Journal of Economic Surveys*, 21(4), 786–826.
4 E. Fama (1970) 'Efficient Capital Markets: A Review of Theory and Empirical Work', *The Journal of Finance*, 25(2), 383–417.
5 C. Carolan (1998) 'Autumn Panics and Calendar Phenomenon', *Market Technicians Association Dow Award*. http://knowledgebase.mta.org/?fuseaction=kb.home&kbDomainID=28C003A0-A674-6EB7-BA133B53DC07BEBC. Visited 31 August 2015.
6 G. A. W. Griffioen (2003) 'Technical Analysis in Financial Markets'. http://www.oocities.org/gerwingriffioen/publications/tta.htm. Visited 31 August 2015.
7 J. D. Schwager (1999) *Getting Started in Technical Analysis* (New York: John Wiley & Sons), p. 2.
8 S. Nison (1991) *Japanese Candlestick Charting Technique*, 2nd edn (Paramus, NJ: Prentice Hall Press), pp. 15–18.
9 S. Nison (1994) *Beyond Candlesticks: New Japanese Charting Techniques Revealed* (New York: John Wiley & Sons), p. 14.
10 Park and Irwin, 'What Do We Know About the Profitability of Technical Analysis?'.
11 J. A. Paulos (2003) *A Mathematician Plays the Stock Market* (New York: Basic Books), pp. 38–39.
12 M. P. Taylor and H. Allen (1992) 'The Use of Technical Analysis in the Foreign Exchange Market', *Journal of International Money and Finance*, 11(3), 304–314.
13 Schwager, *Getting Started in Technical Analysis*.
14 http://www.investopedia.com/articles/trading/11/indicators-and-strategies-explained.asp
15 Elliott published his theory of market behavior in *The Wave Principle* (1938) and summarized it in 1939 in a series of articles published in the *Financial World* magazine; in 1946 he covered it comprehensively in *Nature's Laws: The Secret of the Universe* (New York: Elliott, 1946).
16 S. W. Poser (2003) *Applying Elliott Wave Theory Profitably* (New York: John Wiley & Sons), pp. 2–17.
17 A. J. Frost and R. R. Prechter, Jr. (2005) *Elliott Wave Principle*, 10th edn (Gainesville, GA: New Classics Library), pp. 31, 78–85.
18 Poser, *Applying Elliott Wave Theory Profitably*.

19 R. R. Prechter, Jr. (2003) 'A New Perspective on the Quantitative Self-Affinity of Elliott Waves', *Journal of Technical Analysis*, Summer–Fall, 25–32.
20 J. Casti (2002) 'I Know What You'll Do Next Summer', *New Scientist*, August 31, p. 29.
21 Frost and Prechter, *Elliott Wave Principle*.
22 R. R. Prechter, Jr. (2006) *Elliott Waves, Fibonacci, and Statistics*, p. 2. http://www.socionomics.org/pdf/EW_Fibo_Statistics.pdf
23 R. Batchelor and R. Ramyar (2006) 'Magic Numbers in the Dow', pp. 13, 31. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.135.1144&rep=rep1&type=pdf. Visited 31 August 2015.
24 R. N. Elliott (1994) R.N. *Elliott's Masterwork*, R. R. Prechter, Jr. ed. (Gainesville, GA: New Classics Library), pp. 70, 217, 194, 196.
25 R. Wilkin (2006) 'Riding the Waves: Applying Elliott Wave Theory to the Financial and Commodity Markets', *The Alchemist*. http://www.lbma.org.uk/assets/alch43_elliott.pdf
26 J. Kotick (2005) 'An Introduction to the Elliott Wave Principle', *The Alchemist*. http://www.lbma.org.uk/assets/alch40_elliott.pdf
27 D. Sornette, A. Johansen and J. P. Bouchaud (1996) 'Stock Market Crashes, Precursors and Replicas', *Journal de Physique I*, 6(1), 167–175.
28 M. B. Fisher, (2002) *The Logical Trader* (Hoboken, NJ: John Wiley & Sons), p. x (foreword by Paul Tudor Jones).
29 G. Neely and E. Hall (1990) *Mastering Elliott Wave – Presenting the Neely Method: The First Scientific, Objective Approach to Market Forecasting with the Elliott Wave Theory* (Brightwaters, New York: Windsor Books).
30 T. Landon, Jr. (2007) 'The Man Who Won as Others Lost', *The New York Times*, Visited 26 August 2015. http://www.nytimes.com/2007/10/13/business/13speculate.html?pagewanted=all
31 B. Mandelbrot and R. L. Hudson (2004) *The (mis)Behavior of Markets* (New York: Basic Books), p. 245.
32 D. R. Aronson (2006) *Evidence-Based Technical Analysis* (Hoboken, New Jersey: John Wiley & Sons), p. 61.
33 Professor Roy Batchelor and researcher Richard Ramyar (a former director of the United Kingdom Society of Technical Analysts and formerly global head of research at Lipper and Thomson Reuters Wealth Management and currently professor at Cass Business School), see Batchelor and Ramyar, 'Magic Numbers in the Dow'.
34 Ibid.
35 Prechter, *Elliott Waves, Fibonacci, and Statistics*.
36 L. Bachelier (2006) *Louis Bachelier's Theory of Speculation: The Origins of Modern Finance*, trans. Mark Davis and Alison Etheridge (Princeton, NJ: Princeton University Press).
37 P. Samuelson (1965) 'Proof That Properly Anticipated Prices Fluctuate Randomly', *Industrial Management Review*, 6, 41.
38 E. F. Fama (1965) 'The Behavior of Stock Market Prices', *The Journal of Business*, 38(1), 34–105.
39 E. F. Fama and K. R. French (1992) 'The Cross-Section of Expected Stock Returns', *The Journal of Finance*, 47(2), 427–465.
40 T. Odean (1998) 'Are Investors Reluctant to Realize Their Losses?', *The Journal of Finance*, 53(5), 1775–1798.
41 B. M. Barber and T. Odean (2000) 'Trading is Hazardous to Your Wealth: The Common Investment Performance of Individual Investors', *The Journal of Finance*, 55(2), 773–806.

42 L. Barras, O. Scaillet and R. Wermers (2010) 'False Discoveries in Mutual Fund Performance: Measuring Luck in Estimated Alphas', *The Journal of Finance*, 65(1), 179–216.
43 See http://www.cxoadvisory.com/gurus/. Visited 24 August 2015.
44 *The Economist* (2005) 'In Come the Waves: The Worldwide Rise in House Prices is the Biggest Bubble in History. Prepare for the Economic Pain When It Pops', June 16, http://www.economist.com/node/4079027. Visited 24 August 2015.
45 A. Tversky and D. Kahneman (1974) 'Judgment under Uncertainty: Heuristics and Biases', *Science*, 185, 1124–1130.
46 M. Grinblatt, S. Titman and R. Wermers (1995) 'Momentum Investment Strategies, Portfolio Performance, and Herding: A Study of Mutual Fund Behavior', *American Economic Review*, 85, 1088–1105.
47 R. Wermers (1999) 'Mutual Fund Herding and the Impact on Stock Prices', *Journal of Finance*, 54, 581–622.
48 S. G. Badrinath and S. Wahal (2002) 'Momentum Trading By Institutions', *Journal of Finance*, 57, 2449–2478.
49 J. B. De Long, A. Shleifer, L. H. Summers and R. J. Waldmann (1990) 'Positive Feedback Investment Strategies and Destabilizing Rational Speculation', *Journal of Finance*, 45, 379–395.
50 H. Hong and J. C. Stein (1999) 'A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets', *Journal of Finance*, 54, 2143–2184.
51 N. Jegadeesh and S. Titman (1993) 'Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency', *Journal of Finance*, 48, 65–91.
52 A. Alti, R. Kaniel and U. Yoeli (2012) 'Why Do Institutional Investors Chase Return Trends?', *Journal of Financial Intermediation*, 21(4), 694–721, http://rkaniel.simon.rochester.edu/papers/fundmoment.pdf. Visited 31 August 2015.
53 D. Romer (1993) 'Rational Asset-Price Movements without News', *American Economic Review*, 83, 1112–1130.
54 C. Avery and P. Zemsky (1998) 'Multidimensional Uncertainty and Herd Behavior in Financial Markets', *American Economic Review*, 88, 724–748.
55 H. H. Cao, J. D. Coval and D. Hirshleifer (2002) 'Sidelined Investors, Trading-Generated News, and Security Returns', *Review of Financial Studies*, 15, 615–648.
56 J. Wang (1993) 'A Model of Intertemporal Asset Prices Under Asymmetric Information', *Review of Economic Studies*, 60, 249–282; J. Wang (1994) 'A Model of Competitive Stock Trading Volume', *Journal of Political Economy*, 102, 127–168.
57 D. Hirshleifer, A. Subrahmanyam and S. Titman (1994) 'Security Analysis and Trading Patterns When Some Investors Receive Information Before Others', *Journal of Finance*, 49, 1665–1698.
58 M. J. Brennan and H. H. Cao (1996) 'Information, Trade, and Derivative Securities', *Review of Financial Studies*, 9, 163–208.
59 K. Daniel, D. Hirshleifer and A. Subrahmanyam (1998) 'Investor Psychology and Security Market Under- and Overreactions', *Journal of Finance*, 53, 1839–1884.
60 S. Gervais and T. Odean (2001) 'Learning to Be Overconfident', *Review of Financial Studies*, 14, 1–27.
61 Daniel, Hirshleifer and Subrahmanyam, 'Investor Psychology and Security Market Under- and Overreactions'.
62 Gervais and Odean, 'Learning to Be Overconfident'.
63 G. Caginalp and D. Balenovich (1999) 'Asset Flow and Momentum: Deterministic and Stochastic Equations', *Philosophical Transactions of the Royal Society A*, 357(1758), 2119–2133.

64 G. Caginalp and G. B. Ermentrout (1990) 'A Kinetic Thermodynamics Approach to the Psychology of Fluctuations in Financial Markets', *Applied Math Letters*, 4(4), 17–19; G. Caginalp and M. DeSantis (2011) 'Stock Price Dynamics: Nonlinear Trend, Volume, Volatility, Resistance and Money Supply', *Quantitative Finance*, 11(6), 849–861.

65 G. Caginalp, D. Porter and V. L. Smith (1998) 'Initial Cash/Asset Ratio and Asset Prices: An Experimental Study', *Proceedings of the National Academy of Sciences*, 95(2), 756–761.

66 G. Caginalp, D. Porter and V. L. Smith (2001) 'Financial Bubbles: Excess Cash, Momentum and Incomplete Information', *Journal of Psychology and Financial Markets*, 2(2), 80–99.

67 G. Caginalp, M. DeSantis and D. Swigon (2011) 'Are Flash Crashes Caused by Instabilities Arising from Rapid Trading?', *Wilmott Magazine*, 11, 46–47.

68 'Historical Rhodium Charts', *Kitco*. http://www.kitco.com/charts/historicalrhodium.html. Visited 19 February 2015.

69 C. Chiarella, R. Dieci and L. Gardini (2002) 'Speculative Behaviour and Complex Asset Price Dynamics: A Global Analysis', *Journal of Economic Behavior & Organization*, 49, 173–197.

70 W. A. Brock and C. H. Hommes (1998) 'Heterogeneous Beliefs and Routes to Chaos in a Simple Asset Pricing Model', *Journal of Economic Dynamics and Control*, 22(8–9), 1235–1274.

71 T. Ma and C. H. Yin (2013) 'Chartist Weights and Market Instability: Applications of the Heterogeneous Agent Model in International Stock Markets'. Working paper: http://sfm.finance.nsysu.edu.tw/pdf/2013pdf/098-1542833090.pdf

72 D. M. Cutler, J. M. Poterba and L. H. Summers (1991) 'Speculative Dynamics', *Review of Economic Studies*, 58(3), 529–546.

73 D. M. Cutler, J. M. Poterba and L. H. Summers (1990) 'Speculative Dynamics and the Role of Feedback Traders', *American Economic Review*, 80(2), 63–68.

74 J. DeLong, A. Shleifer, L. Summers and R. Waldmann (1990) 'Noise Trader Risk in Financial Markets', *Journal of Political Economy*, 98(4), 703–738.

75 J. A. Frankeland and K. A. Froot (1991) 'Chartists, Fundamentalists, and the Demand for Dollars', in A. Courakis and M. Taylor, eds, *Private Behaviour and Government Policy in Interdependent Economies* (Oxford: Oxford University Press), 73–126.

76 Hong and Stein, 'A Unified Theory of Underreaction'.

77 T. Lux (1998) 'The Socio-Economic Dynamics of Speculative Markets: Interacting Agents, Chaos and Fat Tails of Return Distributions', *Journal of Economic Behavior and Organization*, 33, 143–165.

78 M. M. Bange (2000) 'Do the Portfolios of Small Investors Reflect Positive Feedback Trading?', *Journal of Financial and Quantitative Analysis*, 35(2), 239–255.

79 D. B. Keim and A. Madhavan (1995) 'Anatomy of the Trading Process: Empirical Evidence on the Behavior of Institutional Investors', *Journal of Financial Economics*, 37(3), 371–398.

80 D. C. Ling (2005) 'A Random Walk Down Main Street: Can Experts Predict Returns on Commercial Real Estate?', *Journal of Real Estate Research*, 27(2), 137–154.

81 R. Kouwenberg and R. C. J. Zwinkels (2011) 'Chasing Trends in the USA Housing Market', February, http://complexsystems.anu.edu.au/w/images/9/96/2010.12.03.zwinkels.pdf. Visited 31 August 2015.

82 D. B. Keim (2003) *The Cost of Trend Chasing and The Illusion of Momentum Profits* (The Wharton School, University of Pennsylvania). Draft July 29.

83 V. L. Smith, G. L. Suchanek and A. W. Williams (1988) 'Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Asset Markets', *Econometrica*, 56(5), 1119–1151.

84 V. Lei, C. N. Noussair and C. R. Plott (2001) 'Nonspeculative Bubbles in Experimental Asset Markets: Lack of Common Knowledge of Rationality vs. Actual Irrationality', *Econometrica*, 69(4), 831.
85 S. S. Levine and E. J. Zajac (2007) 'The Institutional Nature of Price Bubbles'. Available at SSRN 960178. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=960178. Visited 31 August 2015.
86 C. Hommes, J. Sonnemans, J. Tuinstra and H. van de Velden (2005) 'Coordination of Expectations in Asset Pricing Experiments', *Review of Financial Studies*, 18(3), 955–980. Available at doi:10.1093/rfs/hhi003 and http://rfs.oxfordjournals.org/content/18/3/955. Visited 31 August 2015.
87 Levine and Zajac (2007) 'The Institutional Nature of Price Bubbles'.

6. Contagion

1 The World Bank definitions on contagion can be found in the following link: http://econ.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTRESEARCH/EXTPROGRAMS/EXTMACROECO/0%2c%2ccontentMDK:20889756~pagePK:64168182~piPK:64168060~theSitePK:477872%2c00.html
2 R. Dornbusch, Y. Park and S. Claessens (2000) 'Contagion: Understanding How It Spreads', *The World Bank Research Observer*, 15(2), 167–195.
3 K. Forbes and R. Rigobon (2002) 'No Contagion, Only Interdependence: Measuring Stock Market Co-movement', *Journal of Finance*, 57(5), 2223–2261.
4 An additional definition of contagion can be found on: http://www.investorwords.com/1066/contagion.html
5 G. Calvo and C. Reinhart (2000) 'When Capital Inflows Come to a Sudden Stop: Consequences and Policy Options', in P. Kenen and A. Swoboda, eds, *Reforming the International Monetary and Financial System* (Washington, DC: International Monetary Fund), pp. 175–201.
6 G. L. Kaminsky, C. M. Reinhart and C. A. Végh (2003) 'The Unholy Trinity of Financial Contagion', *Journal of Economic Perspectives*, 17(4), 51–74.
7 C. Reinhart and K. Rogoff (2009) *This Time is Different: Eight Centuries of Financial Folly*, (Princeton, NJ: Princeton University Press).
8 D. Bisias, M. D. Flood, A. W. Lo and S. Valavanis (2012) 'A Survey of Systemic Risk Analytics', US Department of Treasury, *Office of Financial Research*, 1.
9 S. Oosterloo and J. de Haan (2003) 'A Survey of Institutional Frameworks for Financial Stability', *Occasional Studies, De Nederlandsche Bank*, 1(4), 11–16.
10 J. Dow (2000) 'What is Systemic Risk? Moral Hazard, Initial Shocks, and Propagation', *Monetary and Economic Studies*, 1–24.
11 S. Eijffinger (2012) 'Defining and Measuring Systemic Risk', in S. Eijffinger and D. Masciandaro, eds, *Handbook of Central Banking, Financial Regulation and Supervision* (Cheltenham: Northampton Edward Elgar), pp. 316–317.
12 F. Allen, T. Beck, E. Carletti, P. R. Lane, D. Schoenmaker and W. Wagner (2011) 'Cross-Border Banking in Europe: Implications for Financial Stability and Macroeconomic Policies', *CEPR Report*, London.
13 Ibid.
14 V. V. Acharya (2009) 'A Theory of Systemic Risk and Design of Prudential Bank Regulation', *Journal of Financial Stability*, 5(3), 224–255; R. Ibragimov, D. Jaffee and J. Walden (2011) 'Diversification Disasters', *Journal of Financial Economics*, 99, 333–348; W. Wagner (2010) 'Diversification at Financial Institutions and Systemic Crises', *Journal of Financial Intermediation*, 19, 333–354.

15 F. Allen and D. Gale (2000) 'Bubbles and Crises', *Economic Journal*, 110, 236–255; A. Dasgupta (2004) 'Financial Contagion through Capital Connections: A Model of the Origin and Spread of Bank Panics', *Journal of the European Economic Association*, 2, 1049–1084; X. Freixas, B. Parigi and J. C. Rochet (2000) 'Systemic Risk, Interbank Relations and Liquidity Provision by the Central Bank', *Journal of Money, Credit and Banking*, 32(2), 611–638.

16 M. Brunnermeier and L. H. Pedersen (2009) 'Market Liquidity and Funding Liquidity', *Review of Financial Studies*, 22(6), 2201–2238.

17 P. E. Mistrulli (2011) 'Assessing Financial Contagion in the Interbank Market: Maximum Entropy versus Observed Interbank Lending Pattern', *Journal of Banking & Finance*, 35(5), 1114–1127.

18 *Telegraph* (2012a) 'Deposit Flight from Spanish Banks Smashes Record in July', August 28, http://www.telegraph.co.uk/finance/financialcrisis/9504911/Deposit-flight-from-Spanish-banks-smashesrecord-in-July.html

Telegraph (2012b) 'Withdrawals at Santander UK amid Spain Fears', May 18, http://www.telegraph.co.uk/finance/personalfinance/consumertips/banking/9276164/Withdrawals-at-Santander-UK-amid-Spain-fears.html

19 http://www.bloomberg.com/news/articles/2015-06-28/greece-closing-banks-as-expiring-bailout-spurs-withdrawal-wave, accessed 30 July 2015. Greece Closing Banks as Expiring Bailout Spurs Withdrawals, Paris Christos Paul, 28 June 2015, CEST Updated on 29 June 2015.

20 C. W. Calomiris and C. M. Kahn (1991) 'The Role of Demandable Debt in Structuring Optimal Banking Arrangements', *American Economic Review*, 81, 497–513.

21 Allen and Gale, 'Bubbles and Crises'.

22 X. Freixas, B. Parigi and J. C. Rochet (2000) 'Systemic Risk, Interbank Relations, and Liquidity Provision by the Central Bank', *Journal of Money, Credit, and Banking*, 32, 611–638.

23 C. Upper (2011) 'Simulation Methods to Assess the Danger of Contagion in Interbank Markets', *Journal of Financial Stability*, doi:10.1016/j.jfs.2010.12.001.

24 E. Nier, J. Yang, T. Yorulmazer and A. Alentorn (2007) 'Network Models and Financial Stability', *Journal of Economic Dynamics and Control*, 31, 2033–2060.

25 D. Diamond and P. Dybvig (1983) 'Bank Runs, Deposit Insurance, and Liquidity', *Journal of Political Economy*, 91, 401–419; I. Goldstein and A. Pauzner (2005) 'Demand Deposit Contracts and the Probability of Bank Runs', *Journal of Finance*, 60, 1293–1327.

26 V. V. Acharya and T. Yorulmazer (2008) 'Cash-in-the Market Pricing and Optimal Resolution of Bank Failures', *Review of Financial Studies*, 21, 2705–2742.

27 P. Aghion, P. Bolton and M. Dewatripont (2000) 'Contagious Bank Failures in a Free Banking System', *European Economic Review*, 44, 713–718; D. Diamond and R. Rajan (2005) 'Liquidity Shortages and Banking Crises', *Journal of Finance*, 60, 615–647.

28 V. V. Acharya and T. Yorulmazer (2007) 'Too Many to Fail: An Analysis of Time-Inconsistency in Bank Closure Policies', *Journal of Financial Intermediation*, 16, 1–31.

29 Y. Chen (1999) 'Banking Panics: The Role of First-Come, First-Served Rule and Information Externalities', *Journal of Political Economy*, 107(5), 946–968.

30 L. Kodres and M. Pritsker (2002) 'A Rational Expectations Model of Financial Contagion', *Journal of Finance*, 57(2), 769–800.

31 A. Dasgupta (2004) 'Financial Contagion through Capital Connections: A Model of the Origin and Spread of Bank Panics', *Journal of the European Economic Association*, 2, 1049–1084.

32 V. V. Acharya, D. Gromb and T. Yorulmazer (2008) 'Imperfect Competition in the Inter-Bank Market for Liquidity as a Rationale for Central Banking', *Mimeograph*.

33 J. C. Rochet and J. Tirole (1996) 'Interbank Lending and Systemic Risk', *Journal of Money Credit and Banking*, 28, 733–762.
34 D. Humphrey (1986) 'Payments Finality and Risk of Settlement Failure', in A. Saunders and L. White, eds, *Technology and the Regulation of Financial Markets: Securities, Futures and Banking* (Lexington, MA: Lexington Books), 97–120.
P. Angelini, G. Mariesca and D. Russo (1996) 'Systemic Risk in the Netting System', *Journal of Banking and Finance*, 20, 853–868.
35 C. A. Northcott (2002) 'Estimating Settlement Risk and the Potential for Contagion in Canada's Automated Clearing Settlement System', Bank of Canada, *Working Paper*, 41.
36 M. Blavarg and P. Nimander (2002) 'Inter-bank Exposures and Systemic Risk', Sveriges Riksbank, *Economic Review*, 2, 19–45.
37 R. Cifuentes, G. Ferrucci and H. S. Shin (2005) 'Liquidity Risk and Contagion', *Journal of the European Economic Association*, 3(2–3), 556–566; F. Fecht (2004) 'On the Stability of Different Financial Systems', *Journal of the European Economic Association*, 2(6), 969–1014.
38 A. Shleifer and R. W. Vishny (1997) 'The Limits of Arbitrage', *Journal of Finance*, 52(1), 35–55.
39 G. Calvo (1998) 'Capital Market Contagion and Recession: An Explanation of the Russian Virus', *Mimeograph*.
40 G. Kaminsky and C. Reinhart (2000) 'On Crises, Contagion, and Confusion', *Journal of International Economics*, 51(1), 145–168.
41 F. Caramazza, L. Ricci and R. Salgado (2000) 'Trade and Financial Contagion in Currency Crises', *IMF Working Paper* WP/00/55; Kaminsky and Reinhart, 'On Crises, Contagion, and Confusion'; C. Van Rijckeghem and B. Weder (2000) 'Financial Contagion: Spillovers through Banking Centers', *Mimeograph*.
42 J. Frankel and S. Schmukler (1998) 'Crises, Contagion and Country Funds: Effects on East Asia and Latin America', in R. Glick, ed., *Managing Capital Flows and Exchange Rates: Perspectives from the Pacific Basin* (New York: Cambridge University Press), pp. 232–266; G. Kaminsky, R. Lyons and S. Schmukler (2000) 'Managers, Investors, and Crises: Mutual Fund Strategies in Emerging Markets', *NBER Working Paper*, 7855.
43 L. Kodres and M. Pritsker (2002) 'A Rational Expectations Model of Financial Contagion', *Journal of Finance*, 57(2), 769–800.
44 M. Goldstein (1998) *The Asian Financial Crisis* (Washington, DC: Institute for International Economics).
45 G. L. Kaminsky, C. M. Reinhart and C. A. Végh (2003) 'The Unholy Trinity of Financial Contagion', *Journal of Economic Perspectives*, 17(4), 51–74.
46 G. L. Kaminsky, R. Lyons and S. Schmukler (2002) 'Liquidity, Fragility, and Risk: The Behavior of Mutual Funds During Crises', *Mimeograph*.
47 B. Eichengreen, A. Rose and Ch. Wyplosz (1996) 'Contagious Currency Crises: First Tests', *Scandinavian Journal of Economics*, 98(4), 463–484.
48 R. Glick and A. Rose (1999) 'Contagion and Trade: Why are Currency Crises Regional?', *Journal of International Money and Finance*, 18(4), 603–617.
49 Kaminsky and Reinhart, 'On Crises, Contagion, and Confusion'.
50 A. Mody and M. Taylor (2002) 'Common Vulnerabilities', *Mimeograph*.
51 D. Hirshleifer and T. Siew Hong (2003) 'Herd Behaviour and Cascading in Capital Markets: A Review and Synthesis', *European Financial Management*, 9(1), 25–66.
52 A. Kirman (1993) 'Ants, Rationality and Recruitment', *Quarterly Journal of Economics*, 108(1), 137–155.

53 Keynes's example was used to describe the choices of rational agents in a market. A newspaper would publish a contest in which the participants were rewarded for selecting, out of six female pictures, the one who would be chosen "most beautiful." A naïve participant would vote for the one she considered the most beautiful face. However, the more sophisticated contestant would rather assess common perceptions of beauty and vote for the one she thought would be chosen by the majority. This strategy could be carried a step further to consider other participants would each have their own ideas of what popular perceptions about beauty are, and so on, each level trying to assess the outcome of the process considering the reasoning of other rational agents.

54 A. Orléan (1990) 'Le rôle des influences interpersonelles dans la determination des cours boursiers', *Revue Economique*, V, 839–868; A. Orléan (1989) 'Mimetic Contagion and Speculative Bubbles', *Theory and Decision*, 27, 63–92; R. Topol (1991) 'Bubbles and Volatility of Stock Prices: Effect of Mimetic Contagion', *Economic Journal*, 101, 786–800.

55 V. Bala and S. Goyal (1998) 'Learning from Neighbors', *Review of Economic Studies*, 65(3), 595–621.

A. Bernardo and I. Welch (2001) 'On the Evolution of Overconfidence and Entrepreneurs', *Journal of Economics and Management Strategy*, 10(3), 301–330; N. Khanna (1997) 'Optimal Contracting with Moral Hazard and Cascading', *Review of Financial Studies*, 11(3), 559–596; M. Ridley (1996) *The Origins of Virtue: Human Instincts and the Evolution of Cooperation* (Hudson Street, NY: Viking/Penguin Books), pp. 182–185; R. J. Shiller (1995) 'Conversation, Information, and Herd Behavior', *American Economic Review*, 85(2), 181–185.

56 S. Bikhchandani, D. Hirshleifer and I. Welch (1998) 'Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades', *Journal of Economic Perspectives*, 12(3), 151–170; A. Devenow and I. Welch (1996) 'Rational Herding in Financial Economics', *European Economic Review*, 40(3–5), 603–615; D. Gale (1996) 'What have We Learned from Social Learning?', *European Economic Review*, 40(3–5), 617–628; S. Grant, S. King and B. Polak (1996) 'Information Externalities, Share-Priced Based Incentives and Managerial Behavior', *Journal of Economic Surveys*, 10(1), 1–21; D. Hirshleifer (1995) 'The Blind Leading the Blind: Social Influence, Fads and Informational Cascades', in K. Ieurulli and M. Tommasi, eds, *The New Economics of Human Behaviour* (Cambridge: Cambridge University Press), pp. 188–215; D. Hirshleifer (1998) 'Information Cascades and Social Conventions', in *The New Palgrave Dictionary of Economics and the Law* in The New Palgrave Dictionary of Economics and the Law, vol. II, 300–306 (Peter Newman, ed., New York: Stockton Press, 1998)

57 Shiller, 'Conversation, Information, and Herd Behavior'; R. J. Shiller (2000) *Irrational Exuberance* (Princeton, NJ: Princeton University Press).

58 A. A. Brandenburger and B. Polak (1996) 'When Managers Cover Their Posteriors: Making the Decisions the Market Wants to See', *Rand Journal of Economics*, 27, 523–541.

59 C. Avery and P. Zemsky (1998) 'Multidimensional Uncertainty and Herd Behavior in Financial Markets', *American Economic Review*, 88, 724–748.

60 D. Scharfstein and J. C. Stein (1990) 'Herd Behavior and Investment', *American Economic Review*, 80, 465–479.

61 B. Trueman (1994) 'Analyst Forecasts and Herding Behavior', *The Review of Financial Studies*, 7, 97–124.

62 J. Zwiebel (1995) 'Corporate Conservatism and Relative Compensation', *Journal of Political Economy*, 103, 1–25.

63 I. Welch (1992) 'Sequential Sales, Learning and Cascades', *Journal of Finance*, 47, 695–732; Devenow and Welch, 'Rational Herding in Financial Economics'; S. Bikhchandani, D. Hirshleifer and I. Welch (1992) 'A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades', *Journal of Political Economy*, 100(5), 992–1026.
64 J. R. Graham (1999) 'Herding Among Investment Newsletters: Theory and Evidence', *Journal of Finance*, 54, 237–268.
65 K. Froot, D. Scharfstein and J. Stein (1992) 'Herd on the Street: Informational Inefficiencies in a Market with Short-Term Speculation', *Journal of Finance*, 47, 1461–1484.
66 D. Hirsleifer, A. Subrahmanyam and S. Titman (1994) 'Security Analysis and Trading Patterns When Some Investors Receive Private Information Before Others', *Journal of Finance*, 49, 1665–1698.
67 J. Golec (1997) 'Herding on Noise: The Case of Johnson Redbook's Weekly Retail Sales Data', *Journal of Financial and Quantitative Analysis*, 32, 367–381.
68 R. J. Shiller (1984) 'Stock Prices and Social Dynamics', *Brooklings Papers*, 2, 457–498.
69 R. J. Shiller and J. Pound (1986) 'Survey Evidence on Diffusion of Interest Among Institutional Investors', *Cowles Foundation Discussion Papers* 794, Cowles Foundation for Research in Economics, Yale University.
70 K. D. West and A. W. Kleidon (1988) 'Bubbles, Fads and Stock Price Volatility Tests: A Partial Evaluation; Discussion', *Journal of Finance*, 43, 639–660.
71 D. Galant (1995) 'Financial Follies', *Institutional Investor*, 29, 139–140.
72 A. Orléan (1989) 'Pour une Approche Cognitive des Conventions Economiques (Toward a Cognitive Approach to Economic Conventions)', *Revue Economique*, 40, 241–272.
73 J. Lessourne (1992) *The Economics of Order and Disorder* (Oxford: Clarendon Press).
74 Graham, 'Herding among Investment Newsletters'.
75 David S. Scharfstein and Jeremy C. Stein (1990) 'Herd Behavior and Investment,' *American Economic Review,* 80(3), 465–479; Graham, 'Herding among Investment Newsletters'; I. Welch (2000) 'Herding among Security Analysts', *Journal of Financial Economics*, 58(3), 369–396.
76 Scharfstein and Stein, 'Herd Behavior and Investment'.
77 Welch, 'Sequential Sales, Learning and Cascades'.
78 Bikhchandani, Hirshleifer and Welch, 'A Theory of Fads'.
79 Trueman, 'Analyst Forecasts and Herding Behavior'.
80 Zwiebel, 'Corporate Conservatism and Relative Compensation'.
81 E. Maug and N. Naik (1996) 'Herding and Delegated Portfolio Management: The Impact of Relative Performance Evaluation on Asset Allocation', *Working Paper*, London Business School.
82 A. Banerjee (1992) 'A Simple Model of Herd Behavior', *Quarterly Journal of Economics*, 107, 797–817.
83 Welch, 'Sequential Sales, Learning and Cascades'.
84 Bikhchandani, Hirshleifer and Welch, 'A Theory of Fads'.
85 E. N. White (1990) *Crashes and Panics, Lessons from History* (New York: Dow Jones Irving).
86 R. Wermers (1999) 'Mutual Fund Herding and the Impact on Stock Prices', *The Journal of Finance*, 54(2), 581–622.
87 Froot, Scharfstein and Stein, 'Herd on the Street'.
88 E. Falkenstein (1996) 'Preferences for Stock Characteristics as Revealed by Mutual Fund Portfolio Holdings', *Journal of Finance*, 51, 111–135.
89 Froot, Scharfstein and Stein, 'Herd on the Street'.

90 Hirsleifer, Subrahmanyam and Titman, ‘Security Analysis and Trading Patterns’.
91 Scharfstein and Stein, ‘Herd Behavior and Investment’.
92 J. M. Keynes (1936) *The General Theory of Employment, Interest, and Money* (London: Macmillan), p. 79.
93 Scharfstein and Stein, ‘Herd Behavior and Investment’.
94 Graham, ‘Herding among Investment Newsletters’.
95 R. Michaely and K. L. Womack (1999) ‘Conflict of Interest and the Credibility of Underwriter Analyst Recommendations’, *Review of Financial Studies*, 12 Special, 653–686.
96 Bikhchandani, Hirshleifer and Welch, ‘A Theory of Fads’.
97 Graham, ‘Herding among Investment Newsletters’.
98 R. Sias and L. Starks (1997) ‘Return Autocorrelation and Institutional Investors’, *Journal of Financial Economics*, 46, 103–131.
99 R. C. Klemkosky (1977) ‘The Impact and Efficiency of Institutional Net Trading Imbalances’, *The Journal of Finance*, 32(1), 79– 86.
100 S. J. Grossman and J. E. Stiglitz (1976) ‘Information and Competitive Price Systems’, *American Economic Review*, 66, 246–254.
101 G. Genotte and H. Leland (1990) ‘Market Liquidity, Hedging and Crashes’, *American Economic Review*, 80, 999–1021.
102 M. J. Brennan and E. S. Schwartz (1989) ‘Portfolio Insurance and Financial Market Equilibrium’, *Journal of Business*, 62(4), 455–472.
103 R. Roll (1989) ‘Price Volatility, International Market Links, and Their Implications for Regulatory Policies’, *Journal of Financial Services Research*, 3, 211–246.
104 Genotte and Leland, ‘Market Liquidity, Hedging and Crashes’.
105 S. Burness, R. Cummings and J. Quirk (1980) ‘Speculative Behavior and the Operation of Competitive Markets Under Uncertainty: A Survey Article’, *Staff Paper* 80-11, Department of Economics, Montana State University, Bozeman, 1980 as cited in the book *Speculative behavior and the Operation of Competitive Markets Under Uncertainty*, by Michael Antony Stephen Guth. (England: Avebury Publishing).
106 C. P. Kindleberger (1989) *Manias, Panics, and Crises: A History of Financial Crisis*, Rev. ed. (New York: Basic Books).
107 D. Friedman and M. Aoki (1992) ‘Inefficient Information Aggregation as a Source of Asset Price Bubbles’, *Bulleting of Economic Research*, 44, 251–279.
108 J. L. Wootton (1998) *Regional Property Market in 1998* (Hong Kong: Jones Lang Wootton), p. 32.
109 J. B. De Long, A. Shleifer, L. H. Summers and R. J. Waldmann (1990) ‘Positive Feedback Investment Strategies and Destabilizing Rational Speculation’, *Journal of Finance*, 45, 379–395;
J. B. De Long, A. Shleifer, L. H. Summers and R. J. Waldmann (1990) ‘Noise Trader Risk in Financial Markets’, *Journal of Political Economy*, 98, 703–738.
110 M. Youssefmir, B. A. Huberman and T. Hogg (1998) ‘Bubbles and Market Crashes’, *Computational Economics*, 12, 97–114.
111 C. Chiarella (1992) ‘The Dynamics of Speculative Behaviour’, *Annals of Operations Research*, 37(1), 101–123.
112 De Long, Shleifer, Summers and Waldmann, ‘Positive Feedback Investment Strategies’; De Long, Shleifer, Summers and Waldmann, ‘Noise Trader Risk in Financial Markets’.
113 J. B. De Long, A. Shleifer, L. H. Summers and R. J. Waldmann (1991) ‘The Survival of Noise Traders in Financial Markets’, *Journal of Business*, 64, 1–19.

114 O. J. Blanchard and M. W. Watson (1982) 'Bubbles, Rational Expectations, and Financial Markets,' in P. Wachtel, ed., *Crises in the Economic and Financial Structure* (Lexington, MA: Lexington Books), 295–316.
115 J. Leach (1991) 'Rational Speculation', *Journal of Political Economy*, 99, 131–144.
116 J. Wang (1993) 'A Model of Intertemporal Asset Prices Under Asymmetric Information', *Review of Economic Studies*, 60, 249–282.
117 Genotte and Leland, 'Market Liquidity, Hedging and Crashes'.
118 P. Krugman and M. Miller (1993) 'Why have a Target Zone?', *Carnegie-Rochester Conference Series on Public Policy*, 38, 279–314.
119 B. O'Flaherty (1987) 'Guessing Better than the Crows How the Crowd will Behave: Conventions, Beauty Contestants and Bubbles', *Discussion Paper* 347, Department of Economics, Columbia University, New York.
120 M. J. Fishman and K. M. Hagerty (1992) 'Insider Trading and the Efficiency of Stock Prices', *Rand Journal of Economics*, 23, 106–122.
121 De Long, Shleifer, Summers and Waldmann, 'Positive Feedback Investment Strategies'.
122 Scharfstein and Stein, 'Herd Behavior and Investment'.
123 R. Gallant, P. Rossi and G. Tauchen (1992) 'Stock Prices and Volume', *Review of Financial Studies*, 5, 199–242.
124 J. Y. Campbell, S. J. Grossman and J. Wang (1993) 'Trading Volume and Serial Correlation in Stock Returns', *Quantitative Journal of Economics*, 108, 905–939.
125 C. Hiemstra and J. Jones (1994) 'Testing for Linear and Nonlinear Granger Causality in the Stock Price-Volume Relation', *Journal of Finance*, 49, 1639–1664.
126 E. Baek and W. Brock (1992) 'A Nonparametric Test for Independence of a Multivariate Time Series', *Statistica Sinica*, 2, 137–156.
127 R. Jennings, L. Starks and J. Fellingham (1981) 'An Equilibrium Model of Asset Trading with Sequential Information Arrival', *Journal of Finance*, 36, 143–161.
128 P. Silvapulle and J. S. Choi (1999) 'Testing for Linear and Nonlinear Granger Causality in the Stock Price-Volume Relation: Korean Evidence', *Quarterly Review of Economics and Finance*, 39, 59–76.
129 J. Lakonishok, A. Shleifer, R. Thaler and R. Vishny (1991) 'Window Dressing by Pension Fund Managers', *American Economic Review*, 81, 226–231.
130 M. Grinblatt, S. Titman and R. Wermers (1995) 'Momentum Investment Strategies, Portfolio Performance and Herding: A Study of Mutual Fund Behavior', *American Economic Review*, 85, 1088–1105
131 Falkenstein, 'Preferences for Stock Characteristics'.
132 Wermers, 'Mutual Fund Herding and the Impact on Stock Prices'.
133 W. Weidlich and G. Haag (1983) *Concepts and Models of a Quantitative Sociology* (New York: Springer-Verlag).
134 Kindleberger, *Manias, Panics, and Crises*.
135 Topol, 'Bubbles and Volatility of Stock Prices'.
136 Kirman, 'Ants, Rationality and Recruitment', 136–156.
137 T. Lux (1995) 'Herd Behaviour, Bubbles and Crashes', *Economic Journal: The Journal of the Royal Economic Society*, 105, 881–896.
138 Shiller, 'Stock Prices and Social Dynamics'.
139 Kindleberger, *Manias, Panics, and Crises*.
140 Lux, 'Herd Behaviour, Bubbles and Crashes'.
141 Topol, 'Bubbles and Volatility of Stock Prices'.
142 Kirman, 'Ants, Rationality and Recruitment'.

143 R. J. Shiller and J. Pound (1986) 'Survey Evidence on Diffusion of Investment among Institutional Investors', *NBER Working Papers 1851*, National Bureau of Economic Research.
144 Banerjee, 'A Simple Model of Herd Behavior'.
145 Devenow and Welch, 'Rational Herding in Financial Economics', 603–615.
146 E. Miller (1977) 'Risk, Uncertainty, and Divergence of Opinion', *Journal of Finance*, 32, 1151–1168.
147 J. A. Scheinkman and W. Xion (2003) 'Overconfidence and Speculative Bubbles', *Journal of Political Economy*, 111, 1183–1219.
148 J. M. Harrison and D. M. Kreps (1978) 'Speculative Investor Behavior in a Stock Market with Heterogeneous Expectations', *Quarterly Journal of Economics*, 92, 323–336.
149 J. Chen, H. Hong and J. Stein (2001) 'Breadth of Ownership and Stock Returns', *Journal of Financial Economics*, 66, 171–206.
150 K. Diether, C. Malloy and A. Scherbina (2002) 'Differences of Opinion and the Cross-Section of Stock Returns', *Journal of Finance*, 57, 2113–2141.
151 N. Jegadeesh and S. Titman (1993) 'Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency', *Journal of Finance*, 48, 65–91.
152 M. Grinblatt, S. Titman and R. Wermers (1995) 'Momentum Investment Strategies, Portfolio Performance and Herding: A Study of Mutual Fund Behavior', *American Economic Review*, 85, 1088–1105.
153 B. Barber, T. Odean and N. Zhu (2009) 'Systematic Noise', *Journal of Financial Markets*, 12, 547–569.
154 H. Hong and J. C. Stein (1999) 'A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets', *Journal of Finance*, 54, 2143–2184.
155 DeLong, Shleifer, Summers and Waldmann, 'Positive Feedback Investment Strategies'.
156 R. J. Shiller (2002) 'Bubbles, Human Judgment, and Expert Opinion', *Financial Analysts Journal*, 58, 18–26.
157 U. Bhattacharya, N. Galpin, R. Ray and X. Yu (2009) 'The Role of the Media in the Internet IPO Bubble', *Journal of Financial and Quantitative Analysis*, 44(3), 657–682.
158 K. D. Daniel, D. Hirshleifer and A. Subrahmanyam (1998) 'Investor Psychology and Security Market Under- and Over-Reactions', *Journal of Finance*, 53, 1839–1885.
159 E. Asem and G. Y. Tian (2010) 'Market Dynamics and Momentum Returns', *Journal of Financial and Quantitative Analysis*, 45, 1549–1562.
160 N. Barberis, A. Shleifer and R. Vishny (1998) 'A Model of Investor Sentiment', *Journal of Financial Economics*, 94, 307–343.
161 P. Klibanoff, O. Lamont and T. A. Wizman (1998) 'Investor Reaction to Salient News in Closed-End Country Funds', *Journal of Finance*, 2, 673–699.
162 V. L. Smith, G. L. Suchanek and A. W. Williams (1988) 'Bubbles, Crashes and Endogenous Expectations in Experimental Spot Asset Markets', *Econometrica*, 56, 1119–1151.
163 M. Dufwenberg, T. Lindqvist and E. Moore (2005) 'Bubbles and Experience: An Experiment', *American Economic Review*, 95, 1731–1737.
164 L. F. Ackert, N. Charupat, B. K. Church and R. Deaves (2002) 'Bubbles in Experimental Asset Markets: Irrational Exuberance No More', Working Paper 2002–24, Federal Reserve Bank of Atlanta.
165 E. Naruvy and C. N. Noussair (2006) 'The Effect of Short Selling on Bubbles and Crashes in Experimental Spot Asset Markets', *Journal of Finance*, 61, 1119–1157.
166 Scharfstein and Stein (1990) 'Herd Behavior and Investment'.

167 H. Hong, T. Lim and J. C. Stein (2000) 'Bad News Travels Slowly: Size, Analyst Coverage, and the Profitability of Momentum Strategies', *The Journal of Finance*, 55(1), 265–295.
168 I. Welch (2000) 'Views of Financial Economists on the Equity Premium and on Professional Controversies', *The Journal of Business*, 73(4), 501–537.
169 Shiller, 'Bubbles, Human Judgment, and Expert Opinion'.
170 O. Lamont and A. Frazzini (2008) 'Dumb Money: Mutual Fund Flows and the Cross Section of Stock Returns', *Journal of Financial Economics*, 88, 299–322.
171 P. M. DeMarzo, R. Kaniel and I. Kremer (2008) 'Relative Wealth Concerns and Financial Bubbles', *Review of Financial Studies*, 21, 19–50.
172 Graham, 'Herding among Investment Newsletters'.
173 F. Allen and G. Gorton (1993) 'Churning Bubbles', *Review of Economic Studies*, 60, 813–836.
174 Allen and Gale, 'Bubbles and Crises'.

7. Bubbles

1 G. McQueen and S. Thorley (1994) 'Bubbles, Stock Returns, and Duration Dependence', *Journal of Financial & Quantitative Analysis*, 29, 379–401.
2 F. X. Russell (1988) 'The Bubble Bandwagon', *InterMarket*, 5, 51–52.
3 S. Salant and D. Henderson (1978) 'Market Anticipation of Government Policy and the Price of Gold', *Journal of Political Economy*, 86, 627–648.
4 O. Blanchard (1979) 'Speculative Bubbles, Crashes and Rational Expectations', *Economic Letters*, 3, 387–389.
5 M. Funke, S. Hall, M. Sola, W. Charemza and K. West (1994) 'Rational Bubbles During Poland's Hyperinflation: Implications and Empirical Evidence; and Comments', *European Economic Review*, 38, 1257–1285.
6 J. Werner (1997) 'Arbitrage, Bubbles, and Valuation', *International Economic Review*, 38, 453–464.
7 J. Tirole (1985) 'Asset Bubbles and Overlapping Generations', *Econometrica*, 53(6), 1071–1100.
8 R. Flood and P. M. Garber (1980) 'Market Fundamentals versus Price-Level Bubbles: The First Tests', *Journal of Political Economy*, 88, 745–770.
9 J. D. Hamilton and C. H. Whiteman (1985) 'The Observable Implications of Self-Fulfilling Expectations', *Journal of Monetary Economics*, 16, 353–373.
10 McQueen and Thorley, 'Bubbles, Stock Returns, and Duration Dependence'.
11 R. P. Flood and P. M. Garber (1994) *Speculative Bubbles, Speculative Attacks, and Policy Switching* (Cambridge and London: MIT Press).
12 J. K. Galbraith (1955) *The Great Crash* (New York: Mariner Books, Reprint ed., 10 September 2009).
13 D. Scharfstein and K. Stein (1990) 'Herd Behavior and Investment', *American Economic Review*, 80, 465–479.
14 K. Froot and M. Obstfeld (1991) 'Intrinsic Bubbles: The Case of Stock Prices', *American Economic Review*, 81, 1189–1214.
15 G. W. Evans (1991) 'Pitfalls in Testing for Explosive Bubbles in Asset Prices', *American Economic Review*, 81, 922–930.
16 Hamilton and Whiteman, 'The Observable Implications of Self-Fulfilling Expectations'.
17 Flood and Garber, 'Market Fundamentals versus Price-Level Bubbles'.
18 K. Sakakibara (1986) 'A Fisher-Brown Proposition for Speculative Bubbles', *Journal of Macroeconomics*, 8, 485–490.

19 Froot and Obstfeld, 'Intrinsic Bubbles'.
20 N. F. Brady (1988) 'Presidential Task Force on Market Mechanisms: Report of the Presidential Task Force on Market Mechanisms', USA Government Printing Office, Washington, DC. Other titles: Brady Commission Report.
21 Sakakibara, 'A Fisher-Brown Proposition for Speculative Bubbles'.
22 C. Gilles (1989) 'Charges as Equilibrium Prices and Asset Bubbles', *Journal of Mathematical Economics*, 18, 155–167.
23 G. Grossman and N. Yanagawa (1993) 'Asset Bubbles and Endogenous Growth', *Journal of Monetary Economics*, 31, 3–19.
24 Gilles, 'Charges as Equilibrium Prices and Asset Bubbles'.
25 Froot and Obstfeld, 'Intrinsic Bubbles'.
26 J. M. Harrison and D. M. Kreps (1978) 'Speculative Investor Behavior in a Stock Market with Heterogeneous Expectations', *Quarterly Journal of Economics*, 92(2), 323–336.
27 Blanchard, 'Speculative Bubbles, Crashes and Rational Expectations'; O. J. Blanchard and M. W. Watson (1982) 'Bubbles, Rational Expectations, and Financial Markets', in P. Wachtel, ed., *Crises in the Economic and Financial Structure* (Lexington, MA: Lexington Books), pp. 295–315; Flood and Garber, 'Market Fundamentals versus Price-Level Bubbles'.
28 J. Tirole (1982) 'On the Possibility of Speculation under Rational Expectations', *Econometrica*, 50, 1163–1181.
29 F. Allen, S. Morris and A. Postlewaite (1993) 'Finite Bubbles with Short Sale Constraints and Asymmetric Information', *Journal of Economic Theory*, 61, 206–229.
30 Tirole, 'On the Possibility of Speculation under Rational Expectations'.
31 Blanchard and Watson, 'Bubbles, Rational Expectations, and Financial Markets'.
32 Froot and Obstfeld, 'Intrinsic Bubbles'.
33 Tirole, 'Asset Bubbles and Overlapping Generations'.
34 Tirole, 'On the Possibility of Speculation under Rational Expectations'.
35 L. B. Alloy and L. Y. Abramson (1979) 'Judgment of Contingency in Depressed and Non-Depressed Students: Sadder but Wiser?', *Journal of Experimental Psychology: General*, 108, 441–485.
36 G. A. Akerlof and W. T. Dickens (1982) 'The Economic Consequences of Cognitive Dissonance', *American Economic Review*, 72, 307–319.
37 R. Aumann (1976) 'Agreeing to Disagree', *The Annals of Statistics*, 4, 1236–1239.
38 J. M. Keynes (1936/1964) *The General Theory of Employment Interest, and Money* (London: Harcourt Brace, originally published 1936), pp. 155–156.
39 D. Cass and K. Shell (1983) 'Do Sunspots Matter?', *Journal of Political Economy*, 91(21), 193–228.
40 W. S. Jevons (1878) 'Commercial Crises and Sun-Spots', *Nature*, xix, 33–37.
41 Cass and Shell, 'Do Sunspots Matter?'.
42 J.-M. Grandmont and P. Malgrange (1986) 'Nonlinear Economic Dynamics: Introduction', *Journal of Economic Theory*, 40, 3–12.
43 C. Azariadis and R. Guesnerie (1986) 'Sunspots and Cycles', *The Review of Economic Studies*, 53(5), 725–737.
44 Cass and Shell, 'Do Sunspots Matter?'.
45 D. Diamond and P. Dybvig (1983) 'Bank Runs, Deposit Insurance, and Liquidity', *Journal of Political Economy*, 91, 401–419.
46 C. Gu (2006) 'Asymmetric Information and Bank Runs', *Mimeograph*, WP 10-05, Cornell University; (2011) published as 'Noisy Sunspots and Bank Runs', in *Macroeconomic Dynamics*.

47 J. M. Grandmont (1986) 'Stabilizing Competitive Business Cycles', *Journal of Economic Theory*, 40, 57–76.
48 M. Woodford (1986) 'Stationary Sunspot Equilibria in a Finance Constrained Economy', *Journal of Economic Theory*, 40, 128–137.
49 T. Keister (1998) 'Money Taxes and Efficiency When Sunspots Matter', *Journal of Economic Theory*, 83, 43–68.
50 W. F. M. De Bondt and R. Thaler (1989) 'Anomalies: A Mean-Reverting Walk Down Wall Street', *Journal of Economic Perspectives*, 3(1), 189–202.
51 S. Schachter, W. Gerin, D. C. Hood and P. Andreassen (1985) 'Was the South Sea Bubble a Random Walk?', *Journal of Economic Behavior and Organization*, 6, 323–329.
52 K. Arrow (1982) 'Risk Perception in Psychology and Economics', *Economic Inquiry*, 20, l–9.
53 C. Camerer (1989) 'Bubbles and Fads in Asset Prices', *Journal of Economic Surveys*, 3(1), 4–41.
54 B. Friedman (1984) 'Stock Prices and Social Dynamics: Comments and Discussion', *Brooking Papers on Economic Activity*, 2, 504–508.
55 K. D. West and A. W. Kleidon (1988) 'Bubbles, Fads and Stock Price Volatility Tests: A Partial Evaluation; Discussion', *Journal of Finance*, 43, 639–660.
56 R. W. Michener (1982) 'Variance Bounds in a Simple Model of Asset Pricing', *Journal of Political Economy*, 90, 166–175.
57 M. Flavin (1983) 'Excess Volatility in the Financial Markets: A Reassessment of the Empirical Evidence', *Journal of Political Economy*, 91, 929–956.
58 T. A. Marsh and R. C. Merton (1986) 'Dividend Variability and Variance Bounds Test for Rationality of Stock Market Prices', *American Economic Review*, 76, 483–498.
59 R. J. Shiller (1981) 'The Use of Volatility Measures in Assessing Market Efficiency', *Journal of Finance*, 36, 291–304.
60 J. Mattey and R. Meese (1986) 'Empirical Assessment of Present Value Relations', (including discussion), *Econometric Reviews*, 5, 17l–234.
61 R. C. Merton (1987) 'On the Current State of the Stock Market Rationality Hypothesis', in S. Fischer et al., eds, *Macroeconomics and Finance: Essays in Honor of Franco Modigliani* (Cambridge: MIT Press), 93–124
62 R. J. Shiller (1986b) 'The Marsh-Merton Model of Managers' Smoothing of Dividends', *American Economic Review*, 76, 499–503.
63 Marsh and Merton, 'Dividend Variability and Variance Bounds Test'.
64 R. A. Meese (1986) 'Testing for Bubbles in Exchange Markets: A Case of Sparkling Rates?', *Journal of Political Economy*, 94, 345–373.
65 A. W. Kleidon (1986) 'Anomalies in Financial Economics: Blueprint for Change?', *Journal of Business*, 59, S469–S499.
66 N. G. Mankiw, D. Romer and M. D. Shapiro (1985) 'An Unbiased Reexamination of Stock Market Volatility', *Journal of Finance*, 40, 677–687.
67 Shiller, 'The Use of Volatility Measures in Assessing Market Efficiency'.
68 J. Y. Campbell and R. J. Shiller (1987) 'Cointegration and Tests of Present Value Models', *Journal of Political Economy*, 95, 1062–1088.
69 D. W. Diamond and R. E. Verrecchia (1981) 'Information Aggregation in a Noisy Rational Expectations Economy', *Journal of Financial Economics*, 9, 221–235.
70 S. J. Grossman and J. E. Stiglitz (1980) 'On the Impossibility of Informationally Efficient Markets', *American Economic Review*, 70, 393–408.
71 P. Dubey, J. Geanakoplos and M. Shubik (1987) 'The Revelation of Information in Strategic Market Games: A Critique of Rational Expectations Equilibrium', *Journal of Mathematical Economics*, 16, 105–137.

72 D. Friedman and M. Aoki (1986) 'Asset Price Bubbles from Poorly Aggregated Information: A Parametric Example', *Economics Letters*, 21, 49–52.
73 C. F. Camerer and K. Weigelt (1991) 'Information Mirages in Experimental Asset Markets', *Journal of Business*, 64, 463–493.
74 G. Huberman and G. W. Schwert (1985) 'Information Aggregation Inflation, and the Pricing of Indexed Bonds', *Journal of Political Economy*, 93, 92–114.
75 K. R. French and R. Roll (1986) 'Stock Return Variances: The Arrival of Information and the Reaction of Traders', *Journal of Financial Economics*, l7, 5–26.
76 Meese, 'Testing for Bubbles in Exchange Markets'.
77 French and Roll, 'Stock Return Variances'.
78 R. Roll (1988) 'R^2', *Journal of Finance*, 43, 541–566.
79 J. Y. Campbell and A. S. Kyle (1993) 'Smart Money, Noise Trading and Stock Price Behavior', *Review of Economic Studies*, 60(1), 1–34.
80 French and Roll, 'Stock Return Variances'.
81 J. B. De Long, A. Shleifer, L. H. Summers and R. J. Waldmann (1987) 'The Economic Consequences of Noise Traders', *Harvard Institute of Economic Research*, Discussion Paper 1348.
82 V. L. Smith, G. L. Suchanek and A. W. Williams (1988) 'Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Asset Markets', *Econometrica*, 56, 1119–1151.
83 M. Dufwenberg, T. Lindqvist and E. Moore (2005) 'Bubbles and Experience: An Experiment', *American Economic Review*, 95, 1731–1737.
84 L. F. Ackert, N. Charupat B. K. Church and R. Deaves (2002) 'Bubbles in Experimental Asset Markets: Irrational Exuberance No More', *Working Paper* 24. https://www.frbatlanta.org/research/publications/wp/2002/02_24.aspx
85 E. Naruvy and C. N. Noussair (2006) 'The Effect of Short Selling on Bubbles and Crashes in Experimental Spot Asset Markets', *Journal of Finance*, 61, 1119–1157.
86 B. Lev (1989) 'On the Usefulness of Earnings and Earnings Research: Lessons and Directions from Two Decades of Empirical Research', *Journal of Accounting Research*, Supplement 27, 153–201.
87 Roll, 'R^2'.
88 P. D. Easton, T. S. Harris and J. A. Ohlson (1992) 'Aggregate Accounting Earnings can Explain Most of Security Returns. The Case of Long Event Windows', *Journal of Accounting and Economics*, 15, 119–142.
89 R. Frankel and C.M.C. Lee (1996) 'Accounting Valuation, Market Expectation, and the Book-to-Market Effect', *Journal of Finance*, 51, (July) 1040–1041.
90 R. Harris (1999) 'The Accuracy, Bias and Efficiency of Analysts' Long-run Earnings Growth Forecasts', *Journal of Business Finance and Accounting*, 26 (June/July 1999), 725–755.
91 Flood and Garber, 'Market Fundamentals versus Price-Level Bubbles'.
92 Hamilton and Whiteman, 'The Observable Implications of Self-Fulfilling Expectations'.
93 McQueen and Thorley, 'Bubbles, Stock Returns, and Duration Dependence'.
94 Tirole, 'Asset Bubbles and Overlapping Generations'.
95 S. A. O'Connell and S. P. Zeldes (1988) 'Rational Ponzi Games', *International Economic Review*, 29, 431–450.
96 B. T. Diba (1990) 'Bubbles and Stock-Price Volatility', *Proceedings of the Thirteenth Annual Economic Policy Conference of the Federal Reserve Bank of St. Louis, G. P. Dwyer; R. W. Hafer, eds, Boston/Dordrecht/London,* 9–29.
97 B. T. Diba and H. I. Grossman (1987) 'On the Inception of Rational Bubbles', *Quarterly Journal of Economics*, 102, 697–700.

98 B. T. Diba and H. I. Grossman (1988) 'The Theory of Rational Bubbles in Stock Prices', *Economic Journal*, 98, 746–754.
99 G. W. Evans and S. Honkapohja (1992) 'On the Robustness of Bubbles in Linear RE Models', *International Economic Review*, 33, 1–14.
100 Allen, Morris and Postlewaite, 'Finite Bubbles with Short Sale Constraints'.
101 J. B. Taylor (1977) 'On the Conditions for Uniqueness in the Solution of Rational Expectations Models', *Econometrica*, 45, 1377–1385.
102 R. J. Shiller (1978) 'Rational Expectations and the Dynamic Structure of Macroeconomic Models', *Journal of Monetary Economic*, 4, 1–44.
103 Blanchard, 'Speculative Bubbles, Crashes and Rational Expectations'.
104 Blanchard and Watson, 'Bubbles, Rational Expectations, and Financial Markets'.
105 D. Quah (1985) 'Estimation of a Nonfundamentals Model for Stock Price and Dividend Dynamics', manuscript MIT.
106 Tirole, 'Asset Bubbles and Overlapping Generations'.
107 O'Connell and Zeldes, 'Rational Ponzi Games'.
108 Diba, 'Bubbles and Stock-Price Volatility'.
109 Shiller, 'The Use of Volatility Measures in Assessing Market Efficiency', 291–311.
110 R. P. Flood and R. J. Hodrick (1986) 'Asset Price Volatility, Bubbles and Process Switching', *Journal of Finance*, 41, 831–842.
111 J. H. Cochrane (1991) 'Volatility Tests and Efficient Markets: A Review Essay', *Journal of Monetary Economics*, 27, 463–485.
112 S. Ikeda and A. Shibata (1992) 'Fundamentals-Dependent Bubbles in Stock Prices', *Journal of Monetary Economics*, 30, 143–168.
113 B. T. Diba and H. I. Grossman (1988) 'Explosive Rational Bubbles in Stock Prices?', *American Economic Review*, 78, 520–530.
114 Campbell and Shiller, 'Cointegration and Tests of Present Value Models'.
115 Diba and Grossman, 'Explosive Rational Bubbles in Stock Prices?'.
116 W. Charemza and D. Deadman (1995) 'Rational and Intrinsic Bubbles: A Reinterpretation of Empirical Results', *Applied Financial Economics*, 5, 199–202.
117 F. Allen and G. Gorton (1988) 'Rational Finite Bubbles', *Mimeograph*, The Wharton School, University of Pennsylvania. Published: F. Allen and G. Gorton (1993) 'Churning Bubbles', *Review of Economic Studies*, 60, 4(205), 813–836. http://www.nber.org/papers/w3707
118 H. L. Han (1996) 'Cointegration and Tests of a Present Value Model in the Stock Market', *Applied Economics*, 28, 267–272.
119 D. Friedman and M. Aoki (1992) 'Inefficient Information Aggregation as a Source of Asset Price Bubbles', *Bulleting of Economic Research*, 44, 251–279.
120 Campbell and Shiller, 'Cointegration and Tests of Present Value Models'.
121 Diba and Grossman, 'Explosive Rational Bubbles in Stock Prices?'.
122 Evans, 'Pitfalls in Testing for Explosive Bubbles in Asset Prices'.
123 K. D. West (1987) 'A Specification Test for Speculative Bubbles', *Quarterly Journal of Economics*, 102, 553–580.
124 R. P. Flood, R. T. Hodrick and P. Kaplan (1986) 'An Evaluation of Recent Evidence on Stock Market Bubbles', *National Bureau of Economic Research*, Working Paper 1971, Cambridge, MA. http://www.nber.org/papers/w1971.pdf
125 West, 'A Specification Test for Speculative Bubbles'.
126 E. Renshaw (1990) 'Some Evidence in Support of Stock Market Bubbles', *Financial Analysts Journal*, 46, 71–73.
127 R. Chirinko and H. Schaller (1996) 'Bubbles, Fundamentals, and Investment: A Multiple Equation Testing Strategy', *Journal of Monetary Economics*, 38, 47–76.

128 West, 'A Specification Test for Speculative Bubbles'.
129 McQueen and Thorley, 'Bubbles, Stock Returns, and Duration Dependence'.
130 Blanchard and Watson, 'Bubbles, Rational Expectations, and Financial Markets'.
131 G. W. Evans (1986) 'A Test for Speculative Bubbles in the Sterling-Dollar Exchange Rate: 1981–84', *American Economic Review*, 76, 621–636.
132 H. Bierman (1995) 'Bubbles, Theory, and Market Timing', *Journal of Portfolio Management*, 22, 54–56.
133 Flood and Garber, 'Market Fundamentals versus Price-Level Bubbles'.
134 A. Shleifer and R. Vishny (1997) 'The Limits of Arbitrage', *Journal of Finance*, 52, 35–55; D. Abreu and M. K. Brunnermeier (2003) 'Bubbles and Crashes', *Econometrica*, 71(1), 173–204.
135 Diba and Grossman, 'On the Inception of Rational Bubbles'.
136 B. S. Bernanke, M. Gertler and S. Gilchrist (1999) 'The Financial Accelerator in a Quantitative Business Cycle Framework', in J. B. Taylor and M. Woodford, eds, *Handbook of Macroeconomics* (Elsevier Science B.V.), Chapter 21, pp. 1341–1393.
137 H. Hong, J. Scheinkman and W. Xiong (2008) 'Advisors and Asset Prices: A Model of the Origins of Bubbles', *Journal of Financial Economics*, 89, 268–287.
138 M. K. Brunnermeier and C. Julliard (2008) 'Money Illusion and Housing Frenzies', *Review of Financial Studies*, 21(1), 135–180.
139 G. L. Kaminsky, C. M. Reinhart and C. A. Végh (2003) 'The Unholy Trinity of Financial Contagion', *Journal of Economic Perspectives*, 17(4), 51–74.
140 K. Daniel, D. Hirshleifer and A. Subrahmanyam (1998) 'Investor Psychology and Security Market Under- and Overreactions', *Journal of Finance*, 53, 1839–1885.
141 Bernanke, Gertler and Gilchrist, 'The Financial Accelerator'.
142 A. Scherbina (2008) 'Suppressed Negative Information and Future Underperformance', *Review of Finance*, 12, 533–565.
143 W. Xiong and J. Yu (2011) 'The Chinese Warrants Bubble', *American Economic Review*, 101, 2723–2753.
144 E. Lefèvre (1923) *Reminiscences of a Stock Operator*, First published January 17th by Doubleday & Company E. Published 2006 by John Wiley & Sons, p. 180.
145 M. Jensen and W. Meckling (1976) 'Theory of the Firm: Managerial Behaviour, Agency Cost and Ownership Structure', *Journal of Financial Economics*, 3, 305–360; J. Stiglitz and A. Weiss (1981) 'Credit Rationing in Markets with Imperfect Information', *American Economic Review*, 71, 393–410.
146 J. B. De Long, A. Shleifer, L. H. Summers and R. J. Waldmann (1990) 'Positive Feedback Investment Strategies and Destabilizing Rational Speculation', *Journal of Finance*, 45, 379–395.
147 Abreu and Brunnermeier, 'Bubbles and Crashes'.
148 M. K. Brunnermeier and S. Nagel (2004) 'Hedge Funds and the Technology Bubble', *Journal of Finance*, 59, 2007–2034.
149 Shleifer and Vishny, 'The Limits of Arbitrage'; W. Xiong (2001) 'Convergence Trading with Wealth Effects: An Amplification Mechanism in Financial Markets', *Journal of Financial Economics*, 62, 247–292; D. Gromb and D. Vayanos (2002) 'Equilibrium and Welfare in Markets with Financially Constrained Arbitrageurs', *Journal of Financial Economics*, 66, 361–407.
150 R. Sadka and A. Scherbina (2010) 'Mispricing and Costly Arbitrage', *Journal of Investment Management*, 7, 1–13.
151 De Long, Shleifer, Summers and Waldmann, 'Positive Feedback Investment Strategies'.
152 R. McKinnon and H. Pill (1998) 'International Over-Borrowing: A Decomposition of Credit and Currency Risks', *SCID Working Paper* 19, Stanford University. https://siepr.stanford.edu/publicationsprofile/1611

153 F. Allen and D. Gale (1998) 'Optimal Financial Crises', *Journal of Finance*, 53, 1245–1284; B. Bernanke (1983) 'Non-Monetary Effects of the Financial Crisis in the Propagation of the Great Depression', *American Economic Review*, 73, 257–276; B. Bernanke and M. Gertler (1989) 'Agency Costs, Collateral, and Business Fluctuations', *American Economic Review*, 79, 14–31; B. Holmstrom and J. Tirole (1997) 'Financial Intermediation, Loanable Funds, and the Real Sector', *Quarterly Journal of Economics*, 112, 663–691.

154 Bernanke and Gertler, 'Agency Costs, Collateral, and Business Fluctuations'; N. Kiyotaki and J. Moore (1997) 'Credit Cycles', *Journal of Political Economy*, 105(2), 211–248; Bernanke, Gertler and Gilchrist, 'The Financial Accelerator', 1341–1393.

155 H. Shin (2010) *Risk and Liquidity*, Clarendon Lecture in Finance (Oxford, UK: Oxford University Press).

156 R. Koo (2008) *The Holy Grail of Macroeconomics: Lessons from Japan's Great Recession* (Singapore: John Wiley & Sons), p. 160.

157 M. K. Brunnermeier and Y. Sannikov (2013) 'A Macroeconomic Model with a Financial Sector', April 17.

158 B. S. Bernanke and M. Gertler (1999) 'Monetary Policy and Asset Price Volatility', *Economic Review*, 84, 17–51.

Author Index

Subject Index