



## UvA-DARE (Digital Academic Repository)

### Complex systems in financial economics: Applications to interbank and stock markets

in 't Veld, D.L.

**Publication date**

2014

**Document Version**

Final published version

[Link to publication](#)

**Citation for published version (APA):**

in 't Veld, D. L. (2014). *Complex systems in financial economics: Applications to interbank and stock markets*. [Thesis, fully internal, Universiteit van Amsterdam].

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.



Daan in 't Veld

# Complex Systems in Financial Economics

Applications to Interbank  
and Stock Markets

Complex systems are characterised by strong interaction at the micro level that can induce large changes at the macro level. This thesis applies the theory of complex systems to the interbank market (Part I) and the stock market (Part II). Evidence found in data from the Netherlands and the US makes clear in what sense these markets are complex systems. The observed phenomena are explained by modelling the adaptive behaviour of financial agents, for example how they form their trading relationships or how they choose investment strategies. The applications help to understand the mechanisms behind the emergence of the financial-economic crisis in 2007 and 2008, and relate to the debate on policy measures aiming to prevent a future crisis of this kind.

Daan in 't Veld (1987) graduated in econometrics and mathematical economics from the University of Amsterdam. In 2010 he started working on his PhD thesis at the same university, in the Center for Nonlinear Dynamics in Economics and Finance (CeNDEF). His research interests include models of behavioural finance, asset pricing and financial networks.

Complex Systems in Financial Economics  
Daan in 't Veld



# Complex Systems in Financial Economics

Applications to Interbank and Stock Markets

ISBN 978 90 361 0402 9

Cover design: Crasborn Graphic Designers bno, Valkenburg a.d. Geul

This book is no. **592** of the Tinbergen Institute Research Series, established through cooperation between Thela Thesis and the Tinbergen Institute. A list of books which already appeared in the series can be found in the back.

# Complex Systems in Financial Economics

Applications to Interbank and Stock Markets

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor  
aan de Universiteit van Amsterdam  
op gezag van de Rector Magnificus  
prof. dr. D.C. van den Boom  
ten overstaan van een door het college voor promoties  
ingestelde commissie,  
in het openbaar te verdedigen in de Agnietenkapel  
op vrijdag 17 oktober 2014, te 14:00 uur

door

**Daan Laurens in 't Veld**

geboren te Amsterdam

**Promotiecommissie:**

**Promotores:** Prof. dr. C.H. Hommes  
Prof. dr. C.G.H. Diks

**Overige leden:** Prof. dr. H.P. Boswijk  
Prof. dr. P.C. De Grauwe  
Dr. M.J. van der Leij  
Dr. I.P.P. van Lelyveld  
Prof. dr. J. Tuinstra

Faculteit Economie en Bedrijfskunde

Dit onderzoek werd mede mogelijk gemaakt door steun van De Nederlandsche Bank (DNB) en de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) in het programma Complexiteit.

*Opgedragen aan  
Karen Laura Kwast  
(1956–2011)*





# Acknowledgements

During the last four years this thesis has emerged from something that used to be a mere dot on the horizon. Many people have supported me on my metaphorical journey through the economic science. Some of them pointed out new directions when I was lost; others got me back on my feet when I had slipped back; they sometimes inspired me with stories where they had been, or simply joined me in looking ahead.

In the first place I would like to thank my main supervisor Cars Hommes. During our conversations I have always felt the great confidence he placed in me. His suggestions – be it in modelling complex systems, relating mathematical models to actual policies, writing economic papers, preparing for conferences, or any other part of the scientific process – have been extremely valuable for me to this day.

As I moved forward during the years, I increasingly realised how lucky I had been to receive my PhD position. Just at the time I came along NWO and DNB had awarded a grant to Cars together with Cees Diks. I would like to thank Cees for giving me the freedom to find my way in the Complexity project. His scrutiny in reading later versions of my work helped me a lot in finishing the thesis.

I have at least two reasons to thank my former lecturer Jan Tuinstra: for introducing me to mathematical economics and for raising my interest in scientific research. Much of the spirit Jan evoked in me has been maintained in writing this thesis. I am very grateful to Iman van Lelyveld for starting and finishing a research project with me in De Nederlandsche Bank. The same goes for my officemate Marco van der Leij, who shared his knowledge about financial

---

networks and lifted the quality of my thesis to a higher level.

In a later stage I came to speak about my research with Peter Boswijk and Paul De Grauwe, and I am honoured that they agreed to complete my PhD committee. I am much obliged to Paul for his hospitality and the interest he showed during my stay at the London School of Economics.

I thank my fellow PhD students and other colleagues during my time at CeNDEF: Te Bao, Jakob Grazzini, Zhenxing Huang, Dávid Kopányi, Maurice Koster, Michiel van de Leur, Tomasz Makarewicz, Domenico Massaro, Anghel Negriu, Marius Ochea, Valentyn Panchenko, Roald Ramer, Isabelle Salle, Florian Sniekers, Tatiana Kiseleva, Florian Wagener, Juanxi Wang, Marcin Wolski and Paolo Zeppini. Also regular lunch discussions with Mark Dijkstra from ACLE are well remembered. A special thanks to Misha Anufriev who, having moved to Australia halfway my four years, kept a close connection during many meetings throughout Europe.

It is always nice to meet new colleagues in unexpected places. My former chess teammate Marcel Canoy turned out to be a thought-provoking, original economist, and showed me how to be yourself in economics.

Outside the professional sphere I am grateful to my friends for sharing their spare time with me, especially to Pieter, Lennart, Devon, Léon and Steven. I thank my father, brother and all others in my peaceful and happy family for always being there. Finally, I would like to thank my dear Dasha for her love and true devotion. She fills my heart and clears my mind.

Amsterdam, August 2014

# Contents

<b>1</b>	<b>General introduction and outline</b>	<b>1</b>
1.1	The financial crisis of 2007-2008 . . . . .	1
1.2	The shift in views on the financial system . . . . .	4
1.3	An introduction to complex systems . . . . .	6
1.4	Complex systems in economics . . . . .	8
1.5	Purpose and methodology . . . . .	10
1.6	Outline of the thesis . . . . .	11
<b>I</b>	<b>Applying complex systems to the interbank market</b>	<b>15</b>
<b>2</b>	<b>Finding the core: Network structure in the Dutch interbank market</b>	<b>17</b>
2.1	Introduction . . . . .	17
2.2	Network models . . . . .	20
2.2.1	Nested split graphs . . . . .	22
2.2.2	Core periphery networks . . . . .	24
2.2.3	Fitting deterministic network models . . . . .	26
2.3	Data description . . . . .	30
2.4	Estimation results . . . . .	32
2.4.1	Estimating the core . . . . .	34
2.4.2	Monte Carlo simulations . . . . .	36
2.5	Network structure in perspective . . . . .	40
2.6	Conclusions . . . . .	43

---

Appendix 2.A	Robustness analysis using the weighted network . . . . .	45
Appendix 2.B	Visualising errors in the CP model . . . . .	51
Appendix 2.C	Testing the CP against the configuration model . . . . .	53
<b>3</b>	<b>The formation of a core periphery structure in heterogeneous financial networks</b>	<b>55</b>
3.1	Introduction . . . . .	55
3.2	Network formation model . . . . .	59
3.2.1	Payoffs in a financial network with intermediators . . . . .	59
3.2.2	Network stability concepts . . . . .	63
3.3	Basic structures . . . . .	65
3.4	Results for homogeneous traders . . . . .	68
3.4.1	Efficient networks . . . . .	68
3.4.2	Stability of basic structures . . . . .	69
3.4.3	A dynamic process . . . . .	75
3.5	Results for heterogeneous traders . . . . .	81
3.6	Calibrating the model to the Dutch interbank market . . . . .	87
3.7	Conclusion . . . . .	89
Appendix 3.A	Trade surpluses in the interbank market . . . . .	91
Appendix 3.B	Siedlarek’s payoff function . . . . .	92
Appendix 3.C	Proofs . . . . .	95
Appendix 3.D	Longer intermediation chains . . . . .	107
<b>II</b>	<b>Applying complex systems to the stock market</b>	<b>111</b>
<b>4</b>	<b>Booms, busts and behavioural heterogeneity in U.S. stock prices</b>	<b>113</b>
4.1	Introduction . . . . .	113
4.2	Model description . . . . .	118
4.2.1	Fundamental value . . . . .	119

4.2.2	Behavioural heterogeneous beliefs . . . . .	120
4.2.3	Econometric form . . . . .	124
4.3	Estimation results for a simple behavioural model . . . . .	126
4.3.1	Estimating the fundamental value of the Gordon model . . . . .	126
4.3.2	Estimation of the heterogeneous agents model . . . . .	130
4.3.3	A representative agent benchmark model . . . . .	135
4.4	Monte Carlo simulations . . . . .	136
4.4.1	Evaluating the power of the main tests . . . . .	137
4.4.2	What are the economic effects of different behavioural regimes? . . . .	139
4.5	Estimation using a consumption-habit time-varying risk premium . . . . .	142
4.5.1	Estimating the Campbell-Cochrane model . . . . .	143
4.5.2	Estimation of the HAM under the consumption-habit fundamental value	149
4.5.3	Evaluating the tests for the model under a consumption-habit FV . . . .	153
4.6	Conclusion . . . . .	155
Appendix 4.A	Deriving the profit function . . . . .	157
Appendix 4.B	Robustness analysis . . . . .	159
<b>5</b>	<b>Adverse effects of demand constraints in a financial market model with</b>	
	<b>heterogeneous agents</b>	<b>163</b>
5.1	Introduction . . . . .	163
5.2	A model with heterogeneous beliefs and demand constraints . . . . .	167
5.2.1	Individual asset demand . . . . .	167
5.2.2	Market equilibrium . . . . .	171
5.3	Stylised adaptive belief systems . . . . .	173
5.3.1	Fundamentalists versus chartists . . . . .	176
5.3.2	Fundamentalists versus optimists and pessimists . . . . .	186
5.4	Discussion . . . . .	196
Appendix 5.A	Proofs . . . . .	200

## CONTENTS

---

<b>6 Summary</b>	<b>209</b>
<b>Bibliography</b>	<b>213</b>
<b>Краткое содержание (Summary in Russian)</b>	<b>225</b>
<b>Samenvatting (Summary in Dutch)</b>	<b>231</b>

# Chapter 1

## General introduction and outline

The financial economy is full of complex systems. This view is becoming increasingly accepted; see Kirman (2010) and Hommes (2013) for recent influential books on complex economic systems. This thesis applies the theory of complex systems to interbank and stock markets – two markets that were in the midst of the financial-economic crisis of 2007-2008. The questions that this thesis intends to answer originate mainly from the financial crisis and its ongoing ramifications. Let us therefore start with discussing the far-reaching events that have become part of our recent economic history.

### 1.1 The financial crisis of 2007-2008

*“When the music stops, in terms of liquidity, things will be complicated. But as long as the music is playing, you’ve got to get up and dance. We’re still dancing.”* – Charles Prince (see Nakamoto and Wighton, 2007).

Brunnermeier (2009) provides a summary of the main events in 2007 and 2008. At the heart of the financial crisis lies a *bubble in house prices*. The speculative behaviour during the last months of the housing bubble has been characterised as a ‘game of musical chairs’, following an earlier analogy for speculation due to Keynes (1936, p. 156). The quote above is from

Citigroup's director Charles Prince on 10 July 2007, when asked about the ongoing commitment of the bank to issue and trade mortgage products. It suggests that many investors were aware that real estate portfolios were overvalued and that the bubble would be unsustainable in the long run, but nevertheless stayed in the market. Only by October 2007, banks were writing off substantial losses in the mortgage markets.

As such the housing bubble was not extraordinary. Only a few years earlier, with the turning of the millennium, stock prices of technological firms had rocketed and subsequently collapsed without a system-wide financial crisis or effects on the real economy. So how could the housing bubble have such a large effect?

Part of the explanation lies in rather technical developments in the financial sector itself. One heavily criticised component of so-called financial innovation is the *securitisation* of mortgages, in which mortgages of different credit quality, including the infamous 'subprime' loans, are pooled in financial products called CDOs. The reselling of CDOs limited monitoring because of the stretched supply chain of loans. With failing oversight, the housing bubble created growing underestimated risks to banks.

When subprime mortgage losses led to liquidity hoarding and a dry-up of funding, large financial institutions in the U.S. started to experience serious problems. First, Bear Stearns experienced a bank run by hedge funds pulling out, and was bought by JPMorgan Chase in March 2008. On Monday 15 September 2008, Lehman Brothers declared bankruptcy. This would become the single most important event in the financial crisis. One day later the large insurance company AIG had to be bailed out by the Federal Reserve after it declared a serious liquidity shortage.

The narrative so far has uncovered the underlying roots of the financial crisis: a real-estate bubble and speculation; securitisation; liquidity shortages; bank runs and bankruptcies. Nevertheless, the crisis would have been comparatively mild if the initial subprime losses – estimated by the International Monetary Fund (2007, October) at about \$200 billion – could have been contained to the financial institutions directly involved. Instead, after the failure of Lehman



Brothers the financial crisis grew out to global proportions. The question remains: how?

The main theme in this thesis is that the financial system as a whole organises itself in such a way that the relatively minor events could grow out to an *endogenously emerging global crisis*. This kind of emergence is typical for what we will call complex systems, in which small exogenous shocks at the micro level may trigger a large change – referred to as a critical transition – at the macro level. We will consider two types of financial markets through which the financial turmoil spread during 2008.

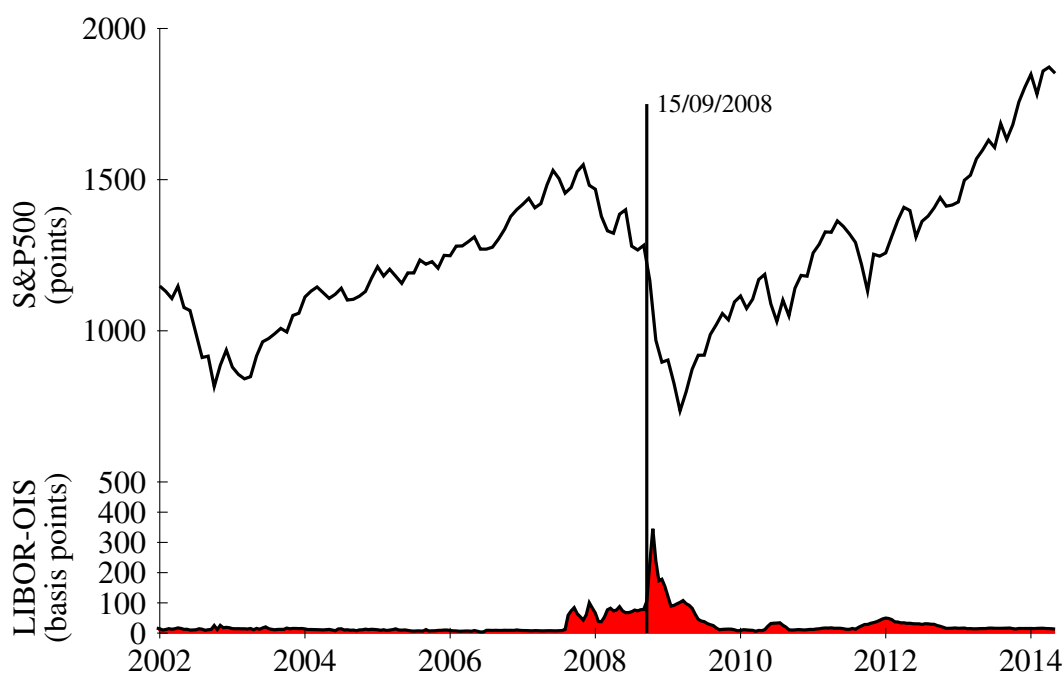


Figure 1.1: The S&P500 stock index and the LIBOR-OIS spread. Sources: Datastream and Bloomberg.

A first channel for amplification of financial instability is the *interbank market*. In such a market banks lend money to each other if, for one reason or another, this is more beneficial than to turn to the central bank. Most of the interbank transactions are ‘over-the-counter’, meaning that only the two banks involved in the loan are aware of its existence and size. At the same time, many financial institutions are both lenders and borrowers in the interbank market.

In this way the interbank market creates a network in which each bank is (unintentionally) connected to almost any other bank in just a few steps. When the news of the subprime losses

had spread in 2007, banks already became more reluctant to lend to each other. See Figure 1.1 for the LIBOR (London Interbank Offered Rate) rate measured in differences from the OIS (overnight indexed swap) rate. The LIBOR-OIS spread measures the risk premium which banks require for counterparty risk in the interbank market. It reached an all-time high of 3.64% on 10 October 2008, worsening the system-wide liquidity crisis.

A second relevant amplification channel is the *stock market*. During 2008 the S&P500 index, the financial bellwether of the U.S. economy, lost almost half of its value. From Figure 1.1, it is clear that the largest drop occurred closely after the Lehman Brother bankruptcy and around the same time of stress in the interbank market.

The total losses to the global economy were even larger. Blanchard (2008) estimates the decline in world stock market valuation from July 2007 to November 2008 at \$26 *trillion*, more than 100 times the initial subprime losses. These numbers suggest how relatively small initial shocks to the mortgage market gradually cascaded through the highly interconnected financial system and strongly amplified into a global financial-economic crisis. The effects on the real economy have been coined the ‘Great Recession’, the largest economic downturn since the Great Depression of the 1930s.

## 1.2 The shift in views on the financial system

The view that the financial system is intrinsically susceptible to instability and crises has been expressed long before the events in recent years. Much of the work of Hyman Minsky is devoted to his ‘financial instability hypothesis’. In even stronger words, Minsky (1986) argued that the fragility of the financial system is one of the fundamental flaws of capitalism, suggesting that strict regulation of financial markets is necessary to stabilise the economy.

Yet despite the warnings of Minsky, the majority of economists used to have an unequivocally positive assessment of the financial system. In February 2004, the future chairman of the Federal Reserve System Ben Bernanke held a well-known speech called ‘The Great Mod-

eration', the period with historically low macroeconomic volatility that had started in 1987 and would abruptly end in 2007. He argued that major reforms in monetary policy had probably been a main stabilising source. Interestingly, he also briefly mentioned the supposedly improved functioning of financial markets – among other examples of structural change – as a possibly contributing factor to the Great Moderation:

*“The increased depth and sophistication of financial markets [...] may have increased macroeconomic flexibility and stability.”* – Bernanke (2004).

By now, it is an understatement to say that the general opinion has changed. Policy makers have been faced with completely unpredictable market transitions and started to develop a different discourse. The same Ben Bernanke – after being appointed chairman of the Fed in 2006 – interpreted the ‘Flash Crash’ of 6 May 2010 as a typical outcome generated by complex financial-economic systems:

*“The brief market plunge was just a small indicator of how complex and chaotic, in the formal sense, these systems have become. **Our financial system is so complicated and so interactive** – so many different markets in different countries and so many sets of rules. What happened in the stock market is just a little example of how things can cascade or how technology can interact with market panic.”* – Ben Bernanke (see Chan, 2010).

Different times call for different approaches. The president of the European Central Bank Jean-Claude Trichet expressed the need for alternative approaches using the famous words:

*“In the face of the crisis, we felt abandoned by conventional tools.”* – Trichet (2010).

Other interesting speeches of central bankers are Haldane (2009), Haldane and Madouros (2012)

and Yellen (2013). Andrew Haldane, Executive Director of Financial Stability at the Bank of England, was one of the first who plotted the *financial network* to capture the complex interactions between financial institutions. Showing the cross-border holdings of financial assets and liabilities of 18 countries in 1985 and 2005, Figure 1.2 illustrates how the financial system has emerged in the 2000s as a dense and highly interconnected network with short paths between large financial centres.

Jackson (2008) and Goyal (2009) are seminal books on modelling economic networks. This literature is bound to grow considerable with the new interest in financial networks and complex financial-economic systems. This thesis proposes a methodology of complex systems for financial economics, which includes the analysis of financial networks, and presents applications to interbank and stock markets.

### **1.3 An introduction to complex systems**

What are complex systems? Weaver (1948) introduces the following metaphor to explain the concept of a complex system. Warren Weaver was a mathematician working with Claude Shannon on information theory, a devoted proponent of raising public interest in science, and an important source of inspiration for the future Nobel Prize Laureate in Economics Friedrich Hayek.

Consider a scientist at a billiard table whose prime interest is in analysing the motion of billiard balls, and not in winning the game. If there is a single ball on the table, the position of the ball after a certain time can be analysed quite accurately by *classical mechanics*, also known as Newtonian mechanics. Collisions between multiple balls make the problem considerably harder, but it is still possible to calculate the motion of two or even three balls.

A billiard table with a small number of balls is an example of a *simple* system for which straightforward yet accurate laws can be developed. However, as soon there are ten or fifteen balls on the table, such as in an actual game of pool, the problem becomes unmanageable. It

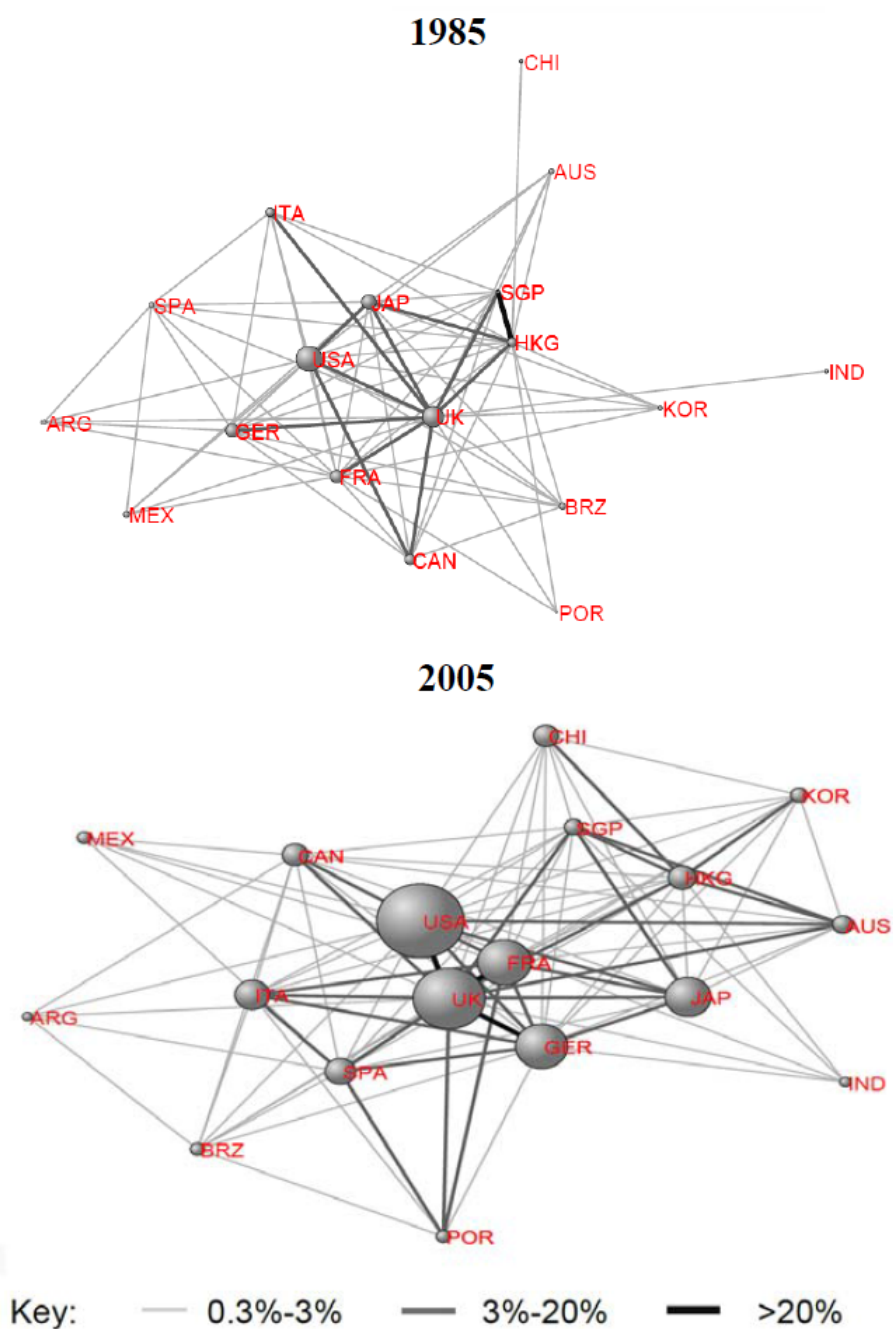


Figure 1.2: The global financial network in 1985 and in 2005. Nodes are scaled in proportion to total external financial stocks, and the thickness of links between nodes is proportional to bilateral external financial stocks relative to GDP. Source: Kubelec and Sá (2010), adapted from Haldane (2009).

becomes surprisingly difficult, not to say entirely impracticable, to deal with so many variables at the same time. This is an example of a *complex system*.

Complex systems necessarily consist of multiple objects, but size is not the only determining property. If many particles interact with each other in a random and uncorrelated way, the behaviour of the system as a whole can be actually quite simple. Imagine for example a large billiard table with thousands of balls rolling randomly over its surface, colliding with one another and with the side rails. The scientist can no longer trace the motion of one particular ball in detail, but can rely on *statistical mechanics*. Statistical mechanical approaches and in particular so-called ‘mean field theory’ make it possible to obtain precise estimates of many interesting properties of this billiard system, such as the average number of collisions of a ball over a period of time.

Statistical techniques have a great potential to simplify large-scale systems, but often rely on strong assumptions such as fully random and disorganised interactions. Returning to the billiard table, if the thousands of balls somehow form cliques or are influenced by global factors such as the ditches in the carpet, mean field theory no longer applies. Indeed, the complex systems considered in this thesis cannot be handled by either classical low-dimensional techniques or statistical mean field techniques. The examples of simple and complex systems in physics will turn out to have clear counterparts in economics and finance.

### **1.4 Complex systems in economics**

Complex systems have been studied already for decades. Early contributions in economics include Simon (1962) and Hayek (1964). These economists pragmatically turned to complex systems when faced with concrete problems of implementing policy. Simon (1962) explains his idea of a complex system as follows:

*“Roughly, by a complex system I mean one made up of a large number of parts that inter-*

*act in a nonsimple way. In such systems, the whole is more than the sum of the parts, not in an ultimate, metaphysical sense, but in the important pragmatic sense that, given the properties of the parts and the laws of their interaction, it is not a trivial matter to infer the properties of the whole."* – Simon (1962, p. 468).

Yet most of the models developed by economists are models of simple systems. Consider a standard oligopoly model for the market of, for example, cola drinks. Assume there are only two producers of cola drink. The firms interact with each other by setting prices. This kind of interaction is well described by *classical game theory*. A starting point of analysing such a model is to impose that the firms are fully homogeneous, i.e. make interchangeable products under identical technologies and for equal costs.

On the other side of the market, there are millions of possible consumers of cola drink. Assuming that their cola-drinking decisions are based on individual utility maximisation, all these consumers can be captured by a *representative agent*, linking the cola prices to aggregate demand of the entire population of cola-drinkers. Because interactions of consumers are disregarded, introducing a representative agent is equivalent to assuming that consumers are homogeneous and all behave as the representative consumer.

Classical game theory and representative agents in economics are counterparts of classical and statistical mechanics in physics. They have greatly advanced the modelling of simple economic systems. For example, models such as the oligopoly model can be used to analyse under which conditions the market is efficient, that is, functioning not worse than a centralised production line organised by a social planner. Moreover, it can be used for comparative statics, as first formally discussed by Samuelson (1947) in his *Foundations of Economics Analysis*.

However, during the financial crisis the economic models that had received so much critical acclaim were clearly misguided. The predominant analytical framework used in central banking has been so-called ‘dynamic stochastic general equilibrium’ (DSGE) modeling. In essence most DSGE models describe simple systems, similar in spirit to the stylised cola market described

above. Large groups of economic agents are forced into a single rational representative agent and explicit interaction is not taken into account. These are the ‘conventional tools’ Trichet and other policy makers have criticised, and for good reason: these models simply cannot explain the emergence of a crisis. This thesis instead models financial-economic systems as complex systems.

### **1.5 Purpose and methodology**

The study of complex systems in economics has resulted in a solid theoretical groundwork. The purpose of this thesis is to show that the theory can be used to explain empirical facts and make policy recommendations in financial economics. In this thesis the theory of complex systems will be applied to interbank markets and stock markets. As these two markets strongly amplified the financial crisis in 2008, the analysis helps to understand the mechanisms behind the emergence of the crisis.

A first methodological step is to establish *empirical relevance* of the complex systems approach. Why exactly are approaches of simple systems insufficient? Even as policy makers observed that the simple systems models were of limited use during the crisis, the burden of proof lies on complex systems to improve empirical performance. This thesis presents evidence for a non-random network structure of interbank exposures showing that banks in this market do *not* interact in a disorganised way. For the stock market, evidence is found for distinct temporal regimes in asset price dynamics that cannot be explained from the aggregate level. These empirical findings indicate that interbank and stock markets are complex systems in the sense discussed above.

The empirical relevance of complex systems leads to new questions for *financial market regulation*. Should banks be allowed to engage freely in financial innovation? What kind of information do regulators need in order to monitor not only individual institutions, but the system as a whole? More generally, what can policy makers do to improve the stability of complex



financial systems? For the relatively unexplored interbank market, policy measures aiming to create a more robust structure require knowledge of why individual optimisation of banks leads to the observed network structure. Market regulation for stock markets is already much more common. The policy analysis in this thesis expands on a small strand of economic literature recognising that financial-economic systems are complex.

The complex systems models that are developed in this thesis evolve around a relatively small number of key economic concepts. As a methodological strategy, the economic models in this thesis abstract as much as possible from financial detail. A small number of economic ideas will be used to explain observed emergent phenomena. One of these ideas is that financial agents establish long-term *trading relationships* along which they prefer to trade. As long as their counterparties meet their expectations, they should have little reason to seek new, unfamiliar partners. A second idea is that investors are susceptible to *market sentiment*, i.e. react sensitively to actions of other market participants, even when there is reason to believe that underlying fundamentals of financial assets have changed little.

A common thread to these economic ideas is that agents in complex financial-economic systems act adaptively in their interactions. They do not have a comprehensive overview of all aspects of the complex systems, but use simple, reasonable strategies to make decisions. For this reason the modelling of investors in this thesis naturally relates to *behavioural finance* (e.g. Shiller, 2005). The interaction of many adaptive agents leads to coordination at the global level that would have been hard to infer from the individual level. The complex system as a whole generates emergent phenomena.

## **1.6 Outline of the thesis**

This thesis contains four research chapters organised in two parts. Part of the research draws upon joint work. Chapter 2 has been written jointly with Iman van Lelyveld, while Chapter 3 is the result of fruitful cooperation with Marco van der Leij and Cars Hommes. Chapter 4 is an

extension of earlier work by Cars Hommes. Each chapter is written as a self-contained research paper, with its own introduction, conclusion and appendices, and could be read independently from other chapters.

Part I (Chapters 2 and 3) is concerned with interbank markets, taking particular notice of the interbank market in the Netherlands. Mainly through its tight connections to American banks suffering from the subprime mortgage crisis, the open Dutch banking system came under serious pressure. In October 2008, ABN AMRO bank was nationalised by the Dutch government, and also ING bank received a large capital injection. An important motivation for these government interventions was reduce the risk of further financial contagion.

Chapter 2 analyses data on interbank exposures in the Netherlands for the period from 1998 up to 2008. The focus is on the *core periphery structure*, which is found to give a better fit to the data than other network models proposed in the literature. The good fit of the core periphery structure suggests the presence of a stable set of densely connected core banks that intermediate between periphery banks. Monitoring core banks, which are systemically more important in passing through shocks, can help reduce systemic risk. Interestingly, the core periphery structure deteriorated considerably as the financial crisis developed during 2008, a finding also consistent with analysis of interbank markets in other countries.

Chapter 3 aims at giving an explanation for the core periphery structure in a network formation model of a financial market with an explicit role for *intermediation* by middlemen. In a core periphery network, core banks receive intermediation benefits for supporting indirect connections between periphery banks. The core periphery network is found to be strategically unstable if agents are homogeneous: there is always at least one agent who has an incentive to change its links. After allowing for heterogeneity among agents, a core periphery network can arise endogenously, with the big banks forming the core. This is an interesting finding especially in combination with the estimated core periphery network in the Netherlands. The Dutch core banks have a higher average size than periphery banks, even though some unexpected medium-sized banks are part of the core as well.

Part II of the thesis (Chapters 4 and 5) is concerned with stock markets. In Chapter 4, a complex systems model for asset pricing is estimated using U.S. stock prices from 1950 up to 2012. Market sentiment is modelled by *behavioural heterogeneity*: traders in the stock market are boundedly rational and disagree about how quickly stock prices return to the fundamental value. Fundamentalist traders believe in mean-reversion, while chartists extrapolate trends. Investors gradually switch between the two rules, based upon their relative performance, leading to self-reinforcing regimes of mean-reversion and trend-following. The estimated behavioural heterogeneity is both statistically and economically significant. The analysis in this chapter suggests that market sentiment switches between different behavioural regimes, which amplify fundamental shocks such as technology shocks and the subsequent dot-com bubble, and the Lehman bankruptcy in 2008 and the associated financial crisis.

Chapter 5 investigates the consequences of bounded rationality and heterogeneous beliefs for financial market regulations constraining large investment positions. It has been noted that the crisis in stock markets occurred at a time when *leverage* ratios had increased to enormous amounts and when simultaneously investors engaged massively in *short-selling*. This chapter shows that leverage and short-selling constraints can have large adverse effects on financial stability if investors have heterogeneous beliefs about future prices. The intuition for these adverse effects is that different beliefs themselves can be stabilising or destabilising. If the market regulations (unintentionally) restrict the stabilising fundamentalist strategy, mispricing and price volatility may in fact increase.



# **Part I**

## **Applying complex systems to the interbank market**



## Chapter 2

# Finding the core: Network structure in the Dutch interbank market

### 2.1 Introduction

Understanding complex interbank markets is crucial for safeguarding financial stability, as became clear during the financial crisis. Whereas the relevance of the network structure prior to the crisis was mentioned only infrequently, it has now caught the attention of both academics (Tirole, 2011) and policy makers (Haldane, 2009; Yellen, 2013). Systemic risk and contagion have become keywords in finance, and debate on their precise definition and implications is ongoing.

Traditionally, the interbank market has been considered without taking the network structure into account. Bhattacharya and Gale (1987) for instance propose that the main reason for interbank trade is to co-insure against idiosyncratic liquidity shocks. For various reasons, banks might not be indifferent with whom they transact, giving rise to non-random networks. One reason might be that interbank lending functions as a peer-monitoring device (Rochet and Tirole, 1996). Recently, some have argued forcefully that relationships matter in the interbank market. Cocco, Gomes, and Martins (2009) show that borrowing banks in Portugal pay a lower interest

rate on loans from banks with whom they have a stronger relationships, and Bräuning and Fecht (2012) analyse relationship lending in Germany during the financial crisis of 2007-2008. Such relationships will be reflected in the network structure.

As the theoretical literature started to consider networks, it became clear that the actual structure of linkages between banks affects the stability of the systems and the possible contagion after large shocks. In a seminal contribution, Allen and Gale (2000) use stylised examples to show that the fragility of the system depends crucially on the structure of interbank linkages. If a network is ‘complete’, i.e. all banks are connected to all other banks, a shock to a single bank can easily be shared and thus the stability of the system is safeguarded. If instead the network becomes clustered, spillovers to some of the banks can become substantial.<sup>1</sup>

Given its importance in theoretical work, the empirical analysis of financial networks is still lagging behind. Datasets are few and far between. Interbank exposures have mainly been used for stress test exercises. As noted by Upper (2011) in an overview of this literature, the estimated contagious effects are limited.<sup>2</sup> This is not surprising as the datasets generally only cover a single market and because behaviour does not change conditional on the state of the world. The limited literature that uses the tools of network theory and comes closest to our approach is discussed in Section 2.2.

The first goal of this chapter is to find a network model that best describes the structure of actual interbank markets. This is important, as the shape of the observed network is a key determinant of the stability of the system, as discussed in Section 2.5. To this end we will analyse a long running panel of bank links in the Netherlands. We will focus on the estimation of a core periphery (CP) model as recently applied by Craig and von Peter (2014). We will define this model more formally below, but loosely speaking this is a structure in which core

---

<sup>1</sup> The examples in Allen and Gale (2000) are clearly simplified and subsequent research has shown that many other aspects are relevant. For instance, Gai, Haldane, and Kapadia (2011) show in a model with unsecured claims and repo activity that systemic liquidity crises as seen in 2007-2008 can arise with funding contagion spreading throughout the network. Castiglionesi (2007) argues that it is not so much the structure of the market but rather the lack of complete contracting. See Chinazzi and Fagiolo (2013) for a survey of research exploiting networks to explain contagion and systemic risk in financial markets.

<sup>2</sup> Studies include Furfine (2003), Upper and Worms (2004), van Lelyveld and Liedorp (2006), Degryse and Nguyen (2007), and Mistrulli (2011)



banks intermediate between dependent periphery banks. Discriminating between these two types of banks is very useful for bank supervision in practice.

On a more technical note, we place more emphasis on model selection than previous studies in proposing the core periphery model for the interbank market. Another structure that recently has been put forward as a good description for interbank networks (e.g. König, Tessone, and Zenou, 2014) is the so-called nested split graph (NSG). We clarify the apparent confusion between the CP and NSG models, and find a better fit of the CP model in terms of the error score. Using Monte Carlo simulations, we show that the error score is a satisfactory measure of fit to discriminate between networks close to either an CP or NSG network. The simulations also show that the good fit of the CP model is highly unlikely to be generated by random processes of other well-known, stochastic network models. Overall, by combining various existing methods to fit networks to the Dutch data, we make clear that the core periphery model gives a better fit to the interbank market than others used in the literature.

For the Netherlands we find a core of around 15 banks. This core – unsurprisingly – almost always includes all the large banks. Some small banks, however, are also part of the core (although we cannot discuss these because of confidentiality reasons). Figure 2.1, a snapshot for a typical moment in time, shows that the model produces only few errors: there are few missing links in the core, and also relatively few existing links in the periphery.

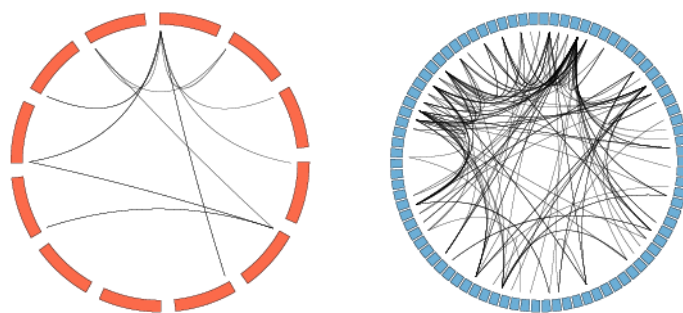


Figure 2.1: Errors of the CP model for the Dutch interbank market on 2005Q1. Banks are plotted on the circle. For the core (left), the errors are the missing links between any pair of the 12 core banks: 9 out of the 132 possible links. For the periphery (right), the errors are links between any two periphery banks (87 in total): 157 out of 7482 possible links.

The second goal is to place the network structure results in context. Are core banks different in terms of (regulatory) risk measures or business models? We will amongst other findings, show that core banks have, on average, lower capital buffers. So, even though they might be seen as more systemic, their buffers are lower. Taking the network structure into account in theoretical modelling and stress testing will be a crucial next step in developing systemic risk assessments. In the core periphery model a small set of core banks is highly connected, while periphery banks are not connected with each other but only to the core. The core banks are thus systemically more important in passing through shocks and should hence be monitored carefully to maintain financial stability.

The remainder of this chapter is structured as follows. In Section 2.2 we discuss different network models that have been applied to the interbank market with an emphasis on the core periphery network. Section 2.3 describes our dataset and the market network structure. Section 2.4 contains our main estimation results including a discussion on testing and model selection. In Section 2.5, we relate our results to financial stability. Section 2.6 concludes.

## 2.2 Network models

Mathematical network or graph theory has been applied to many different fields including biology, technological networks, and information science (see Newman, 2010, for a comprehensive review). In this section we will discuss both *stochastic* (in which link formation is determined by some random process such as the Erdős-Rényi random graph and the Barabási-Albert scale-free model) and *deterministic* models (in which a structure is postulated such as nested split graphs and core periphery network).<sup>3</sup> After describing the Dutch interbank market data in the next section, we will then, in Section 2.4, compare the relevant network models and will find

---

<sup>3</sup> We define a ‘network model’ as a simple representation of a structure of bilateral relations (in our case interbank loans). Thus the word ‘model’ does not refer here to a formal microeconomic characterisation of banks’ decisions. For the four models considered, however, it will become clear that latent microeconomic models exist that capture the concepts of random interaction, preferential attachment, counterparty reliability, and intermediation.

that the CP model best explains the data.

One of the earliest theoretical models of a network was introduced by Erdős and Rényi (1959, henceforth ER). In their random graphs, each possible link between any two nodes can occur with a certain independent and identical probability. The benchmark ER model has a severe limitation for our empirical application: it is not likely to produce a network with wide ranging degrees. The *degree* of a bank in a directed interbank network is the number of banks to which it lends (*outdegree*) or from which it borrows (*indegree*) and is denoted by  $k_i$  (and  $k_i^{out}$ ,  $k_i^{in}$ ). The property of a fat tail in the degree distribution has been observed in many types of networks and has led to the development of scale-free models by Barabási and Albert (1999), in which the probability of forming links increases proportionally with the degree (i.e. *preferential attachment*, PA). Preferential attachment by banks could result from the wish to interact with the most reliable counterparties. Trustworthiness is in this model associated with popularity: banks who initially have the largest number of interactions will attract more linkages over time. In particular, preferential attachment distributes the probability of a connection proportionately to its links  $k_i$ .

Until recently the empirical work on the interbank exposures almost exclusively built on the scale-free model. Theoretically, a scale-free network displays a power law in the degree distribution. In one of the earliest descriptions of interbank network topology, Boss, Elsinger, Summer, and Thurner (2004) fit power laws to two different regions in the degree distribution for Austria. Using data from Brazil, Cont, Mousa, and Bastos e Santos (2013) connect a systemic risk measure from a stress test exercise with local network characteristics, after calculating various properties of the network including the scale-free parameter. Martínez-Jaramillo, Alexandrova-Kabadjova, Bravo-Benítez, and Solórzano-Margain (2014) also find large degree heterogeneity in the Mexican interbank market. In sum, the empirical literature points to the apparent scale-free distributions as its main finding.<sup>4</sup>

<sup>4</sup> Fricke and Lux (2012) summarise more findings of the interbank network literature. In reviewing this literature, it is useful to contrast interbank exposures analysed in this chapter with overnight interbank transactions. Whereas interbank exposures refer to stock variables on the balance sheet, interbank transactions are flow variables. This chapter is concerned with interbank exposures.

However, the statistical support for many of the claims to scale-freeness of networks has been called into question (Stumpf and Porter M. A, 2012). The problem boils down to the fact that scale-freeness is an asymptotic property, and even in the largest datasets there are few observations of extreme degrees. The limited and mostly *fixed size* of banking networks makes the scale-free network of Barabási and Albert (1999), that relies on continuing growth, inappropriate.

A different class of network models consists of *deterministic models*. An important advantage of such models is that they fit well within economic theories of profit optimising banks. Banks do not interact randomly with each other; interbank exposures are the complex result of agents making well-considered decisions to borrow from or lend to one another. In the following two sections, we discuss two deterministic network models that recently have been applied to interbank markets: nested split graphs and core periphery networks. To our knowledge we are the first to compare the two models.

The two deterministic network models have some features in common, and in fact a core periphery structure is sometimes informally presented as indicative for nestedness. Indeed, we will see that both models impose the large heterogeneity in degree that is observed in the data. Moreover, they are *single-centred*: there are no distinct communities of banks that cluster together (for geographical and historical reasons). Within the small banking sector of the Netherlands, communities are not expected to play an important role. Despite their similarities, we will show below that nested split graphs and core periphery networks are actually based on different economic concepts. It is therefore useful to be able to discriminate between them in empirical interbank networks.

### **2.2.1 Nested split graphs**

*Nested split graphs* (NSG) are networks with a simple structure based on the degrees of nodes. To see how nested split graphs can be relevant for interbank networks, assume that banks can be ranked on the basis of their reliability and that every bank takes this ranking into account when

choosing counterparties. Each bank initially interacts with the most reliable counterparty; if it needs a second counterparty, it chooses the second most reliable bank, and so on. In this sense the network structure of counterparties exhibits nestedness.<sup>5</sup> Nested split graphs can be defined by the following strong condition (cf. König, Tessone, and Zenou, 2014, p.13):

**Condition NSG:** For any bank, the set of its counterparties is contained in the set of counterparties of any other bank with at least as many linkages.

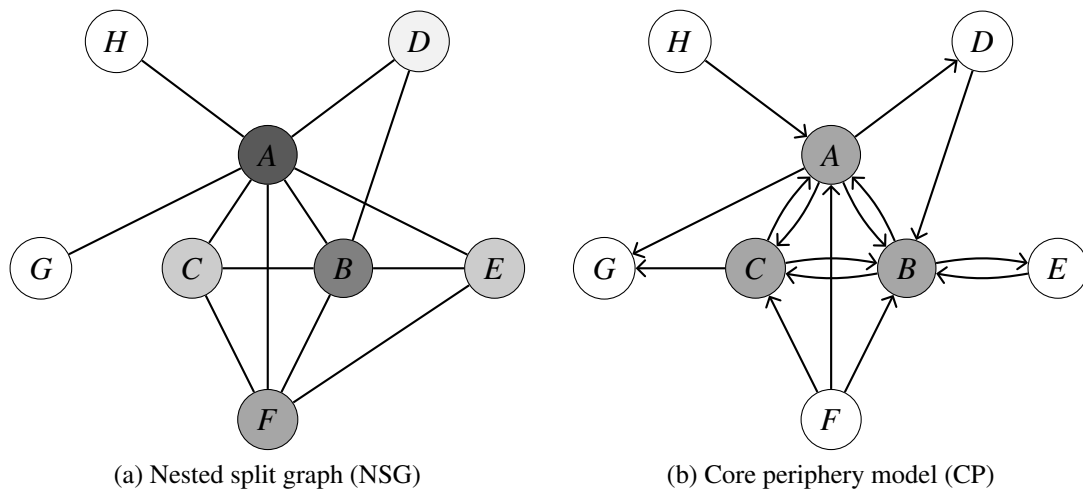


Figure 2.2: Example of perfect models

Figure 2.2a gives an example of a nested split graph that is perfect, i.e. it fulfills Condition NSG exactly. The shading of the nodes corresponds with the more preferable counterparties (having a higher degree). Because any set of counterparties should be contained in the set of counterparties of the bank with the largest number of linkages, any (non-empty) NSG contains at least one bank that is connected to all other banks in the system. In this case bank A has 7 links, one to every other bank. There are some banks that require more than one link, for example bank D which has an additional link to B. Condition NSG requires that all banks with more than one link have one counterparty in common: this counterparty is bank B. Two banks have three counterparties, namely C and E, and have the same set of counterparties  $\{A, B, F\}$ , consistent with Condition NSG.

<sup>5</sup> Nested split graphs are related to  $k$ -cores: the sets of nodes having at least  $k$  connections. Imakubo and Soejima (2010) investigate  $k$ -cores in the Japanese interbank market.

Cohen-Cole, Patacchini, and Zenou (2011) and König, Tessone, and Zenou (2014) apply nested split graphs to interbank market, and argue that they give a good fit to the datasets they consider. The idea of the nested split graph model is the strict ordering accepted by all banks with the single most preferable counterparty on top.<sup>6</sup> The alternative core periphery network instead suggests that there are some equally important core banks, and that periphery banks select one or more banks from this core as counterparties.

### 2.2.2 Core periphery networks

In the literature so far, there is some apparent confusion between nestedness and core periphery structures in interbank networks. We present the *core periphery* (CP) network as an alternative for nested split graphs. The idea is that core banks that intermediate between periphery banks, which in turn do not interact with each other. In Chapter 3 we will propose a financial market model in which banks can trade indirectly through intermediation and a core periphery network forms endogenously. The formal notion of core periphery networks was introduced in the social sciences by Borgatti and Everett (1999). The concept has recently been applied to the interbank market by Craig and von Peter (2014) and subsequently by others using different datasets.<sup>7</sup>

To test the concept of an interbank core periphery network in a quantitative way, Craig and von Peter (2014) introduce a formal definition. In particular, the relation between the core and the periphery is more clearly specified than in Borgatti and Everett (1999). In a perfect core periphery network, the following three conditions are satisfied:

**Condition CP1:** Core banks are all bilaterally linked with each other.

**Condition CP2:** Periphery banks do not lend to each other.

---

<sup>6</sup> Consistent with the prominent place of league tables in trade journals such as *The Banker*.

<sup>7</sup> The ideas of ‘money-centre banks’ or a two-tiered banking systems was previously described qualitatively (and graphically) by Upper and Worms (2004), Degryse and Nguyen (2007) and Soramäki, Bech, Arnold, Glass, and Beyeler (2007). The core periphery model captures these ideas quantitatively, and provides a way to test them in the data.

**Condition CP3:** Core banks both lend to and borrow from at least one periphery bank.

This definition of a core periphery model imposes a strict distinction between core and periphery banks. In Appendix 2.A we consider and estimate a version of the model in which each bank receives a continuous measure of its ‘coreness’, following Fricke and Lux (2012).

These requirements are illustrated in Figure 2.2b. Banks A, B and C are core banks and all lend to each other bilaterally. These banks also intermediate between the remaining periphery banks: for example, bank A intermediates between H and D. Note that Condition CP3 is violated if the link from C to G is removed (as C would not lend to the periphery), but not by removing the link from A to G (as A lends to D). Some periphery banks, like bank D and E, may lend and borrow at the same time, as long as they are not connected to other periphery banks.

The two figures taken together show the key differences between both deterministic network models. Most importantly, the core periphery network is defined as a directed graph. The directionality in the core periphery model stems from the fact that Conditions CP1 and CP3 require links in both directions, and follows from the underlying intermediation function of the core to the periphery. This directionality gives the CP model a potential advantage over the nested split graph in explaining the data. We will show that the CP model gives a better fit to the data, suggesting that the direction of links (i.e. borrowing or lending) is indeed important in capturing the global structure of interbank relationships.

Despite the few differences in connections between the two panels in Figure 2.2, the models attach a different importance to the most well-connected banks. The nested split graph orders the banks in partitions  $\{A\}$ ,  $\{B\}$ ,  $\{F\}$ ,  $\{C,E\}$ , and so on, based on their degree. In the core periphery model, the core of systemically important banks is formed by  $\{A,B,C\}$  but leaves some freedom for the exact relation between the core and the periphery, provided Condition CP3 is fulfilled. There is a jump in the degree distribution from periphery banks with degree up to 3 and the least connected core bank C with degree 6. A striking difference of the CP model is that bank F is placed in the periphery while it is the third bank in the ranking of the NSG.

In sum, NSG and CP networks rely on different economic concepts (counterparty reliability

and intermediation, respectively), and identify different sets of bank as most important in the system. The next section introduces our methodology to determine which of the two gives the best fit to observed networks.

### 2.2.3 Fitting deterministic network models

Deterministic network models impose strong restrictions on networks: either a network satisfies the conditions or it doesn't. Of course, in practice the data will rarely coincide exactly with a nested split graph. Similarly, the core periphery model will generally have both missing links within the core as well as links existing within the periphery. To evaluate the fit of the models, we count the number of *errors*, i.e. the number of links that have to be added or deleted to obtain a perfect deterministic model. An alternative measure of fit would be the Pearson correlation between the observed network and a perfect CP network, as discussed by Boyd, Fitzgerald, and Beck (2006). We follow previous studies in choosing the number of errors as the fitness measure, sometimes known as the Hamming distance, and motivate it below.

For the core periphery model, minimising the number of errors determines the estimated set of core banks, and thus the fit of the model to network data. For any chosen set of core banks, the errors with respect to the (perfect) CP model can be counted by checking the Conditions CP1, CP2 and CP3. The *optimal core* is defined by the core producing the smallest number of errors, and finding the optimal core is similar to running a regression.<sup>8</sup> The domain of the minimisation problem contains all possible cores, so that both the size and the composition of the set of core banks can be determined optimally from an observed network. For our dataset the found solutions intuitively correspond to core periphery networks, as can easily be checked by the high density of the core.

Fitting a nested split graph can be done in a similar way by counting errors for different orderings of banks and choosing the ordering with the lowest number of errors. However,

---

<sup>8</sup> Following Craig and von Peter (2014), violating Condition CP3 does not count as one error but is punished more strongly, producing errors for every periphery bank which it could have lent to or borrow from. This way of counting errors is convenient, as it ensures that the optimal core satisfies Condition CP3.

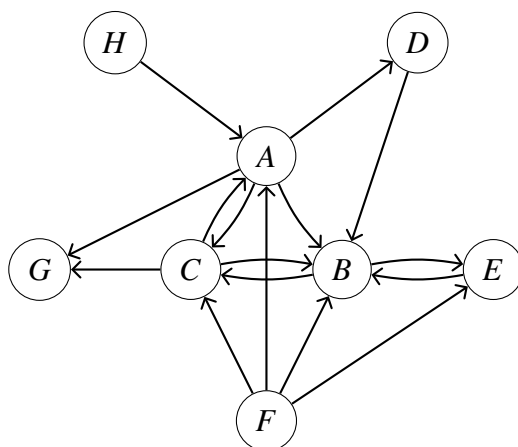


because nested split graphs are defined in terms of undirected links, the directionality in the network data first has to be transformed. This is done by defining an undirected link between a pair of banks if it has a lending relationship in at least one direction. Information contained in the network is thus reduced, but this is necessary for fitting a nested split graph.

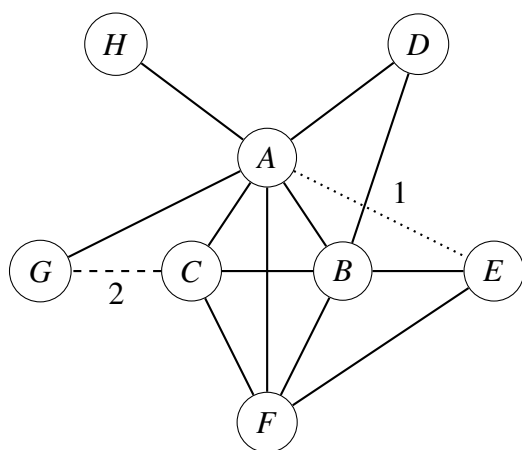
Figure 2.3 provides an example of a network with 8 banks. As in the interbank exposure data, links are directed. The two bottom plots of Figure 2.3 show how this network can lead to the perfect structures as shown in the panels in Figure 2.2 by making some small changes in the connections. For the fit of the NSG in the bottom left plot, the directionality of the network is first removed as discussed above. To have a perfect core periphery structure, the missing link between core banks A and B should be added (error 1) and the link between periphery banks E and F should be removed (error 2). To have a perfect nested split graph, bank A with the largest degree should link to all other banks including to E (error 1), while bank G is not allowed to link to C, as C is not the second bank in the ordering (error 2). Fitting the network to other orderings of banks for the NSG model or other sets of core banks for the CP model requires more changes in the structure. Hence, for this example network, the distance to both the closest CP structure and the closest NSG is two errors.

It is easy to see how slight changes in the network in Figure 2.3 could lead to a further distance from one deterministic model without changing the fit to the other. For the the CP network, for example, additional errors by removing a link between A and C or by adding a reciprocal link from E to F leaves the undirected network unchanged, and therefore the fit of the NSG is unchanged as well. For the NSG, removing the link from A to D or adding a link from C to H further limits the extent of nestedness, but does not conflict with Condition CP3 about the relation between core and periphery. These modifications to the example network show that having a lower number of errors gives a good indication about which structure gives a better fit to the data.

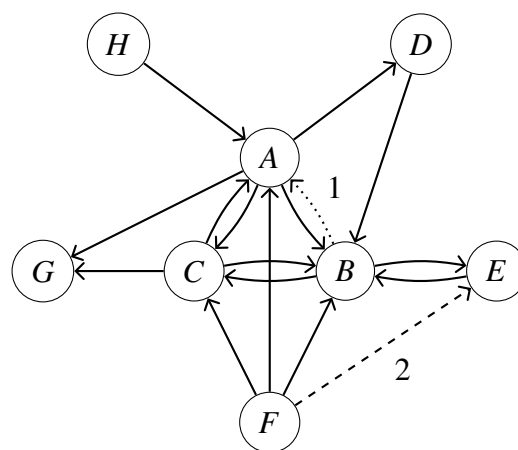
Nevertheless the number of errors is a raw measure of distance, as for example the maximum number of errors that can occur depends on the number of banks. Moreover, undirected



(a) Example of a directed network



(b) Fit to a nested split graph



(c) Fit to a core periphery network

Figure 2.3: Example of a directed network with 8 banks (a), its fit of the (projected and hence) undirected network to a nested split graph (b) and the fit to a core periphery network (c). The numbers indicate the two errors in both fits: one dotted link that should be added (1) and one dashed link that should be removed (2) to fulfill the conditions of the respective models.

networks generally consist of fewer links than directed networks, driving down the number of errors when fitting undirected networks. A improved measure of distance is therefore the *error score* (Craig and von Peter, 2014):

$$e = \frac{E}{L}. \quad (2.1)$$

The error score expresses the number of errors  $E$  as a proportion of the number of actual links  $L$ . Continuing with the example in Figure 2.3, we find an error score for the CP model equal to  $e^{CP} = \frac{2}{16}$  versus an error score for the NSG model of  $e^{NSG} = \frac{2}{13}$ , that is, a better fit for the CP model using the error score.

In Section 2.4, the error score will be used to compare the fit of different network structures to one particular observed network, as well as between the fit to different observed networks over time. Using Monte Carlo simulations, we argue in Section 2.4.2 that the error score is a satisfactory measure of fit in the statistical sense: when the real network is close to an CP, the error score of fitting an CP will be shown to be much lower than the error score of fitting an NSG, and vice versa.

A convenient property of the error score is that it is bounded by 1. To see that this holds for the CP network and NSG, note that these network models include an empty network without links as a possible perfect structure. The empty network satisfies Condition NSG, as the set of counterparties is empty for every bank; and it also satisfies Conditions CP1, CP2 and CP3 if the set of core banks is empty. Compared to an empty network model, any observed network has as many errors as links, so  $e = \frac{L}{L} = 1$ . The fitted CP network and the fitted NSG are the result of optimising over all possible structures and by definition cannot have higher error score than the empty network. Having an error score considerably below 1 is therefore a first requirement in evaluating the fit of a deterministic model.

For small networks, fitting deterministic models can easily be done by an exhaustive search. However, because the size of this domain rises exponentially with the number of banks, there exist ‘greedy’ algorithms for both the NSG and the CP model with a reduced, linear running

time. The algorithm iteratively improves the outcome by shuffling banks that generate most errors. Craig and von Peter (2014) developed such an algorithm for the CP model and show its robustness to the initial partition of core and periphery. For nested split graphs, a similar greedy algorithm is provided by Mannila and Terzi (2007).

## 2.3 Data description

Financial institutions interact on many levels. Sometimes these interactions are very short-lived with contractual obligations expiring before the end of the trading day. Other contracts, such as for instance swaps, can be long-running and can last up to 30 years. Our dataset combines all of these interactions on a quarterly reporting frequency, where we are limiting the sample to exposures up to one year. We use prudential reporting of balance sheet positions of Dutch banks. Each quarter banks have to report the interbank assets and liabilities to the market as a whole. In addition, banks have to report exposures to their largest counter parties. Assuming that the distribution of interbank exposures is equal to the distribution of claims in general, given by the large exposure reporting, we construct a matrix of interbank exposures for the Dutch banking market.

Data are available from 1998Q1 to 2008Q4 with the number of reporting banks varying between 91 and 103. This number includes Dutch banks, foreign subsidiaries, branches of foreign banks and investment firms. To get a feel for the data we show the overall domestic market volume in Figure 2.4.<sup>9</sup> In the long run the volume has been growing, dropping markedly in the financial crisis. The two vertical lines indicate 1) the beginning of the crisis in August 2007 once the subprime mortgages made the headlines, and 2) the bankruptcy of Lehman Brothers on 15 September 2008. Computed network measures such as path length tell a similar story: the network became more connected over time with a marked reversal due to the crisis.

In order to apply unweighted or binary network models discussed in Section 2.2 to our

---

<sup>9</sup> The dataset has been used before in Liedorp, Medema, Koetter, Koning, and van Lelyveld (2010). Unfortunately, information about relations with foreign counterparties is not available to us.

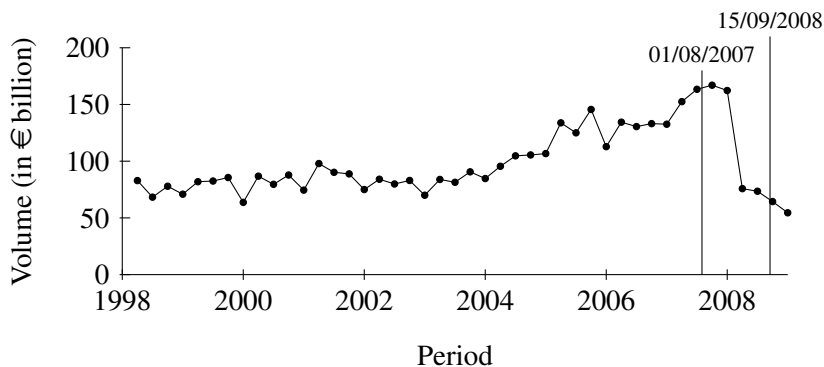


Figure 2.4: Domestic market volume over time

dataset, we reduce our weighted links to binary links. We use the same € 1.5 million threshold as Craig and von Peter (2014) to determine the most important links. For every period a sparse network remains of about 8% of the possible number of links. This makes the Dutch interbank network much denser than the German network (where less than 1% of possible links materialise). It seems that the main difference of the German market compared to the Netherlands is the large number of small banks, with relatively less links above the threshold.

Imposing a threshold necessarily involves some arbitrariness. In the present context, a higher threshold will result in fewer qualifying links and consequently a smaller core. However, this dependency does not change the key feature of the interbank market consisting of two groups. A more relevant question is how much information is lost by using the threshold. In Appendix 2.A we estimate a different version of the CP model using the weighted network of exposures and show that the estimation results, at least for the CP model, do not depend on the threshold.

As the first property of our binary network we first check the distribution of the degrees  $\{k_i^{out}\}$  and  $\{k_i^{in}\}$ . The distribution of the number of links per bank is an important feature of different network models and was also analysed by Boss, Elsinger, Summer, and Thurner (2004) and later interbank studies. We show selected distributions of in- and outdegree in Figure 2.5 (i.e. the first period, the last period and the overall aggregate over all periods). These distributions are plotted as the inverse cumulative distribution (showing the proportion of banks

with degree higher than  $k$ ) on a log-log scale. Scale-free networks in the limit of infinitely many banks display a power law in the degree distribution, corresponding to a straight line in such log-log plots.

The plots already give a first feel for the relevance of the different network models. There clearly exist some banks which are highly connected, with the number of counterparties reaching up to more than 90 counterparties. The ER model seems therefore inappropriate as it typically produces a much lower maximum degree. As the log-log plots show no linear relations, we did not try to fit a power law to these degree distributions. The indegree distribution looks like a concave decreasing function that could be present in both the PA, NSG and CP models. However, for many periods such as 1998Q1 shown in Figure 2.5, there is a jump in the outdegree. This discontinuity in the degree distribution is a first indication of a clear distinction between core and periphery. For the last period 2008Q4, there is no jump in the degree distribution.

## 2.4 Estimation results

Given the empirical interbank network in the Netherlands, we will evaluate the fit of the CP model in relation to the alternative models. The models are either stochastic (ER, PA) or deterministic (NSG, CP), but all four are in principle suitable candidates for the single-centred, fixed-size Dutch interbank market. As we will see, the core periphery structure provides the best description of the data. In Section 2.4.1 we estimate the core periphery model, finding the number of estimated core members and the error score (as defined in Section 2.2.3) for every period. We compare the error score of the CP model with the error score of the NSG model, and also relate our results with previous analysis of interbank markets in other countries. In Section 2.4.2, we pay further attention to selecting the CP model using Monte Carlo simulations.

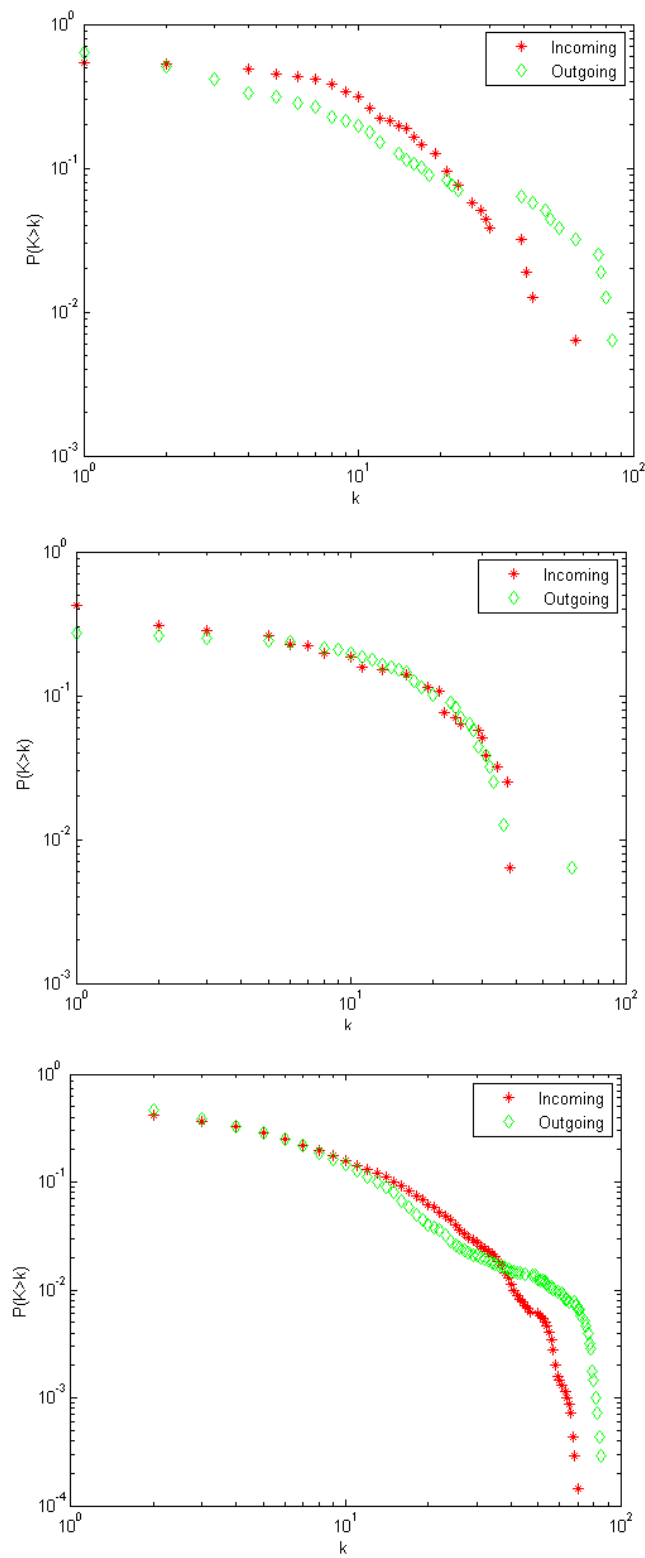


Figure 2.5: In- and outdegree distributions for 1998Q1 (top), 2008Q4 (middle) and over all periods (down)

### 2.4.1 Estimating the core

To estimate the CP model, we use the sequential optimisation algorithm developed by Craig and von Peter (2014). In our dataset with around 100 banks, the optimal core varies between 10 and 19 banks. Figure 2.6 plots the core size per period. Although in the first two periods a relatively large core is found of respectively 18 and 19 banks, the core size thereafter stays stable and close to the average of around 13 banks.

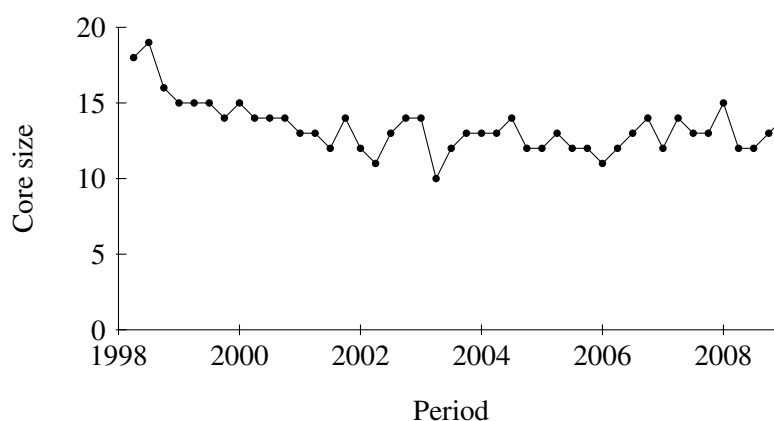


Figure 2.6: Number of core banks over time

There are two important points with respect to the core size. First, the relative core size is closely related to the density of the network. As the density of the network clearly depends heavily on the data construction, one should therefore be careful not to take the core size too literally. Second, there is no correlation with the total volume in the interbank market: although volumes almost doubled, the higher number of links above the threshold stayed relatively stable which leads to a stable core size.

Figure 2.7 shows the error scores of the CP and NSG models. The fit of the CP model is between 0.21 and 0.38, with an average of 0.29. We also calculate the error scores of NSG model using the algorithm of Mannila and Terzi (2007). Because the NSG model is fitted to the undirected version of the empirical network, the number of actual links by which the number of errors is divided is different for both models. The error scores of the NSG model turn out to be much higher for most periods, reaching up to 0.64, more than twice as high as the error score



of the CP model for some periods.<sup>10</sup>

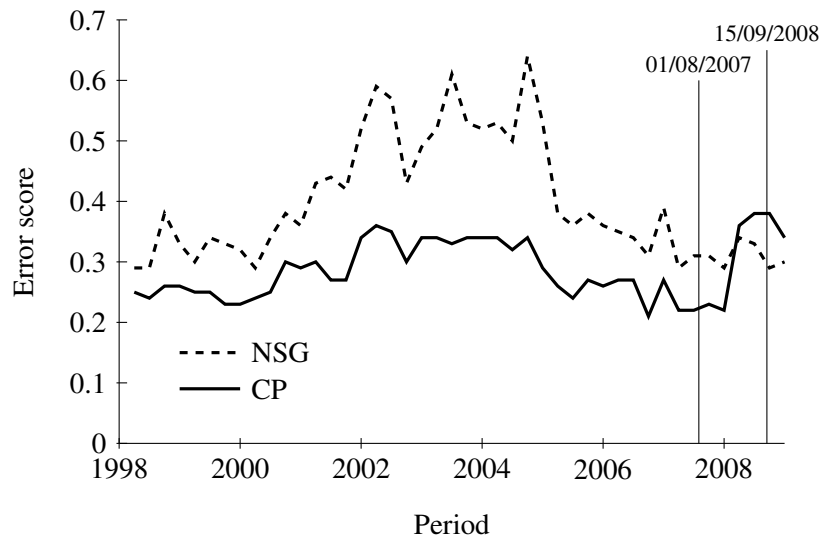


Figure 2.7: Error score of the CP and NSG models over time, 1998Q1-2008Q4

Interestingly, the fit of the CP model is clearly worse in the last four quarters, after the subprime crisis had started to spread. In fact, the fit of the NSG model is better in these periods. This suggests that the CP structure, as far as it is a good description of the Dutch market, dissolved during the crisis. Further indications for the deterioration of the core periphery structure are the more regular degree distribution in the last period (Figure 2.5) and the large number of errors in the core, particularly in period 2008Q4. In Appendix 2.A we estimate a different version of the core periphery model on the weighted network of exposures, and find a similar deteriorated fit in 2008Q1-Q4.

The finding that the core periphery structure dissolved during the financial crisis seems to be confirmed by analysis of other datasets. Using a dataset up to 2010 for the Italian market, Fricke and Lux (2012) established a significant structural break in the crisis, and a deteriorated fit of the model afterwards. Analysing different network metrics (i.e. “motifs”), Squartini, Van Lelyveld, and Garlaschelli (2013) find early signs for a changing network as well.

<sup>10</sup> When the CP model is estimated on the undirected network, we find even lower error scores. The conclusion that the CP model gives a better fit does therefore not depend on the directionality of the model, but rather on the nature of Conditions NSG and CP1-3.

In Matrix 2.1 we calculate the transition probabilities between the states of being in the core and in the periphery, and being inactive in the interbank market ('Exit'). Most importantly, the transition from core to core indicates that on average 83% of the core banks stay in the core the next period. As we found that the number of core banks is quite stable, the flows from and to the core are almost equal in absolute terms. The higher persistence in the periphery merely reflects that it consists of many more banks.

$$\begin{pmatrix} & \text{Core} & \text{Periphery} & \text{Exit} \\ \text{Core} & 83\% & 16\% & 1\% \\ \text{Periphery} & 2\% & 96\% & 2\% \\ \text{Exit} & 0\% & 2\% & 98\% \end{pmatrix}$$

Matrix 2.1: Transition matrix for different states in the Dutch interbank market, 1998Q1-2008Q4.

In Table 2.1, the empirical results of the CP model for the Netherlands are summarised and compared to Germany, Italy, the UK and Mexico. The fit of 0.29 in the Dutch market is worse than found in the German and Mexican markets (0.12 and 0.25)<sup>11</sup>, but better than in the Italian and UK equivalent markets (0.42 and 0.47). In Appendix 2.B we visualise the fit of the CP model in the first period, and we also plot all errors in both the core and the periphery for all 44 periods. These plots show that most of the errors occur in the periphery, where a perfect CP structure would have no links (i.e. an empty subgraph).

## 2.4.2 Monte Carlo simulations

We have shown that the error score of CP model was lower than that of the NSG model in the sample from 1998Q1 until 2008Q1. Now we will assess the fit of the core periphery structure

---

<sup>11</sup> The better fit in the German market can partly be explained by the lower density. We tried a higher threshold of €1 billion, in which case the density of the Dutch interbank market is 0.5%, very close to the density in the German market. The average error score decreases to 0.15.

	Netherlands	Germany	Italy	UK	Mexico
<i>Description</i>					
Total number of banks	100	1800	±120	176	46
Network density	8%	0.4%	± 15%	3.2%	26%
Average number of core banks	± 15	± 45	± 30	16	± 16
Average core size	± 15%	± 2.5%	± 25%	9.1%	± 35%
<i>Fit</i>					
Error frequency, as % of links	29%	12%	42%	47%	25%
Transition prob. core→core	83%	94%	83%	NA	94%

Table 2.1: Comparing the CP model for the Dutch interbank market to Germany (Craig and von Peter, 2014), Italy (Fricke and Lux, 2012), the UK (Langfield, Liu, and Ota, 2012) and Mexico (Solis-Montes, 2013). Langfield, Liu, and Ota (2012) only have one period and thus cannot compute a transition probability.

using Monte Carlo simulations. First, we simulate ER and PA networks using a data generating process that takes into account the size and density of the actual network. For each of these simulated networks, we calculate the errors of the core periphery model. In this way we can use the error score as a test statistic to evaluate whether the actual number of errors is likely to be drawn from the simulated distribution functions. Second, we make similar Monte Carlo simulations for data generating processes that are close to the fitted CP and NSG networks. This way, we check whether the error score is a well-behaved measure of fit to discriminate between the two models. Finally, we also investigate the relation between the fit of the CP model and the observed degree distributions  $\{k_i^{out}\}$  and  $\{k_i^{in}\}$ .

Starting with the explanatory power of the CP model over stochastic network models, we show the distribution function of ER and PA networks for a typical period (1998Q1) in Figure 2.8. We generate 1000 replications of ER random graphs and PA networks with 103 banks and 1135 links as in 1998Q1, and fit the CP model on every replication. We find that the CP model gives a much better description of the Dutch interbank data than the ER random graphs. If the probabilities of links are independent of each other, the generated networks have a barely skewed degree distribution. Fitting such random networks to a CP model gives a core that is too small. Note that although it is theoretically possible to end up in a perfect CP network (zero errors) this is extremely unlikely for a finite number of replications.

We also reject the hypothesis that the actual network is drawn from a distribution of PA networks in favour of the CP mechanism, as the actual error score is much lower than the entire distribution of generated PA networks. Note that the precise distribution of errors as shown in Figure 2.8 depends on our choice of the Xie, Zhou, and Wang (2008)-procedure which is appropriate given the fixed size of around 100 banks.<sup>12</sup>

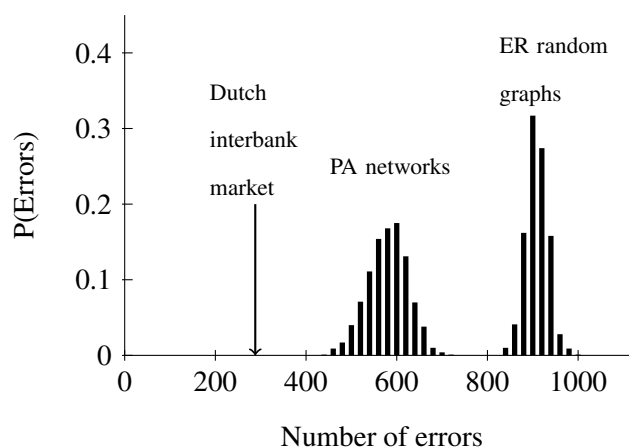


Figure 2.8: Simulated distribution function of number of errors for 1998Q1, 1000 PA networks and 1000 ER random graphs.

Using similar Monte Carlo simulations, we now argue that the error score is a satisfactory measure of fit to discriminate between CP and NSG networks. In particular, we start from perfect CP and NSG structures and simulate randomised networks using a small probability that each link is rewired. To mimic the empirical problem where we do not know a priori which model is better, we fit both the CP and NSG model on every simulated network, irrespective of whether it was generated from a perfect CP or NSG structure. For example, we generate randomised networks from a perfect CP network with a low probability of rewiring, and fit both

<sup>12</sup> Xie, Zhou, and Wang (2008) suggest the following procedure to generate a PA network: starting from any (random) initial network, iteratively (1) remove a random link, and (2) add a new link between two banks that are both selected with probabilities  $\frac{k_i}{\sum_j k_j}$ . As noted by Zhong and Liu (2012), the model implicitly allows self-loops, multiple links and many isolated nodes. We exclude self-loops and multiple links, and let the procedure run until we attain a maximum of 10% isolated links. The resulting degree distribution no longer follows a power law, but does generate a large asymmetry in degrees. Fricke and Lux (2012) follow a different procedure by Goh, Kahng, and Kim (2001), but their procedure imposes the power law degree distribution, rather than deriving it from preferential attachment. Finally, Craig and von Peter (2014) use another procedure that does not generate stochastic scale-free networks, so they do not find an approximately normal Monte Carlo distribution.

models. Of course we expect to find much fewer errors when fitting a CP model than when fitting a NSG model, but depending on the relations between the core and periphery and on the random outcomes of rewiring, it is possible that the NSG model gives a better fit.

This Monte Carlo procedure has three steps. First, to have reasonable choices of perfect CP and NSG structures, we start from the fits of these models to the first quarter in our dataset, 1998Q1. The fit of the CP model gives a directed network satisfying Conditions CP1, CP2 and CP3, while the fit of the NSG model is an undirected network satisfying Condition NSG. Second, we randomise these networks using the procedure of Watts and Strogatz (1998). This procedure rewires each link from bank  $i$  to bank  $j$  with a probability  $p$  to a different bank  $k \neq i, j$ . We generate  $R = 1000$  replications of simulated networks per underlying generating process. Third, we estimate both the NSG and the CP model on the simulated network.<sup>13</sup>

Table 2.2 presents the simulation results. In the first column the true type of simulated network is indicated, with either the fitted CP or the fitted NSG and the rewiring probability  $p$ . The top row indicates which network model has been fitted to these simulated networks. The results clearly show that if the underlying network is an NSG the average error score of fitting an NSG is much lower than for fitting an CP, and vice versa a much lower error score is achieved when fitting an CP structure when the actual underlying network is a rewired CP. We conclude that the error score rightfully discriminates between CP and NSG networks and works well as a measure of distance.

Finally, as typical in the complex network literature, we have also tested whether the network outcomes are significant under the so-called *configuration model*. Such a configuration model uses the full degree distributions and is thus a very different approach from the one presented so far (see Appendix 2.C for further elaboration). This alternative approach shows that, given the degree distribution, the CP model does not add anything. Note that this has no bearing

<sup>13</sup> As discussed in Section 2.2.3, we have to make sure that the directionality of the network and the model coincide. To fit the CP model on the undirected NSG simulations, we add directionality to undirected links before the rewiring in step two. We choose probabilities to transform an undirected link in either two reciprocal directed links or one directed link that ensure that the number of directed links is close to the observed value of 1135 links in 1998Q1. To fit the NSG model, we remove the directionality of the generated networks after the rewiring.

Simulated network	Error scores			
	$e^{CP}$		$e^{NSG}$	
CP fit, $p = 1\%$	0.5%	(0.4%)	16.2%	(0.5%)
CP fit, $p = 5\%$	2.6%	(0.7%)	20.4%	(1.0%)
NSG fit, $p = 1\%$	12.7%	(0.7%)	1.6%	(0.5%)
NSG fit, $p = 5\%$	14.6%	(0.8%)	7.5%	(1.1%)

Table 2.2: Mean of the error score (standard deviation in brackets) from fitting CP and NSG models to  $R = 1000$  simulated networks based on fits of the Dutch interbank market in 1998Q1.

on whether the CP model is plausible in itself.

In summary of the estimation results, we favour the core periphery structure over random graphs (ER), preferential attachment networks (PA), or nested split graphs (NSG), because the error score in the data is much lower for the CP model than for ER, PA or NSG models. The better fit can be fully attributed to the degree distribution, with a small subset of banks having many links, removed sharply from the majority of banks having only a few.

## 2.5 Network structure in perspective

In this final section, we place our discussion of the interbank network structure in a wider context. We will first describe how core/periphery membership relates to various measures of bank health and business models. This is followed by a discussion of how these results might be relevant for financial stability assessment.

Core banks tend to be the larger banks (cf. Craig and von Peter, 2014) and authorities might have already identified these as systemic. More interestingly, however, some smaller, less obvious banks also enter the core (although we can't discuss these for confidentiality reasons). Figure 2.9 plots the distribution of total asset size (in log scale) for all core and periphery banks, over all periods. While the core has a higher average size than the periphery, we observe that the group of core banks can be divided in the small set of the largest banks, and an additional group of medium-sized banks of a size similar to many periphery banks.

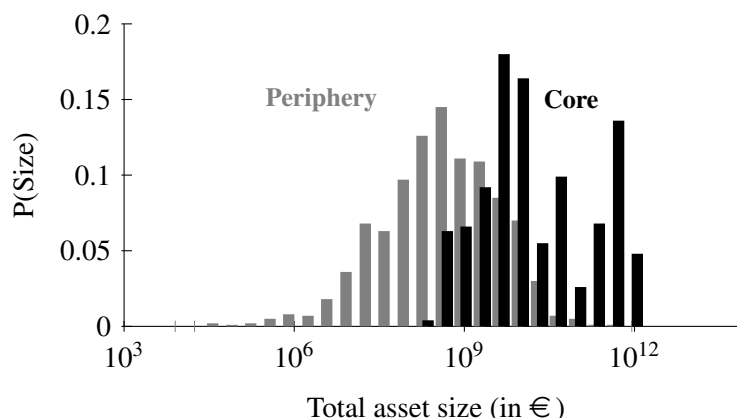


Figure 2.9: Distribution of total asset size for core and periphery banks over all periods.

Core and periphery banks might not only differ in size but also on other dimensions. An often used regulatory classification of bank soundness are the CAMELS indicators. These indicators rate a bank on Capital, Asset quality, Management capability, Earnings, Liquidity and Sensitivity to market risks. As discussed in detail in Liedorp, Medema, Koetter, Koning, and van Lelyveld (2010), there are various well known proxies and we have investigated them all. For almost all of these variables there does not seem to be a difference between the two groups of banks. In Figure 2.10 we show a selection of these variables. Panel (a) plots the average buffer of the core and periphery members over time. The difference is 10 percentage points on average and is statistically significant. Core banks thus tend to have a *lower* buffer, defined as the Tier1 capital over the total assets.<sup>14</sup> Panel (b) shows the returns on assets, which are largely similar except for the crisis period where core banks (which we showed to be larger and more thinly capitalised) faced significant losses. The third panel (c) plots the cost to income ratio, an often used proxy for management quality, which is typical for many of the other indicators: both the core and the periphery banks are largely similar. The final panel (d) shows that the net interest margin, another measure of banks' business models, gives also no difference between the two groups.

<sup>14</sup> Alternative measures condition the buffer on risk sensitivity, e.g. using Basel Risk Weighted Assets (RWA), ratings or CDS spreads. Unfortunately, both the regulatory (RWAs) and the public measures are not available for the vast majority of the periphery banks. However, simple measures, such as the leverage ratio shown, have recently been demonstrated to outperform more complex, risk sensitive measures as predictors of financial distress (cf. Hilscher and Wilson, 2011; Haldane and Madouros, 2012).

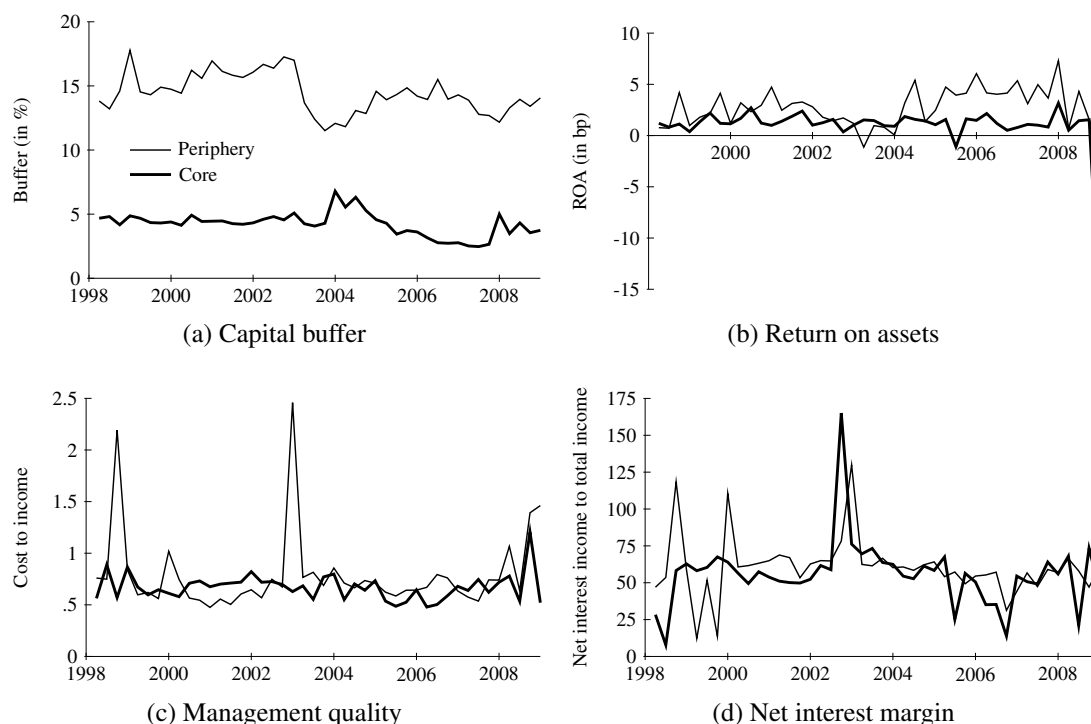


Figure 2.10: Selected CAMELS and business model indicators

We have shown that the core periphery model gives the best fit out of the network models examined, although the structure deteriorated during the financial crisis. In any case, the way interbank linkages are formed is far from random. In all periods, the degree distribution shows that there are certain banks with many more counterparties than expected based on the benchmark ER random graph model. This persistent asymmetry in degrees has not yet been incorporated in the theoretical literature on financial contagion and bankruptcy cascades, which follows Allen and Gale (2000) in assuming homogeneous banks. Freixas, Parigi, and Rochet (2000) do have – limited – heterogeneity as they allow for one bank to be the clearing bank.

Taking this heterogeneity seriously could start with an analysis as in Nier, Yang, Yorulmazer, and Alentorn (2007) who investigate how connectivity affects the diffusion of shocks within an ER random graph. If such simulation studies would be performed in a core periphery network, the failure of a core bank would lead to much more severe contagion throughout the system, while a shock to a periphery bank can more easily be absorbed by its counterparties in the core. The fact that these core banks have lower buffers would exacerbate this. Georg



(2013) takes the results of Nier, Yang, Yorulmazer, and Alentorn (2007) further by comparing different interbank network structures and shows that money-centre networks are more stable than random networks.<sup>15</sup>

## 2.6 Conclusions

The interbank market plays an important role for Dutch banks in the management of their liquidity positions. While foreign counterparties are easily found in the small and open banking environment of the Netherlands, there exists a core of domestic banks that are connected to many periphery banks. Following studies in other countries, we described the network that arises in the interbank market. We identify the most important banks based on the network structure. We also made efforts to visualise the network, which aims at helping authorities to supervise complex market structures.

Our main methodological contribution lies in combining new methods for comparing different network models. Our comparison shows that the core periphery model fits better empirically than random graphs, preferential attachment networks, and nested split graphs. This implies that the simple and economically intuitive division between two tiers of banks is justified. In our dataset of Dutch interbank connections from 1998Q1 until 2008Q4, we find a core of roughly 15 out of the 100 active banks. The distinction between core and periphery banks is found to be stable over the entire sample. Thus the hierarchical structure seems to be a ‘new stylised fact’ for interbank markets (cf. Fricke and Lux, 2012).

Our findings raise the question why individual banks’ decisions on loan contracts align with a structure of strongly connected core banks and dependent periphery banks. In the line of reasoning of Chapter 3, we have discussed the core periphery model in terms of loan intermediation. This seems like a fruitful idea in developing an improved understanding of the mechanisms generating interbank networks. Fricke and Lux (2012) discuss different explana-

---

<sup>15</sup> Georg (2013) also provides evidence that the central bank only stabilises interbank markets in the short run. See van Lelyveld, Liedorp, and Kampman (2011) for an analysis of how market structure affects stability in reinsurance markets.

tions of a core periphery structure and emphasise the comparative advantage of core banks in gathering information (e.g. about counterparty risk), an advantage mostly driven by bank size. We have shown that although core membership is indeed correlated with total asset size, there are several medium-sized and unexpected banks in the core of the Dutch interbank market.

Irrespective of what determines the structure, we have shown that there exists a core of systemically important banks. Moreover, core banks' difficulties will arguably have larger effects to the system as a whole than the failure of periphery banks. Authorities might therefore consider to require the core members to hold higher buffers. In contrast, however, the core banks in our sample have lower buffers.

The core periphery structure deteriorated considerably during the financial crisis, as was also found by Fricke and Lux (2012) for the interbank market of Italy. More research is required to explain the deteriorated fit by individual banks' behaviour and to provide further guidelines for supervisory authorities during financial distress. The methodology proposed in this chapter opens up new opportunities for systemic risk assessments of the interbank market, especially as more granular data is becoming available.

## Appendix 2.A Robustness analysis using the weighted network

In this appendix we make use of the size of the exposure data and check the robustness of the results on the CP model. We denote by  $\mathbf{D}$  the exposure matrix at some particular point in time, where element  $\mathbf{D}_{i,j} > 0$  represents the size of an exposure of bank  $i$  on bank  $j$ . First of all, we consider the loan size distribution over all periods. Figure 2.11 plots this inverse cumulative distribution (of loans in €) in log-log scale. The plot shows that a quite high number of the links (around 30%) is estimated below € 1000. As these small loans are due to the data construction (see below) rather than actual banking agreements, it is important to use some sort of threshold value in our analysis of unweighted links.

Figure 2.11 shows a wide dispersion of loans within the sample. We could estimate a power law distribution on the right tail to see whether loan sizes in the Netherlands display scale free behaviour. In an early analysis of interbank network structure, Boss, Elsinger, Summer, and Thurner (2004) fit power laws to the Austrian loans.

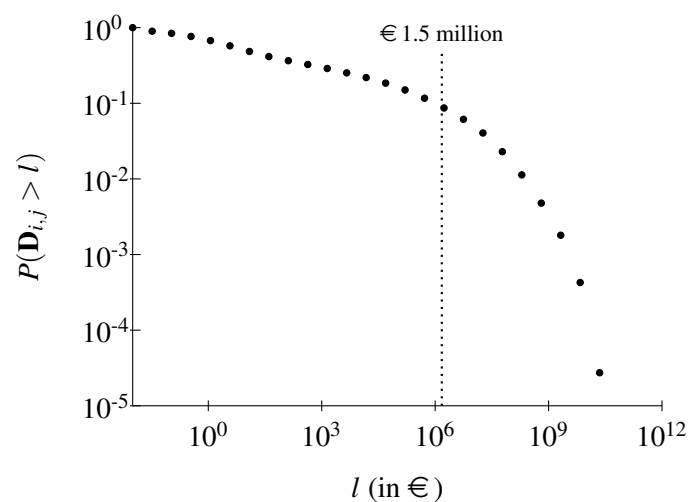


Figure 2.11: Loan size distribution over all periods

In the analysis of Section 2.4 we have used a threshold of € 1.5 million to construct the un-

weighted network  $\mathbf{A}$  of binary links (right of the dotted line in Figure 2.11):

$$\mathbf{A}_{i,j} = \begin{cases} 1 & \text{if } \mathbf{D}_{i,j} \geq \text{€ 1.5 million} \\ 0 & \text{if } \mathbf{D}_{i,j} < \text{€ 1.5 million.} \end{cases} \quad (2.2)$$

In Figure 2.12, both the total number of links and the ‘significant’ links above the threshold are presented over time. Interestingly, the total number of links shows a sharp drop in 2008Q1 when almost half of the links disappear permanently. This might be driven by the way the data is constructed. For the smaller loans, it is assumed that the totals of the interbank assets and liabilities for a given bank (observable from the balance sheet) follow the same distribution over other banks as the payment system. After the beginning of the financial crisis in August 2007, the payment system between banks was already under stress, and this might have caused a direct effect on the estimation of the small loans. See Liedorp, Medema, Koetter, Koning, and van Lelyveld (2010) for a full description of the construction of the dataset.

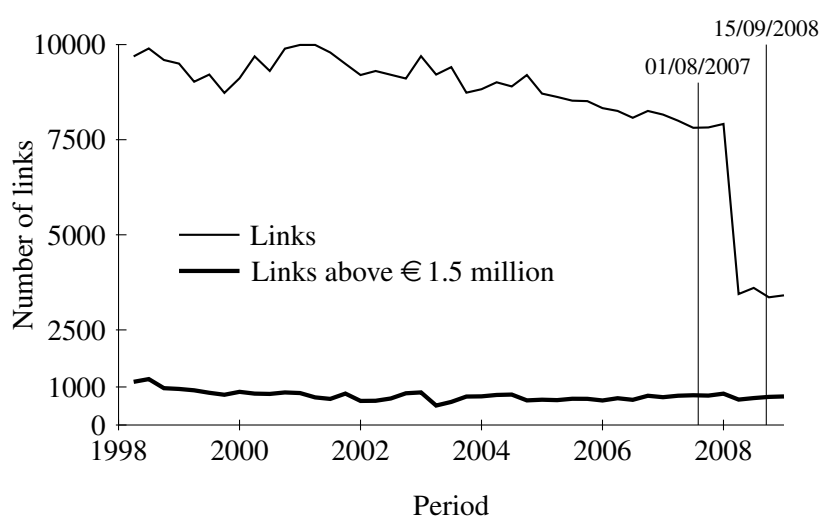


Figure 2.12: Number of links over time

While the small loans seem to drop in number for reasons unrelated to actual balance sheet management of the banks, the number of links above the threshold shows a much milder decrease during of the crisis. The steady number of ‘significant’ links is therefore a more reliable

source of information than considering all links with nonzero weight. Also note that while only around 1000 links remain, the remaining links cover practically all market volume (more than 99.5%). This observation is in accordance with the wide dispersion of link size as noted above. Given that a threshold is necessary for the (discrete) network models in Section 2.2, the choice of € 1.5 million seems not too restrictive.

To investigate whether the choice of the threshold has an effect of our results for the core periphery structure, we estimate the model using the different methodology of Fricke and Lux (2012). They consider a continuous version of a core periphery model where is each bank is assigned a level of ‘coreness’. The optimisation procedure chooses the ‘coreness’ of each bank such that the implied ‘pattern matrix’  $\mathbf{P}$  is closest to the observed matrix of interbank exposures  $\mathbf{D}$ , as will be specified below. Apart from the advantage that no threshold has to be imposed on the exposures  $\mathbf{D}_{i,j}$ , this method can be used to detect differences between bank in terms of their coreness, both within the core and between core and periphery.

Fricke and Lux (2012) distinguish between two versions of the continuous model depending on the specification of the pattern matrix  $\mathbf{P}$ . In the symmetric version, there is one vector  $c$  where  $0 \leq c_i \leq 1$  is a measure of coreness for bank  $i$  and  $\mathbf{P} = cc'$ . The discrete model can now be seen as a special case with  $c_i \in \{0, 1\}$ , which leads to Conditions CP1 and CP2.<sup>16</sup> In the asymmetric version, there are two vectors  $u$  and  $v$  representing the ‘out-coreness’ and ‘in-coreness’ of each bank, and  $\mathbf{P} = uv'$ . The asymmetric version can be useful if there is lot of difference between lending and borrowing behaviour of core banks.

For the continuous models we use the *proportional reduction of error (PRE)* as the measure of fit, defined as:

$$PRE = 1 - \frac{SS(\mathbf{D} - \mathbf{P})}{SS(\mathbf{D} - \langle \mathbf{D} \rangle)}, \quad (2.3)$$

---

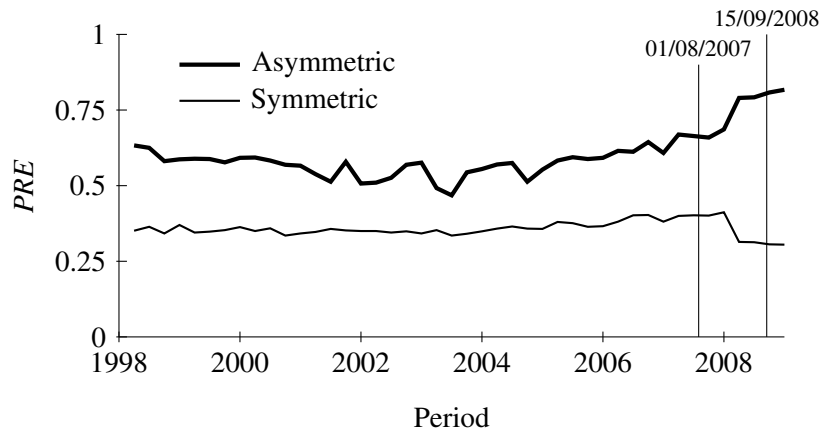
<sup>16</sup> Condition CP3 would not follow from this example, because  $c_i c_j = 0$  if  $i$  is in the core and  $j$  in the periphery. Instead, Condition CP3 leaves the structure of links between the core and periphery for the most part unspecified, indicating just a minimum of one link to and one from the periphery. Fricke and Lux (2012) estimate the discrete model with and without Condition CP3, and conclude that the differences are negligible in their dataset.

with  $\langle D \rangle$ ) the global average of  $\mathbf{D}$  (across all elements, excluding the diagonal), and  $SS(\cdot)$  the sum of squared deviations of the off-diagonal elements of the input matrix. Thus, maximising the *PRE* is equivalent to minimising the distance between  $\mathbf{P}$  and  $\mathbf{D}$ . Note that the *PRE* is similar in spirit as the error score for the discrete model. Both measures scale the amount of errors of the optimal structure to that of a trivial structure of just a periphery, constructed by only using the information about the number of banks and the number of links.

Figures 2.13 and 2.14 presents the main results of the estimation of the continuous core periphery models. Following Fricke and Lux (2012), we have log-transformed the data in the form  $\tilde{\mathbf{D}} = \log(1 + \mathbf{D})$  in order to adjust for the skewness of the data. As also found by these authors, fitting the continuous models on the raw network matrix  $\mathbf{D}$  leads to a very poor fit that is not comparable to the fit of the discrete model. The fit in terms of the *PRE* (Figure 2.13) shows that the symmetric continuous model has a deteriorated fit during the four quarters of 2008, similar to the discrete model.

Surprisingly, the asymmetric version of the model performs better during the financial crisis than before 2008. The fit of the asymmetric version should always be better than the fit of the symmetric model, as it has twice as many parameters. In the period 1999Q1-2007Q4 the ratio of the *PREs* is around 1.6, but in 2008Q1-Q4 it almost doubles to around 2.6. As it turns out, the reduced fit of a model with one set of systemically important banks for the Dutch interbank market during the crisis can be improved by making the distinction between important lending banks and important borrowing banks. In contrast, Fricke and Lux (2012) found that for their data both model versions gave a worse fit during the crisis.

Figure 2.14 shows how the fit is achieved for the asymmetric model and relates the results back to those of the discrete model. It plots the level of in-coreness  $v_i$  and out-coreness  $u_i$  for all banks  $i$  and over all periods. The colour of the dots indicates whether the bank was assigned to the core or to the periphery in the discrete model. The asymmetric continuous model confirms that these binary assignments to core or periphery are sensible: core banks have higher in-coreness and out-coreness. While there is not a very complete separation between the two

Figure 2.13: Fit of the continuous models in *PRE*

groups, many core banks have distinctly higher coreness measures. We find that core banks have relatively balanced positions (in the sense of similar in- and out-coreness) and are thus important for both lending and borrowing in the interbank market.

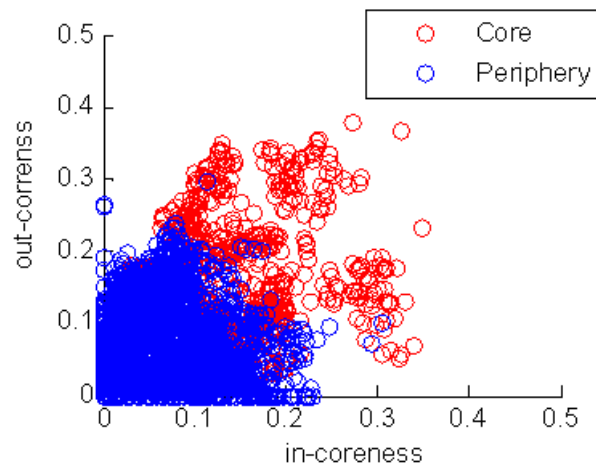


Figure 2.14: Out-coreness vs. in-coreness over all periods, labelled by the discrete model

In general, we find that different models of the core periphery structures give comparable results, confirming the results of the discrete model presented in Section 2.4. To make the relation between the different model versions more explicit, we follow Fricke and Lux (2012) in calculating correlations between the different versions of the model in Matrix 2.2. The correla-

tions are based on the (stacked) coreness vectors over all periods, i.e. the binary core vectors for the discrete model, and the vectors  $c$ ,  $u$  and  $v$  for the continuous models. The lowest correlation found is 30% between in-coreness and out-coreness, indicating that there is some heterogeneity within the core of the interbank market. The overall core of the discrete model is somewhat higher correlated with out-coreness than with in-coreness.<sup>17</sup> Nevertheless, we can safely say that the discrete approach gives satisfactory results for our dataset.

$$\begin{pmatrix} & \text{Symmetric} & \text{Out-coreness} & \text{In-coreness} \\ \text{Discrete} & 74\% & 60\% & 53\% \\ \text{Symmetric} & & 81\% & 77\% \\ \text{Out-coreness} & & & 30\% \end{pmatrix}$$

Matrix 2.2: Correlation matrix for different versions of the core periphery model.

---

<sup>17</sup> Interestingly, Fricke and Lux (2012) found that overall coreness is almost completely driven by the outgoing links of core banks, as shown by correlations of 73% between discrete and out-coreness, and of 26% between discrete and in-coreness. In their dataset the asymmetric version of the continuous CP model leads to a considerably improved fit.



## Appendix 2.B Visualising errors in the CP model

This appendix shows two visualisations of the core periphery model. Note that even for a relative small number of 100 banks, network plots very quickly become inscrutable. Figure 2.15 shows the interbank linkages of Dutch banks in 1998Q1, where the estimated core banks are placed in an inner circle within the circle of periphery banks.<sup>18</sup> It is visible that the fitted network incorporates a clear structure, while retaining the general nature of the data. Figure 2.16 focuses on the errors in all periods. For the core banks, the errors are the missing links, whereas for the periphery the links between any two periphery banks are drawn.

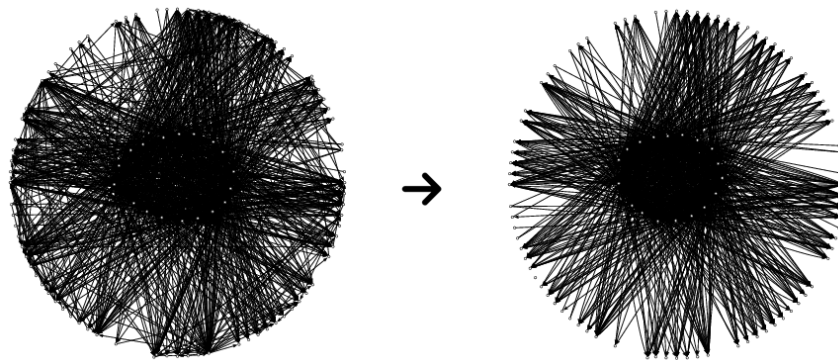


Figure 2.15: Plot of the Dutch interbank market on 1998Q1 (left) and its fit on the CP model (right). The outer circle represents periphery banks; the inner circle of small white dots core banks.

<sup>18</sup> These plots were made with the Pajek program as described in de Nooy, Mrvar, and Batagelj (2011).

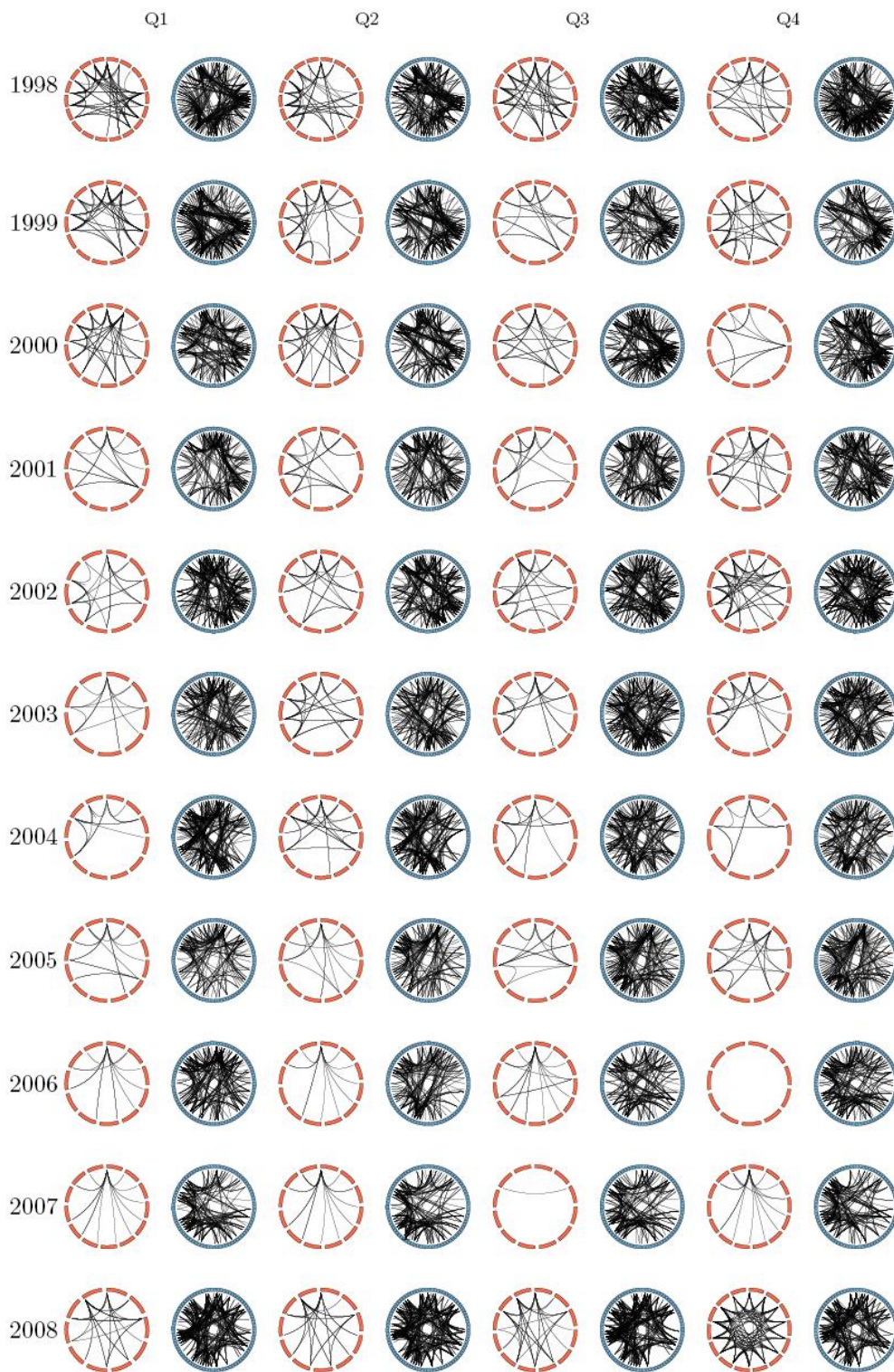


Figure 2.16: Errors in the CP model: core (red, left) and periphery (grey, right). The optimal core is established using the sample of all banks per period. The core (periphery) banks are then placed on the circle and the lines shown are those between banks that miss (have) a link that is (not) supposed to be there.

## Appendix 2.C Testing the CP against the configuration model

As typical in the complex network literature, we also test whether the network outcomes are significant given the same degree distribution, under a so-called *configuration model*. Such a model is very different from the models discussed so far, as it makes use of the full degree distributions  $\{k_i^{out}\}$  and  $\{k_i^{in}\}$  instead of just the density of the network. A useful way to think of this model is by considering a rewiring algorithm, where two distinct links in the actual network are broken and the nodes are reconnected reversely.<sup>19</sup> The distribution of the error score under the configuration model indicates whether or not, given the degree distribution, there is additional evidence for the CP structure in the ‘ordering’ of the links.

Instead of a rewiring algorithm, we use the analytical maximum likelihood method of Squartini and Garlaschelli (2011) which is computationally much less demanding, so that we can test the configuration model in all periods. As a first easy exercise, we use the method on the random graph model, so that only the density is included in the information set. The best estimator for the probability of a link  $p$  is then simply equal to the density of the network. For the configuration model, the linking probabilities  $p_{ij}$  are such that the degrees match the observations in expectation:

$$k_i^{out} = \sum_{j \neq i} p_{ij} \quad (2.4)$$

$$k_i^{in} = \sum_{j \neq i} p_{ji}. \quad (2.5)$$

Squartini and Garlaschelli (2011) show that the linking probabilities should be of the form:

$$p_{ij} = \frac{x_i y_j}{1 + x_i y_j}. \quad (2.6)$$

The analytical method amounts to solving the equation for the parameters  $\{x_i\}$  and  $\{y_i\}$ .

Given the maximum likelihood estimations on the probability of linkages between banks,

---

<sup>19</sup> This procedure is sometimes called “Maslov-Sneppen” rewiring (cf. Squartini and Garlaschelli, 2011).

the z-score is defined as the observed number of errors expressed in standard deviations from the expected value under the null model. Figure 2.17 shows the z-score over all periods for both the random graph and the configuration models. As the z-scores of the random graphs are very negative and large, and those of the configuration model indistinguishable from 0, the observed CP structure can be completely explained by the in- and outdegree sequences.

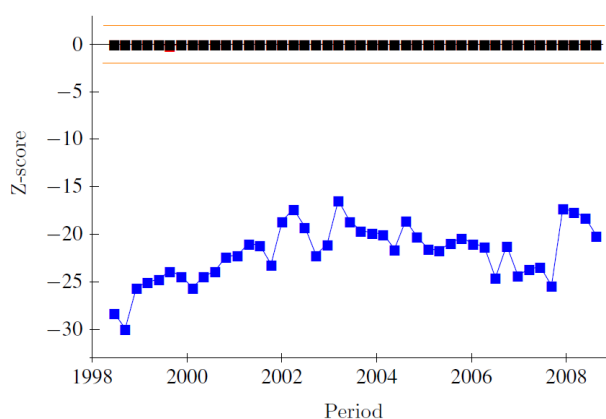


Figure 2.17: Z-scores for the number of errors under the ER model (blue, down) and the configuration model (black, up).

The z-scores very close to 0 might seem conspicuous, but they do have a meaning. Note that the errors in the CP model equal the number of existing links in the periphery plus the number of missing links in the core up to a constant (assuming Condition CP3 holds). The observed error score can in expectation be reproduced exactly by the degree distributions  $\{k_i^{out}\}$  and  $\{k_i^{in}\}$  and the core membership for all banks  $i$ . The variance of this number is small yet non-zero, as the number of links within the core or the periphery can change according to the probabilities  $p_{ij}$ . This results in a 0 z-score.

From this test we can conclude that there is no additional evidence for the CP structure apart from the degree distribution. It is important to note that in a perfect CP structure, there are no rewirings possible that would generate errors from this network. The simulated distribution function would therefore consist of a single point, and z-scores would be undefined. Testing the CP structure against the configuration model can therefore not give any indication of whether the CP structure is plausible in itself.

# Chapter 3

## The formation of a core periphery structure in heterogeneous financial networks

### 3.1 Introduction

The extraordinary events of 2007 and 2008 in which the financial system almost experienced a global meltdown, have led to an increased interest in the role of the interconnectedness of financial institutions and markets on systemic risk, the risk that liquidity or solvency problems in one financial institution spreads to the whole financial sector. This interest followed on earlier studies of financial networks that arose after the bailout of LTCM in 1998, in which one hedge fund, LTCM, had such a large trading position that its failure threatened to bring down the whole financial sector.

Starting with Allen and Gale (2000), a number of authors have argued that there is a non-monotonic relation between the interconnectedness of the financial system and systemic risk. For low connectivity a bank default leads to a local cascade, but not a global one. For high connectivity, diversification has a dampening effect on contagion, but for intermediate connectivity

the default of one bank can lead to the failure of the whole financial system. This non-monotonic effect of connectivity has been extended to different types of financial contagion mechanisms, see e.g. Gai and Kapadia (2010), Gai, Haldane, and Kapadia (2011), Amini, Cont, and Minca (2012), Battiston, Delli Gatti, Gallegati, Greenwald, and Stiglitz (2012) and Georg (2013). In a survey on systemic risk and financial contagion, Chinazzi and Fagiolo (2013) summarises the main finding in this literature as the ‘robust-yet-fragile’ property of the financial system. Connections can serve at the same time as shock-absorbers and shock-amplifiers.

Apart from theoretical and simulation analysis, there has been increasing research on stress tests of actual financial systems, see Upper (2011) for a review. However, both the theoretical, simulation and the stress test analysis assumes that the network of financial interconnections is *exogenously fixed*. This assumption has helped to focus on the causal relation between the financial network structure, that is, the pattern of financial interconnections, and systemic risk. Nevertheless, the assumption ignores the fact that financial networks do not come out of the blue. Financial interconnections are formed consciously by financial institutions who borrow, lend and trade financial assets which each other in order to maximise profits.

This raises the question what kind of network structure financial networks have, and how financial networks are formed. Regarding the first question, recent empirical evidence suggests that interbank markets have a *core periphery structure*, as found in Chapter 2 for the Netherlands and by Craig and von Peter (2014) and Fricke and Lux (2012) for other countries. In such networks the core is highly interconnected, while peripheral banks have only few connections, mostly with banks in the core. In this chapter we aim to give insight into the second question: why does the financial network have a core periphery network structure? That is, we aim to give theoretical foundations for the empirical finding of a core periphery network structure, in order to increase our understanding of the economic forces shaping financial networks.

We present a network formation model of a financial market with an explicit role for *intermediation* or *brokerage*. The model extends the results of Goyal and Vega-Redondo (2007) and Babus (2011). Unlike those papers, we follow Siedlarek (2012a) in the possibility of imperfect

competition between intermediators, and allow for free entry of intermediators. The aim is to explore whether the core periphery structure can be explained by interbank intermediation of core banks between periphery banks.

We identify which networks arise in equilibrium if banks optimally form links with other banks for interbank trading. We first consider the case in which banks are homogeneous. We find that the efficient ‘star’ network with one central counterparty becomes unstable if the costs for linking decrease, because peripheral banks have an incentive to replicate the position of the central player in order to receive intermediation benefits. We show that best-response dynamics converge to a unique stable outcome that ranges intuitively from an empty network (if linking costs are high) to a complete network (if linking costs are low).

Although the set of stable networks under different parameter values is rich, we find that the core periphery network cannot be stable. The intuition for this result is that the core periphery network creates an unsustainable *inequality between core and periphery banks*. Core banks receive large benefits from intermediating between periphery banks, especially if the number of banks in the system is large. Periphery banks have therefore often an incentive to enter the core, even if competition between intermediators reduces their benefits. Higher linking costs may deter periphery banks from replicating the core banks’ position, but at the same time diminish the incentives of core banks to hold relationships with all other banks in the core.

We then introduce heterogeneity in our model. We consider two types of banks, big banks and small banks, and allow big banks to have more frequent trading opportunities. We find that for sufficiently large differences between big and small banks, it becomes beneficial for large banks to have direct lending relationships with all other large banks in the core, such that the core periphery network becomes a stable structure.

This chapter is related to two main strands of literature. First the literature on network formation, see Goyal (2009) for a general overview. Most closely related to this chapter is Goyal and Vega-Redondo (2007). They consider a model in which each every pair of players undertakes a transaction, possibly through intermediation by other players. It is shown that

a star network arises with a single broker intermediating between all other traders. However, Goyal and Vega-Redondo (2007) use a limit case in which two or more brokers cannot earn intermediation rents simultaneously from a pair of traders: hence a core periphery network with multiple brokers is not found to be stable. Babus (2011) specifies a more detailed model for the interbank market, but also finds as the main outcome a star network with one central counterparty. These models mimic the observed inequality in the distribution of trades, but are too stylised to come close to empirically observed structures.<sup>1</sup>

Other network formation papers have found more complex structures, but typically ones that are (unrealistically) regular, that is, where all banks have (almost) the same number of links. Buskens and van de Rijt (2008) and Kleinberg, Suri, Tardos, and Wexler (2008) both find multipartite graphs using two different underlying mechanisms. Acemoglu, Ozdaglar, and Tahbaz-Salehi (2013) use an alternative formation scheme in which bank post contingent debt contracts to all other potential counterparties, and find that equilibrium networks have rings.<sup>2</sup>

The second relevant branch of literature is work on trading networks. Corominas-Bosch (2004) and Kranton and Minehart (2001) are early examples that assume a bipartite network. This approach rules out possible intermediation. Gale and Kariv (2007) are the first to model trading networks that do take intermediation into account. Blume, Easley, Kleinberg, and Tardos (2009) is a related paper that uses a similar distribution of trading surpluses as Goyal and Vega-Redondo (2007), and finds that equilibrium outcomes are always efficient. Gofman (2011) however shows that equilibrium pricing outcomes for an exogenous network are generally not efficient, unless players make a take-it-or-leave-it offer (as in Gale and Kariv, 2007), or the

---

<sup>1</sup> Bedayo, Mauleon, and Vannetelbosch (2013) extend the result of Babus (2011) by heterogeneous discount rates: weak (or impatient) players have a lower discount rate than strong players. The authors find that any stable network is a core periphery network in which the core consists of all weak players. In this chapter we focus on heterogeneity in size, and assume that big banks have more frequent trading opportunities. We show under which conditions a core periphery network with a core of big banks can form.

<sup>2</sup> In the social science literature some papers explain core periphery networks. These network formation models are typically concerned with optimal effort levels to account for peer effects. Galeotti and Goyal (2010) and Hiller (2012) provide conditions under which core periphery networks are the only stable network structure. See also Persitz (2009) who adopts heterogeneity in the connections model of Jackson and Wolinsky (1996). This literature cannot so easily be translated to financial networks, because of the different interpretation of links as (channels for) financial transactions. See Cohen-Cole, Patacchini, and Zenou (2011) and König, Tessone, and Zenou (2014) who propose related models with nestedness for interbank networks.



network is complete.<sup>3</sup>

The rest of this chapter is organised as follows. In Section 3.2 we introduce our model with the pay-off function and the stability concepts. Section 3.3 treats the basic structures and includes the formal definition of a core periphery network. Our main results are presented separately for homogeneous traders (Section 3.4) and heterogeneous traders (Section 3.5). In Section 3.6 we calibrate our model to the interbank market in the Netherlands. Section 3.7 concludes.

## 3.2 Network formation model

Consider an undirected network  $g$  between a set of agents  $N$ , having cardinality  $n = |N|$ . Denote by  $g_{ij} = g_{ji} = 1$  the existence of a trading relationship, and by  $g_{ij} = g_{ji} = 0$  the absence of it. Network  $g$  results in a payoff vector  $\{\pi_i(g)\}_{i \in N}$  to each agent  $i$  in  $N$ , as follows.

### 3.2.1 Payoffs in a financial network with intermediators

Each pair  $ij$  in the network creates a *potential trade surplus* of  $\alpha_{ij}$ . This trade surplus can only be realised if  $i$  and  $j$  are connected directly, or indirectly through a path with one middleman. It is possible that multiple of these short paths (of length 2) exist: the number of middlemen linking  $i$  and  $j$  is denoted as  $m_{ij}$ . The trade surplus  $\alpha_{ij}$  does not depend on whether  $i$  and  $j$  are directly or indirectly connected, or on the number of middlemen  $m_{ij}$ .

If the path length between  $i$  and  $j$  is more than 2, that is, if  $i$  and  $j$  can only be connected through a chain of at least two middlemen, we assume that trade *cannot* be realised. This assumption keeps the analysis tractable. Also, intermediation over longer routes in financial networks seems empirically and intuitively less relevant.<sup>4</sup> Our main result – core periphery networks are generally unstable under homogeneity ( $\alpha_{ij} = 1$  for all  $i, j \in N$ ), but can be stable if agents are heterogeneous – does not depend on this assumption as shown in Appendices 3.C

<sup>3</sup> See also Choi, Galeotti, and Goyal (2013)

<sup>4</sup> The same assumption was made by Kleinberg, Suri, Tardos, and Wexler (2008) and Siedlarek (2012b).

and 3.D.

We assume that, if realised, the trade surplus of  $i$  and  $j$  is divided as follows. If  $i$  and  $j$  are directly connected, then  $i$  and  $j$  each receive half of the surplus. If  $i$  and  $j$  are indirectly connected by  $m_{ij}$  middlemen, then the endnodes  $i$  and  $j$  each receive a share of  $f_e(m_{ij}, \delta)$ , whereas each of the  $m_{ij}$  middlemen receives a share of  $f_m(m_{ij}, \delta)$ . Note that by definition:

$$2f_e(m_{ij}, \delta) + m_{ij}f_m(m_{ij}, \delta) = 1. \quad (3.1)$$

Apart from the number of middlemen,  $m_{ij}$ , the division also depends on a parameter  $\delta \in [0, 1]$ . This variable captures the level of *competition* between the number of middlemen  $m_{ij}$  of a certain trade between  $i$  and  $j$ , and leads to the following payoffs.

If there is one middleman,  $m_{ij} = 1$ , then  $i$ ,  $j$ , and the middleman split the surplus evenly, hence,

$$f_e(1, \delta) = f_m(1, \delta) = \frac{1}{3}. \quad (3.2)$$

If there is more than only one middleman, then the share that the agents get depends on the amount of competition. If there is no competition,  $\delta = 0$ , the middlemen collude, and  $i$  and  $j$  obtain the same share as if there was one intermediary. The middlemen share the intermediation benefits evenly, that is, for all  $m_{ij} \in \{2, \dots, n-2\}$ :

$$f_e(m_{ij}, 0) = \frac{1}{3} \text{ and } f_m(m_{ij}, 0) = \frac{1}{3m_{ij}}. \quad (3.3)$$

In the case of perfect competition,  $\delta = 1$ , between  $m_{ij} > 1$  intermediaries, their payoffs will always disappear, hence

$$f_e(m_{ij}, 1) = \frac{1}{2} \text{ and } f_m(m_{ij}, 1) = 0. \quad (3.4)$$

The same happens in the limit of the number of middlemen to infinity. A further straightforward assumption is monotonicity with respect to the competitiveness and number of middlemen.

Table 3.1 summarises the assumed dependencies of the functions to the parameters. Fig-

Share of payoffs for:	endnodes	middlemen
Level of competition $\delta$	$f_e(\cdot, 0) = \frac{1}{3}$	$f_m(\cdot, 0) = \frac{1}{3m}$
	$\frac{\partial f_e}{\partial \delta} > 0$	$\frac{\partial f_m}{\partial \delta} < 0$
	$f_e(\cdot, 1) = \frac{1}{2}$	$f_m(\cdot, 1) = 0$
Number of middlemen $m$	$f_e(1, \cdot) = \frac{1}{3}$	$f_m(1, \cdot) = \frac{1}{3}$
	$\frac{\partial f_e}{\partial m} > 0$	$\frac{\partial f_m}{\partial m} < 0$
	$\lim_{m \rightarrow \infty} f_e(m, \cdot) = \frac{1}{2}$	$\lim_{m \rightarrow \infty} f_m(m, \cdot) = 0$

 Table 3.1: Assumptions about the payoff shares  $f_e(m_{ij}, \delta)$  and  $f_m(m_{ij}, \delta)$ .

Figure 3.1 gives examples of payoff shares received by different agents involved in a trade between  $i$  and  $j$  for different levels of competition  $\delta$  and different numbers of middlemen  $m_{i,j}$ .

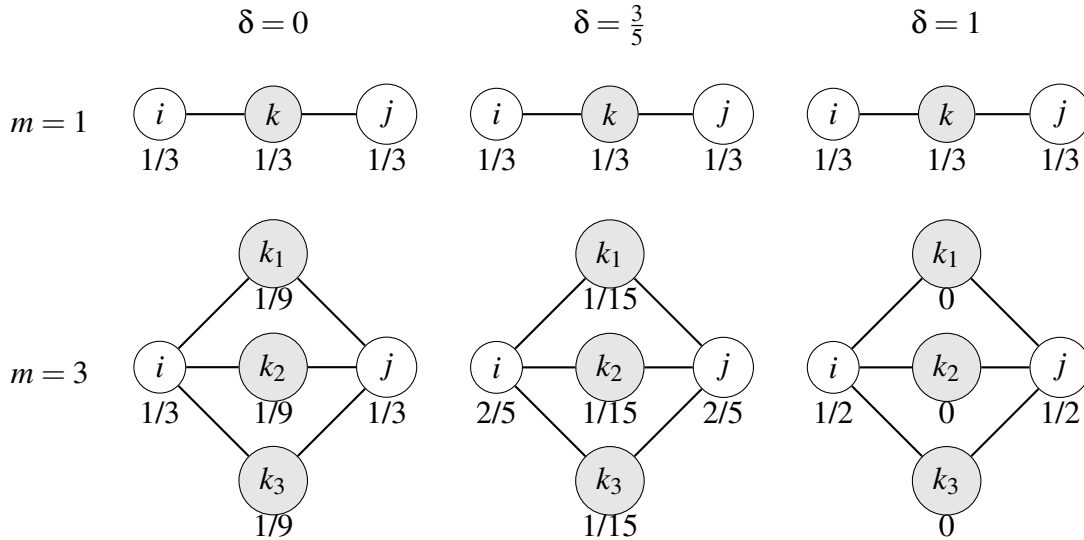


Figure 3.1: Examples of payoff shares received by endnodes  $i$  and  $j$  and intermediaries  $k$ , depending on the parameters  $\delta$  and  $m$  under the specification of  $f_e$  and  $f_m$  in equation (3.6), as in Siedlarek (2012a).

Finally, we assume that each link between  $i$  and  $j$  involves a cost  $c$  to both  $i$  and  $j$ . The payoff function for an agent  $i$  in graph  $g$  is then the sum of benefits agent  $i$  obtains from the

trades it is involved in, minus the cost of linking:

$$\pi_i(g) = \sum_{j \in N_i^1(g)} \left( \frac{1}{2} \alpha_{ij} - c \right) + \sum_{j \in N_i^2(g)} \alpha_{ij} f_e(m_{ij}, \delta) + \sum_{j, l \in N_i^1(g) | g_{jl} = 0} \alpha_{jl} f_m(m_{jl}, \delta), \quad (3.5)$$

where  $N_i^r(g)$  denotes the set of nodes at distance  $r$  from  $i$  in network  $g$ . Here the first term denotes the benefits of direct trades minus the cost of maintaining a link, the second term denotes the benefits of indirect trades through middlemen, and the third term denotes intermediation benefits.<sup>5</sup>

We now comment on the interpretation of these payoffs, in particular how trade surpluses  $\alpha_{ij}$ , links  $g_{ij}$ , linking costs  $c$  and surplus distributions  $f_e$  and  $f_m$  can be translated back to an interbank market setting. First, trade possibilities in the interbank market arise when one bank has a liquidity shortage and another bank has a liquidity surplus, relative to its targeted liquid asset holdings. These shortages and surpluses arise expectedly and unexpectedly to the banks through their daily operations. The trade surpluses in our model should be seen as the net present values of the benefits from the realisation of arising trade opportunities. In Appendix 3.A we formalise this interpretation of trade surpluses.

Second, the undirected links in this model should be interpreted as established preferential trading relationships. The existence of such preferential trading relationships has been shown by Cocco, Gomes, and Martins (2009) in the Portuguese interbank market, by Bräuning and Fecht (2012) in the German interbank market and by Afonso and Lagos (2012) in the U.S. federal funds market. We assume that trade opportunities can only be realised if the agents are linked directly or indirectly through mutual trading relationships. This is a strong, but not implausible assumption. A bank that attempts to borrow outside its established trading relationships may signal that it is having difficulties to obtain liquidity funding and, as a consequence, may face higher borrowing costs. Hence, banks have incentives to use their established trading relationships.

---

<sup>5</sup> The payoff function of Goyal and Vega-Redondo (2007) is a special case of ours with homogeneity ( $\alpha_{ij} = 1 \forall i, j$ ) and perfect competition ( $\delta = 1$ ).

Third, it is assumed that the preferential trading relationship comes at a fixed cost. This cost follows from establishing mutual trust and from monitoring, i.e. assessing the other bank's risks. In principle, it is possible that these costs are not constant over banks, e.g. economies of scale may decrease linking costs in the number of relationships that are already present. The possible heterogeneity in linking costs is most likely smaller than the heterogeneity in trading surpluses, and we therefore assume that all banks pay an equal(ly small) cost for a trading relationship.

Fourth, the division of the trade surplus and intermediation benefits in our model follow from a bargaining protocol. We do not model this bargaining process explicitly. However, our assumptions on  $f_e(\cdot)$  and  $f_m(\cdot)$  generalise an explicit Rubinstein-type of bargaining process developed by Siedlarek (2012a). From his bargaining protocol, in which  $\delta$  has the interpretation of a discount factor, the distribution of the surplus is given by

$$f_e(m, \delta) = \frac{m - \delta}{m(3 - \delta) - 2\delta} \text{ and } f_m(m, \delta) = \frac{1 - \delta}{m(3 - \delta) - 2\delta}. \quad (3.6)$$

It is easily checked that (3.6) satisfies the assumptions we made on  $f_e(\cdot)$  and  $f_m(\cdot)$ . We use this explicit function to illustrate our (more general) results by Figures 3.7, 3.8 and 3.9. In Appendix 3.B, we give a detailed explanation of the specification of payoffs by Siedlarek (2012a).

### 3.2.2 Network stability concepts

Given the setup of the payoffs discussed above, we analyse which networks arise if agents form links strategically. Here we assume that for the establishment of a link between two agents, both agents have to agree, and both agents face the cost of a link, a version of network formation that is called *two-sided network formation*. The tools for this analysis come from network formation theory, see Goyal (2009) for a text book discussion. Network formation theory has developed stability or equilibrium concepts to analyse the *stability* of a network. Here stability does not refer to systemic risk, but to the question whether an agent or a pair of agents has an incentive and the possibility to modify the network in order to receive a higher payoff.

There are many stability concepts, which differ in the network modifications allowed. For an overview of these stability concepts, we refer to Jackson (2005) or Goyal (2009). For our purposes we consider two stability concepts.

The first concept is *pairwise stability*, originally introduced by Jackson and Wolinsky (1996). A network is pairwise stable if for all the links present, no player benefits from deleting the link, and for all the links absent, one of the two players does not want to create a link. Denote the network  $g + g_{ij}$  as the network identical to  $g$  except that a link between  $i$  and  $j$  is added. Similarly, denote  $g - g_{ij}$  as the network identical to  $g$  except that the link between  $i$  and  $j$  is removed. Then the definition of pairwise stability is as follows:

**Definition 3.1.** *A network  $g$  is pairwise stable if for all  $i, j \in N, i \neq j$ :*

$$(a) \text{ if } g_{ij} = 1, \text{ then } \pi_i(g) \geq \pi_i(g - g_{ij}) \wedge \pi_j(g) \geq \pi_j(g - g_{ij});$$

$$(b) \text{ if } g_{ij} = 0, \text{ then } \pi_i(g + g_{ij}) > \pi_i(g) \Rightarrow \pi_j(g + g_{ij}) < \pi_j(g).$$

The concept of pairwise stable networks only allows for deviations of one link at a time. This concept is often too weak to draw distinguishable conclusions, i.e. in many applications including ours, there are many networks that are pairwise stable.

In our application, we consider it relevant that agents may consider to propose many links simultaneously in order to become an intermediary and establish a client base. The benefits from such a decision may only become worthwhile if the agent is able to create or remove more than one link. This leads us naturally to the concept of *unilateral stability*, originally proposed by Buskens and van de Rijt (2008).

A network is unilaterally stable if no agent  $i$  in the network has a profitable unilateral deviation: a change in its links by either deleting existing links such that  $i$  benefits, or proposing new links such that  $i$  and all the agents to which it proposes a new link benefit. Denote  $g^{iS}$  as the network identical to  $g$  except that all the links between  $i$  and every  $j \in S$  are altered by  $g_{ij}^{iS} = 1 - g_{ij}$ , i.e. are added if absent in  $g$  or are deleted if present in  $g$ .

**Definition 3.2.** A network  $g$  is unilaterally stable if for all  $i$  and for all subsets of players  $S \subseteq N \setminus \{i\}$ :

- (a) if  $\forall j \in S : g_{ij} = 1$ , then  $\pi_i(g) \geq \pi_i(g^{iS})$ ;
- (b) if  $\forall j \in S : g_{ij} = 0$ , then  $\pi_i(g^{iS}) > \pi_i(g) \Rightarrow \exists j : \pi_j(g^{iS}) < \pi_j(g)$ .

Note that unilateral stability implies pairwise stability, that is, a network that is unilaterally stable is also pairwise stable, but not vice versa. This can be easily verified by considering subsets  $S$  that consist of only one node  $j \neq i$ .<sup>6</sup>

Apart from analysing the stability of networks, for policy considerations it is also relevant to consider efficient networks. As usual in the literature, we define a network efficient, if it maximises the total sum of payoffs of the agents.

**Definition 3.3.** A network  $g$  is efficient if there is no other network  $g'$ , such that

$$\sum_{i \in N} \pi_i(g') > \sum_{i \in N} \pi_i(g). \quad (3.7)$$

### 3.3 Basic structures

Our analysis revolves around a couple of relevant network structures, which we define here. Denote the *empty network*,  $g^e$ , as the network without any links, i.e.  $\forall i, j \in N : g_{ij} = 0$ , and the *complete network*,  $g^c$ , as the network with all possible links, i.e.  $\forall i, j \in N : g_{ij} = 1$ . A *star network*,  $g^s$  (see Figure 3.2), has a single player, the centre of the star, that is connected to all other nodes, and no other links exist:  $\exists i$  such that  $\forall j \neq i : g_{ij} = 1$  and  $\forall j, k \neq i : g_{jk} = 0$ .

A *core periphery network* is a network, in which the set of agents can be partitioned in a core and a periphery, such that all agents in the core are completely connected within and are

<sup>6</sup> Our definition of unilateral stability is slightly less restrictive than the definition in Buskens and van de Rijt (2008). While we are following Buskens and van de Rijt (2008) in the simultaneity in deleting *or* proposing multiple links, we do not allow simultaneous deleting *and* proposing of multiple links. This adaptation eases the exposition, but does not affect our results qualitatively.

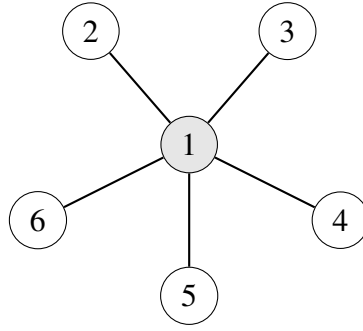


Figure 3.2: A star network with  $n = 6$  players

linked to some periphery agents, and all agents in the periphery have at least one link to the core, but no links to other periphery agents.

**Definition 3.4.** A network  $g$  is a core periphery network, if there exists a set of core agents  $K \subset N$  and periphery agents  $P = N \setminus K$ , such that:

- (a)  $\forall i, j \in K : g_{ij} = 1$ , and  $\forall i, j \in P : g_{ij} = 0$ ;
- (b)  $\forall i \in K : \exists j \in P$  with  $g_{ij} = 1$ , and  $\forall j \in P : \exists i \in K$  with  $g_{ij} = 1$ .

See Figure 3.3 for an example. A special case of a core periphery network is the *complete core periphery network*, where each agent in the core  $K$  is linked to *all* agents in the periphery  $P$ :  $\forall i \in K$  and  $\forall j \in P$  it holds that  $g_{ij} = 1$ . See Figure 3.4 for an example. We denote a complete core periphery network with  $k = |K|$  agents in the core as  $g_{com}^{CP(k)}$ .<sup>7</sup>

Finally, a *complete multipartite network* is a network, in which the agents can be partitioned into  $q$  groups,  $N = \{K_1, K_2, \dots, K_q\}$ , such that nodes do not have links within their group, but are completely connected to all nodes outside their own group. Formally, in a complete multipartite network it holds that  $\forall m \in \{1, 2, \dots, q\} : \forall i \in K_m$  we have  $\forall j \in K_m : g_{ij} = 0$  and  $\forall j \notin K_m : g_{ij} = 1$ . Complete multipartite networks will be denoted as  $g_{k_1, k_2, \dots, k_q}^{mp(q)}$ , where  $k_m \equiv |K_m|$  is the size of the  $m$ -th group. Multipartite networks are called *balanced* if the group sizes are as equal as

<sup>7</sup> By Definition 3.4, empty, star and complete networks are special cases of core periphery networks with cores of size  $k = 0$ ,  $k = 1$  and  $k = n$  respectively. A complete core periphery networks with  $k = n - 1$  is also identical to a complete network. In discussing our results we will make clear when we are speaking of non-trivial core periphery networks with  $k \in \{2, 3, \dots, n - 2\}$ .



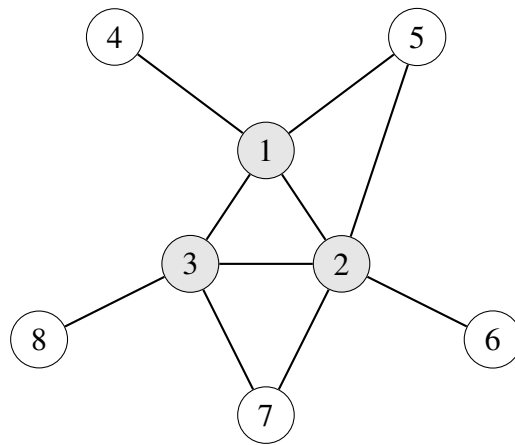


Figure 3.3: A core periphery network with  $n = 8$  players, of which  $k = 3$  are in the core

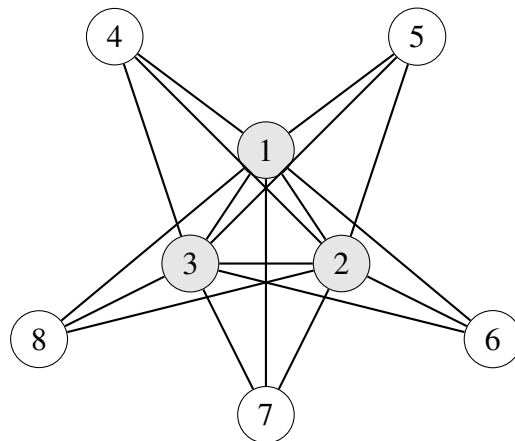


Figure 3.4: A complete core periphery network with  $n = 8$  players, of which  $k = 3$  are in the core

possible, i.e.  $|k_m - k_{m'}| \leq 1$  for all  $m, m'$ . Figure 3.6 in Section 3.4.3 presents examples of complete multipartite networks that arise in our model.

## 3.4 Results for homogeneous traders

We now analyse the model for the baseline homogeneous case where all pairs generate the same trade surplus,  $\alpha_{ij} = 1$  for all  $i, j \in N$  with  $i \neq j$ .<sup>8</sup> After briefly discussing efficient networks, we analyse the model using the two stability concepts described above, pairwise and unilateral stability. In Section 3.4.2 we develop the main result that core periphery networks are not unilaterally stable under the assumption of homogeneity. To find what structures arise in this homogeneous case, if not core periphery networks, we investigate a best-response dynamic process in Section 3.4.3.

### 3.4.1 Efficient networks

We start with a description of the efficient network. Following the results of Goyal and Vega-Redondo (2007), minimally connected networks, i.e. with  $n - 1$  links, are efficient for  $c < \frac{n}{4}$ . Networks are efficient if all trade surpluses are realized irrespective of the distribution of these trading surpluses. For higher  $c$  it is efficient to have no network at all, i.e. an empty network. Because of the assumption that two agents only trade if they are at distance 1 or 2, the network should not have a maximal distance higher than 2, leaving the star as the unique efficient minimally connected network. This is summarised in the following theorem.

**Theorem 3.1.** *If the payoff function  $\pi(g)$  is given by equation (3.5) with  $\alpha_{ij} = 1$  for all  $i, j \in N$ , then:*

- (a) *If  $c \geq n/4$ , then the empty network is efficient.*
- (b) *If  $c \leq n/4$ , then the star network is efficient.*

---

<sup>8</sup>The normalisation  $\alpha = 1$  in the homogeneous case goes without loss of generality. If  $\alpha_{ij} = \alpha$  for all  $i, j$  with  $\alpha > 0$ , then the payoffs in equation (3.5) are proportional to  $\alpha$  when costs  $c$  are considered as a fraction of  $\alpha$ .

(c) *No other network structure than the empty or star network is efficient.*

Theorem 3.1 implies that core periphery networks with  $k \geq 2$  are not efficient.

### 3.4.2 Stability of basic structures

A natural starting point of the analysis is the stability of the empty, star and complete networks. The easiest way to find under which conditions these networks are unilaterally stable is by introducing the best feasible (unilateral) action of a certain player  $i$ . An action of player  $i$  is feasible if its proposed links to every  $j \in S$  are accepted by every  $j \in S$ , or if  $i$  deletes all links with every  $j \in S$ . This action changes the network  $g$  into  $g^{iS}$ . The best feasible action is formally defined as follows.

**Definition 3.5.** *A feasible action for player  $i$  is represented by a subset  $S \subseteq N \setminus \{i\}$  with:*

- (a)  $\forall j \in S : g_{ij} = 1$ , or:
- (b)  $\forall j \in S : g_{ij} = 0$  and  $\pi_j(g^{iS}) \geq \pi_j(g)$ .

*The best feasible action  $S^*$  is the feasible action that gives  $i$  the highest payoffs:*

$$\forall S \subseteq N \setminus \{i\} : \pi_i(g^{iS^*}) \geq \pi_i(g^{iS}).$$

In an empty network, the best feasible action for a player is to connect either to all other players or to none, as shown in the following lemma.

**Lemma 3.1.** *In an empty network that is not unilaterally stable, the best feasible action for a player is to add links to all other nodes.*

**Proof.** The marginal benefits<sup>9</sup> of adding  $l$  links to an empty network are:

$$M_i(g^e, +l) = l\left(\frac{1}{2} - c\right) + \binom{l}{2} \frac{1}{3}. \quad (3.9)$$

As this function is convex in  $l$ , maximising the marginal benefits over  $l$  results in either  $l^* = 0$  or  $l^* = n - 1$ .  $\square$

The star is unstable for low costs  $c$ , because periphery members gain incentives to add links. The gain of having a direct (rather than an intermediated) trade are in this case  $\frac{1}{2} - \frac{1}{3} = \frac{1}{6}$ . However, it is even better for one periphery member to propose multiple links, thereby creating more chains of neighbours through which it intermediates.

**Lemma 3.2.** *Consider a star network that is not unilaterally stable because a peripheral player  $i$  can deviate by adding one or more links. Then the best feasible action is to add links to all  $l = n - 2$  other periphery players. In other words, it is never a best feasible action to add  $0 < l < n - 2$  links to some other peripheral nodes.*

**Proof.** Consider the deviation of a peripheral player  $i$  to add  $l$  links. The marginal benefits consist of making trades direct instead of intermediated by the centre of the star, and of creating middlemen benefits:

$$M_i(g^s, +l) = l\left(\frac{1}{6} - c\right) + \binom{l}{2} f_m(2, \delta). \quad (3.10)$$

As in Lemma 3.1 this function is convex in  $l$ , so the maximum is reached at  $l^* = 0$  or  $l^* = n - 2$ .  $\square$

---

<sup>9</sup> Throughout this chapter, marginal benefits of an action  $S$  by player  $i$  for player  $j$  (not necessarily equal to  $i$ ) are defined as:

$$M_j(g, S) \equiv \pi_j(g^{iS}) - \pi_j(g). \quad (3.8)$$

To simplify the notation, we will informally denote the action  $S$  as  $+l$  or  $-l$ , where  $l = |S|$  is the number of links that is either added or deleted in the network. From the text it will be clear which player  $i$  and which set  $S$  are considered.

Lemmas 3.1 and 3.2 imply that to find the stability conditions of empty and star networks, it is sufficient to check whether adding or deleting all possible links is beneficial in these networks. Proposition 3.1 below specifies the stability conditions for empty, star and complete networks under homogeneous agents. Figure 3.7 in Section 3.4.3 depicts the stability regions I, II, and III of the empty, star and complete networks in the  $(\delta, c)$ -space for  $n = 4$  and  $n = 8$ .

**Proposition 3.1.** *In the homogeneous baseline model with  $\alpha_{ij} = 1$  for all  $i, j \in N$  the following networks are stable:*

I: *The empty network is unilaterally stable if and only if  $c \geq \frac{1}{2} + \frac{1}{6}(n-2)$ .*

II: *The star network is unilaterally stable if and only if*

$$c \in \left[ \frac{1}{6} + (n-3) \min \left\{ \frac{1}{2} f_m(2, \delta), f_e(2, \delta) - \frac{1}{3} \right\}, \frac{1}{2} + \frac{1}{6}(n-2) \right].$$

III: *The complete network is unilaterally stable if and only if  $c \leq \frac{1}{2} - f_e(n-2, \delta)$ .*

**Proof.** As in Goyal and Vega-Redondo (2007), the centre of the star earns more than the periphery members, but receives less profit per link. The star therefore becomes unilaterally unstable if the centre of the star would benefit from deleting all its links, i.e.

$$\begin{aligned} M_i(g^s, -(n-1)) = -\pi_i(g^s) &< 0 \\ \Rightarrow c &> \frac{1}{2} + \frac{1}{6}(n-2). \end{aligned} \quad (3.11)$$

If this condition holds the empty network is unilaterally stable.

Lemma 3.2 states that if a peripheral player  $i$  in a star network wants to deviate by adding links, this player will become a member of a newly formed core. In other words, a complete core periphery network arises with  $k = 2$ . Player  $i$  becomes directly connected to all other players, and also receives intermediation gains which depend on  $\delta$ . For the new core member  $i$

to have positive marginal benefits of supporting all new links, it is required that

$$\begin{aligned} M_i(g^s, +(n-2)) &> 0 \\ \Rightarrow c &< \frac{1}{6} + \frac{1}{2}(n-3)f_m(2, \delta), \end{aligned} \quad (3.12)$$

and every remaining peripheral player  $j \neq i$  requires

$$\begin{aligned} M_j(g^s, +(n-2)) = \frac{1}{6} - c + (n-3)(f_e(2, \delta) - \frac{1}{3}) &\geq 0 \\ \Rightarrow c &\leq \frac{1}{6} + (n-3)(f_e(2, \delta) - \frac{1}{3}). \end{aligned} \quad (3.13)$$

Conversely, the star network is stable if the deviation of player  $i$  is *not* beneficial to  $i$  and/or a peripheral player  $j \neq i$ , i.e. if  $c$  exceeds the minimum of the two values in equations (3.12) and (3.13):

$$c \geq \frac{1}{6} + (n-3) \min\left\{\frac{1}{2}f_m(2, \delta), f_e(2, \delta) - \frac{1}{3}\right\}. \quad (3.14)$$

A complete network can become unstable when any two nodes decide to remove a mutual link. Note that the trade between those two links is sustained by intermediation through all other nodes. Formally we have therefore a core periphery network with  $k = n - 2$  core players. The complete network is therefore stable if:

$$c \leq \frac{1}{2} - f_e(n-2, \delta). \quad (3.15)$$

□

Given the stability conditions for empty, star and complete networks in Proposition 3.1, we see that these networks cannot be stable at the same time, except for borderline cases such as  $c = \frac{1}{2} + \frac{1}{6}(n-2)$ . In Section 3.4.3, we will show that if these networks are stable, they are also

the unique outcome of a best-response dynamic process.

We now consider the stability of core periphery networks, starting with complete core periphery networks. The profits in a complete core periphery network where  $k$  core members are connected to all other nodes are:

$$\pi_i(g_{com}^{CP(k)}) = \begin{cases} (n-1)(\frac{1}{2}-c) + \binom{n-k}{2}f_m(k, \delta) & \text{if } i \in C \\ k(\frac{1}{2}-c) + (n-k-1)f_e(k, \delta) & \text{if } i \in P. \end{cases} \quad (3.16)$$

The reason we focus on this special type of complete core periphery networks is that they are formed naturally as peripheral agents try to improve their position by avoiding payments to core players and linking to other peripheral players. As it was optimal in an empty graph and in a star network to link immediately to all players (Lemmas 3.1 and 3.2), if a peripheral agent will add a link it is optimal to connect to all other players, thereby entering in the core himself.

**Lemma 3.3.** *Consider a complete core periphery network with  $k \in \{2, 3, \dots, n-2\}$  that is not unilaterally stable because a peripheral player  $i$  can deviate by adding one or more links. Then the best feasible action is to add links to all  $l = n - k - 1$  other periphery players. In other words, it is never a best feasible action to add  $0 < l < n - k - 1$  links to some other peripheral nodes.*

**Proof.** The marginal benefits for replicating the position of the centre of the star was given in Lemma 3.2. It can easily be generalised to complete core periphery networks with  $k$  core members:

$$M_i(g_{com}^{CP(k)}, +l) = l(\frac{1}{2} - f_e(k, \delta) - c) + \binom{l}{2}f_m(k+1, \delta). \quad (3.17)$$

The maximum of this function is reached at  $l^* = 0$  or  $l^* = n - k - 1$ . □

Lemmas 3.2 and 3.3 show that peripheral players have an incentive to replicate the position of core players to raise their profits. Stable networks must therefore exhibit limited inequality be-

tween players in different network positions. This intuitive economic phenomenon, introduced in our model by using unilateral stability as stability concept, will drive the results for the core periphery network below.

Although there is an incentive for periphery players to become a new core member if the cost  $c$  are low enough, the resulting complete core periphery network is not stable. We now show that under fairly general conditions in the baseline homogeneous model *any* core periphery network (complete or incomplete) is not stable.

**Proposition 3.2.** *Let the payoff function be homogeneous ( $\alpha_{ij} = 1$  for all  $i, j \in N$ ) and let  $c > 0$  and  $0 < \delta < 1$  be given. Then:*

- (a) *complete core periphery networks with  $k \geq 2$  core members are not pairwise stable, and thus also not unilaterally stable, for all  $n > k + 2$ ;*
- (b) *incomplete core periphery networks with  $k \geq 1$  core member(s) are not unilaterally stable if  $n$  is sufficiently large. More precisely, there is a function  $F(c, \delta, k)$  such that if  $n > F(c, \delta, k)$ , the result holds.*

**Proof.** See Appendix 3.C. □

We emphasise that this result does not depend on the assumption that trade surpluses between  $i$  and  $j$  are only realised if the path length between  $i$  and  $j$  is less than 3, as shown in 3.C. This assumption is made to keep the analysis of our dynamic process tractable (see Section 3.4.3).

The intuition for Proposition 3.2 is that for homogeneous agents core periphery networks create an unsustainable inequality between core and periphery players. The proof of Proposition 3.2 relies on the idea that the differences in payoff between the core and periphery cannot be too large. Consider for example the payoffs for complete core periphery networks in equation (5.54). As the network size  $n$  increases for a fixed core size  $k$ , core members get more and more intermediation benefits (at a quadratic rate), while periphery payoffs fall behind (as it grows at a linear rate).

We now compare this result with that of Goyal and Vega-Redondo (2007). They show that



for  $c > \frac{1}{6}$ ,  $\delta = 1$  and  $n$  sufficiently large, the star network (i.e. a core periphery network with  $k = 1$ ) is the unique non-empty *bilaterally* stable network, stressing the importance of the star structure. Proposition 3.2 refines this result and indicates that for  $\delta < 1$  core periphery networks with  $k$  core players *including* star networks with  $k = 1$  are *not unilaterally* stable, even for values of  $\delta$  arbitrary close but below 1.

We argue that the crucial assumption for our result is imperfect competition, that is,  $\delta < 1$ . Imperfect competition creates profitable deviations for peripheral players to circumvent (large) intermediation fees to core players and to start receiving intermediation benefits. A further important difference in this chapter compared to Goyal and Vega-Redondo (2007) is the stability concept: in the proof of Proposition 3.2, we make explicit use of the fact that periphery players may add multiple links at the same time, which is not allowed in Goyal and Vega-Redondo (2007). Because we allow for unilateral deviations, peripheral players can replicate the position of core players in order to benefit from intermediation. Inequality between core and periphery players makes that core periphery networks cannot be stable for sufficiently large  $n$ .

### 3.4.3 A dynamic process

So far, we have analysed the unilateral stability of empty, star, complete and core periphery architectures. Core periphery networks were shown to be generally unstable under homogeneous agents. To investigate what other networks can plausibly be formed, we consider round-robin best-response dynamics as in Kleinberg, Suri, Tardos, and Wexler (2008), starting from an empty graph. We order nodes  $1, 2, \dots, n$  and in this order nodes consecutively try to improve their position by taking best feasible actions. After node  $n$  the second round starts again with node 1. The process is stopped if  $n - 1$  consecutive players can not improve their position, thereby agreeing to the deviation made by the last player.

The assumptions about the dynamic process are made primarily to ease the exposition and to simplify the analytical analysis. The advantage of starting in an empty network is that initially nodes can only add links to the network. A fixed round-robin order limits the number

of possibilities that has to be considered for every step in the process.<sup>10</sup> The outcome of the dynamic process must be a unilaterally stable network. We find that the best-response dynamics converge to a unique unilaterally stable network for every choice of the model parameters (except for borderline cases).<sup>11</sup>

Theorem 3.2 below shows which network structures result from the best-response dynamics. The parameter regions for which empty, star and complete networks are the attracting steady state coincide with conditions I, II and III in Proposition 3.1 for which these networks are stable. In the remaining parameter regions the attracting steady state cannot be a core periphery network (in line with Proposition 3.2) and turns out to be a multipartite network. The theorem singles out one special type of multipartite networks, called maximally unbalanced bipartite networks, for parameters satisfying condition IV, while parameters under condition V can lead to various types of multipartite networks. Figure 3.7 plots the parameter regions I to IV for  $n = 4$  (for which  $V = \emptyset$ ) and I to V for  $n = 8$  in the  $(\delta, c)$ -space.

**Theorem 3.2.** *Consider the homogeneous baseline model with  $\alpha_{ij} = 1$  for all  $i, j \in N$ . From an empty graph, the round-robin best-response dynamics converge to the following unilaterally stable equilibria:*

I: for  $c > \frac{1}{2} + \frac{1}{6}(n - 2)$

*the empty network,*

II: for  $c \in \left( \frac{1}{6} + (n - 3) \min \left\{ \frac{1}{2} f_m(2, \delta), f_e(2, \delta) - \frac{1}{3} \right\}, \frac{1}{2} + \frac{1}{6}(n - 2) \right)$

*the star network,*

III: for  $c < \frac{1}{2} - f_e(n - 2, \delta)$

*the complete network.*

---

<sup>10</sup> Houy (2009) considers a dynamic framework for the model of Goyal and Vega-Redondo (2007), in which randomly drawn agents make best responses to the network. Simulations for our model indicate that the results in Theorem 3.2 also hold for a random order of agents.

<sup>11</sup> In principle, a dynamic process may lead to cycles of improving networks (cf. Jackson and Watts, 2002; Kleinberg, Suri, Tardos, and Wexler, 2008), but Theorem 3.2 shows that this is not the case in our model.

- IV: for  $c \in \left( \frac{1}{2} - f_e(2, \delta) + (n-4) \min\{\frac{1}{2}f_m(3, \delta), f_e(3, \delta) - f_e(2, \delta)\}, \frac{1}{6} + (n-3) \min\{\frac{1}{2}f_m(2, \delta), f_e(2, \delta) - \frac{1}{3}\} \right)$   
 the maximally unbalanced bipartite network  $g_{2, n-2}^{mp(2)}$
- V: for  $c \in \left( \frac{1}{2} - f_e(n-2, \delta), \frac{1}{2} - f_e(2, \delta) + (n-4) \min\{\frac{1}{2}f_m(3, \delta), f_e(3, \delta) - f_e(2, \delta)\} \right)$   
 a multipartite network  $g_{k_1, k_2, \dots, k_q}^{mp(q)}$  with  $q \geq 2$  and  $|k_m - k_{m'}| < n-4$  for all  $m, m' \in \{1, 2, \dots, q\}$ .

**Proof.** See Appendix 3.C. □

Figure 3.5 shows the possible routes of best-response dynamics for  $n = 4$ , depending on the remaining parameters  $c$  and  $\delta$ . For such a low  $n$ , the only possible multipartite network is one that consists of  $q = 2$  groups of 2 nodes, which coincides with a ring of all 4 players. It is noteworthy that for large  $n$  the set of attained multipartite networks is quite diverse. Possible outcomes for  $n = 8$  are the maximally unbalanced bipartite network  $g_{2,6}^{mp(2)}$ , but also, for example, a less than maximally balanced bipartite network  $g_{3,5}^{mp(2)}$  or a balanced multipartite network  $g_{2,3,3}^{mp(3)}$  consisting of 3 groups. Figure 3.6 presents these three examples of multipartite networks and the values of  $\delta$  and  $c$  for which they are formed.

Figure 3.7 illustrates the parameter regions specified by Theorem 3.2 for  $n = 4$  and  $n = 8$  under the specification of  $f_e$  and  $f_m$  in equation (3.6), as in Siedlarek (2012a). The possible network outcomes range intuitively from empty to complete networks as the cost of linking decreases. The star is an important outcome in between empty and complete networks, but cannot be a unilaterally stable outcome for intermediate competition  $\delta$  and relatively low  $c$ . As explained in Section 3.4.2, for intermediate  $\delta$  the incentives to enter the core and the incentives for periphery members to accept the proposal of the new core member are both high. The parameter area between complete networks and stars gives multipartite networks as the stable outcomes, and this area increases with  $n$ . Different types of multipartite networks can arise including the maximally unbalanced bipartite network with group sizes of 2 and  $n - 2$ .

A comparison with the results of Goyal and Vega-Redondo (2007) can easily be made by setting  $\delta = 1$ . We observe that their assumption of perfect competition  $\delta = 1$  is a special case for which a complete network is not stable. In their model agents in a complete network always

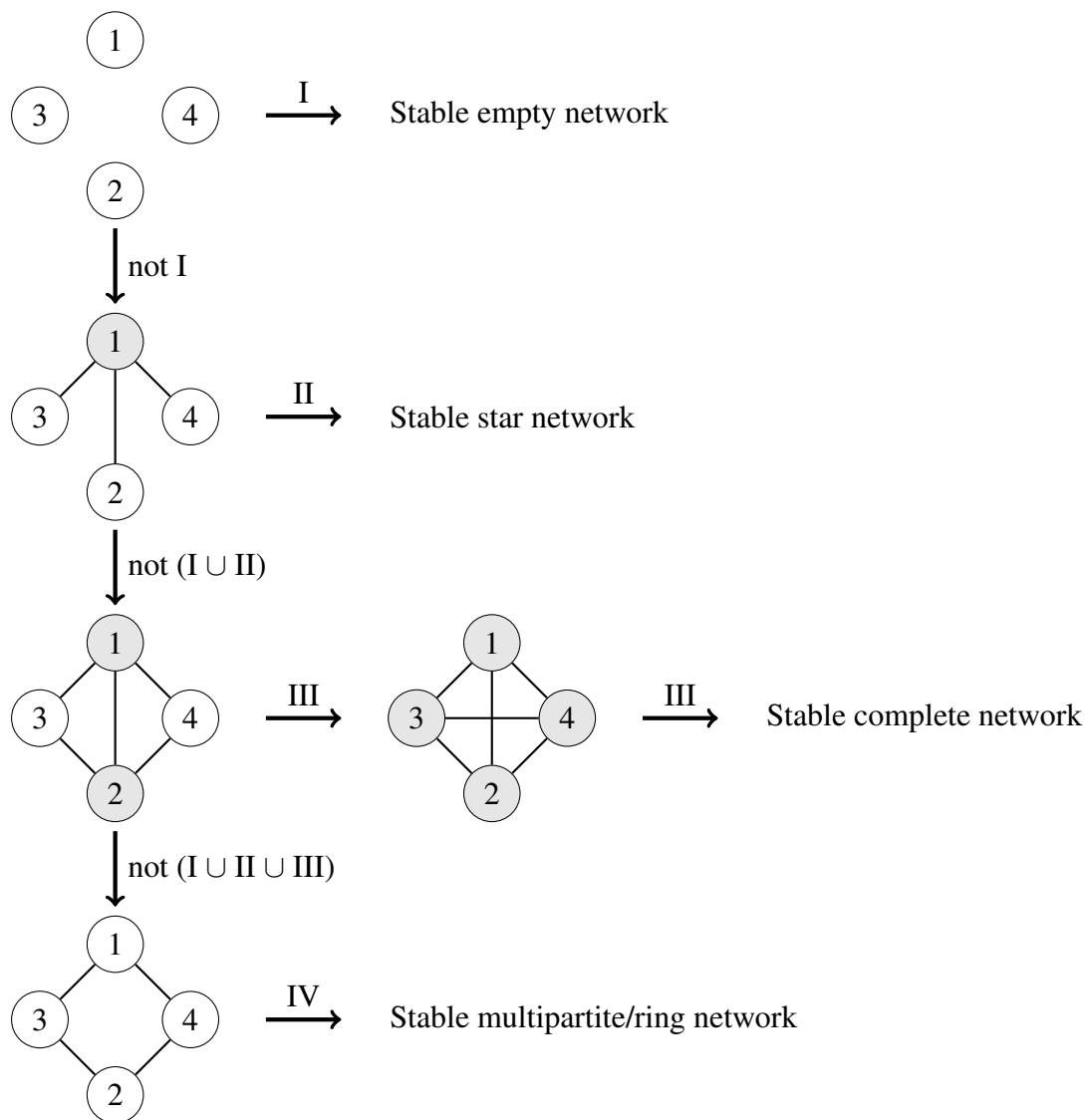
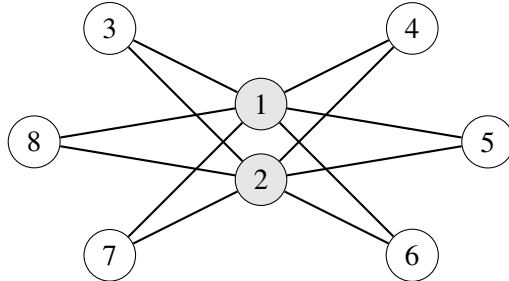
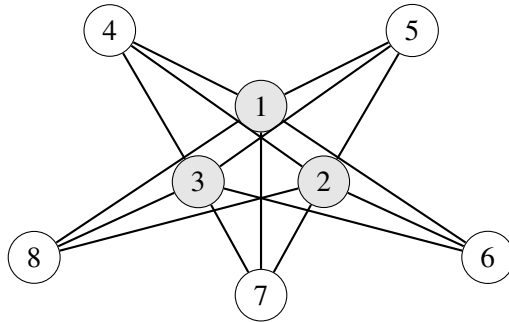


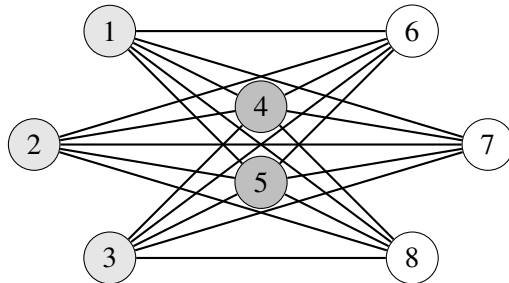
Figure 3.5: Map of best-response dynamics from an empty graph for  $n = 4$  leading to one out of 4 stable possible structures. The roman numbers correspond with the conditions in Theorem 3.2 under which the best-response route follows the direction of the arrows.



(a) Black square:  $(\delta, c) = (0.8, 0.3)$ . The maximally unbalanced bipartite network:  $g_{2,6}^{mp(2)}$ .



(b) Black circle:  $(\delta, c) = (0.4, 0.15)$ . A less than maximally unbalanced bipartite:  $g_{3,5}^{mp(2)}$ .



(c) Black star:  $(\delta, c) = (0.8, 0.06)$ . A balanced multipartite network:  $g_{2,3,3}^{mp(3)}$ .

Figure 3.6: Examples of attained complete multipartite networks after best-response dynamics from an empty graph for  $n = 8$  players. The symbols (black square, black circle and black star) correspond to locations in the  $(\delta, c)$ -space in Figure 3.7.

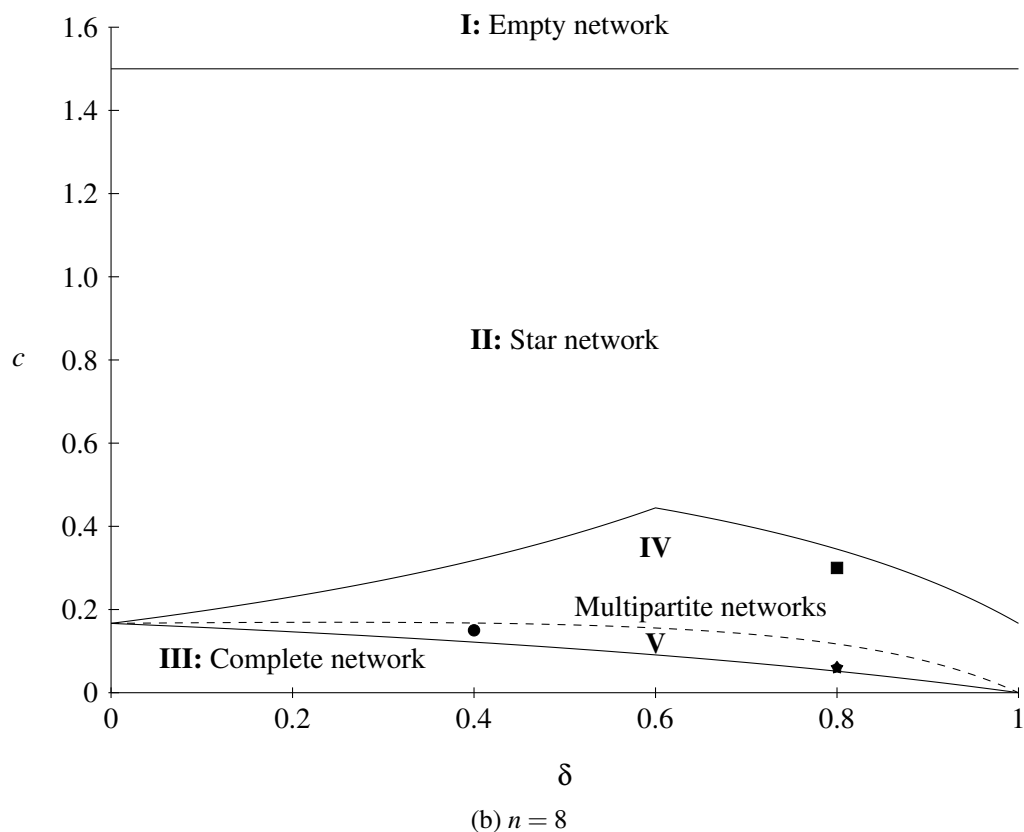
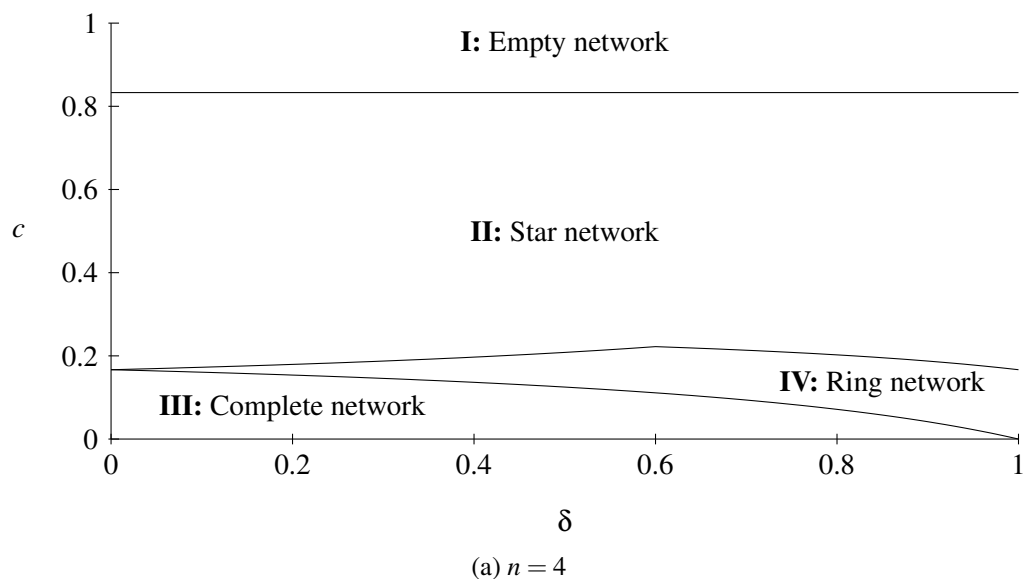


Figure 3.7: Attained equilibria after best-response dynamics from an empty network in  $(\delta, c)$ -space for  $\alpha_{ij} = 1$  for all  $i, j \in N$  and  $n \in \{4, 8\}$  under the specification of  $f_e$  and  $f_m$  in equation (3.6). The roman numbers correspond with those in Theorem 3.2. The symbols (black square, black circle and black star) correspond to examples of multipartite networks in Figure 3.6.

have incentives to remove links even for arbitrary low linking costs, because intermediated trades along multiple middlemen gives them exactly the same share of the surplus.

More importantly, we show that the star is less stable when competition is imperfect compared to the case of Goyal and Vega-Redondo (2007). For  $\delta < 1$  and a relatively low  $c$  multipartite networks arise instead of stars. Multipartite networks arise for larger regions of parameter choices if  $n$  increases; see Figure 3.9a for the results for  $n = 100$ . Interestingly, multipartite networks were found by Buskens and van de Rijt (2008) and Kleinberg, Suri, Tardos, and Wexler (2008) as the main equilibrium architecture. The current analysis shows that empty, star, complete and multipartite networks can all arise within a network formation model with intermediation and imperfect competition. Our model thus reproduces earlier results of homogeneous network formation models and places them in a more general perspective.

The results indicate a conflict between stability and efficiency in our network formation model with intermediation benefits. Whenever the star network is attained, it is the efficient network. However, for low costs under conditions III, IV or V, attracting networks are either multipartite or complete networks, and these structures are denser than is efficient. The upper-bound on  $c$  for which stable networks are not efficient is increasing in  $n$  as shown in Figure 3.9a. So for relatively large network sizes, stable networks can be expected to be overconnected.

### **3.5 Results for heterogeneous traders**

We found that complete core periphery networks were not stable, not because such networks are inefficient, but because the large inequalities in payoffs between core and periphery players were not sustainable. Instead other overconnected networks are formed, namely multipartite networks, in which banks are connected to members of other groups, but not within their own group. In real interbank markets the core banks are typically the biggest banks, and they have a strong incentive to have tight connections within the core as well as to the periphery for intermediation reasons.

For this reason we analyse the consequences of exogenous heterogeneity in the size of banks within our model.<sup>12</sup> We introduce two types of banks,  $k$  big banks and  $n - k$  small banks. The difference in size is captured by a parameter  $\alpha \geq 1$  quantifying the relative size of a big bank. Trade is assumed to be proportional to size as motivated in Appendix 3.A. The trade surplus between two big banks  $i$  and  $i'$  will be weighted as  $\alpha_{ii'} = \alpha^2$ , and the trade surplus between big bank  $i$  and small bank  $j$  with  $\alpha_{ij} = \alpha$ . The trade surplus between two small banks  $j$  and  $j'$  remains  $\alpha_{jj'} = 1$ . Our model does not explain why some banks are big or small, but rather takes their size as exogenously given. Given heterogeneity in the size of banks, we will show that a stable core periphery network can form.

The number of big banks,  $k$ , has become a new exogenous parameter, and we consider a complete core periphery network where the core consists of the  $k$  big banks. To investigate the stability of such a structure, we start with possible deviations of a peripheral player by adding links to other periphery players. Note that the change in payoffs by these deviations do not depend on  $\alpha$ , because they only concern trade surpluses between small banks. Hence Lemma 3.3 holds also in this heterogeneous setting: if adding one link improves the payoff of a small periphery player, the best feasible action is to connect to all other small banks.

If this action is feasible and executed, a complete core periphery network arises with  $k + 1$  core players,  $k$  big banks and 1 small bank. For the new core member  $i$  to have positive marginal benefits of supporting all new links, it is required that

$$\begin{aligned} M_i(g_{com}^{CP(k)}, +(n - k - 1)) &> 0 \\ \Rightarrow c &< \frac{1}{2} - f_e(k, \delta) + \frac{1}{2}(n - k - 2)f_m(k + 1, \delta) \end{aligned} \quad (3.18)$$

---

<sup>12</sup> In doing so we also add to a small literature analysing network formation with heterogeneous agents, e.g. Galeotti, Goyal, and Kamphorst (2006), Persitz (2009) and Bedayo, Mauleon, and Vannetelbosch (2013).



and every remaining peripheral player  $j \neq i$  requires

$$M_j(g_{com}^{CP(k)}, +(n-k-1) \geq 0$$

$$\Rightarrow c \leq \frac{1}{2} - f_e(k, \delta) + \frac{1}{2}(n-k-2) + (n-k-2)(f_e(k+1, \delta) - f_e(k, \delta)). \quad (3.19)$$

Conversely, the deviation of player  $i$  is *not* a best feasible action if  $c$  exceeds the minimum of the two values in equations (3.18) and (3.19):

$$c \geq \frac{1}{2} - f_e(k, \delta) + (n-k-2) \min\left\{\frac{1}{2}f_m(k+1, \delta), f_e(k+1, \delta) - f_e(k, \delta)\right\}. \quad (3.20)$$

If this condition is fulfilled, the complete core periphery network can be unilaterally stable for large enough heterogeneity  $\alpha$ , as stated in the following proposition.

**Proposition 3.3.** *If  $c$  exceeds the minimum level to restrain periphery links given by (3.20), there exists an  $\bar{\alpha}$ , such that for all  $\alpha > \bar{\alpha}$  the complete core periphery network with  $k$  big banks is unilaterally stable. In other words, for a sufficiently large level of heterogeneity  $\alpha > 1$ , the complete core periphery network with  $k$  big banks is unilaterally stable under the following condition:*

$$\text{VII: } c \in \left( \begin{array}{l} \frac{1}{2} - f_e(k, \delta) + (n-k-2) \min\left\{\frac{1}{2}f_m(k+1, \delta), f_e(k+1, \delta) - f_e(k, \delta)\right\}, \\ \min\left\{ \alpha^2\left(\frac{1}{2} - f_e(n-2, \delta)\right), \alpha\left(\frac{1}{2} - f_e(k-1, \delta)\right) + \frac{1}{2}(n-k)(n-k-1)f_m(k, \delta), \right. \\ \left. \min_{l \leq k} \left\{ \alpha\left(\frac{1}{2} - f_e(k-l, \delta)\right) + \frac{n-k-1}{l}(f_e(k, \delta) - f_e(k-l, \delta)) \right\} \right\} \end{array} \right).$$

**Proof.** See Appendix 3.C. □

For limit values of the parameters  $\delta$  and  $k$ , the required level of heterogeneity  $\bar{\alpha}$  is arbitrarily close to 1, as shown in 3.C. In combination with Proposition 3.2, this shows that heterogeneity is crucial in understanding core periphery networks. Complete core periphery networks are never stable under homogeneous players, but can be stable for arbitrary small levels of heterogeneity.

Heterogeneity also affects the results of the best-response dynamics. Starting from an empty

graph, complete core periphery networks can arise because players replicate the position of the central players and become intermediators for periphery players. If the relative size  $\alpha$  is sufficiently large and condition VII of Proposition 3.3 is fulfilled, the result will be unilaterally stable. Theorem 3.3 below specifies the seven possible attracting stable networks for the special case of  $n = 4$  and  $k = 2$  big banks. In this dynamic process, it is assumed that the  $k$  large banks are the first in the round-robin order. Given  $n = 4$ ,  $k = 2$  and any choice of the other parameters  $c$ ,  $\delta$  and  $\alpha$  (except for borderline cases), the theorem shows that the dynamics converge to a unique unilaterally stable network.<sup>13</sup>

**Theorem 3.3.** *Consider the model with  $n = 4$ , of which  $k = 2$  big banks having size  $\alpha > 1$  and  $n - k = 2$  small banks having size 1. From an empty graph, the round-robin best-response dynamics starting with the two big banks converge to the following unilaterally stable equilibria:*

I: for  $c > \max \left\{ \frac{1}{2}\alpha^2, \frac{1}{6}\alpha^2 + \frac{5}{9}\alpha + \frac{1}{9} \right\}$   
*the empty network,*

II: for  $c \in \left( \frac{1}{6}\alpha + \min \left\{ \frac{1}{2}f_m(2, \delta), f_e(2, \delta) - \frac{1}{3} \right\}, \min \left\{ \frac{5}{6}\alpha + \frac{1}{6}, \frac{1}{6}\alpha^2 + \frac{5}{9}\alpha + \frac{1}{9} \right\} \right)$   
*the star network,*

III: for  $c < \frac{1}{2} - f_e(2, \delta)$   
*the complete network,*

IV: for  $c \in \left( \min \left\{ \alpha^2 \left( \frac{1}{2} - f_e(2, \delta) \right), \frac{1}{6}\alpha + f_m(2, \delta), \frac{1}{6}\alpha + f_e(2, \delta) - \frac{1}{3} \right\}, \frac{1}{6}\alpha + \min \left\{ \frac{1}{2}f_m(2, \delta), f_e(2, \delta) - \frac{1}{3} \right\} \right)$   
*the multipartite (ring) network  $g_{2,2}^{mp(2)}$*

V: *(other multipartite networks do not exist for  $n = 4$ ),*

VI: for  $c \in \left( \min \left\{ \frac{5}{6}\alpha + \frac{1}{6}, \frac{1}{6}\alpha^2 + \frac{5}{9}\alpha + \frac{1}{9} \right\}, \max \left\{ \frac{1}{2}\alpha^2, \frac{1}{6}\alpha^2 + \frac{5}{9}\alpha + \frac{1}{9} \right\} \right)$   
*the ‘single pair’ network  $g$  with  $g_{12} = 1$  and  $g_{ij} = 0$  for all  $(i, j) \neq (1, 2)$ ,*

---

<sup>13</sup> Under heterogeneous players the results of the dynamic process depend on the order in which agents make best responses. If agents are randomly drawn to make best responses, multiple attracting steady states may exist. Using simulations, we found that for certain parameter values in VII a multipartite ring network can arise. However, for a large subset of the parameter region VII, the complete core periphery remains the unique attracting steady state even under a random order of agents.

VII: for  $c \in \left( \frac{1}{2} - f_e(2, \delta), \min \left\{ \alpha^2 \left( \frac{1}{2} - f_e(2, \delta) \right), \frac{1}{6} \alpha + f_m(2, \delta), \frac{1}{6} \alpha + f_e(2, \delta) - \frac{1}{3} \right\} \right)$   
 the complete core periphery network  $g_{com}^{CP(2)}$   
 (other core periphery networks do not exist for  $n = 4$ ).

**Proof.** See Appendix 3.C. □

Figure 3.8 illustrates the different network outcomes for  $n = 4$ ,  $k = 2$  and two levels of heterogeneity  $\alpha \in \{1.5, 2\}$ . Theorem 3.3 introduces a new (simple) type of network, called a ‘single pair’ network, in which only the two big banks are linked and the small banks have no connections. The parameter region VI is nonempty if the level of heterogeneity is sufficiently high:  $\alpha > \frac{1}{6}(5 + \sqrt{37}) \approx 1.85$ . As observed in Figure 3.8a, for  $\alpha = 1.5$  the regions of empty and star networks share a border in the  $(\delta, c)$ -diagram:  $c = \frac{1}{6}\alpha^2 + \frac{5}{9}\alpha + \frac{1}{9}$ . For a larger value like  $\alpha = 2$  in Figure 3.8b, the single pair network arises under condition VI. This type of network structure in which only part of the nodes is connected intuitively arises because some connections are more worthwhile.

Let us now discuss the main network outcome of interest, namely core periphery networks. The condition under which the complete core periphery network is the attracting steady state is in Figure 3.8 indicated by the shaded regions. As expected, this region increases with the level of heterogeneity  $\alpha$ . Also observe that for complete core periphery networks to arise it is necessary that competition is less than fully perfect, i.e.  $\delta < 1$ . For the special case of  $\delta = 1$  (as considered by Goyal and Vega-Redondo, 2007) core periphery networks are never unilaterally stable, not even under large heterogeneity. Finally, complete core periphery networks arise for a larger set of parameters if the number of players is larger than the minimal network size of  $n = 4$ . See Figure 3.9b for the shaded region VII given  $n = 100$ ,  $k = 15$  and  $\alpha = 10$ .

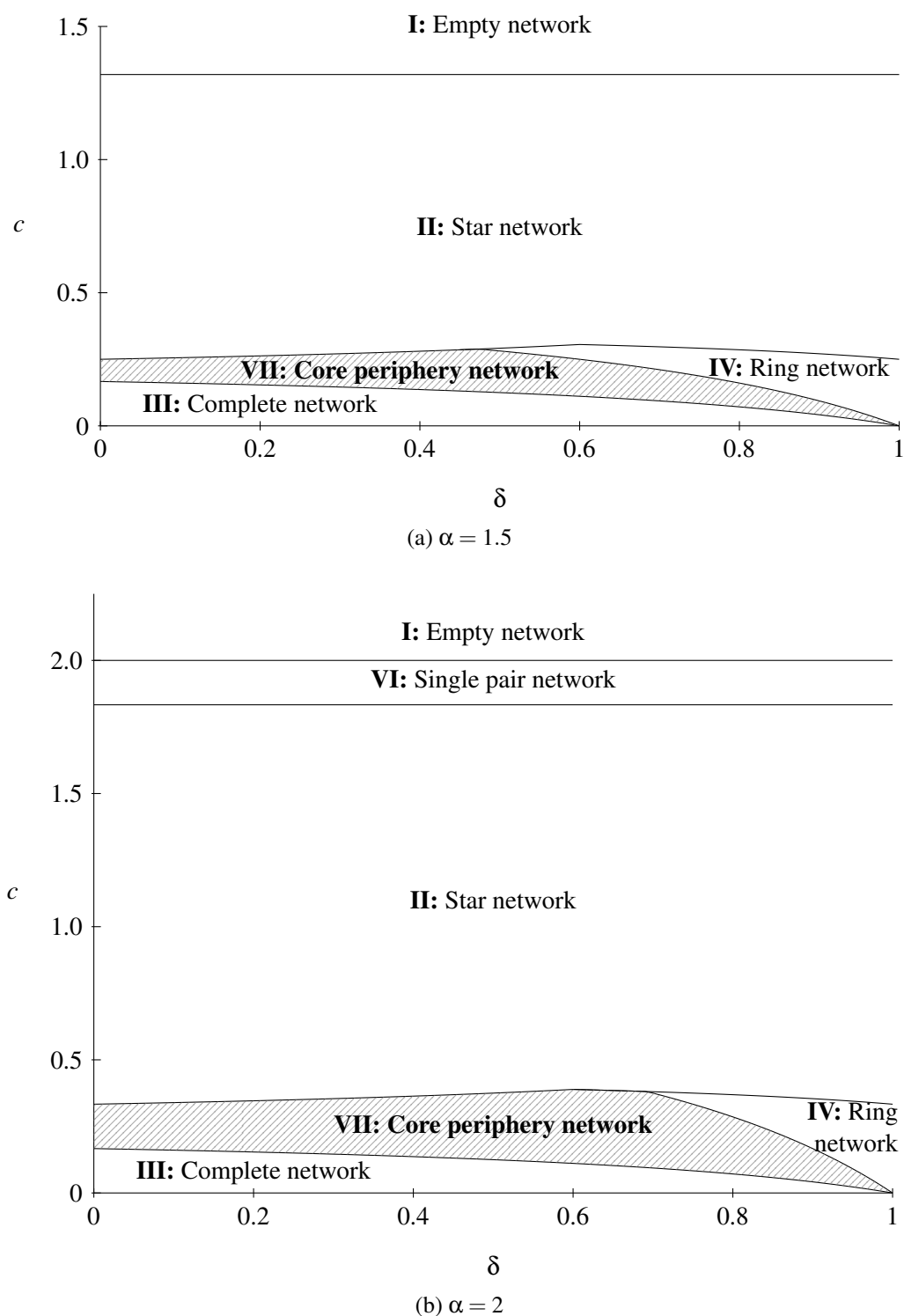


Figure 3.8: Attained equilibria after best-response dynamics from an empty graph in  $(\delta, c)$ -space for  $n = 4$ ,  $k = 2$  and  $\alpha \in \{1.5, 2\}$  under the specification of  $f_e$  and  $f_m$  in equation (3.6). The roman numbers correspond with those in Theorem 3.3. In the shaded area, the complete core periphery network with  $k = 2$  big banks is the unique unilaterally stable outcome of the dynamics.

### 3.6 Calibrating the model to the Dutch interbank market

To gain some insight into the general applicability of the model, we calibrate the model to the Dutch interbank market. In Chapter 2 we have investigated the network structure of the interbank market in the Netherlands, a relatively small market with approximately a hundred financial institutions. The attracting networks in the dynamic homogeneous model with  $n = 100$  are indicated in Figure 3.9a. Under homogeneous banks the model predicts multipartite networks for most parameter values. In contrast, in Chapter 2 it was found that the observed network contains a very densely connected core with around  $k = 15$  core banks.

We choose a level of heterogeneity  $\alpha = 10$  to capture in a stylised way the heterogeneity of banks in the Netherlands. In reality banks in the core as well as in the periphery of the Dutch banking system are quite diverse. A few very large banks reach a total asset size of up to €1 trillion, while the asset value of some investment firms active in the interbank market may not be more than a few million euro. The median size of a core bank in Chapter 2 lies around €8 billion. For periphery banks the median size is approximately €300 million (see Figure 2.9 in Chapter 2 for the plotted distribution of total asset size over core and periphery banks). Ignoring the exceptionally large size of some core banks, a relative difference of  $\alpha = 10$  seems a reasonable order of magnitude.

Figure 3.9b shows the parameter region given by Proposition 3.3 for which a complete core periphery network of the fifteen big banks is a unilaterally stable network. The stability of the complete network does not depend on  $\alpha$  as also indicated in Figure 3.9b. These complete and complete core periphery networks would also be the outcome of best-response dynamics starting with the big banks.<sup>14</sup> The observed core periphery structure in the Netherlands can be reproduced for many reasonable choices of linking costs  $c$  and competitiveness  $\delta$ .

This calibration suggests that our model is very suitable to explain stylised facts of national

---

<sup>14</sup> A full description of the outcomes of these best-response dynamics would require a generalisation of Theorem 3.3 for all  $n \geq 4$ . Simulations show that for linking costs so high that CP networks are not stable, many different structures arise. As none of them are core periphery networks, we restrict ourselves to Proposition 3.3 and Theorem 3.3.

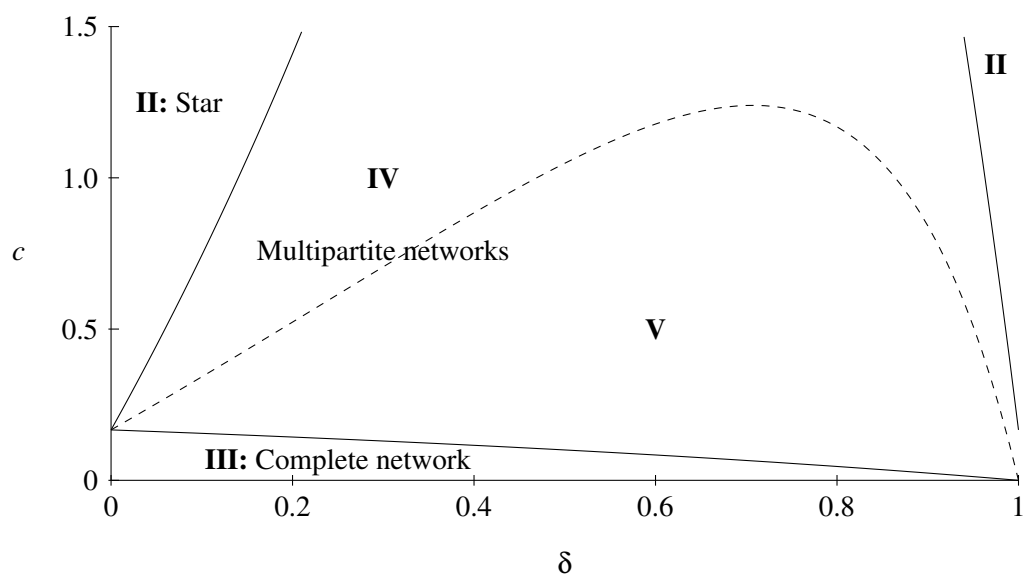
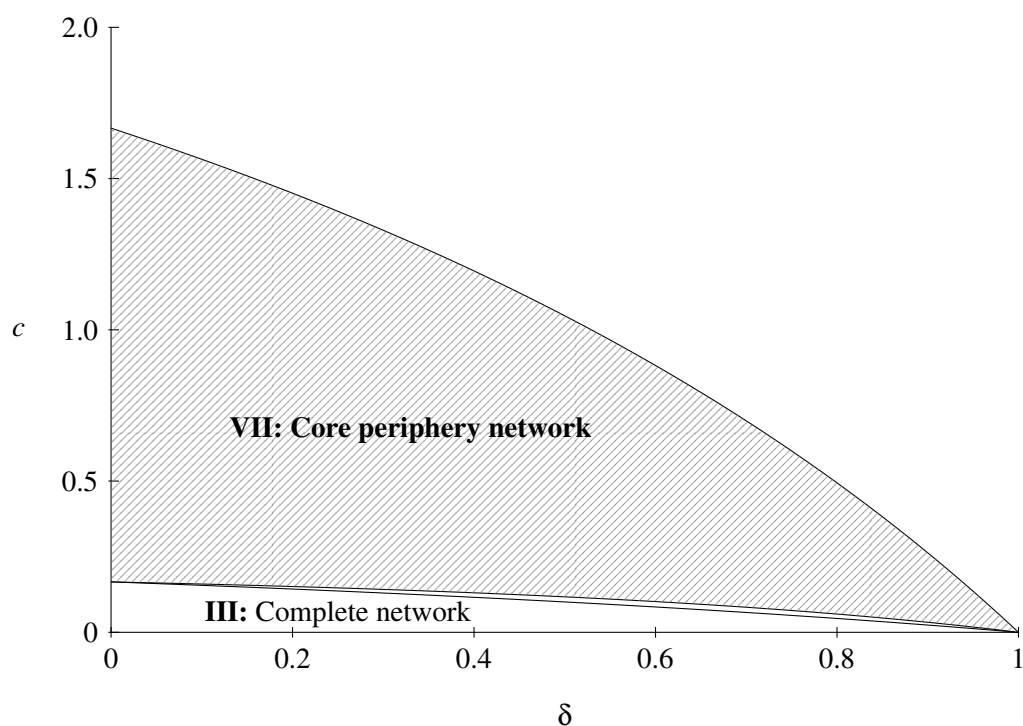

 (a) Homogeneous banks ( $\alpha_{ij} = 1$  for all  $i, j \in N$ )

 (b) Heterogeneous banks:  $k = 15$  big banks (with relative size  $\alpha = 10$ ) and  $n - k = 85$  small banks

Figure 3.9: Calibration of the model to the Dutch interbank market. Attained equilibria after best-response dynamics from an empty graph in  $(\delta, c)$ -space for  $n = 100, k = 15$  and  $\alpha \in \{1, 10\}$  under the specification of  $f_e$  and  $f_m$  in equation (3.6). The roman numbers correspond with those in Theorem 3.2 and Proposition 3.3. In the shaded area, the complete core periphery network with  $k = 15$  big banks is the unique unilaterally stable outcome of best response dynamics starting with the large banks.

or perhaps even international interbank networks. It should be noted that observed core periphery networks are not necessarily complete core periphery networks, which can be explained as follows. Empirical studies of interbank markets often rely on either balance sheet data measuring the total exposure of one bank on another, or overnight loan data specifying the actual trades. We interpret the undirected links in our model as established preferential lending relationships, which are typically difficult to observe directly in practice. Given a theoretically complete core periphery network of lending relationships, trades and exposures are executed on the same structure of connections; see Appendix 3.A for the interpretation of trade surpluses in our model. The empirically observed core periphery structure can have less than complete connections between core and periphery depending on the realisations of trade opportunities. In any case, the densely connected core of a subset of the banks is a well-documented empirical fact.

### **3.7 Conclusion**

In this chapter we have proposed a way to explain the formation of financial networks by intermediation. We have focused on the core periphery network because it is found to give a fair representation of the complex empirical structures, while at the same time being relatively simple and intuitively appealing. In our model brokers strive to intermediate between their counterparties and compete with each other. Our results suggest that heterogeneity is crucial, that is, the core periphery structure of the interbank network cannot be understood separately from the heterogeneity and inequality in the intrinsic characteristics of banks.

Naturally our analysis can be extended in many ways. It seems interesting to endogenise heterogeneity in our model by updating the size of each bank with the payoffs received from trades in the network. Better-connected banks receive higher payoffs by intermediation, which could feed back on the balance sheet and thus on future trade opportunities of these banks. This is consistent with a recent empirical finding by Akram and Christophersen (2010) that banks

with many financial linkages in the Norwegian interbank market face lower interest rates. We can expect that core periphery networks also arise endogenously in the extended model with ex ante homogeneous agents and feedback of the network structure on trade surpluses.

Another important extension is to introduce default probabilities in order to understand network formation in stress situations. This seems relevant in light of findings that the fit of the core periphery network in the interbank market deteriorated during the recent financial crisis as found in Chapter 2 and in Fricke and Lux (2012). The chapter provides an economic model for financial networks with high empirical relevance, and suggests a potentially fruitful road of finding policies to reduce systemic risk.



## Appendix 3.A Trade surpluses in the interbank market

This chapter uses a stylised model where transactions are executed between every pair of agents. We now give a formal interpretation of the trade surpluses  $\alpha_{ij}$  in a simple two-period framework. In the first period, banks establish relationships in the interbank market to insure themselves against liquidity shocks. In the second period, after the shocks have been revealed, each bank uses these relationships to rebalance its liquidity position. If it is unable to do so in the interbank market, the bank has to trade with the central bank under less favourable conditions. The value  $\alpha_{ij}$  is the ex ante surplus of the trade relationship between  $i$  and  $j$ .

We consider a bank  $i$  that has a preferential lending relationship with bank  $j$ . Each bank  $i$  faces exogenous future liquidity shocks  $x_i$ :  $x_i$  positive implies bank  $i$  has excess liquidity and  $x_i$  negative implies that bank  $i$  has a liquidity shortage. Assume that the CB offers lending and borrowing facilities to banks at rates  $\bar{p}$  and  $\underline{p}$ . Without interbank transactions, the values attributed to shocks are:

$$\Pi_i^{CB} = \begin{cases} \underline{p}x_i > 0 & \text{if } x_i \geq 0 \\ \bar{p}x_i < 0 & \text{if } x_i < 0. \end{cases} \quad (3.21)$$

These CB facilities represent the outside options of banks.

Using its relationship with  $j$ , bank  $i$  can try to receive more favourable conditions for lending and borrowing, by agreeing on a rate  $p > \underline{p}$  when he had a positive liquidity shock, or  $p < \bar{p}$  in case he is short of liquidity. Conversely, bank  $j$  would also be happy to trade for some  $p \in (\underline{p}, \bar{p})$ , as long as the shocks are of opposite sign. In this case the ex post value of the trade surplus between  $i$  and  $j$  is therefore:

$$TS_{ij} = (\bar{p} - \underline{p}) \min\{|x_i|, |x_j|\}. \quad (3.22)$$

For simplicity, assume that two random banks receive a shock of size 1 and that these shocks are of opposite sign. When the shocks hit banks  $i$  and  $j$ , these two banks can trade and create a

total surplus of  $(\bar{p} - p)$ . Assume that the probability of receiving a shock for  $i$  is  $\lambda_i$ . The total ex ante value of a trading relationship between  $i$  and  $j$  is then:

$$\alpha_{ij} = \lambda_i \lambda_j (\bar{p} - p). \quad (3.23)$$

In the baseline model with homogeneous banks, we have imposed that probabilities of receiving liquidity shocks are equal for all banks ( $\lambda_i = \lambda_j \forall i, j$ ), and normalised  $\alpha_{ij}$  to 1. For heterogeneous banks, the assumptions on trade surpluses  $\alpha_{ij}$  made in Section 3.5 boils down to assuming that a big bank  $i$  has a higher probability of receiving liquidity shocks than a small bank  $j$  in a fixed proportion  $\lambda_i = \alpha \cdot \lambda_j$ . Normalising the surplus of two small banks  $j$  and  $j'$  to  $\alpha_{jj'} = 1$ , gives that the higher trade surplus between one big bank  $i$  and one small bank  $j$  is  $\alpha_{ij} = \alpha$  and between big banks  $i$  and  $i'$  is  $\alpha_{i'i} = \alpha^2$ .

### Appendix 3.B Siedlarek's payoff function

Siedlarek (2012a) applies the bargaining protocol introduced by Merlo and Wilson (1995) to derive the distribution of a surplus for a trade that is intermediated by competing middlemen. The underlying idea is that one involved agent is selected to propose a distribution of the surplus. If the proposer succeeds in convincing the other agents to accept his offer, the trade will be executed. Otherwise, the trade will be delayed and another (randomly selected) agent may try to make a better offer. A common parameter  $0 \leq \delta \leq 1$  is introduced with which agents discount future periods, that results in the level of competition as used in this chapter. As a special case, for  $\delta = 1$  the surplus is distributed equally among the essential players as in Goyal and Vega-Redondo (2007).

More formally, assume that the set of possible trading routes is known to all agents (i.e. complete information). Each period a route is selected on which the trade can be intermediated, and additionally one player (the 'proposer') on this route that proposes an allocation along the

entire trading route.<sup>15</sup> Any state, which is a selection of a possible path and a proposer along that path, is selected with equal probability and history independent. The question now is: what is the equilibrium outcome, i.e. the expected distribution of the surplus, taking into account that every agent proposes optimally under common knowledge of rationality of other possible future proposers? Siedlarek (2012a) shows that the unique Markov perfect equilibrium is characterized by the following payoff function for any player  $i$  in a certain state:

$$f_i = \begin{cases} 1 - \sum_{j \neq i} \delta E_j[f_j] & \text{if } i \text{ is the proposer in this state} \\ \delta E_i[f_i] & \text{else if } i \text{ is involved in this state.} \end{cases} \quad (3.24)$$

This equation shows that the proposer can extract all surplus over and above the outside option value given by the sum of  $E_j[f_j]$  over all other players  $j \neq i$ . All (and only) the players along the same route have to be convinced by offering exactly their outside option.

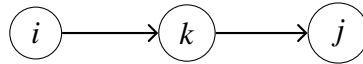


Figure 3.10: A trading route with one intermediary

In the simple example of intermediary  $k$  connecting  $i$  and  $j$  (see Figure 3.10), each of the three players has equal probability of becoming the proposer and proposes an equal share to the other two, so the equilibrium distribution of the surplus will simply be the equal split  $(f_i, f_k, f_j) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  for all  $\delta$ .

In case  $m$  intermediaries compete for the trade between  $i$  and  $j$  (as in Figure 3.11), the

---

<sup>15</sup> Siedlarek (2012a) assumes that only *shortest* paths are considered, an assumption not explicitly made by Goyal and Vega-Redondo (2007). However, in the model Goyal and Vega-Redondo (2007) only essential players, who by definition are part of shortest paths, receive a nonzero share of the surplus.

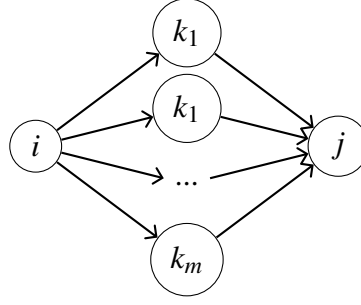


Figure 3.11: A trading network with  $m$  competing intermediaries

extracted intermediation rents are reduced. Siedlarek (2012a) shows that:

$$f_i = f_j = f_e(m, \delta) \equiv \frac{m - \delta}{m(3 - \delta) - 2\delta} \quad (3.25)$$

$$f_k = f_m(m, \delta) \equiv \frac{1 - \delta}{m(3 - \delta) - 2\delta}. \quad (3.26)$$

As mentioned before, this distribution of payoffs satisfies our assumptions on  $f_e(m, \delta)$  and  $f_m(m, \delta)$ . The distribution of Siedlarek (2012a) is used as the leading example in the results. All results hold for general  $f_e(m, \delta)$  and  $f_m(m, \delta)$  satisfying the assumptions of Table 3.1.

In a general network with longer intermediation chains  $d > 2$ , the payoffs cannot be calculated directly, but depend on specifics of (part of) the network. The vector of payoffs  $\vec{F}$  can be calculated indirectly from the network  $g$  using the following notation:

- $s$ : The number of shortest paths that support the trade,
- $d$ : The length of the shortest paths (i.e. the number of required intermediaries plus one),
- $\vec{P}$ :  $(n \times 1)$ -vector of shortest paths that run through any player,
- $K$ :  $(n \times n)$ -diagonal matrix of times that any player receives an offer,
- $S$ :  $(n \times n)$ -off-diagonal matrix of times that two players share a path.

As in any given state the distribution of profits is as given in equation (3.24), the vector of

payoffs that averages over all possible states of the world is:

$$\vec{F} = \frac{1}{s \cdot d} (\vec{P} - \delta(S - K) \vec{F}). \quad (3.27)$$

And if  $\delta < 1$  this can be solved to:

$$\vec{F} = (s \cdot d \cdot I_n + \delta(S - K))^{-1} \vec{P}. \quad (3.28)$$

In Appendix 3.D we use this general formula to investigate whether core periphery networks can be stable when intermediation paths of lengths  $d = 3$  are allowed.

## Appendix 3.C Proofs

### *Proof of Proposition 3.2.*

The intuition of the second proof is *limited inequality between core and periphery players*. We will provide sufficient conditions under which at least one agent will be able improve its position in any (complete or incomplete) core periphery network.

(a) Complete core periphery networks. Consider the deviation of a core player  $i$  to delete one link to another core member. The marginal benefit of doing so is:

$$M_i(g_{com}^{CP(k)}, -1) = c - \left(\frac{1}{2} - f_e(n - 2, \delta)\right). \quad (3.29)$$

The marginal benefit for a peripheral player  $j$  to add a link in the periphery is:

$$M_j(g_{com}^{CP(k)}, +1) = \left(\frac{1}{2} - f_e(k, \delta)\right) - c. \quad (3.30)$$

For the network to be pairwise stable, the two necessary conditions are  $M_i(g_{com}^{CP(k)}, -1) < 0$  and

$M_j(g_{com}^{CP(k)}, +1) < 0$ , so:

$$\begin{aligned} \frac{1}{2} - f_e(k, \delta) < c < \frac{1}{2} - f_e(n-2, \delta) \\ f_e(n-2, \delta) < \frac{1}{2} - c < f_e(k, \delta). \end{aligned}$$

However, as  $\frac{\partial f_e}{\partial k} > 0$  and  $k < n-1$ , this interval of  $c$  is empty. Complete core periphery networks are never pairwise stable, and therefore also not unilaterally stable.

(b) Incomplete core periphery networks. We will determine the benefits of a periphery player to propose links to all other periphery players. In Lemma 3.3 we showed that this is a best feasible action, if there exists a profitable deviation by adding links at all, in *complete* core periphery networks. In a general core periphery network, adding to all other periphery players is not necessarily a best feasible action, but it will help to find the sufficient condition for instability of the networks, as claimed in the Proposition.

Consider a periphery player  $i$  proposing  $l_i = n - k - 1$  links in order to reach all other periphery players. The minimal marginal benefits for  $i$  of this action are bounded by:

$$M_i(g^{CP(k)}, +l_i) > -(n-k-1)c + \binom{n-k-1}{2} f_m(k+1, \delta). \quad (3.31)$$

This lower bound for the marginal benefits consists of two parts. The first part,  $-(n-k-1)c$ , denotes the direct costs for adding the  $l_i$  links and is linear in  $n$ . The second part,  $\binom{n-k-1}{2} f_m(k+1, \delta)$ , denotes the minimal intermediation benefits from becoming a new intermediary between the  $n-k-1$  remaining periphery members and is quadratic in  $n$ . In general these intermediation benefits for  $i$  can be higher if there are fewer  $m_{jl} < k$  intermediators in the original network  $g^{CP(k)}$  between  $j, l \in P$ . Benefits from trade (direct or indirect) do not have to be considered because, obviously, adding links to the network can only increase the access from  $i$  to other players, and the benefits from trade therefore weakly increase.

Let the parameters  $c > 0$ ,  $0 < \delta < 1$  and  $k \geq 1$  be given. Then the marginal benefit function (3.31) is strictly convex in  $n$  and will always become bigger than 0 for large enough  $n$ : the

quadratic intermediation part of the expression will always dominate the linear cost part for strictly positive intermediation benefits  $f_m(k+1, \delta) > 0$ . If  $c$  is large  $M_i$  may be negative for small  $n$ , but eventually the intermediation benefits will become so large that the deviation will become beneficial. More precisely, positive marginal benefits for  $i$  imply:

$$M_i(g^{CP(k)}, +l_i) > 0 \quad (3.32)$$

$$\Rightarrow c < \frac{1}{2}(n-k-2)f_m(k+1, \delta) \quad (3.33)$$

$$\Leftrightarrow n > F_i(c, \delta, k) \equiv k+2 + \frac{c}{\frac{1}{2}f_m(k+1, \delta)}. \quad (3.34)$$

For the deviation of  $l_i$  links to be executed, the peripheral players  $j \in (P \setminus i)$  also have to agree with the addition, that is, they should not receive a lower payoff. The marginal benefits for  $j$  depend on its number of connections  $n_j$ , but are bounded by:

$$M_j(g^{CP(k)}, +l_i) > -c + (n-k-2) \min_{n_j \leq k} \{f_e(m+1, \delta) - f_e(m, \delta)\}. \quad (3.35)$$

The second part of these marginal benefits indicates that the number of intermediation paths from  $j$  to any other periphery member  $l \in (P \setminus i, j)$  strictly increases, and therefore the benefits from trade strictly increase.<sup>16</sup> This function is similarly convex, and positive marginal benefits for  $j$  imply:

$$M_j(g^{CP(k)}, +l_i) \geq 0 \quad (3.36)$$

$$\Rightarrow c < (n-k-2) \min_{n_j \leq k} \{f_e(m+1, \delta) - f_e(m, \delta)\} \quad (3.37)$$

$$\Leftrightarrow n > F_j(c, \delta, k) \equiv k+2 + \frac{c}{\min_{n_j \leq k} \{f_e(m+1, \delta) - f_e(m, \delta)\}}. \quad (3.38)$$

Combining the conditions for  $i$  and  $j$ , a sufficient lower bound for the network size  $n$  for any

<sup>16</sup> It is possible that player  $i$  pays intermediation benefits to  $j \in P$  if it can access some  $l \in C | g_{il} = 0$  better. The lower bound on  $M_j$  given in (3.35) is sufficient for the proof.

core periphery network to be unstable is:

$$n > F(c, \delta, k) \equiv k + 2 + \frac{c}{\min\left\{\frac{1}{2}f_m(k+1, \delta), \min_{n_j \leq k}\{f_e(m+1, \delta) - f_e(m, \delta)\}\right\}}. \quad (3.39)$$

□

**Remarks on Proposition 3.2.** In the derivation of this sufficient lower bound on  $n$ , we have not made any assumptions about the path lengths on which trade is allowed. It is sufficient to consider the parts of the marginal benefits for  $i$  and  $j$  that depend on the constant costs  $c$  and the new intermediation route is formed between  $j, l \in (P \setminus i)$ .

A core periphery network is generally unstable because the inequality between core and periphery becomes large for increasing  $n$ , and a periphery player can always benefit by adding links to all other players. An important assumption for this result is therefore that multiple links can be added at the same time. It is crucial that  $\delta > 0$  because for  $\delta = 0$  intermediated trade always generates  $f_e(m, 0) = \frac{1}{3} \forall m$ , i.e. additional intermediation paths do not increase profits for endnodes. Moreover it is crucial to impose  $\delta < 1$  because for  $\delta = 1$  intermediation benefits disappear, i.e.  $f_m(m, 1) = 0 \forall m > 1$ .

**Proof of Theorem 3.2.**

By Lemma 3.1, after the move of player  $i = 1$  the network is either empty or a star. Obviously, in case it is empty it is stable, as all other nodes face the same decision as  $i = 1$ . If it is a star network and none of the other nodes wants to add a subset of links, this is the final stable outcome. Otherwise, by Lemmas 3.2 and 3.3, each next node  $i = 2, \dots, k$  adds links to all nodes not yet connected to  $i$ . There is a third and final possibility that the dynamic process is ended before the second round, namely if node  $i = n - 1$  fulfills the complete network by linking to node  $n$ . The conditions I, II and III for convergence to the empty, star or complete network coincide with the stability conditions of these three networks as in Proposition 3.1.



For parameters not satisfying I, II or III, the first round of best responses results in a complete core periphery network with  $1 < k < n - 1$  core members. This complete core periphery network is not stable by the part (a) of Proposition 3.2. Because the  $k + 1$ -th node did not connect to other periphery nodes, adding links in the periphery cannot be beneficial. Therefore the first core bank  $i = 1$  must have an incentive to delete at least one within-core link. The marginal benefit of deleting  $l$  core links in the network is:

$$M_i(g_{com}^{CP(k)}, -l) = l(c - \frac{1}{2} + f_m(n - l - 1, \delta)). \quad (3.40)$$

As  $M_i(g_{com}^{CP(k)}, -0) = 0$  and  $M_i(g_{com}^{CP(k)}, -1) > 0$ , there is a unique choice  $l_1^* > 0$  of the optimal number of core links to delete. It will become clear that for parameters outside  $I \cup II \cup III$ , attracting networks are multipartite networks of various sorts, depending on the choice  $l_1^*$ .

An important case is that a complete core periphery network with  $k = 2$  has arisen after the first round. For  $k = 2$  the only possible solution is  $l_1^* = 1$ . A complete bipartite network arises with a small group of 2 players and a large group of  $n - 2$  players, the maximal difference in group size possible for bipartite networks. We denote such a network as  $g_{2,n-2}^{mp(2)}$ .

The case of  $k = 2$  happens if the third player  $i = 3$  does not enter the core, either because entering is not beneficial for himself or because some other periphery player  $j$  does not accept the offer of  $i$ . For  $k = 2$  a positive marginal benefit of entering the core implies:

$$\begin{aligned} M_i(g_{com}^{CP(2)}, +(n-2)) &> 0; \\ \Rightarrow c &< \frac{1}{2} - f_e(2, \delta) + \frac{1}{2}(n-4)f_m(3, \delta), \end{aligned} \quad (3.41)$$

and the periphery player  $j$  has an incentive to accept the offer if

$$\begin{aligned} M_j(g_{com}^{CP(2)}, +(n-2)) &\geq 0; \\ \Rightarrow c &\leq \frac{1}{2} - f_e(2, \delta) + (n-4)(f_e(3, \delta) - f_e(2, \delta)); \end{aligned} \quad (3.42)$$

So if after the first round the result is  $k = 2$  and the third player has not entered the core, it must be the case that (compare equation (3.14)):

$$c \geq \frac{1}{2} - f_e(2, \delta) + (n-4) \min\left\{\frac{1}{2}f_m(3, \delta), f_e(3, \delta) - f_e(2, \delta)\right\}. \quad (3.43)$$

The parameter values for which the maximally unbalanced network  $g_{2, n-2}^{mp(2)}$  is the attracting steady state are given by condition IV.

Alternatively, under the remaining condition V, the first round has resulted in a complete core periphery network with  $k > 2$ . This network cannot be stable, and the first core bank  $i = 1$  has an optimal choice of  $0 < l_1^* \leq k - 1$  links to delete depending on the parameters. First consider that the core bank deletes all its within-core links, i.e.  $l_1^* = k - 1$ . This action necessarily implies that the linking costs exceed the loss in surplus associated with having an indirect connection to other core banks via the periphery rather than a direct connection:

$$c > \frac{1}{2} - f_e(n - k, \delta). \quad (3.44)$$

Given this high level of linking costs, the next core banks  $i = \{2, \dots, k\}$  have the same incentive to delete all their within-core links. The attracting network is therefore a complete multipartite network  $g_{k_1, k_2}^{mp(2)}$  with two groups of size  $k_1 = k$  and  $k_2 = n - k$ .

Finally, consider the case  $0 < l_1^* < k - 1$ .<sup>17</sup> Denote  $K_1 \subset K$  as the set of core banks with at least one missing within-core link after the best response of  $i = 1$ , including players  $i = 1$  itself. The size of this set of banks is  $k_1 = l_1^* + 1$ . The best response of  $i = 1$  necessarily implies:

$$c > \frac{1}{2} - f_e(n - k_1, \delta). \quad (3.45)$$

Given this high level of linking costs, all next banks  $i \in K_1$  with connections to some other  $j \in K_1$

---

<sup>17</sup> This is the only case that the best response is not unique in a parameter region with nonzero measure: the first core banks  $i = 1$  has  $\binom{k}{l_1^*}$  best responses deleting  $l_1^*$  links from the core. The resulting multipartite networks can consist of different sets  $K_1, K_2, \dots, K_{q-1}$ , but all are isomorphic.

with  $j > i$  have the same incentive to delete every link  $g_{ij}$ . These banks do not have an incentive to remove any further link, because the number of intermediators for indirect connections to banks  $j \in K_1$  would become less than  $n - k_1$ . If the latter were worthwhile,  $i = 1$  would have chosen a higher number  $l_1^*$ . Therefore  $K_1$  becomes a group of players not connected within their group, but completely connected to all players outside their group.

Players  $i \in (K \setminus K_1)$  are connected to all other players  $j \in N$ . Because of the lower bound on  $c$  in equation (3.45), these players have an incentive to remove links to some  $j$  as long as the number of intermediators for the indirection connections to  $j$  stays above  $n - k_1$ . This will lead to other groups  $K_2, \dots, K_{q-1}$  of players not connected within their group, but completely connected to all players outside their group. The remaining group  $K_q$  was the original periphery at the end of the first round. The result of the best-response dynamics if  $0 < l_1^* < k - 1$  is a multipartite network  $g_{k_1, k_2, \dots, k_q}^{mp(q)}$  with  $q \geq 3$ .

For all parameters under condition V, the resulting network is multipartite with  $q = \lfloor \frac{k}{k_1} \rfloor + 1$  groups. These multipartite networks are more balanced than  $g_{2, n-2}^{mp(2)}$ , i.e. have group sizes  $|k_m - k_{m'}| < n - 4$  for all  $m, m' \in \{1, 2, \dots, q\}$ .  $\square$

***Proof of Proposition 3.3.***

We will derive a  $\bar{\alpha}$ , such that for all  $\alpha > \bar{\alpha}$  the complete core periphery network with  $k$  big banks is unilaterally stable. First observe that no player wants to add any links. In a complete core periphery network only periphery players can add links, but given a sufficiently high  $c$  satisfying (3.20) this is not beneficial. We therefore only need to consider the possible deletion of one or multiple links by either core or periphery players.

First, consider a core player  $i$ . Player  $i$  can delete links with other core players and/or links

with periphery players. The marginal benefit of deleting  $l^c$  core links and  $l^p$  periphery links is:

$$\begin{aligned}
 M_i(g_{com}^{CP(k)}, -(l^c + l^p)) &= l^c \left( c - \alpha^2 \left( \frac{1}{2} - f_e(n - l^c - l^p - 1, \delta) \right) \right) \\
 &\quad + l^p \left( c - \alpha \left( \frac{1}{2} - f_e(k - l^c - 1, \delta) \right) - (2n - 2k - l^p - 1) f_m(k, \delta) \right) \\
 &\equiv M_i^c + M_i^p.
 \end{aligned} \tag{3.46}$$

The marginal benefit of deleting  $l^c + l^p$  links can be separated in benefits from deleting links with the core  $M_i^c$  and benefits from deleting links with the periphery  $M_i^p$ . The cross-over effects of deleting links with both groups of banks are negative:  $M_i^c$  is decreasing in  $l^p$  and  $M_i^p$  is decreasing in  $l^c$ . To find the conditions under which the core player does not want to delete any link, it is therefore sufficient to consider deletion of links in each group separately.

The marginal benefit of deleting  $l^c$  core links is:

$$M_i(g_{com}^{CP(k)}, -l^c) = l^c \left( c - \alpha^2 \left( \frac{1}{2} - f_e(n - l^c - 1, \delta) \right) \right). \tag{3.47}$$

If  $M_i(g_{com}^{CP(k)}, -l^c) > 0$ , it must hold that  $M_i(g_{com}^{CP(k)}, -1) > 0$ , because  $f_e(n - l^c - 1, \delta)$  decreases in  $l^c$ . Player  $i$  thus has a beneficial unilateral deviation if

$$\begin{aligned}
 &M_i(g_{com}^{CP(k)}, -l^c) > 0 \\
 \Rightarrow &M_i(g_{com}^{CP(k)}, -1) > 0 \\
 \Leftrightarrow &\alpha < \sqrt{c / \left( \frac{1}{2} - f_e(n - 2, \delta) \right)}.
 \end{aligned} \tag{3.48}$$

The marginal benefit for a core player  $i$  of deleting  $l^p$  links with the periphery is :

$$M_i(g_{com}^{CP(k)}, -l^p) = l^p \left( c - \alpha \left( \frac{1}{2} - f_e(k - 1, \delta) \right) - (2n - 2k - l^p - 1) f_m(k, \delta) \right). \tag{3.49}$$

If  $M_i(g_{com}^{CP(k)}, -l^p) > 0$ , it must be a best feasible action to choose  $l^p = n - k$  and delete all links with the periphery, as the function is convex in  $l^p$ . Player  $i$  thus has a beneficial unilateral

deviation if

$$\begin{aligned}
 & M_i(g_{com}^{CP(k)}, -l^p) > 0 \\
 \Rightarrow & M_i(g_{com}^{CP(k)}, -(n-k)) > 0 \\
 \Leftrightarrow & \alpha < \frac{c - \frac{1}{2}(n-k)(n-k-1)f_m(k, \delta)}{\frac{1}{2} - f_e(k-1, \delta)}. \tag{3.50}
 \end{aligned}$$

Second, consider a player  $i \in P$  in the periphery. This marginal benefit for a periphery bank  $i \in P$  to delete  $l$  links is positive if:

$$\begin{aligned}
 & M_i(g_{com}^{CP(k)}, -l) = l \left( c - \alpha \left( \frac{1}{2} - f_e(k-l, \delta) \right) \right) - (n-k-1)(f_e(k, \delta) - f_e(k-l, \delta)) > 0 \\
 \Leftrightarrow & \alpha < \frac{c - \frac{n-k-1}{l}(f_e(k, \delta) - f_e(k-l, \delta))}{\frac{1}{2} - f_e(k-l, \delta)}. \tag{3.51}
 \end{aligned}$$

The network is unilaterally stable if neither of these three deviations is beneficial, i.e. if  $\alpha$  exceeds all values given in equations (3.48), (3.50) and (3.51):

$$\begin{aligned}
 \bar{\alpha} = \max \left\{ \right. & \sqrt{c / \left( \frac{1}{2} - f_e(n-2, \delta) \right)}, \frac{c - \frac{1}{2}(n-k)(n-k-1)f_m(k, \delta)}{\frac{1}{2} - f_e(k-1, \delta)}, \\
 & \left. \max_{l \leq k} \left( \frac{c - \frac{n-k-1}{l}(f_e(k, \delta) - f_e(k-l, \delta))}{\frac{1}{2} - f_e(k-l, \delta)} \right) \right\}. \tag{3.52}
 \end{aligned}$$

By rewriting we get given a sufficiently large level of heterogeneity  $\alpha > 1$  the following condition for unilaterally stable core periphery networks in terms of linking costs:

$$\begin{aligned}
 c \in \left( \right. & \frac{1}{2} - f_e(k, \delta) + (n-k-2) \min \left\{ \frac{1}{2} f_m(k+1, \delta), f_e(k+1, \delta) - f_e(k, \delta) \right\}, \\
 & \min \left\{ \alpha^2 \left( \frac{1}{2} - f_e(n-2, \delta) \right), \alpha \left( \frac{1}{2} - f_e(k-1, \delta) \right) + \frac{1}{2}(n-k)(n-k-1)f_m(k, \delta), \right. \\
 & \left. \min_{l \leq k} \left\{ \alpha \left( \frac{1}{2} - f_e(k-l, \delta) \right) + \frac{n-k-1}{l}(f_e(k, \delta) - f_e(k-l, \delta)) \right\} \right\} \left. \right).
 \end{aligned}$$

□

**Remarks on Proposition 3.3.** Notice that, for given  $c$  and  $\delta$  and for  $n$  sufficiently large, the

level of heterogeneity  $\bar{\alpha}$  as given in (3.52) equals

$$\sqrt{c / \left( \frac{1}{2} - f_e(k-1, \delta) \right)}. \quad (3.53)$$

For such a value of  $\alpha$ , the complete core periphery network with  $k$  big banks in the core is unilaterally stable if condition (3.20) is satisfied. To minimise  $\bar{\alpha}$  we take the smallest value of  $c$  satisfying (3.20). A lower bound for  $\bar{\alpha}$ , given values of  $\delta$ ,  $k$  and sufficiently large  $n$ , is thus given by:

$$\bar{\alpha} \geq \bar{\alpha}_{min} \equiv \sqrt{\frac{\frac{1}{2} - f_e(k, \delta) + (n-k-2) \min\{\frac{1}{2} f_m(k+1, \delta), f_e(k+1, \delta) - f_e(k, \delta)\}}{\frac{1}{2} - f_e(k-1, \delta)}}. \quad (3.54)$$

Using the assumptions we made about the distribution of intermediated trades in Section 3.2.1, one can verify that

$$\lim_{k \rightarrow \infty} \bar{\alpha}_{min} = \lim_{\delta \downarrow 0} \bar{\alpha}_{min} = \lim_{\delta \uparrow 1} \bar{\alpha}_{min} = 1, \quad (3.55)$$

showing that an arbitrary small level of heterogeneity can be sufficient to have an unilaterally stable core periphery network.

***Proof of Theorem 3.3.***

Starting in an empty network, the first big bank  $i = 1$  has two relevant options: either connect to all players or connect only to big bank 2. Linking to only part of the small banks cannot be a best feasible action, because linking to an additional small bank pays off positive additional intermediation benefits (similar to the proof of Lemma 3.1). Player 1's payoffs depending on

its action  $S$  are:

$$\pi_1(S) = \begin{cases} 0 & \text{if } S = \emptyset \\ \frac{1}{2}\alpha^2 - c & \text{if } S = \{2\} \\ \frac{1}{2}\alpha^2 - c + 2(\frac{1}{2}\alpha - c) + 2\frac{1}{3}\alpha + \frac{1}{3} \\ = \frac{1}{2}\alpha^2 + \frac{5}{3}\alpha + \frac{1}{3} - 3c & \text{if } S = \{2, 3, 4\}. \end{cases} \quad (3.56)$$

Under condition I, player 1's best feasible action is not to add any links. The resulting network is empty. In this case the network must be unilaterally stable. The reason is that second big bank faces the same decision as  $i = 1$ , and small banks have strictly lower payoffs from adding links, so no player will decide to change the network structure.

Under condition VI, the best feasible action is to add only one link to big bank 2. The second big can choose from the same resulting networks as  $i = 1$  could, so does not add or remove links. Small banks have strictly lower payoffs from adding links and also do not change the structure. So under condition VI the resulting network with only one link, namely  $g_{12} = 1$ , is stable.

If the first player adds links to all three other players, a star network is formed. Then the best response for the second big bank  $i = 2$  is either to do nothing or to add links to both periphery players. If 2 does nothing, the periphery nodes 3 and 4 will likewise decide not to add links, and the star network is the final, stable outcome. For 2 to have positive marginal benefits of supporting two links, it is required that

$$\begin{aligned} M_2(g^s, +2) &> 0 \\ \Rightarrow c &< \frac{1}{6}\alpha + \frac{1}{2}f_m(2, \delta), \end{aligned} \quad (3.57)$$

and each peripheral player  $j \in \{3, 4\}$  requires

$$\begin{aligned} M_j(g^s, +2) &\geq 0 \\ \Rightarrow c &\leq \frac{1}{6}\alpha + f_e(2, \delta) - \frac{1}{3}. \end{aligned} \quad (3.58)$$

Conversely, the star network is stable if the deviation of player 2 is *not* beneficial to 2 and/or a peripheral player  $j$ , i.e. if  $c$  exceeds the minimum of the two values in equations (3.57) and (3.58):

$$c \geq \frac{1}{6}\alpha + \min\left\{\frac{1}{2}f_m(2, \delta), f_e(2, \delta) - \frac{1}{3}\right\}, \quad (3.59)$$

leading to condition II.

Consider the case that both big banks have added all possible links. Then there is a possibility that the dynamic process lead to a complete network if node  $i = 3$  links to 4. This happens if  $c$  is so small that the share  $f_e(2, \delta)$  from intermediated trade between 3 and 4 can be raised to  $\frac{1}{2}$  by creating a direct link. The complete network is therefore stable under condition III.

For parameters not satisfying I, II, III or VI, the first round of best responses results in a complete core periphery network with 2 core members. This complete core periphery network can be stable because of the heterogeneity between core and periphery banks with  $\alpha > 1$ . As stated in Proposition 3.3, this occurs under condition VII. For  $n = 4$  and  $k = 2$  condition VII reduces to:

$$\begin{aligned} c &\in \left( \frac{1}{2} - f_e(k, \delta) + (n - k - 2) \min\left\{\frac{1}{2}f_m(k + 1, \delta), f_e(k + 1, \delta) - f_e(k, \delta)\right\}, \right. \\ &\quad \min\left\{ \alpha^2\left(\frac{1}{2} - f_e(n - 2, \delta)\right), \alpha\left(\frac{1}{2} - f_e(k - 1, \delta)\right) + \frac{1}{2}(n - k)(n - k - 1)f_m(k, \delta), \right. \\ &\quad \left. \left. \min_{l \leq k} \left\{ \alpha\left(\frac{1}{2} - f_e(k - l, \delta)\right) + \frac{n - k - 1}{l}(f_e(k, \delta) - f_e(k - l, \delta)) \right\} \right\} \right) \\ &\in \left( \frac{1}{2} - f_e(2, \delta), \min\left\{ \alpha^2\left(\frac{1}{2} - f_e(2, \delta)\right), \frac{1}{6}\alpha + f_m(2, \delta), \frac{1}{6}\alpha + f_e(2, \delta) - \frac{1}{3} \right\} \right). \end{aligned}$$

Notice that when this condition is fulfilled, the second player always adds the two links to the



periphery (cf. condition (3.59)).

Finally, in the remaining region IV, the complete core periphery network cannot be stable. Because the node  $i = 3$  node did not connect to the last periphery node, adding links cannot be a best response. Therefore the first core bank  $i = 1$  must have an incentive to delete the link with 2. The attracting steady state is the multipartite (ring) network consisting of the groups  $\{1, 2\}$  and  $\{3, 4\}$ .  $\square$

## Appendix 3.D Longer intermediation chains

In this appendix we generalise the model to allow for intermediation paths of lengths longer than two. For simplicity, we will analyse the general model for lengths up to distance three and under homogeneity (i.e.  $\alpha_{ij} = 1$  for all  $i, j$ ). These simplifications make it possible to check that our main result – a core periphery structure can form in *heterogeneous* financial networks only – holds in the most general model.

Before rewriting a more general form of the payoff function (3.5) formally, we repeat that the proof of Proposition 3.2 does not require any assumptions on the path lengths on which trade is allowed; see the remarks in 3.C on this proposition. Proposition 3.2 states that complete core periphery networks are not pairwise (or unilaterally) stable, and that incomplete core periphery networks are not unilaterally stable if  $n$  is sufficiently large. Also Proposition 3.3, specifying a level of heterogeneity sufficient for a complete core periphery network to be unilaterally stable, holds for longer intermediation paths.

We introduce new, more general notation for intermediation over longer path lengths: the shares in the trade surplus from  $i$  and  $j$  are denoted by  $F_e(g, \{i, j\}, \delta)$  for the endnodes, and  $F_m(g, \{i, j, k\}, \delta)$  for middleman  $k$ . In 3.B it is explained how such a distribution can be derived for long intermediation chains in the example of Siedlarek (2012a). Note that if  $i$  and  $j$  at length three are assumed to generate a surplus, middlemen involved in a trade between  $i$  and  $j$  do not

necessarily earn the same: if  $k$  lies on more of the shortest paths than  $k'$ ,  $k$  will earn more than  $k'$ . For this reason (part of) the graph  $g$  must be given as an argument in the function  $F_e$  and  $F_m$ . The payoff function becomes:

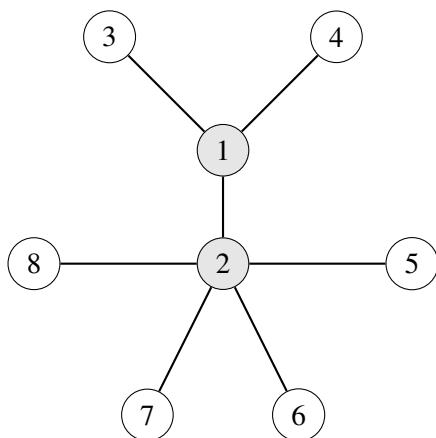
$$\pi_i(g) = n_i\left(\frac{1}{2} - c\right) + \sum_{j \in N_i^3(g)} F_e(g, \{i, j\}, \delta) + \sum_{\substack{k, l \in N_i^2(g) \\ g_{kl} = 0, d_{kl} \leq 3}} F_m(g, \{k, l, i\}, \delta), \quad (3.60)$$

where  $N_i^r(g)$  denotes the set of nodes at distance  $r$  from  $i$  in network  $g$ ,  $n_i = |N_i^1(g)|$  the number of direct connections of  $i$ , and  $d_{kl}$  the distance between nodes  $k$  and  $l$ .

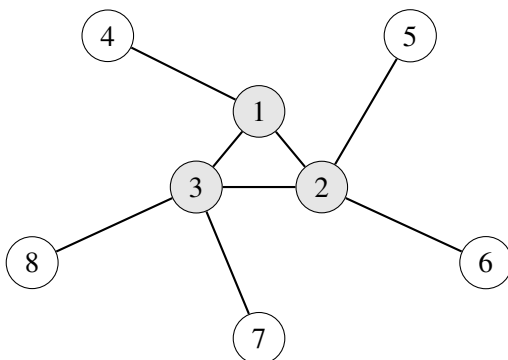
First note that the results for the stability of the star network does not depend on the assumption of maximum intermediation chains of two. Obviously, in the star all pairs are at distance one or two of each other. If links to one or more periphery players are deleted, the distance to them becomes non-defined (because they become isolated nodes outside the main component). If links are added, the distance between pairs of nodes can only decrease. So considerations of intermediation chain of length three do not play a role for the star.

In incomplete core periphery networks, shortest paths of three may exist between some periphery players, which were previously assumed not to generate any trading surplus. By allowing intermediation chains of three we found that some core periphery networks can become unilaterally stable. For  $n = 8$ , we found that  $k = 2$  and  $k = 3$  are the only possibly stable core sizes. See Figure 3.12 for two examples of networks that are stable for the given parameter values when paths of three are allowed. These two network structures are not stable in the baseline model.

We can safely interpret these examples as low-dimensional exceptions to the rule that the core periphery structure in homogeneous networks is generally unstable. The examples in Figure 3.12 show that for  $n = 8$  incomplete core periphery networks with core sizes of  $k = 2$  and  $k = 3$  can be stable. For higher network sizes, however, core periphery networks are always unstable as stated by Proposition 3.2.



(a)  $(\delta, c) = (0.8, 1.1)$ . A minimally connected core periphery network with  $k = 2$ .



(b)  $(\delta, c) = (0.5, 0.9)$ . A minimally connected core periphery network with  $k = 3$ .

Figure 3.12: Examples of unilaterally stable core periphery networks after allowing for intermediation chains of length 3, for  $n = 8$  players and given  $(\delta, c)$ .

Moreover, even though exceptionally for small  $n$  core periphery networks can be unilaterally stable, they are never the outcome of a dynamic process as described in Subsection 3.4.3. For  $n = 8$ , the core periphery networks with  $k = 2$  and  $k = 3$  were found to be stable in parameter regions where the star networks is stable, cf. region II in Figure 3.7b. Exploring the parameter space by simulations, we found that this was always the case for such stable core periphery networks. By Lemma 3.1, the star is created as a first step in the dynamic process whenever the initial empty network is not stable. Therefore the star network is the outcome of a dynamic process even when exceptional (low-dimensional) networks are unilaterally stable as well given the parameter values. This implies that the dynamic results of Theorem 3.2 do not depend on the assumption of maximal intermediation paths of length two, as was already shown for the static results of Propositions 3.2 and 3.3.

## **Part II**

# **Applying complex systems to the stock market**



# Chapter 4

## Booms, busts and behavioural heterogeneity in U.S. stock prices

### 4.1 Introduction

Economic reality shows the limitations of standard asset pricing models with a representative rational agent only concerned with economic fundamentals. In 2008 the S&P500 stock index, the financial bellwether of the dominant market economy of the U.S., and many other stock indices, lost around one half of their total value. While the bankruptcy of Lehman Brothers amounted to a clear fundamental shock to the economy, it is not plausible that all of this loss can be attributed to rational re-evaluation of fundamentals. Other behavioural explanations need to be considered. In this chapter we present evidence from S&P500 data that market sentiment switches between different behavioural regimes, which amplified shocks such as the Lehman bankruptcy, and more generally amplifies booms and busts of the economy.

We first apply the idea of switching market sentiment to a basic framework that provides a fundamental value of the price-dividend ratio: the standard Gordon solution based on a constant risk premium. Within this framework we introduce a simple behavioural model with some agents believing in mean-reversion of stock prices (called fundamentalists) and others (called

chartists) who expect a continuation of the trend. Agents gradually switch between the two rules, based upon their relative performance, so they learn and adapt their behaviour if the market situation changes and the losses of their strategy become too large. Because of the positive expectations feedback in asset markets, self-reinforcing behavioural regimes of mean-reversion and trend-following arise endogenously in the model, explaining large and persistent deviations of the S&P500 from the Gordon fundamental value.

A convenient feature of our model is that it is formulated in deviations from a fundamental price, so that it can be tested against any suitable fundamental benchmark. Behavioural heterogeneity can therefore complement the mainstream financial literature on stock market fluctuations by providing an *amplification mechanism* to explain excess volatility (Shiller, 1981). To this end we combine our model with the consumption-habit asset pricing model of Campbell and Cochrane (1999). They argue in a standard representative-agent framework that booms and busts in asset prices are driven by countercyclical variation in risk premia, which in turn are inversely related to consumption relative to a slow-moving habit level. We show that even if part of the variation in the price-dividend ratio can be explained by consumption-driven variation in risk premia, our model still gives significant parameter estimates and adds explanatory power due to behavioural heterogeneity. Overall, we argue that there is strong evidence for heterogeneous beliefs amplifying booms and busts in the stock market.

Standard asset pricing models do not take heterogeneity into account as these models assume the expectations of individual investors are rational and can be described by a representative agent. Asset prices should in this view equal the *fundamental value* of expected discounted sum of future cashflows, or more specifically dividend payoffs (Campbell and Shiller, 1988a).<sup>1</sup> Various reasons have been proposed why this fundamental value could change over time, as in Campbell and Cochrane (1999). Bansal and Yaron (2004) argue for the effects of long-run economic uncertainty on asset prices. Pástor and Veronesi (2006) and Ofek and Richardson (2003)

---

<sup>1</sup> There is also a large literature of modeling asset prices based on book values or earnings, e.g. Campbell and Shiller (2001). Boswijk, Hommes, and Manzan (2007) estimate an asset pricing model with behavioural heterogeneity both for fundamental valuation based on dividends and earnings, and find robust results.



give particular (but very different) explanations for the high valuations of technology firms in the late 1990s. Nevertheless, these explanations may not be sufficient to fully explain stock market fluctuations. More specifically, we show that for the consumption-habit model of Campbell and Cochrane (1999), behavioural heterogeneity is a significant amplification mechanism.

With the contention that the financial crisis cannot be sufficiently explained by economic fundamentals, this chapter fits within the *behavioural* finance literature. Departing from the strongest form of rationality opens up the alternative view that stock prices may have been overpriced. The behavioural finance literature is surveyed in e.g. Hirshleifer (2001) and Barberis and Thaler (2003). Barberis and Thaler (2003) stress the finding that traders with flawed expectations can not always be driven away from the market. As these traders distort supply and demand based on fundamentals, assets can be partly mispriced. In their words: “One of the biggest successes of behavioral finance is a series of theoretical papers showing that in an economy where rational and irrational traders interact, irrationality *can* have a substantial and long-lived impact on prices.” (Barberis and Thaler, 2003, p. 1053, their emphasis).

Barberis and Thaler (2003) also state that careful empirical analysis remains the main challenge for behavioural models. As one recent example, Branch and Evans (2010) develop a framework with agents learning the parameters of their underparameterised forecasting models and reproduce regime-switching returns and volatilities in monthly U.S. stock data. Adam and Marcet (2011) and Adam, Beutel, and Marcet (2013) provide another example where investors’ subjective beliefs are shown to drive booms and busts in the S&P 500’s price-dividend ratio. In these examples however the model is calibrated to replicate certain characteristics in the data. Our simpler behavioural model with heterogeneous agents contains few parameters and can be estimated directly.

Our boundedly rational traders will be modeled within the *heterogeneous agents* asset pricing framework of Brock and Hommes (1997, 1998)<sup>2</sup>. The literature on heterogeneous agents

---

<sup>2</sup> Other related early heterogeneous agents models include the noise trader models of DeLong, Shleifer, Summers, and Waldmann (1990a,b) and the model with ‘newswatchers’ versus momentum traders of Hong and Stein (1999). These models also assume bounded rationality of (at least one type of) agents, but do not allow for switching between different strategies.

models (HAMs) has been growing in the last decades and is extensively reviewed in e.g. Hommes (2006), LeBaron (2006) and Lux (2009). For example, HAMs have been applied to stock prices empirically in Boswijk, Hommes, and Manzan (2007), Chiarella, He, and Zwinkels (2014) and Lof (2012, 2014). Switching models with heterogeneous agents have also been applied to other financial markets, in particular exchange rates (Kirman and Teysnière, 2002; Westerhoff and Reitz, 2003; Alfarano, Lux, and Wagner, 2005; de Jong, Verschoor, and Zwinkels, 2010), but also for example to option prices (Frijns, Lehnert, and Zwinkels, 2010) and oil prices (ter Ellen and Zwinkels, 2010). This empirical literature is growing fast, see e.g. Cheng, Chang, and Du (2012) for an overview.

From a methodological viewpoint, this chapter contributes to the empirical literature on heterogeneous agents models in several ways. Most importantly, we generalise the asset pricing model with heterogeneous agents and test it against the Campbell-Cochrane consumption-habit model as well as the dynamic Gordon model. A novelty in the empirical HAM literature is that we introduce agents' memory of earlier realised excess returns. This will lead to gradual (rather than instant) switching and makes the model applicable to relatively high frequency (i.e. quarterly) data while a simple economic interpretation remains. We run Monte Carlo simulations to clarify two difficulties in estimating HAMs: the stationarity of the time series and the significance of the switching intensity. We also look in greater detail at the price dynamics in the recent turbulent years in terms of fundamentals and amplification mechanisms. With simulated time series starting before the dot-com bubble and the global financial crisis, we show that these events were not particularly 'extreme events' from a behavioural heterogeneous agents perspective.

Many factors have contributed to the rising interest in behavioural heterogeneity. First, laboratory experiments with human subjects have been performed to study individual expectations and aggregate outcomes, e.g. Hommes, Sonnemans, Tuinstra, and van de Velden (2005, 2008). Experimental studies have the benefit that the underlying asset market fundamentals can be fully controlled; for an overview of the use of laboratory experiments to test for heterogeneous expect-

tations, see Hommes (2011). Anufriev and Hommes (2012) find in experimental asset pricing data that subjects switch between different forecasting rules, consistent with the theoretical model of Brock and Hommes (1998). An interesting finding from these laboratory experiments is that under *positive expectations feedback* coordination on trend-following strategies amplifies asset market fluctuations (Heemeijer, Hommes, Sonnemans, and Tuinstra, 2009).

Second, empirical evidence has shown that switching based on past performance is relevant for real financial markets. For example, Ippolito (1989), Chevalier and Ellison (1997), Sirri and Tufano (1998) and Karceski (2002) found in mutual funds data that money flows out of past poor performers into good performers. Pension funds also switch away from bad performers (Del Guercio and Tkac, 2002). Investors in the stock market can be expected to display similar switching behaviour when choosing between different strategies.

Third, there is growing interest in survey data on expectations of financial specialists, which can be traced back to Frankel and Froot (1987). Comparing six different data sources, Greenwood and Shleifer (2013) show that surveys of stock market investors are highly positively correlated with each other, supporting the idea that they do reflect actual beliefs. The heterogeneity in price expectations also changes over time, as shown in Shiller (1987, 2000) and Vissing-Jorgensen (2004). All three forms of microlevel evidence of actual traders shifting between simple behavioural rules motivate our aggregate model.

We emphasise that the behaviourally heterogeneous expectations of our investors are not model-consistent as in the traditional rational expectations framework. Yet agents are boundedly rational in the sense that they switch to better performing rules, which then become almost self-fulfilling. We show that the data supports self-reinforcing temporary coordination on either mean-reversion or trend-following. Still, on which type of behaviour agents will coordinate is difficult to foresee in advance: the market is unpredictable in the short run. Fundamentals play a complementary role in explaining mean-reversion in stock market fluctuations and make prices predictable in the long run. Overall, strategy switching serves as an amplification mechanism for booms and busts.

The chapter is organised as follows. Section 5.2 develops the general asset pricing model with heterogeneous agents. In Section 4.3 we present our main estimation results under the standard Gordon fundamental value. Section 4.4 provides Monte Carlo simulations to test the robustness of our results, as well as simulated time series generated by our model that illustrate the endogenous behavioural regimes. In Section 4.5 we combine our model with the consumption-habit model of Campbell and Cochrane (1999) that has a time-varying risk premium. Section 4.6 concludes.

## 4.2 Model description

We derive a stylised asset pricing model with heterogeneous agents, generalising Brock and Hommes (1998) and Boswijk, Hommes, and Manzan (2007) to a model that allows for time-variation in dividends and discount rates. We assume that investors have perfect knowledge of the underlying fundamental process, and are therefore able to calculate the ‘fundamental value’, which in this section will be derived in general terms. In Section 4.3, we will specify the present value model with a constant risk premium based on Gordon (1962) for the fundamental value. In Section 4.5, we consider another benchmark fundamental value with a time-varying risk premium, the consumption-habit model of Campbell and Cochrane (1999).

Even though agents know the fundamental value, they have different beliefs about what the price of the asset will be in the next time period. The changes of agents’ beliefs will lead to fluctuating market sentiment. For both benchmark fundamental value models in Sections 4.3 and 4.5, we will show that there is significant evidence for behavioural heterogeneity in the data.

Consider a risky financial asset that pays a random dividend payoff  $D_t$  at time  $t$ . The opportunity cost for investing in the risky asset is captured by the discount rate  $R_{t+1}$  which may in

general vary over time. The standard pricing equation (cf. e.g. Cochrane, 2001, p. 10) is:

$$P_t = \bar{E}_t \left[ \frac{P_{t+1} + D_{t+1}}{R_{t+1}} \right], \quad (4.1)$$

where  $\bar{E}_t[\cdot]$  denote average expectations over all (possibly heterogeneous) agents. Today's price is the expected discounted sum of tomorrow's price and of tomorrow's dividend payoff.

Within this simple and general asset pricing framework, we will derive a heterogeneous agents model that can be estimated given specifications of the underlying fundamental process and agents' beliefs. In Section 4.2.1 we discuss the fundamental value and in Section 4.2.2 the behavioural expectation rules of the agents. Finally, in Section 4.2.3 we present our model in its econometric form. In each section the pricing equation (4.1) is specified in more detail.

### 4.2.1 Fundamental value

We focus on possible belief disagreement about future prices, and assume agreement about fundamentals. This approach reflects the idea of investors that prices are determined endogenously and partly depend on expectations about the next period's price, while the fundamentals follow an exogenous stochastic process. Thus, all agents have identical beliefs about the dividend  $D_{t+1} = (1 + g_{t+1})D_t$  and its discounted value:

$$E_t \left[ \frac{D_{t+1}}{R_{t+1}} \right] = E_t \left[ \frac{1 + g_{t+1}}{R_{t+1}} \right] D_t. \quad (4.2)$$

We can therefore rewrite the pricing equation (4.1) in terms of the price-dividend ratio (PD) ratio  $\delta_t \equiv P_t/D_t$  as:

$$\delta_t = \bar{E}_t \left[ \frac{1}{R_{t+1}} \frac{D_{t+1}}{D_t} (\delta_{t+1} + 1) \right]. \quad (4.3)$$

In the special case when all agents have rational expectations  $E_t$  the pricing equation (4.1)

simplifies to:

$$P_t = E_t \left[ \frac{P_{t+1} + D_{t+1}}{R_{t+1}} \right]. \quad (4.4)$$

The *fundamental value*  $P_t^*$  is obtained under rational expectations from the present value of all future cash flows (see e.g. Boswijk, Hommes, and Manzan, 2007, p. 1965):

$$\begin{aligned} P_t^* &= E_t \left[ \sum_{j=1}^{\infty} \prod_{k=1}^j \left( E_{t+k-1} \left[ \frac{1}{R_{t+k}} \right] \right) D_{t+j} \right] \\ &= E_t \left[ \sum_{j=1}^{\infty} \prod_{k=1}^j \left( E_{t+k-1} \left[ \frac{1+g_{t+k}}{R_{t+k}} \right] \right) D_t \right], \end{aligned} \quad (4.5)$$

and the fundamental PD ratio equals:

$$\delta_t^* = E_t \left[ \sum_{j=1}^{\infty} \prod_{k=1}^j E_{t+k-1} \left[ \frac{1+g_{t+k}}{R_{t+k}} \right] \right]. \quad (4.6)$$

The fundamental values of the price and PD ratio are presented here in the most general form, but will simplify in subsequent sections to the special cases of Gordon (1962) and Campbell and Cochrane (1999) by assumptions on  $g_{t+1}$  and  $R_{t+1}$ .

We assume that each agent  $i$  knows the fundamental price  $P_t^*$ , but is uncertain about whether the next periods price will be equal to  $P_{t+1}^*$  because other investors may use non-fundamental trading decisions. Therefore agent type  $h$  tries to predict the next period's  $P_{t+1}$  by its subjective expectations  $E_{h,t}[P_{t+1}]$  which may differ from  $P_{t+1}^*$ . To allow for a stationary underlying proces, we focus on the subjective expectation about the price-dividend ratio  $E_{h,t}[\delta_{t+1}]$ , which consequently also possibly differ from  $\delta_{t+1}^*$ .

### 4.2.2 Behavioural heterogeneous beliefs

To model heterogeneity we consider  $H$  types of investors using different expectation rules. Average expectations  $\bar{E}_t[\cdot]$  then becomes a weighted sum of beliefs. The fractions or weights

of agents using a particular belief  $E_{h,t}$  are denoted  $n_{h,t}$ . The PD pricing equation (4.3) therefore becomes:

$$\delta_t = \sum_{h=1}^H n_{h,t} E_{h,t} \left[ \frac{1 + g_{t+1}}{R_{t+1}} (\delta_{t+1} + 1) \right]. \quad (4.7)$$

An important idea of the model is to separate behavioural factors influencing prices from fundamental factors. More precisely, we assume common beliefs on fundamental factors such as growth rates  $g_{t+1}$  and discount rates  $R_{t+1}$ :

$$\begin{aligned} E_{h,t} \left[ \frac{1 + g_{t+1}}{R_{t+1}} (\delta_{t+1} + 1) \right] &= E_t \left[ \frac{1 + g_{t+1}}{R_{t+1}} \right] E_{h,t} [\delta_{t+1} + 1] \\ &= \frac{1}{R^*} E_{h,t} [\delta_{t+1} + 1]. \end{aligned} \quad (4.8)$$

We define the *unconditional* expectation  $\frac{1}{R^*} \equiv E \left[ E_t \left[ \frac{1 + g_{t+1}}{R_{t+1}} \right] \right]$  as the *expected effective discount factor* for pricing stocks in terms of the PD ratio. Equation (4.7) becomes:

$$\delta_t = \frac{1}{R^*} \sum_{h=1}^H n_{h,t} E_{h,t} [\delta_{t+1} + 1]. \quad (4.9)$$

It will be convenient to formulate the model in deviations from the fundamental value  $x_t \equiv \delta_t - \delta_t^*$ . We assume that all agents have common and rational beliefs about the fundamental value:

$$E_{h,t} [\delta_{t+1}^*] = E_t [\delta_{t+1}^*] = R^* \delta_t^* - 1. \quad (4.10)$$

Furthermore, agents' beliefs about the next period's PD ratio are assumed to be separable from this belief about the fundamental:

$$\begin{aligned} E_{h,t} [\delta_{t+1}] &= E_t [\delta_{t+1}^*] + E_{h,t} [x_{t+1}] \\ &= E_t [\delta_{t+1}^*] + f_h(x_{t-1}, \dots, x_{t-L}), \end{aligned} \quad (4.11)$$

where  $E_{h,t}[x_{t+1}]$  represents the expected deviation of the PD ratio from the fundamental value and is expressed as a function  $f_h(\cdot)$  of the  $L$  last observed deviations. Note that agents at time  $t$  do not observe the contemporaneous price and react to past realised prices only. This assumption is common in the literature and for example used by Hong and Stein (1999) to model momentum traders.

Under these assumptions about heterogeneous expectations, price deviations from the fundamental are priced as:

$$\begin{aligned} x_t = \delta_t - \delta_t^* &= \frac{1}{R^*} \sum_{h=1}^H n_{h,t} E_{h,t}[\delta_{t+1} + 1] - \delta_t^* \\ &= \frac{1}{R^*} \sum_{h=1}^H n_{h,t} E_{h,t}[x_{t+1}]. \end{aligned} \quad (4.12)$$

The standard asset pricing model based on future stock prices and dividends (4.1) has now been reformulated as a dynamic HAM (4.12) in which price deviations from the fundamental value depend only on discounted expected future price deviations.

We stress that as the model is formulated in deviations from a fundamental PD ratio, it can be used with different benchmark fundamentals. The crucial assumption that has been made is that agents have common beliefs about fundamental factors. Also note that the fundamental benchmark with a rational representative agent is nested as a special case of our model when all agent types believe  $E_{h,t}[x_{t+1}] = 0$ , for all  $h = 1, \dots, H$ . The model for price fluctuations around the fundamental value (12) holds for any choice of agent types and for any choice of the fundamental value  $\delta_t^*$ . This setup is convenient to test empirically whether any deviations from a benchmark fundamental are significant.

For the empirical estimation, we consider the simplest form of heterogeneity in belief types which are linear in the last observation:

$$E_{h,t}[x_{t+1}] = \phi_h x_{t-1}. \quad (4.13)$$



Choosing  $H = 2$  types is sufficient to capture an essential difference between agents. Some agents (called fundamentalists) believe in *mean-reversion* of the stock price to its fundamental value and have a parameter  $(0 <) \phi_1 < 1$ . Other agents (called chartists) believe that the price (in the short run) will move away from the fundamental value and have  $\phi_2 > 1$ . Chartists expect a continuation of the trend and will be a destabilising factor in the model when their impact becomes large.

The behavioural finance literature has a long tradition of models with fundamentalists and chartists; see Hommes (2006) and LeBaron (2006) for extensive surveys. In a recent overview of the empirical HAM literature, Cheng, Chang, and Du (2012) classify the broad variety of agent-based economic papers and underline that the simple fundamentalist-chartist opposition is often sufficient to explain stylised facts from asset price data that seem ‘puzzles’ in a rational representative agent framework. Aoki (2002) argues with a theoretical model that the behaviour of many different market participants can often be clustered in just two groups. Another recent example is Lof (2014), who applies the VAR approach to a heterogeneous asset pricing model with fundamentalists and contrarians.<sup>3</sup>

Agents use simple rules to predict future prices, but switch to other strategies if their predictions become too far-off from actual prices: investors learn from their mistakes. Our agents learn through reinforcement learning or evolutionary selection based upon the relative performance of their forecasting rule. The fractions of agents belonging to one of the two types are updated with a multinomial logit model as in Brock and Hommes (1997) with *intensity of choice*  $\beta^4$ :

$$n_{h,t+1} = \frac{e^{\beta U_{h,t}}}{\sum_{j=1}^H e^{\beta U_{j,t}}}. \quad (4.14)$$

In order to specify the performance measure of belief type  $h$ ,  $U_{h,t}$ , we need to consider the profits of agent types. Following Brock and Hommes (1998) and Boswijk, Hommes, and

<sup>3</sup> Others have proposed models with two types that have more advanced, time-varying adaptive learning beliefs, for example Branch and Evans (2010).

<sup>4</sup> Again, our choice for fluctuating fractions is founded in earlier behavioural finance literature. Cheng, Chang, and Du (2012) state that “*evolving fractions* have been considered to be a [main] cause of many stylised facts” (p. 15, their emphasis).

Manzan (2007), we consider the following profit function in price deviations  $x_t$ :

$$\pi_{h,t+1} = (E_{h,t}[x_{t+1}] - R^* x_t)(x_{t+1} - R^* x_t). \quad (4.15)$$

This expression for realised profits has the intuitive property that it is proportional to agents' demand (depending on the expectations  $E_{h,t}[x_{t+1}] - R^* x_t$ ) times the realised excess return (depending on realisations  $x_{t+1} - R^* x_t$ ). In Appendix 4.A we explicitly derive this profit function from mean-variance maximising investors.<sup>5</sup>

It will turn out that for our application of the model to quarterly data, realised profits of more than one period in the past should be accounted for in the performance measure. To this end we introduce in the performance measure a *memory* parameter  $\omega$ :

$$U_{h,t} = (1 - \omega)\pi_{h,t} + \omega U_{h,t-1}, \quad (4.16)$$

so that the most recent observed profit receives weight  $(1 - \omega)$ . The relative weight of the  $j$ -th lag of realised profits is thus  $\omega^j(1 - \omega)$  and decreases in  $j$ .<sup>6</sup>

We now have a complete specification of the fluctuating fractions of the  $H = 2$  belief types  $n_{1,t}$  and  $n_{2,t}$ . As profits in equation (4.42) for a certain belief  $h$  increase and its performance measure exceeds that of the other belief, more agents will choose this belief, according to the multinomial logit model (5.15). Thus there is a positive relation between realised profits and fractions of the agents' belief types.

### 4.2.3 Econometric form

The HAM in equation (4.12) with the additional assumptions about expectation formation can be written as an econometric AR(1) model with a time-varying coefficient after adding an error

---

<sup>5</sup> Strictly speaking, the profit function is proportional to the expression in (4.42) by a constant factor  $C$  which is captured in the estimate  $\beta^* = \beta C$ .

<sup>6</sup> In contrast, Boswijk, Hommes, and Manzan (2007) define the performance measure as the last period's profit:  $\tilde{U}_{h,t} = \pi_{h,t}$ .

term:

$$\begin{aligned} x_t &= \frac{1}{R^*} (n_{1,t} \phi_1 + n_{2,t} \phi_2) x_{t-1} + \varepsilon_t, \\ &\equiv \varphi_t x_{t-1} + \varepsilon_t. \end{aligned} \quad (4.17)$$

The error terms are assumed to be independently and identically distributed:  $\varepsilon_t \sim IID(0, \sigma^2)$ .<sup>7</sup>

The time-varying coefficient  $\varphi_t$  replaces the constant parameter in a regular AR(1)-model and is interpreted here as the *average market sentiment*. As market sentiment rises, prices stay for a longer number of periods away from the fundamental value. Combining equations (5.15), (4.42) and (4.16), fractions depend nonlinearly on the four parameter and all past realisations:<sup>8</sup>

$$n_{1,t} = f_{n1}(\phi_1, \phi_2, \beta, \omega; x_{t-1}, x_{t-2}, \dots, x_1), \quad (4.18)$$

$$n_{2,t} = 1 - n_{1,t}. \quad (4.19)$$

Equations (4.17), (4.18) and (4.19) summarise our heterogeneous agents model for price deviations from any fundamental value. The key idea of the model is *positive expectations feedback*. Initially there is some distribution of fundamentalists and chartists. As shocks are feeded into realised prices, one of the two strategies may receive a higher payoff and attracts more followers given an intensity of choice  $\beta > 0$ . Because the price is determined by market clearing, these fractions affect the next period price: if overall market sentiment is higher, the next period's price will also be higher. This leads to almost self-fulfilling expectations. For example, when chartists (with  $\phi_2 > 1$ ) dominate, temporary bubbles may arise triggered by

<sup>7</sup> Strictly speaking, we do not require the error terms to be fully IID, but they should be uncorrelated.

<sup>8</sup> The exact mathematical formula for the fraction of fundamentalists (not important for our line of argument) equals:

$$n_{1,t} = (1 + \exp[\beta(\phi_1 - \phi_2) \sum_{j=0}^{t-4} [\omega^j (1 - \omega) x_{t-3-j} (x_{t-1-j} - R^* x_{t-2-j})]])^{-1},$$

In this equation the index  $j = 1, 2, \dots$  corresponds to the  $j$ -th lag of realised profits that enters the performance measure through memory in (4.16). At  $j = t - 4$ , the first observation  $x_1$  is used in determining the fraction  $n_t$ , which puts an upper bound on the memory of realised profits. It should be clear that this formula can only be used for  $t \geq 4$ .

fundamental shocks and amplified by trend-following expectations.

The goal of estimating our model is to quantify the effect of positive feedback and switching regimes of market sentiment. This obviously depends on the estimations of the parameters  $\phi_1$  and  $\phi_2$ . If these parameters are closer to each other, the effects of switching decreases. We will show that the difference between  $\phi_1$  and  $\phi_2$  is statistically significant, and also economically significant, in the sense that the heterogenous agents model produces substantially different market predictions than a representative agent model.

### 4.3 Estimation results for a simple behavioural model

The estimation follows a two-step procedure. First we estimate the fundamental value of the Gordon model based on dividends and a constant risk premium. Second, we estimate the heterogeneous agents model summarised by equations (4.17), (4.18) and (4.19) with nonlinear least squares. In Section 4.3.3 we introduce a linear representative agent benchmark model to compare our estimation with.

#### 4.3.1 Estimating the fundamental value of the Gordon model

In this section we specify the fundamental values of stock prices  $P_t^*$  and PD ratios  $\delta_t^*$  using the standard model based on Gordon (1962). The textbook Gordon solution for the fundamental PD ratio under discrete time is constant and equal to:

$$\delta^* = \frac{1+g}{r-g}, \quad (4.20)$$

where  $g$  is the expected growth rate of dividends, and  $r = i + RP$  is the sum of the expected risk free rate  $i$  and the risk premium on stocks  $RP$ , both assumed to be constant. This follows immediately from substituting  $g_{t+1} = g$  and  $R_{t+1} = 1 + r$  in equation (4.6).

We follow Boswijk, Hommes, and Manzan (2007) in using the *dynamic* Gordon model instead of the standard (static) Gordon model. In the dynamic Gordon model, agents can extract

possible changes in the future parameters  $g_{t+1}$  and  $r_{t+1}$  from data on dividend growth rates and interest rates available at time  $t$ . This approach is more flexible and allows for time variation in the fundamental PD ratio around  $\delta^*$ . We will show, however, that the time variation in the fundamental PD ratio of the dynamic Gordon model is relatively small. Notice that the dynamic Gordon model presupposes a fixed risk premium.

Agents use a simple AR(1) rule to update their beliefs with the last observation in the risk free rate and growth rate:

$$E_t[r_{t+j}] = r + \rho^j(r_t - r) \quad E_t[g_{t+j}] = g + \tau^j(g_t - g). \quad (4.21)$$

Boswijk, Hommes, and Manzan (2007) show, using the approach of Poterba and Summers (1988), that the time-varying fundamental PD ratio is to a first-order approximation given by:

$$\delta_t^* = \frac{1+g}{r-g} + \frac{\rho(1+g)}{(r-g)(1+r-\rho(1+g))}(r_t - r) + \frac{\tau(1+r)}{(r-g)(1+r-\tau(1+g))}(g_t - g). \quad (4.22)$$

To estimate the static part  $\delta^*$  in equation (4.22), we use updated data on the S&P500 prices and dividends originally provided by Shiller (2005), with  $T = 252$  end-of-quarter observations from 1950Q1 until 2012Q4. For an easier interpretation, we will focus on the *yearly* price dividend ratio even though it is based on quarterly observations. We also estimate the parameters using yearly data for comparison. See Table 4.1 for the results.

Our findings are close to previous estimates on the postwar period, such as a real yearly dividend yield of around 3.5% and a yearly risk premium of around 2.5%. We find some small differences when we estimate the dynamic Gordon model directly on yearly data, because more data points are used in the quarterly estimation. Ignoring deviations in the dividend growth and risk free rates, we find a yearly constant value of  $\delta^* = 29.9$  for the quarterly estimation.

Next, we estimate the AR(1) rules used by the agents to update the fundamental value. We find quarterly values of  $\rho = 0.40$  for the persistence in the risk free rate and  $\tau = 0.50$  for the growth rate, and we calculate the fundamental value  $\delta_t^*$  according to equation (4.22). Figure 4.1

Frequency	$\pi$	$d/p$	$g$	$r$	$i$	$RP$	$R^*$	$\delta^*$
Quarterly	0.91	0.84	0.40	1.24	0.54	0.70	1.008	119.5
Yearly equivalent	3.71	3.40	1.62	5.06	2.19	2.87	1.034	29.9
Yearly	3.71	3.37	1.31	4.69	2.25	2.43	1.033	30.1

Table 4.1: Estimation of the fundamental value. Values used for estimating the static Gordon solution  $\delta^* = (1 + g)/(r - g)$ :  $\pi$  is the average inflation rate,  $d/p$  is the average dividend yield  $D_t/P_{t-1}$ ,  $g$  is the average dividend growth rate,  $r = d/p + g$  equals the risk free rate plus the required risk premium on stocks,  $i$  is the average real return on T-notes with a 10-year maturity,  $RP = r - i$  is the risk premium and  $R^* = (1 + r)/(1 + g)$  is the expected effective discount rate. We use the CPI index to deflate the nominal variables. All numbers except  $R^*$  and  $\delta^*$  are multiplied by 100. The estimation is done using both quarterly and yearly data from 1951Q1-2012Q4; yearly equivalent estimates based on the quarterly estimation are presented using geometric progression.

plots the fundamental value based on the price-dividend ratio  $\delta_t^*$  and dividends  $D_t$  next to the observed value of the S&P500 index and observed PD ratios. As displayed in the bottom panel of Figure 4.1,  $\delta_t^*$  fluctuates in the interval of  $[20.3, 40.8]$ , but for most periods it stays close to the constant average value of  $\delta^* = 29.9$ .

The S&P500 clearly exhibits excess volatility, that is, it fluctuates much more than its underlying fundamentals; an important point already made by Shiller (1981). Until 1990, the S&P500 is seen to fluctuate relatively quietly around the value that was to be expected from future dividends. After 1995 this changes: stock prices, most notably of firms in the internet and information technology sector, rose much more than was justified by the dividend pay-outs. At the top of the dot-com bubble in 2000, when the PD ratio reached almost 90 compared to the fundamental value of around 30, the S&P500 was by a factor three overpriced relative to the dividend-based fundamental.

It should be noted that we find excess volatility and overpricing in Figure 4.1 because we do not assume a time-varying risk premium. In Section 4.5, we follow the more standard approach that at least part of the variation in asset prices is due to variation in risk premia, using the consumption-habit asset pricing model of Campbell and Cochrane (1999) as fundamental benchmark. Irrespective of the underlying prices, all agents in our behavioural model are aware that prices differ from fundamentals, but do not believe that they can use this knowledge to gain

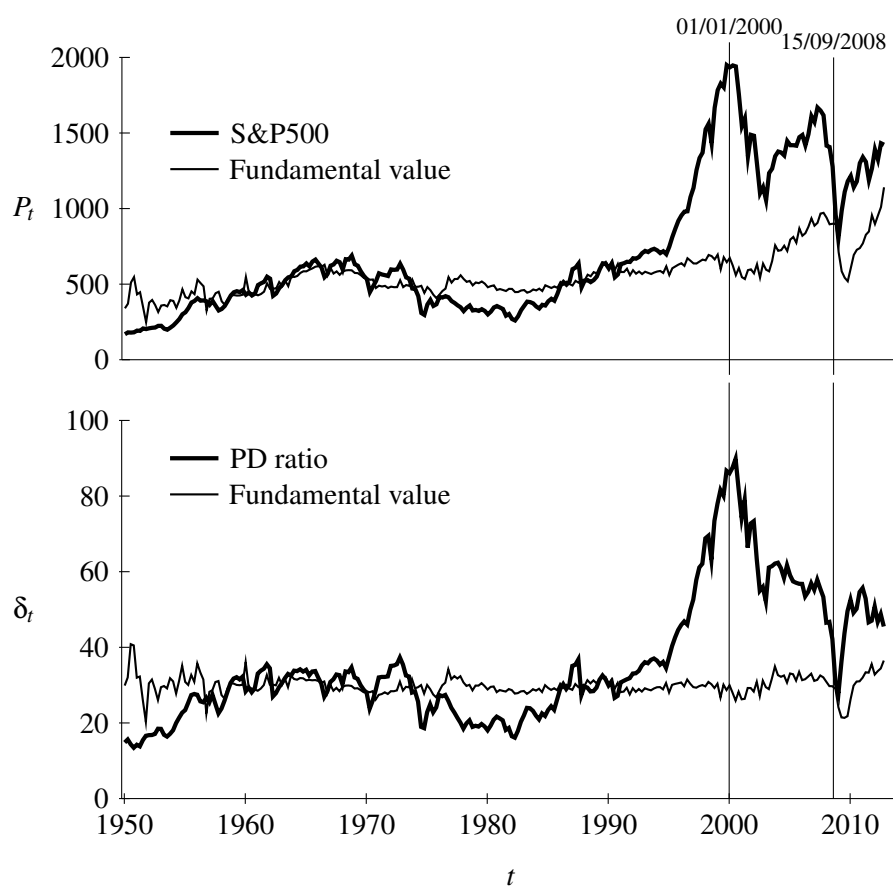


Figure 4.1: The S&P500 index with its fundamental value, corrected for inflation (top panel), and the realised and fundamental yearly PD ratio (bottom panel).

higher profits. This behavioural element is supported by survey data, in particular for the period around the turn of the millennium. Both Shiller (2000) and Vissing-Jorgensen (2004) found that the majority of respondent investors in 2000 were aware of the overvaluation of stock prices, but did not expect that the mispricing would be corrected within a period of a year.

Another observation is that, in our model with a fixed risk premium, the global financial crisis is of a quite different nature than the burst of the dot-com bubble. After 2000 prices went down for three years, but stayed above the dynamic Gordon fundamental value. Partly driven by the securitisation activities of large investment banks, dividends rose steadily from 2003 to 2008, which in turn drove prices up again, perhaps more than justified by fundamentals. After the bankruptcy of Lehman Brothers on 15 September 2008, the stock market crashed and prices returned very closely to the fundamental value. Only afterwards, when the market already started to recover, dividends started to fall.

### 4.3.2 Estimation of the heterogeneous agents model

This section is devoted to the estimation of the heterogeneous agents model using the time series  $x_t = \delta_t - \delta^*$  of price deviations from the Gordon benchmark. From Figure 4.1 it is clear that there is quite some structure in  $x_t$ . It is highly persistent and interrupted by phases of mean-reversion. We perform nonlinear least squares to estimate the heterogeneous agents model and interpret the asset price fluctuations by different behavioural regimes.

Table 4.2 shows the estimation results for our four parameters  $\phi_1$ ,  $\phi_2$ ,  $\beta$  and  $\omega$ , under four different model specifications (A), (B), (C) and (D). For model specification (A), we fix the intensity of choice  $\beta = 1$  for reasons discussed below, and estimate the remaining three parameters. For model (B), we estimate all four parameters simultaneously, while model (C) fixes the intensity of choice at a different, high value of  $\beta = 10$ . Model (D) fixes  $\beta = 1$  as in model (A) and has no memory, i.e. the memory parameter  $\omega = 0$ .

In all estimations, the belief parameters  $\phi_1$  and  $\phi_2$  are significantly different from zero, emphasising that the fundamental value is not very informative in the short run. In the model



### 4.3. ESTIMATION RESULTS FOR A SIMPLE BEHAVIOURAL MODEL

	(A)	(B)	(C)	(D)
$\phi_1$	0.936*** (0.027)	0.947*** (0.023)	0.940*** (0.020)	0.981 (-)
$\phi_2$	1.026*** (0.021)	1.017*** (0.021)	1.026*** (0.020)	0.981 (-)
$\beta$	1	2.443 (7.268)	10	1
$\omega$	0.824*** (0.154)	0.800*** (0.152)	0.852*** (0.190)	0
$\Delta\phi$	0.090** (0.039)	0.070** (0.033)	0.087** (0.029)	0 (-)
$T$	252	252	252	252
$s^2$	14.12	14.09	13.87	14.08
$R^2$	0.952	0.953	0.954	0.951
$AIC$	2.579	2.583	2.552	2.592

\*, \*\*, \*\*\* denote significance at the 10%, 5% and 1% level, respectively.

Standard errors are denoted in brackets.

The  $R^2$  denotes the proportion variation in  $\delta_t$  explained by the model.

Table 4.2: Estimation of the belief coefficients  $\phi_1$  and  $\phi_2$ , the intensity of choice  $\beta$  and the memory parameter  $\omega$  in the heterogeneous agents model  $x_t = \frac{1}{R^*}(n_{1,t}\phi_1 + n_{2,t}\phi_2)x_{t-1} + \varepsilon_t$ , with  $R^* = 1.008$  and the fractions  $n_{1,t}$  and  $n_{2,t}$  updated according to (4.18) and (4.19). All specifications are estimated with nonlinear least squares except for specification (D), which is found by a grid search.

specifications A, B and C with memory, the belief coefficients furthermore show the essential difference between fundamentalism and chartism:  $\phi_1 < 1$  and  $\phi_2 > 1$ . If the regression includes memory the estimated difference  $\Delta\phi \equiv \phi_2 - \phi_1$  is significant and around 0.07 to 0.09: the hypothesis  $\Delta\phi = 0$  is rejected at the 5% level. This is our main evidence for heterogeneity in the S&P500 data.

The estimated intensity of choice  $\beta$  is not found to be significant in model version (B), where all four parameters are estimated simultaneously. This is a typical finding in nonlinear “smooth transition” AR models, as changes in  $\beta$  have often very little effect on the fit of the model. For a detailed general discussion of this issue, see Teräsvirta (1994). In Section 4.4.1, we will use Monte Carlo (MC) simulations and show that the insignificance should not be a concern, as the  $t$ -test for  $\beta$  simply lacks power given the size of our sample. These MC simulations show that for a sample size of 252, even if our estimated model (A) with a  $\beta = 1 > 0$  would be the true underlying data generating process, we would not reject the null hypothesis of  $\beta = 0$ . The MC simulations also show that for sufficiently large sample sizes, this null hypothesis is correctly rejected. Stated differently, the test for no switching ( $\beta = 0$ ) has low power against the HAM model in small samples. This result for nonlinear smooth transition models is similar in spirit as for example the low power of the Dickey-Fuller unit root test for linear near unit-root AR(1) processes, a well-known empirical problem for small samples.

A solution for the insignificant  $\beta$  is to fix it at, e.g. at  $\beta = 1$ . In the robustness checks in Appendix 4.B, we show that the explanatory power of the model (as measured by the  $R^2$ ) is much less sensitive to the value of  $\beta$  than to the values of the other parameters (see in particular Figure 4.7). Notice also that with  $\beta = 10$  fixed at a higher value, as in model (C), the remaining estimated parameters hardly change. Because a non-zero  $\beta$  is necessary to identify different regimes as well as the level of memory  $\omega$ , the model (A) with a value of 1 is our main focus in the rest of this chapter. Considering the model with only the three relevant parameters  $\phi_1$ ,  $\phi_2$  and  $\omega$  makes economic sense and is also the choice based on the lowest Akaike’s information criterion (AIC). Taking  $\beta = 0$  reduces the model to a linear AR(1) model and seriously reduces

the fit of the model, but any positive value from say  $\beta \geq 1$  can be used in the HAM and yields similar estimation results.

The memory parameter  $\omega$  is strongly significant. This means that shocks that are observed more than one period ago are also taken into account in the switching between beliefs. While the value around 0.8 in model (A) might seem high, it implies that more than half of the information is extracted from observations in the last year ( $1 - 0.824^4 \approx 54\%$ ). So while memory is important in estimating HAMs on higher frequency data, it does not require unrealistic processing abilities from the agents. The model remains consistent with the behavioural background of bounded rationality. We also estimate model (D) without memory ( $\omega = 0$ ). In this case the two estimated regimes become identical ( $\phi_2 = \phi_1$ ) and the nonlinear least squares estimation is no longer identified; the results for model (D) in Table 4.2 were found by a grid search. Because of the equal belief parameters, the estimation of model (D) is identical to that of a linear AR(1) model, as we will show in Section 4.3.3. This underlines our statement that memory is important to explain the regime shifts in the given quarterly dataset of the stock index.

From the estimation of the HAM we can infer the estimated fraction of fundamentalists  $n_{1,t}$  over time (Figure 4.2, top panel). This plot points at a structural break in 1995.<sup>9</sup> After some initial large shocks in the beginning of the sample, the fraction remained within the interval  $[0.4, 0.7]$  for most of the periods. Starting from 1995, however, two successive regimes of trend-following and mean-reversion are evident. Trend-following dominated in the 1990s, amplifying the stock price run-up during the dot-com bubble. The mean-reversion regime continued until prices came just below the Gordon fundamental value in 2009, but fundamentalists remained to dominate the markets for most of the periods in recent years. The inclusion of a memory parameter is needed to distinguish these transitions from the relatively high noise levels.<sup>10</sup>

The fractions can be translated directly into the estimated market sentiment  $\phi_t$  over time

<sup>9</sup> In the estimation of the HAM in Section 4.5 we allow for a structural break in 1995Q1, a jump in the risk premium, in the Campbell and Cochrane (1999) fundamental benchmark. Allowing for a structural break in the dynamic Gordon model as is done in Appendix 4.B yields similar results and in particular significant behavioural heterogeneity.

<sup>10</sup> In Boswijk, Hommes, and Manzan (2007), the estimated fractions fluctuated heavily over the whole period, sometimes with large swings from close to 0 to close to 1.

(Figure 4.2, bottom panel). The system is locally stable, as for the market sentiment with equal distribution of beliefs it holds that  $\frac{\phi_1 + \phi_2}{2R^*} = 0.973 < 1$ . However, the plot shows that temporary destabilisation with explosive market sentiment, i.e.  $\phi_t > 1$ , is possible when the market is dominated by chartists. This happened for a consecutive number of quarters during 1995-2000, and chartists strongly amplified the magnitude of the dot-com bubble. After the bubble burst, market sentiment remained low for a relatively long period and slowly recovered, at which point the financial crisis hit in 2008. The fact that the model generates these genuinely different and intuitive regimes makes it economically of interest.

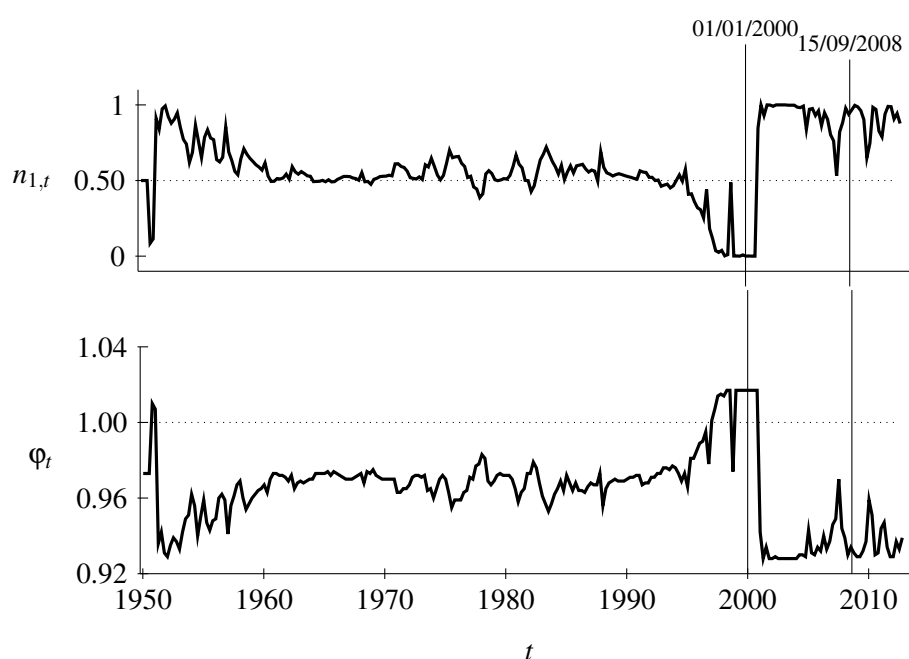


Figure 4.2: Estimated fraction of fundamentalists (top panel) and the corresponding market sentiment (bottom panel) for the HAM (A) under the Gordon model.

Under the assumption of a constant risk premium, the financial crisis was merely a correction back to fundamentals. The relatively small fluctuations in market sentiment since 2001Q1 indicate two points at which some investors moved away from the fundamentalists belief. First, in 2006 and 2007 the recovery of the stock market increased the fraction of chartists to almost 50% and the market sentiment up to 0.970; but already by 2008Q2, *before* the bankruptcy of Lehman Brothers, fundamentalists constituted already almost 100% of the market. According

Variable	Coefficient	(Std. Err.)
$\phi$	0.973***	(0.013)
$DF$	-2.414	$p$ -value=0.402
$T$	252	
$s^2$	14.08	
$R^2$	0.951	
$AIC$	2.584	

\*, \*\*, \*\*\* denote significance at the 10%, 5% and 1% level, respectively.

The  $R^2$  denotes the proportion variation in  $\delta_t$  explained by the model.

Table 4.3: Estimation of the AR(1) model  $x_t = \phi x_{t-1} + \varepsilon_t$

to our HAM around the Gordon fundamental benchmark, the financial crisis therefore has been strongly amplified as a correction back to fundamentals. In 2010 a second temporary upheaval can be observed, but again most agents turned back to mean-reversion.

### 4.3.3 A representative agent benchmark model

The variable  $\phi_t$  in equation (4.17) is a time-varying AR(1)-parameter within the interval  $[\frac{\phi_1}{R^*}, \frac{\phi_2}{R^*}]$ . As an obvious linear benchmark, we also estimate an AR(1) model with a constant parameter, see Table 4.3. Notice that this is *not* the rational representative agent benchmark, which would coincide with the fundamental benchmark, but rather a representative agent believe in constant, linear mean reversion. The estimated coefficient of the linear model is 0.973, which is close to 1. This points to the possibility of non-stationarity, implying that the estimated standard errors should be interpreted with care. The Dickey-Fuller test shows that the null hypothesis of a unit root (i.e.  $\phi = 1$ ) can not be rejected. In Section 4.4.1 we address the result of this test in more detail by Monte Carlo simulations, and find that it is most probably caused by the Dickey-Fuller test having low power in our relatively small sample of  $T = 252$ .

From comparing the  $R^2$ , the heterogeneous agents models with memory are seen to be an improvement over the AR model. The result of a partial F-test shows that the improvement of model (A) over the AR model is significant: the F-statistic is 4.90 and exceeds the 5%-critical value of 3.84. Also, the heterogeneous agents model (A) is preferred over the AR model because of its lower  $AIC$ . Admittedly, in absolute terms the explanatory power is roughly 95% for both

models. The mild improvement of the explanatory power should be considered in relation with the unpredictability of stocks. The estimations on the shorter horizon suffer from considerable noise levels and different models are therefore inevitably more alike. Campbell and Shiller (1988b) already noted that stock prices become less predictable when they are measured over intervals of less than a year rather than over intervals of several years.

We conclude that the HAM statistically outperforms the AR model. Another interesting feature of our heterogeneous agents model is that it gives an intuitive economic interpretation of medium-run bubbles that is lacking in representative agent benchmark models.<sup>11</sup> The next section will discuss the differences between the linear and the nonlinear switching models in more depth by Monte Carlo simulations. It will become clear that the two models are economically very different.

## 4.4 Monte Carlo simulations

To gain understanding about the properties of the HAM and evaluate differences with the simple representative agent AR(1) model, we use the estimated equations as Data-Generating Processes (DGPs) in our Monte Carlo simulations. For these two DGPs we draw shocks from a normal distribution with zero mean and variance equal to the sample variance of the errors  $s^2$ . For the HAM version (A) the DGP is:

$$x_t = \varphi_t x_{t-1} + \varepsilon_t, \varepsilon_t \sim N(0, \sigma^2 = 14.12), \quad (4.23)$$

where  $\varphi_t \in [\frac{\phi_1}{R^*}, \frac{\phi_2}{R^*}] = [0.928, 1.017]$  is updated according to (4.17), (4.18) and (4.19) with  $\phi_1 = 0.936$ ,  $\phi_2 = 1.026$ ,  $\beta = 1$  and  $\omega = 0.824$ . For the model with a fixed AR(1)-coefficient the DGP

---

<sup>11</sup> It is not straightforward to improve the (extremely) simple AR(1) model in a way other than we propose. For example, in estimations of more general AR(p) model the higher order autocorrelation terms are typically not significant. Another possibility is to allow for time-varying volatility using GARCH-errors (Bollerslev, 1986), a method that is successful in explaining daily stock returns. We estimated an AR(1)-GARCH(1,1) model, in which the dynamics after 1995 can be interpreted with high clustered volatility. This interpretation, however, misses the different regimes that seem present in the data. We find that the HAM also dominates the AR-GARCH model in explaining the data.

is:

$$x_t = 0.973x_{t-1} + \varepsilon_t, \varepsilon_t \sim N(0, \sigma^2 = 14.08). \quad (4.24)$$

We first evaluate the power and size of the tests we have used in the estimation in Section 4.3, by running these two DGPs for some number of periods  $T$  and considering the outcomes of these tests. In Section 4.4.2 we use the DGPs to generate time series starting at three observed points in time, in order to illustrate the potential differences between the nonlinear HAM and linear benchmark models in predicting PD ratios.

#### 4.4.1 Evaluating the power of the main tests

Our tests of interest are the test for homogeneity ( $H_0 : \phi_1 = \phi_2$ ) and for no switching ( $H_0 : \beta = 0$ ) for the HAM, and the Dickey-Fuller test for a unit root ( $H_0 : \phi = 1$ ) for the AR model. Since the true DGP is unknown, we perform these three tests under both DGPs, resulting in six combinations. For example, we estimate a HAM (including a free parameter  $\beta$ ) on data generated by a simple AR(1) model, and investigate whether we find significant switching or heterogeneity. In this case it is unlikely that we reject the null hypotheses of no switching and homogeneity, but by pure coincidence of the error realisations, it is possible.

For each of the six combinations of DGP and test, we make  $B$  simulation runs of sample size  $T$ , make for every run estimations on the simulated time series and check whether the  $p$ -value of the particular test is below the nominal significance level  $\alpha$ . If the null hypothesis is false, e.g. when homogeneity is tested on HAM-generated data, the proportion of times we reject it measures the *power* of the test. If the null hypothesis is true, e.g. when homogeneity is tested on AR-generated data, the rejection probability measures the *size* of the test. In Tables 4.4 and 4.5 the Monte Carlo-estimations of the power are shown, and between brackets the size.

To check the asymptotic properties of the tests, we take a large sample size of  $T = 5,000$  and  $B = 1,000$  simulation runs (see Table 4.4). The DF test is asymptotically working correctly: in

True model:	DGP \	Test:	unit root	homogeneity	no switching
		$H_0$	$\phi = 1$	$\phi_1 = \phi_2$	$\beta = 0$
heterogeneous representative	HAM		100%	100%	88%
	AR		100%	(2%)	(0%)

Table 4.4: Rejection probabilities: nominal significance level  $\alpha = 5\%$ ,  $T = 5,000$ ,  $B = 1,000$ , parameters for HAM version (A) in Table 4.2 and for the AR model as in Table 4.3.

True model:	DGP \	Test:	unit root	homogeneity	no switching
		$H_0$	$\phi = 1$	$\phi_1 = \phi_2$	$\beta = 0$
heterogeneous representative	HAM		6.6%	46.2%	0.0%
	AR		11.1%	(4.2%)	(0.0%)

Table 4.5: Rejection probabilities: nominal significance level  $\alpha = 5\%$ ,  $T = 252$ ,  $B = 10,000$ , parameters for HAM version (A) in Table 4.2 and for the AR model as in Table 4.3.

large samples it successfully identifies the true underlying processes to be stationary, and always rejects the null hypothesis of a unit root. Similarly the power of the homogeneity test is 100% for a large sample size. However, the test for no switching fails to reject the null hypothesis in 12% of the cases, even for this large sample size. This already indicates that rejecting the null of no switching may be difficult even when the true DGP is the HAM.

We repeat the Monte Carlo evaluation for the actual sample size of the data ( $T = 252$ ) and  $B = 10,000$  simulation runs. Table 4.5 shows the results. The DF test performs poorly in small samples. While the low small-sample power of the DF test for near unit-root AR processes is well-known in the literature, our results show that the power decreases even further if the true model is a nonlinear HAM. In particular, for our estimated HAM parameters the power is reduced to almost half of the power if the true model an AR process. Therefore, the fact that we failed to reject a unit root in the data is not surprising.

The test for no switching is utterly uninformative in small samples, as it does not reject in any of the replications when the true DGP is a HAM. In other words, the reason that we do not find a significant  $\beta$  in the data is because the corresponding  $t$ -test lacks power. The test for homogeneity on the other hand remains to have a relatively high power (46% rejections for the HAM versus 4% for the AR model). This result underlines that the detected heterogeneity in



the data is much more likely to be driven by real heterogeneity in the underlying process than by pure chance in a representative agent world.<sup>12</sup>

#### 4.4.2 What are the economic effects of different behavioural regimes?

To understand the economic mechanism that is at play in our heterogeneous agents model, we examine some simulated time series in greater depth. As an example, we simulate the model after three periods: 1997Q3 ( $t = 191$ ), 2002Q2 ( $t = 210$ ) and 2004Q4 ( $t = 220$ ). At each of these points in time the PD ratio was close to 60, relatively far above the fundamental value. Using the DGP given in (4.23), we calculate different quantiles of the simulated distribution over a rolling horizon. To start the simulations from the initial market situation at  $t = 191$ ,  $t = 210$  and  $t = 220$ , the observed fractions  $n_{h,t}$  and performances  $U_{h,t}$  of both rules are feeded into the simulation. In the left panels of Figure 4.3, the median prediction and the 5%-, 30%-, 70%- and 95% quantiles for the heterogeneous agents model are presented.<sup>13</sup>

It is striking that the heterogeneous agents model allows for the possibility of a large bubble after 1997Q3, but at the same time generates strong mean-reversion after 2002Q2. After 2004Q4 the median prediction is also decreasing quite quickly to the fundamental value. These differences can be explained by the key mechanism in the model: positive expectations feedback. In 1997Q3, the estimated fraction of fundamentalists is low ( $n_{1,t} = 0.04$ , see Figure 4.2) and the market sentiment parameter exceeds 1 ( $\phi_t = 1.01$ ). With partly self-fulfilling expectations and many investors believing in a trend, stock prices typically move further away from fundamentals, which also happened during the dot-com bubble. Our model suggests that the bubble could have been even more pronounced: the top of the 95%-quantile is 106.6 and is reached after 26 quarters in 2004Q1. Note though that bubbles end endogenously: the increasingly high expectations of chartists are bound to overshoot the realised prices, leading to a

<sup>12</sup> Under the Data-Generated Process by the AR(1) model, the tests for homogeneity and for no switching very rarely reject the null hypothesis, indicating that the size of these tests is not controlled at the nominal level of 5%, probably due to the nonlinearity of the estimated model. This does not affect our conclusions.

<sup>13</sup> For simplicity we ignore here small deviations in the fundamental value of the dynamic Gordon model and consider the static Gordon solution  $\delta^* = 29.9$ .

fundamentalist mean-reverting regime. Our model thus explains the stock market boom in the late 1990s as a temporary bubble triggered by fundamentals and strongly amplified by trend-following behaviour.

For comparison, we also simulate the representative agent AR model (right panels of Figure 4.3) starting from the same periods. Note that the linear AR model gives almost identical predictions in 1997Q3, 2002Q2 and 2004Q4, and completely misses the dot-com bubble. The AR simulations are symmetrically distributed around the median and return slowly to the fundamental value with a constant coefficient  $\phi < 1$ . The HAM is *nonlinear*: it can allow for a bubble in the short run, but also generates a faster return to the fundamental value if fundamentalists become dominant. This rapid decline, as for example in the simulations after 2002Q2, occurs if after some negative shocks the fundamentalists belief keeps attracting more followers, which decreases market sentiment and consequently prices.

The heterogeneous agents model, built upon positive expectations feedback, generates simulated time series of prices that are economically quite different from linear representative agent models. Our model is suitable for making medium-run projections of future prices when predictions of rational representative agent models are unreliable. The financial crisis is within a representative agent world typically perceived as an extreme event. The PD ratio of 27.8 in 2009Q1 is below its 5% quantile of the AR model simulated after 2004Q4, more than four years before the crisis (bottom right panel of Figure 4.3). Because of the high estimated fraction of fundamentalists, the heterogeneous agents model predicts lower prices after 2004Q4 than the AR model, and its upper 95% confidence interval (above the 5% quantile) does contain the possibility of the large drop in stock prices during the financial crisis.

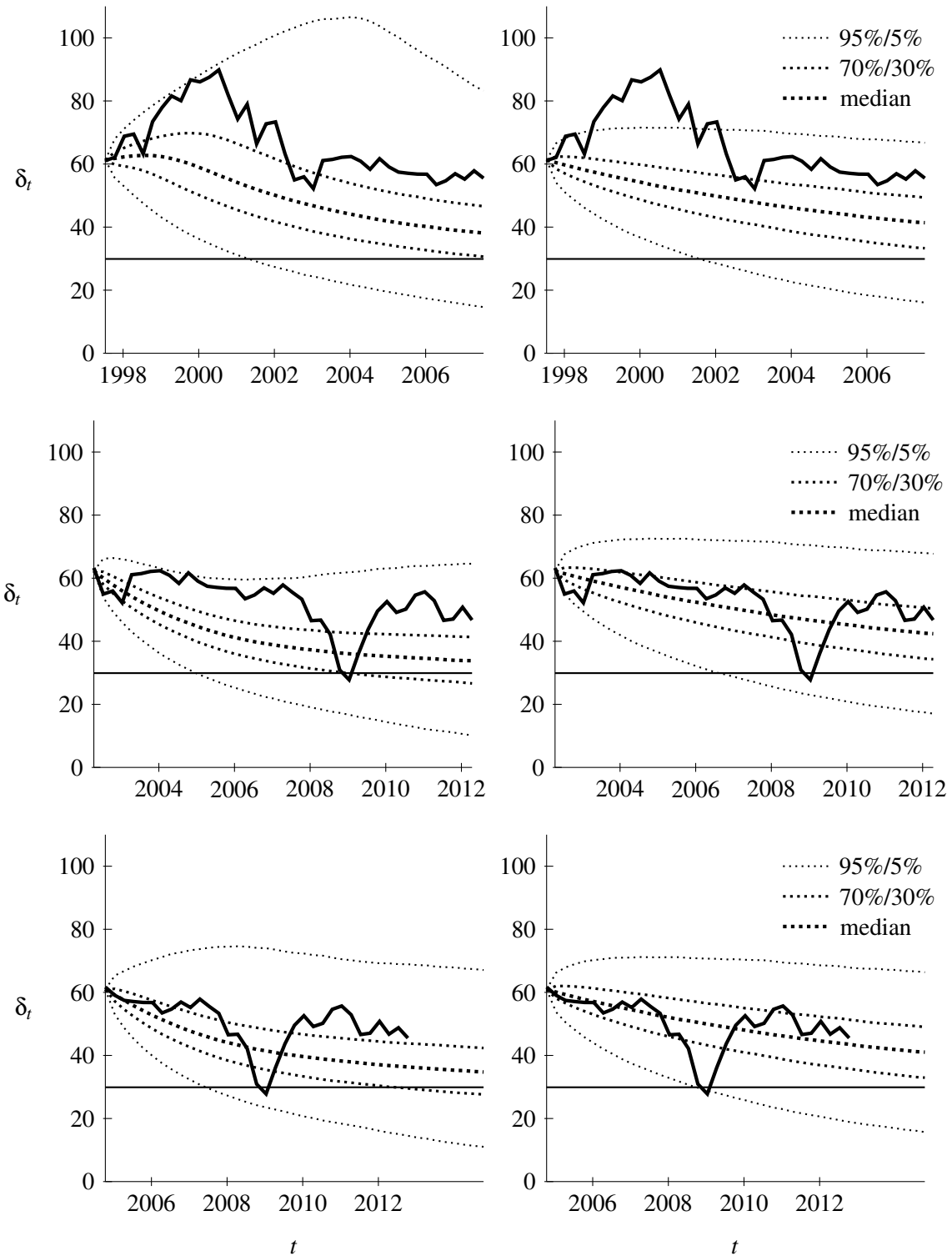


Figure 4.3: Realised and simulated PD ratio after 1997Q3 (top panels), after 2002Q2 (middle panels) and after 2004Q4 (bottom panels). The left panels show the simulated distributions for the heterogeneous agents model, fed with last observed fractions and performances, and the right panels show the simulated distributions for the AR model.

## 4.5 Estimation results using a consumption-habit time-varying risk premium

We have seen that under a constant risk premium stock prices exhibit considerable deviations from their fundamentals and in particular there is large overpricing after 1990. The heterogeneous agents model explains this overpricing as being triggered by fundamental shocks and strongly amplified by a long regime of trend-following behaviour up to the end of 2000. A different explanation from mainstream finance is that the discount rate changed to very low values, such that the same expected future payoffs were valued higher. In Section 4.3 we allowed the discount rate to vary, but only with predictable variation in the risk free rate; the risk premium was assumed to be constant. In this section we relax the assumption of a constant risk premium in order to study the robustness of our results. We will show that even after introducing considerable time-variation in the risk premium significant evidence for behavioural heterogeneity in beliefs remains.

We follow the consumption-habit model for stock price fluctuations by Campbell and Cochrane (1999). This approach is well-known and recently summarised and advocated in Cochrane (2011). The main idea of their model is that investors demand a higher risk premium as consumption decreases during a recession, and conversely become less risk averse when consumption goes up during an economic boom. In order to translate continuing rising consumption levels to a stationary level of risk aversion, Campbell and Cochrane (1999) define a slow moving “habit” or moving average consumption level, and consider the relative distance between consumption and this habit, called “surplus consumption”. Using this surplus consumption, the model predicts the evolution of price-dividend ratios over time.

Our two-step methodology in search for evidence of heterogeneous agents can be applied to any benchmark model for the fundamental value, and here we use the consumption-habit model as the fundamental stock index value. In the first step, we will use the specification of Campbell and Cochrane (1999) for surplus consumption and fit their model on actual PD ratios.

Inspired by the line of thought in Cochrane (2011), we allow for one structural break in surplus consumption to capture high asset price values after the 1990s. In the second step, we will estimate the heterogeneous agents model on deviations from the fitted price-dividend ratios, and test whether time-varying risk premia make a significant difference.

### 4.5.1 Estimating the Campbell-Cochrane model

Below we give a short summary of the habit consumption asset pricing model. We start by recalling the two-period equation for the PD ratio described in Section 5.2:

$$\delta_t = E_t \left[ \frac{1}{R_{t+1}} \frac{D_{t+1}}{D_t} (\delta_{t+1} + 1) \right]. \quad (4.25)$$

To specify the discount rate  $R_{t+1}$ , in this case stochastic, Campbell and Cochrane (1999) stipulate that agents maximise the utility function:

$$U = E_t \sum_{j=1}^{\infty} \kappa^j u(C_{t+j}, H_{t+j}) = E_t \sum_{j=1}^{\infty} \kappa^j \frac{(C_{t+j} - H_{t+j})^{1-\gamma} - 1}{1-\gamma}, \quad (4.26)$$

where  $C$ ,  $H$ ,  $\kappa$  and  $\gamma$  are respectively consumption, the habit level, the subjective time discount factor and the utility curvature parameter. The *surplus consumption ratio*  $S_t$  is defined as the relative difference between consumption and the habit level:

$$S_t = \frac{C_t - H_t}{C_t}. \quad (4.27)$$

From the first-order conditions of (4.26), the discount rate  $R_{t+1}$  equals the inverse of the intertemporal marginal rate of substitution  $M_{t+1}$ :

$$\frac{1}{R_{t+1}} = M_{t+1} = \kappa \frac{u_c(C_{t+1}, H_{t+1})}{u_c(C_t, H_t)} = \kappa \left( \frac{S_{t+1} C_{t+1}}{S_t C_t} \right)^{-\gamma}. \quad (4.28)$$

Combining equations (4.25) and (4.28), Campbell and Cochrane (1999) show that the fundamental PD ratio can be expressed as a function of surplus consumption ratio  $S_t$  only, as it is the only state variable.

We now focus on the calculation of the surplus consumption  $S_t$ . Campbell and Cochrane (1999) assume that the log surplus consumption ratio  $s_t \equiv \log S_t$  evolves as a heteroskedastic AR(1) process:

$$s_{t+1} = (1 - \phi^s)\bar{s} + \phi^s s_t + \lambda(s_t)(c_{t+1} - c_t - \mu_c). \quad (4.29)$$

The symbols  $\phi^s$ ,  $\mu_c$  and  $\bar{s}$  denote parameters for the function of the surplus consumption ratio and  $\lambda(s_t)$  is the sensitivity function of  $s_{t+1}$  to the deviation of consumption growth from its long run average. Consumption growth is modeled as an i.i.d. lognormal process on  $c_t \equiv \log C_t$ :

$$c_{t+1} = c_t + \mu_c + v_{t+1}^c, \quad v_{t+1}^c \sim IID(0, \sigma_c^2). \quad (4.30)$$

Campbell and Cochrane (1999) find the steady-state surplus consumption ratio  $\bar{S} \equiv \exp(\bar{s})$ :

$$\bar{S} = \sigma_c \sqrt{\frac{\gamma}{1 - \phi}}, \quad (4.31)$$

where  $\gamma$  is a parameter of utility curvature. They define the sensitivity function as:

$$\lambda(s_t) = \begin{cases} (\bar{S})^{-1} \sqrt{1 - 2(s_t - \bar{s})} - 1 & \text{if } s_t \leq s_{max} \\ 0 & \text{if } s_t > s_{max}, \end{cases} \quad (4.32)$$

where  $s_{max} \equiv \log S_{max}$  is the value of  $s_t$  at which the upper expression 4.32 becomes zero:

$$s_{max} = \bar{s} + \frac{1}{2}(1 - \bar{S}^2). \quad (4.33)$$

Using the expressions above, we can relate observed consumption levels  $C_t$  to surplus consumption ratios  $S_t$  by four free parameters:  $\mu_c$ ,  $\sigma_c$ ,  $\phi^s$  and  $\gamma$ . We use updated data on real per capita U.S. consumption of nondurable goods and services originally provided by Chen and

#### 4.5. ESTIMATION USING A CONSUMPTION-HABIT TIME-VARYING RISK PREMIUM

<b>Parameter</b>	<b>Variable</b>	<b>Value (quarterly)</b>	<b>Value (yearly)</b>
<i>Matched:</i>			
Mean consumption growth (%)	$\mu_c$	0.46	1.84
S.d. of consumption growth (%)	$\sigma_c$	0.47	0.94
<i>Assumed:</i>			
Persistence coefficient	$\phi^s$	0.97	0.87
Utility curvature	$\gamma$	2.00	2.00
<i>Implied:</i>			
Steady-state surplus consumption	$\bar{S}$	0.038	0.037
Maximum surplus consumption	$S_{max}$	0.063	0.061

Table 4.6: Parameter values for the habit consumption model.

Ludvigson (2009) with  $T = 243$  observations from 1952Q1 to 2012Q3.<sup>14</sup> We match the mean and standard deviation of log consumption growth  $\mu_c$  and  $\sigma_c$  to the data. The parameters  $\phi^s$  and  $\gamma$  are taken identical to Campbell and Cochrane (1999).

Table 4.6 presents the estimates of the parameters  $\mu_c$  and  $\sigma_c$  with values of other parameters that are either assumed or implied by the model. Our steady-state surplus consumption ratio  $\bar{S}$  is around 0.037 and somewhat lower than the value of Campbell and Cochrane (1999) of 0.057, because we find smaller variation in consumption growth ( $\sigma_c = 0.94$  rather than 1.50, on yearly basis). There are also some small differences between quarterly and yearly surplus consumption ratios; we will use the quarterly values in the further analysis.

Figure 4.4 shows consumption in logs and relative to consumption at the start of the sample, and the habit level implied by the model under the assumption that the surplus consumption ratio starts at the steady state at the beginning of the sample 1952Q1. The lower panel of Figure 4.4 shows the surplus consumption ratio. Similar to the estimations of Campbell and Cochrane (1999), the estimated surplus consumption ratio tracks the macroeconomic trends, such as the consumption boom in the 1960s (though with a lag) and the boom in the 1980s. The most recent part of the time series shows the ongoing Great Recession during which consumption dropped down abruptly and persistently to a level very close to the habit.

In the Campbell-Cochrane model the price-dividend ratio is a nearly log-linear function of

<sup>14</sup> Note that no data is available for the period 1950Q1-1951Q4, which reduces our sample for the PD ratio by 8 observations. At the time of writing also the last observation 2012Q4 was not available.

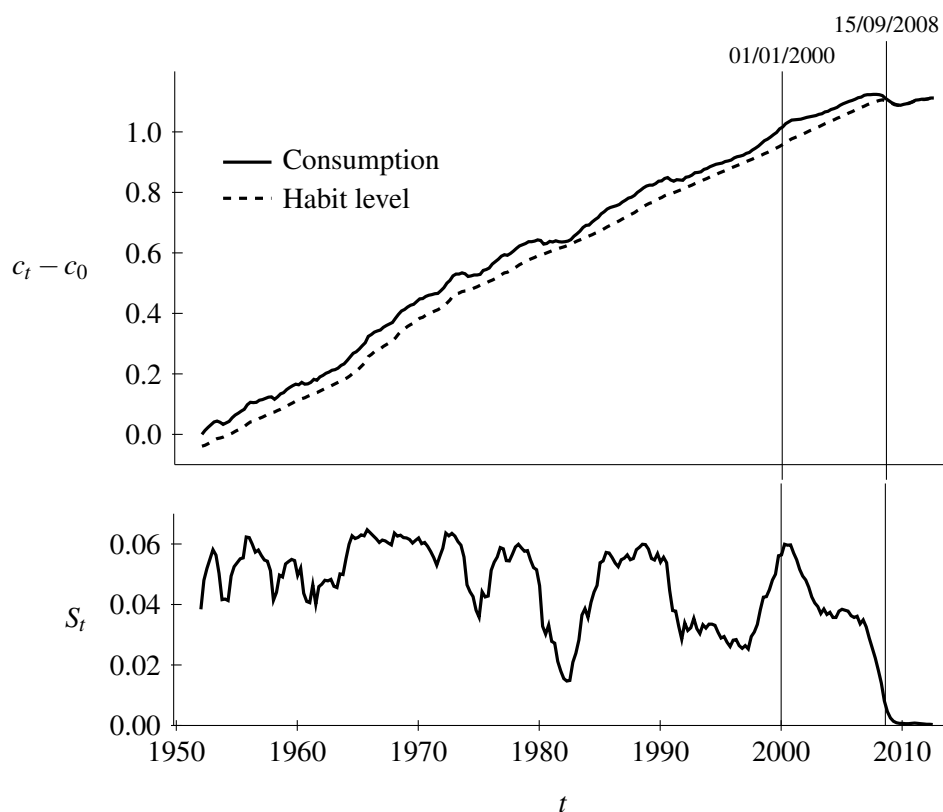


Figure 4.4: Log consumption per capita of nondurable goods and services and the habit level (top panel), and surplus consumption ratio (bottom panel), under the assumption that the surplus consumption ratio starts at the steady state in 1952Q1.



#### 4.5. ESTIMATION USING A CONSUMPTION-HABIT TIME-VARYING RISK PREMIUM

the surplus consumption ratio. We will estimate the log-linear relationship between the surplus consumption ratio and the PD ratio  $\delta_t$  as follows:

$$\delta_t = b(S_t)^p + u_t, \quad (4.34)$$

where  $b > 0$  and  $p > 0$  are the parameters specifying the log-linear relationship, and  $u_t$  is the error term. Instead, Campbell and Cochrane (1999) search for a numerical solution of the PD ratio (4.25) by plugging in discount rates (4.28) and using a numerical integrator to evaluate the conditional expectation over the normally distributed consumption shocks  $v_{t+1}^c$ . Direct estimation is perhaps less precise but much simpler and leads to a graphical representation similar to Cochrane (2011, p. 1073).

It turns out that the statistical fit of this relationship is rather poor, as seen in Table 4.7, with an  $R^2$  of 4%. The estimate of  $p$  is below 0 (pointing to a negative relationship) but insignificant.<sup>15</sup> The main reason for the low fit is that the model does not capture the large stock market boom in the 1990s. Cochrane (2011) circumvents this problem by focusing on PD ratios after 1990 only. As Campbell and Cochrane (1999) note: "Growth in consumption of non-durables and services was surprisingly low in the early 1990s, so our model predicts a fall in price/dividend ratios rather than the increase we see in the data." They list (exogenous) reasons, such as shifts in corporate financial policy and shifts in consumption due to rising income inequality or demographic effects, why the model based on the time series  $C_t$  might underestimate the PD ratio in this period.

To capture these "shifts", we consider as a fundamental benchmark a consumption-habit model with one structural break in the 1990s. More precisely we allow for a structural break in

---

<sup>15</sup> In fact, the fitted PD ratios of the consumption-habit model without structural break are almost constant and very close to those of the (dynamic) Gordon model. Estimation of the HAM using these fundamental PD ratios leads therefore to similar results as in Section 4.3.

Variable	Coefficient	(Std. Err.)	Coefficient	(Std. Err.)
$b$	28.29	(589.3)	303.9***	(78.59)
$p$	-0.074	(1.18)	0.830***	(0.149)
$S_{break}$	0		0.093***	(0.007)
$T$	243		243	
$s^2$	260.5		58.43	
$R^2$	0.040		0.785	
$AIC$	5.540		4.054	

\*, \*\*, \*\*\* denote significance at the 10%, 5% and 1% level, respectively.

Table 4.7: Estimation of the log-linear relation between the surplus consumption ratio and PD ratio.

surplus consumption ratio in quarter 1995Q1 (corresponding to  $t = 173$ )<sup>16</sup>:

$$\delta_t = \begin{cases} b(S_t)^p + u_t & \text{if } t < 173 \\ b(S_t + S_{break})^p + u_t & \text{if } t \geq 173, \end{cases} \quad (4.35)$$

The last two columns of Table 4.7 present the results of this log-linear regression with one structural break. All three coefficients are significant and of the expected sign. We observe that a very large break in surplus consumption ratio of 0.093 is required, almost one and a half times the maximum value  $\bar{S}$ . Figure 4.5 plots the fitted PD ratios resulting from this model,  $\delta_t^{CC}$ . Given the large break, the model does track some of the variation in PD ratios, also after 1995, as argued by Cochrane (2011). Allowing for a structural break in the consumption-habit fundamental value improves the  $R^2$  from 4% to 79%.

The inclusion of one structural break seems, in principal, reasonable because we have a relatively long time series.<sup>17</sup> For example, Pástor and Stambaugh (2001) establish multiple structural breaks in the equity premium of the CSRP NYSE value-weighted portfolio from 1840 to 1999. Still, for the estimates  $b = 303.9$  and  $p = 0.830$  that are consistent with the whole sample, a lower risk premium due to high surplus consumption ratios cannot fully explain the

<sup>16</sup> The results do not depend qualitatively on the exact timing of the structural break. For example, changing the structural break to 1990Q1 ( $t = 153$ ) leads to similar parameter estimates:  $b = 269.7^{***}$ ,  $p = 0.773^{***}$ ,  $S_{break} = 0.091^{***}$ ;  $R^2 = 0.647$ .

<sup>17</sup> In Appendix 4.B, we also estimate our model using a Gordon fundamental value as in Section 4.3 with one structural break in 1995Q1. These estimation results are very similar and support our main result of significant behavioural heterogeneity.

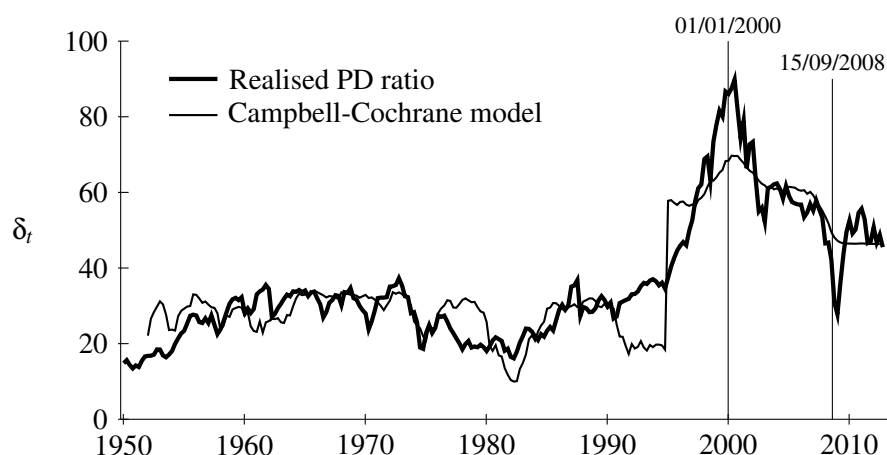


Figure 4.5: Realised PD ratio and its fitted values based on the Campbell-Cochrane model with a structural break in 1995Q1.

dot-com bubble as can be seen from Figure 4.5. The consumption-habit model also fails to explain why the stock market fell so deep in 2008Q3 and recovered so quickly afterwards, because the surplus consumption ratio stayed low for the whole period after 2008Q3. So even after allowing for one structural break, the fit of the Campbell-Cochrane model is far from perfect.

#### 4.5.2 Estimation of the HAM under the consumption-habit fundamental value

The Campbell-Cochrane model, despite being able to reproduce some general patterns in asset prices by a time-varying risk premium, does not fully explain asset price movements. Even if we disregard the large unexplained structural break in the 1990s, the model fails to account sufficiently for the two biggest events after 1990, namely the dot-com bubble and the financial crisis. We therefore extend the model with behavioural heterogeneity between agents as we did for the Gordon model. In other words, we re-estimate the model summarised by equations (4.17), (4.18) and (4.19) using the time series of PD ratios in deviation from the consumption-

Variable	(A)	(B)	(C)	(D)
$\phi_1$	0.699*** (0.071)	0.733*** (0.058)	0.759*** (0.052)	0.796*** (0.100)
$\phi_2$	1.017*** (0.068)	0.985*** (0.052)	0.980*** (0.050)	0.933*** (0.098)
$\beta$	1	3.768 (6.035)	10	1
$\omega$	0.686*** (0.129)	0.719*** (0.118)	0.817*** (0.016)	0
$\Delta\phi$	0.318*** (0.121)	0.253*** (0.085)	0.221*** (0.078)	0.137 (0.186)
$T$	243	243	243	243
$s^2$	13.94	13.72	13.89	14.40
$R^2$	0.948	0.949	0.948	0.946
$AIC$	2.664	2.654	2.657	2.687

\*, \*\*, \*\*\* denote significance at the 10%, 5% and 1% level, respectively.

Standard errors are denoted in brackets.

The  $R^2$  denotes the proportion variation in  $\delta_t$  explained by the model.

Table 4.8: Estimation of the belief coefficients  $\phi_1$  and  $\phi_2$ , the intensity of choice  $\beta$  and the memory parameter  $\omega$  in the HAM  $x_t^{CC} = \frac{1}{R^*}(n_{1,t}\phi_1 + n_{2,t}\phi_2)x_{t-1}^{CC} + \varepsilon_t$ , with  $R^* = 1.008$  and the fractions  $n_{1,t}$  and  $n_{2,t}$  updated according to (4.18) and (4.19), using PD ratios  $\delta_t^{CC}$  fitted to the Campbell-Cochrane model as in the last two columns of Table 4.7. All specifications are estimated with nonlinear least squares.

habit ‘fundamental value’:

$$x_t^{CC} \equiv \delta_t - \delta_t^{CC}. \quad (4.36)$$

The results for the HAM estimated on  $x_t^{CC}$  are presented in Table 4.8.<sup>18</sup>

The parameter estimates are, perhaps surprisingly, in many ways similar to those in Section 4.3 under a constant risk premium. Most importantly, we find that the belief parameters  $\phi_1$  and  $\phi_2$  are significantly different from each other. Comparing the estimations of model version (A) with three parameters on  $x_t^{CC}$  and  $x_t$ , we find that the estimate of  $\Delta\phi$  is larger - 0.32 instead of 0.08 - and highly significant. Even when taking time-varying risk premia into account, there

<sup>18</sup> In this estimation, we maintain the constant value of the expected effective discount rate at  $R^* = 1.008$  from the Gordon model for simplicity. See the pricing equation of the HAM in deviations from the fundamental value (4.12) in Section 4.2.2.

#### 4.5. ESTIMATION USING A CONSUMPTION-HABIT TIME-VARYING RISK PREMIUM

is significant evidence for different behavioural regimes in the data.

There is also significant memory for switching between the two behavioural rules, for estimation of the model on the quarterly data. Setting  $\omega = 0$  as is done in model version (D) leads to a lower fit and an insignificant  $\Delta\phi$ . The estimated intensity of choice is again not significant, because the model fit is very insensitive to the exact value of  $\beta > 0$  and its  $t$ -test has low power; see Section 4.5.3 for a more detailed discussion. Therefore our preferred model specification is (A), which raises the fit from an  $R^2$  of 79% to 95% for the consumption-habit model with one structural break and heterogeneous expectations.<sup>19</sup>

The main differences with the predictions of the HAM estimation in Section 4.3 lie in the values of the two beliefs, and make sense if we consider the different fundamental benchmark we have used. The value of  $\phi_1$  is much lower, because the fundamental value is generally closer to the actual PD ratio. For example, after the burst of the dot-com bubble, prices came within a few years back to the fundamental price based on a large surplus consumption ratio and a low risk premium. Also the parameter  $\phi_2$  is lower and in fact very close to 1, less than one standard deviation away; slightly above 1 in (A) and slightly below 1 in (B) and (C). Hence, while the ‘chartist’ belief is still significantly different from the mean-reverting fundamentalist belief, it is not significantly different from predicting the last observed price, that is, ‘naive expectations’, consistent with a belief that the PD deviation from the fundamental benchmark follows a random walk.

Figure 4.6 plots the estimated fraction of fundamentalists  $n_{1,t}$  and the market sentiment  $\phi_t$  with the time-varying risk premium. The most striking difference is that the fraction of fundamentalists is fluctuating heavily during the entire sample. Before 1995, the proportion of chartists rarely becomes very high, and asset prices stay relatively close to fundamentals. In the 1990s and around the year 2000 market sentiment comes close to 1 (and sometimes exceeds it),

---

<sup>19</sup> Note that in the present case the  $AIC$  slightly favours version (B) over (A), even though the estimation of the intensity of choice  $\beta$  is insignificant. Another note is that the HAM using a time-varying risk premium has a slightly lower  $R^2$  (94.8% instead of 95.2%) while having more parameters. The reason lies in the two-step estimation procedure of first the fundamental value and then the heterogeneous agents part on price deviations from this fundamental. Estimating all parameters simultaneously is possible and gives a higher  $R^2$  for the HAM using the consumption-habit time-varying risk premium.

destabilising the market and amplifying the dot-com stock boom.

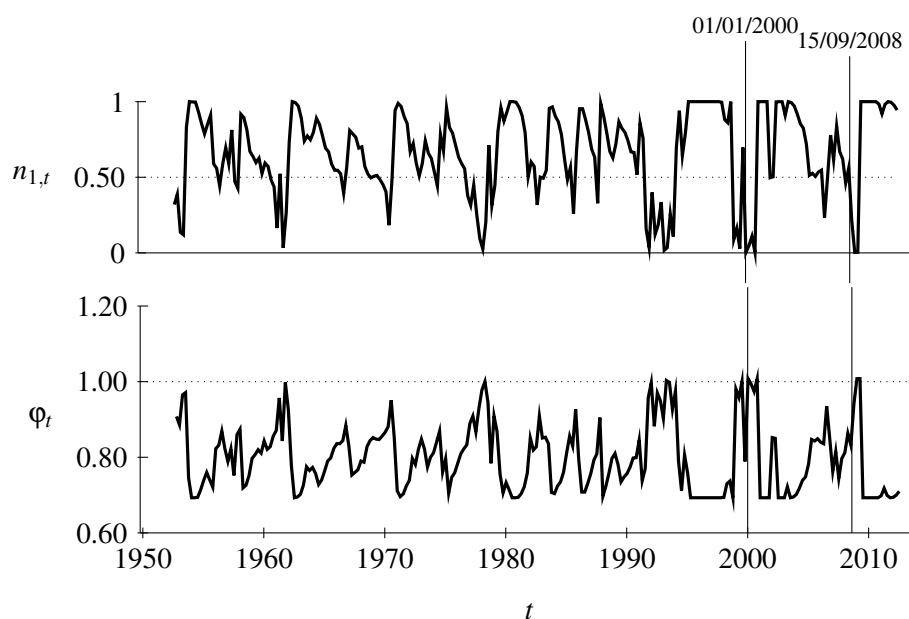


Figure 4.6: Estimated fraction of fundamentalists (top panel) and the corresponding market sentiment (bottom panel) for the HAM (A) using the Campbell-Cochrane model as fundamental benchmark.

The heterogeneous agents model with a fundamental benchmark with time-varying risk premium based on surplus consumption gives an interesting explanation for the depth of the financial crisis. The model predicts that the fundamental price around 2008 is close to 50, based on the higher surplus consumption ratio after the structural break in 1995 and the lower implied risk premium (see Figure 4.5). The bankruptcy of Lehman Brothers lead to a large negative shock at a time when a majority of investors was following a chartist strategy, as the fundamentalist belief was unrewarding in the past periods. This behavioural overreaction of investors to the Lehman Brothers shock, with the market sentiment exceeding 1, strongly amplified the financial crisis. Only starting from 2009Q3 price-dividend ratios returned to the fundamentals. During this period dividends increased, so prices started increasing again.

### 4.5.3 Evaluating the tests for the model under a consumption-habit fundamental value

As a final robustness check for the model with a time-varying risk premium, we run Monte Carlo simulations using the estimated HAM and a benchmark AR model as Data-Generating Processes.<sup>20</sup> We first estimate the linear benchmark model, see Table 4.9. The constant value  $\varphi$  is as expected halfway the beliefs of the two heterogeneous groups  $\phi_1$  and  $\phi_2$ :  $0.863 \approx (0.698 + 1.107)/(2 * 1.008)$ . The Dickey-Fuller test rejects the null hypothesis of a unit root in PD ratios in deviation from the consumption-habit fundamental value with one structural break.

Variable	Coefficient	(Std. Err.)
$\varphi$	0.863***	(0.032)
$DF$	-4.284***	$p$ -value < 0.01
$T$	243	
$s^2$	14.22	
$R^2$	0.946	
$AIC$	2.684	

\*, \*\*, \*\*\* denote significance at the 10%, 5% and 1% level, respectively.

The  $R^2$  denotes the proportion variation in  $\delta_t$  explained by the model.

Table 4.9: Estimation of the AR(1) model  $x_t^{CC} = \varphi x_{t-1}^{CC} + \varepsilon_t$ , using PD ratios fitted to the Campbell-Cochrane model as in the last two columns of Table 4.7.

Using Monte Carlo simulations, we evaluate the Dickey-Fuller test for a unit root ( $\phi = 1$ ), the test for homogeneity ( $\phi_1 = \phi_2$ ) and the test for no switching ( $\beta = 0$ ) given the estimated parameters in the model with a time-varying risk premium (similar as in Section 4.4.1 for the model with a constant risk premium). In Table 10 this is done for a large sample size of  $T = 5000$  and in Table 11 for the actual sample size of 243. In both cases we generate a large number of  $B$  simulated time series under the different DGPs and check the outcome of the three tests for each simulation.

The results are qualitatively the same as in the case with a constant risk premium. The

<sup>20</sup> It is not straightforward to make out-of-sample predictions for PD ratios with these DGPs as in Section 4.4.2, because we would then have to make predictions about future consumption growth. This is outside the scope of this chapter.

True model:	DGP \	Test: $H_0$	unit root	homogeneity	no switching
			$\phi = 1$	$\phi_1 = \phi_2$	$\beta = 0$
heterogeneous representative	HAM		100%	100%	100%
	AR		100%	(2%)	(0%)

Table 4.10: Rejection probabilities: nominal significance level  $\alpha = 5\%$ ,  $T = 5,000$ ,  $B = 1,000$ , parameters for HAM version (A) in Table 4.8 and for the AR model as in Table 4.9.

True model:	DGP \	Test: $H_0$	unit root	homogeneity	no switching
			$\phi = 1$	$\phi_1 = \phi_2$	$\beta = 0$
heterogeneous representative	HAM		80.3%	84.0%	0.0%
	AR		98.2%	(3.2%)	(0.0%)

Table 4.11: Rejection probabilities: nominal significance level  $\alpha = 5\%$ ,  $T = 243$ ,  $B = 10,000$ , parameters for HAM version (A) in Table 4.8 and for the AR model as in Table 4.9.

Dickey-Fuller test is asymptotically correct in rejecting the unit root, but has lower power for a smaller sample size of  $T = 243$ . Because the actual AR(1) parameter is here well below 1 ( $\phi = 0.863$ ) the power is still 98%; much higher than than the low power of 11% found in Table 4.3 (for  $\phi = 0.973$ ). If the actual underlying DGP is a nonlinear HAM, the power of the DF-test decreases to 80%.

When estimating a heterogeneous agents model one should be looking for significantly different beliefs (i.e. a significant  $\Delta\phi$ ), and not necessarily for significant switching (a significant  $\beta$ ). Both tests work asymptotically correct (having a high power and low size), but the test for no switching hardly ever rejects (0.0%) in small sample and is thus uninformative. Again, this explains why the estimated intensity of choice  $\beta$  is not significant. The test for homogeneity has large power (84%) even in small sample. The detected heterogeneity in the S&P500 data is therefore most probably not driven by random shocks in a representative agent model (as the attributed p-value is 3.2%).



## 4.6 Conclusion

In this chapter we investigate the value added of explaining asset pricing movements using a heterogeneous agents model. Motivated by the global financial crisis, our focus is on booms and busts that take place in the medium run, and we estimate our model on quarterly data. A convenient feature of our HAM is that it can be formulated around any benchmark fundamental. We use two well-known models as fundamental benchmark values for the price-dividend ratio, the dynamic Gordon model with a constant risk premium and the Campbell-Cochrane consumption-habit model with a time-varying risk premium. Using deviations from these fundamental benchmarks, we find a statistically significant improvement of the heterogeneous agents model over these standard representative agent models, and significant evidence for behavioural heterogeneity.

The global financial crisis displayed events that were sometimes inconceivable in a representative agent world. Our heterogeneous agents approach can shed light upon unexpected swings in asset prices driven by positive expectations feedback. Investors switch between strategies based upon their relative performance. When the group of fundamentalist traders gains momentum, prices return more quickly to the fundamental value. Temporary bubbles, triggered by small fundamental shocks, are strongly amplified when a majority of investors coordinates on chartist beliefs. The endogenous transitions between the regimes driven by relative profitability makes the model attractive in helping to explain booms and busts in asset pricing data.

In our simple heterogeneous agents model with fundamentalists and chartists, we find differences in expectations – measured quarterly and in deviations from the fundamental value – of 10 to 30 percentage points depending on the model that is used for the fundamental value. We show (using Monte Carlo simulations) that this large and significant effect cannot be expected to arise from standard representative agent models. Our model with different behavioural regimes of market sentiment is also economically significant in the sense that it gives better predictions than representative agent models in periods before the dot-com bubble and the financial crisis in 2008.

A limitation of our methodology is that we only use aggregate data to estimate behavioural rules of individual investors. The simple linear behavioural rules of our model are also found in laboratory experimental markets (e.g. Hommes, Sonnemans, Tuinstra, and van de Velden, 2005) and describe individual forecasting behaviour quite well. A recent and promising line of research combines stock market data with increasingly reliable survey measures of investors' expectations; e.g. Adam, Beutel, and Marcet (2013) present a model where agents are 'internally rational' but hold subjective beliefs about stock prices and calibrate it on data from the S&P500 and the UBS Gallup Survey, showing robustness to other survey data. These general patterns in survey data seem to be consistent with the main conclusions in this chapter. Our model predicts that during the dot-com bubble around 2000 investors were aware that prices were too high compared to measures of true fundamental value, and surveys as in e.g. Shiller (2000) support this claim. An analysis of behavioural heterogeneity combining aggregate price data and survey measures is an interesting possibility for future research.

Our analysis adds to evidence that behavioural heterogeneity and strategy switching plays an important role in asset price dynamics. Boundedly rational traders can be expected to survive in financial markets and amplify booms and busts. In particular, we show that agents switched between fundamentalist and chartist beliefs, and strongly reinforced the decline in asset prices after the shock of the Lehman Brothers bankruptcy. Uncertainty due to behavioural heterogeneity has important implications for risk management. Policy makers should therefore not focus exclusively on rational representative agent models, but should take behavioural heterogeneity into account in assessing the systemic risk in financial markets.

## Appendix 4.A Deriving the profit function

In this appendix we derive the profit function of an agent with belief type  $h$ ,  $\pi_{h,t}$ , in the general asset pricing model with heterogeneous agents. First we need to specify the demand and profits of agents concerned with the asset's excess returns  $\rho_{t+1}$ . Denote  $E_{h,t}[\cdot]$  as the expectation,  $V_{h,t}[\cdot]$  as the belief of the conditional variance and  $a_h$  as the risk aversion, all of agent type  $h$ . The demand for assets (or, if negative, the supply)  $z_h$  of type  $h$  is derived from investors being myopic mean-variance maximisers:

$$z_{i,t} = \frac{E_{i,t}[\rho_{t+1}]}{a_i V_{i,t}[\rho_{t+1}]} \quad (4.37)$$

From demand it is straightforward to derive realised profits  $\pi_{h,t+1}$ :

$$\pi_{h,t+1} = z_{i,t} \rho_{t+1} = \frac{E_{h,t}[\rho_{t+1}]}{a V_t[\rho_{t+1}]} \rho_{t+1} \quad (4.38)$$

In order to write realised profits in terms of the PD ratio, we make two more simplifying assumptions, similar to the assumptions made by Boswijk, Hommes, and Manzan (2007). The first assumption is that agents focus exclusively on the variability of prices when choosing prediction rules. In particular, agents replace the stochastic variables  $\frac{R_{t+1}}{1+g_{t+1}}$  and  $g_{t+1}$  by their unconditional expectations  $R^*$  and  $g \equiv E[E_t[g_{t+1}]]$ . Under this assumption, the performance measure is not affected by stochastic dividend growth and discount rates. It implies that the expect excess return is:

$$\begin{aligned} E_{h,t}[\rho_{t+1}] &= E_{h,t}[P_{t+1} + D_{t+1} - R_{t+1}P_t] \\ &= E_{h,t}[(\delta_{t+1} + 1)(1 + g_{t+1})D_t - R_{t+1}P_t] \\ &= (E_{h,t}[\delta_{t+1}] + 1 - R^* \delta_t)(1 + g)D_t. \end{aligned} \quad (4.39)$$

The second assumption concerns the beliefs about the conditional variance of excess returns. Recall that we already assumed these beliefs to be the same over all agents. The additional

assumption is that the beliefs about the conditional variance are equal to fundamental beliefs about the conditional variance.

$$\begin{aligned}
 V_{h,t}[\rho_{t+1}] &= V_{h,t}[P_{t+1} + D_{t+1} - R_{t+1}P_t] \\
 &= V_t[P_{t+1}^* + D_{t+1} - R_{t+1}P_t^*] \\
 &= V_t[(\delta_{t+1}^* + 1)D_{t+1} + R_{t+1}\delta_t^*D_t] \\
 &= V_t[(\delta_{t+1}^* + 1)(1 + g_{t+1})D_t + R_{t+1}\delta_t^*D_t] \\
 &= V_t[(\delta_{t+1}^* + 1)(1 + g_{t+1}) + R_{t+1}\delta_t^*]D_t^2 = \eta^2 D_t^2. \tag{4.40}
 \end{aligned}$$

These two assumptions provide sufficient conditions for the following simple expression of realised profits, as in Brock and Hommes (1998) and Boswijk, Hommes, and Manzan (2007), to hold in the most general setting with a non-stationary fundamental and time-varying risk premia.

$$\begin{aligned}
 \pi_{h,t+1} &= \frac{E_{h,t}[\rho_{t+1}]}{aV_t[\rho_{t+1}]} \rho_{t+1} \\
 &= \frac{(E_{h,t}[\delta_{t+1}] + 1 - R^*\delta_t)(1 + g)D_t}{a\eta^2 D_t^2} (\delta_{t+1} + 1 - R^*\delta_t)(1 + g)D_t \\
 &= \frac{(1 + g)^2}{a\eta^2} (E_{h,t}[\delta_{t+1}] + 1 - R^*\delta_t)(\delta_{t+1} + 1 - R^*\delta_t), \tag{4.41}
 \end{aligned}$$

and in deviations from the fundamental value:

$$\pi_{h,t+1} = \frac{(1 + g)^2}{a\eta^2} (E_{h,t}[x_{t+1}] - R^*x_t)(x_{t+1} - R^*x_t). \tag{4.42}$$

This expression for realised profits has the intuitive property that it is proportional to the number of shares bought (depending on the expectations  $E_{h,t}[x_{t+1}] - R^*x_t$ ) times the realised excess return (depending on realisations  $x_{t+1} - R^*x_t$ ).

## Appendix 4.B Robustness analysis

In this appendix we analyse the robustness of our estimation results of the heterogeneous agents model of the PD ratios in deviations from the Gordon benchmark model.

In Figure 4.7, we show that the explanatory power of the model (as measured by the  $R^2$ ) is much less sensitive to the value of  $\beta$  than to that of the other parameters  $\phi_1$ ,  $\phi_2$  and  $\omega$ . For  $\beta \geq 1$  the  $R^2$  is essentially flat. This is the reason why we often find a nonsignificant  $\beta$  coefficient. Because a non-zero  $\beta$  is necessary to identify different regimes as well as the level of memory  $\omega$ , we have fixed  $\beta$  to the (arbitrary) value of 1 in most of our estimations. Fixing  $\beta$  to some larger value yields very similar results.

Next, we try to improve the fundamental value of the Gordon model by adding a structural break in the risk premium starting from 1995Q1, just as we did for the Campbell Cochrane consumption-habit fundamental model in Section 4.5. We find that before 1995Q1, the risk premium was high and around 0.85% per quarter; from 1995Q1 the risk premium was much lower at around 0.32% per quarter. This leads to higher fundamental prices after 1995:  $\delta_{pre95}^* = 25.3$  versus  $\delta_{post95Q1}^* = 54.2$ . We consider for the Gordon model with one structural break again both the static and the dynamic model. For the dynamic Gordon model, there is a higher impact of deviations from the average values  $g$  and  $r$ . See Figure 4.8.

We reestimate model versions (A) with  $\beta = 1$ , (B) with all four parameters and (D) with  $\beta = 1$  and  $\omega = 0$ , in Table 4.12 for the static Gordon and in Table 4.13 for the dynamic Gordon model, both with a break in 1995Q1. The main result is robust for the structural break: if the model includes memory, the two beliefs are significantly different from each other. The estimated value of  $\Delta\phi$  is between the estimate for the Gordon model without break (around 0.09) and the estimate for the Campbell Cochrane consumption-habit model (up to 0.30), namely 0.15 to 0.20. The estimates of  $\phi_2$  are very close to 1 (naive expectations), slightly above 1 for the static and slightly below 1 for the dynamic Gordon model. In general, our results are robust and our conclusions do not depend on our choice of the benchmark fundamental value.

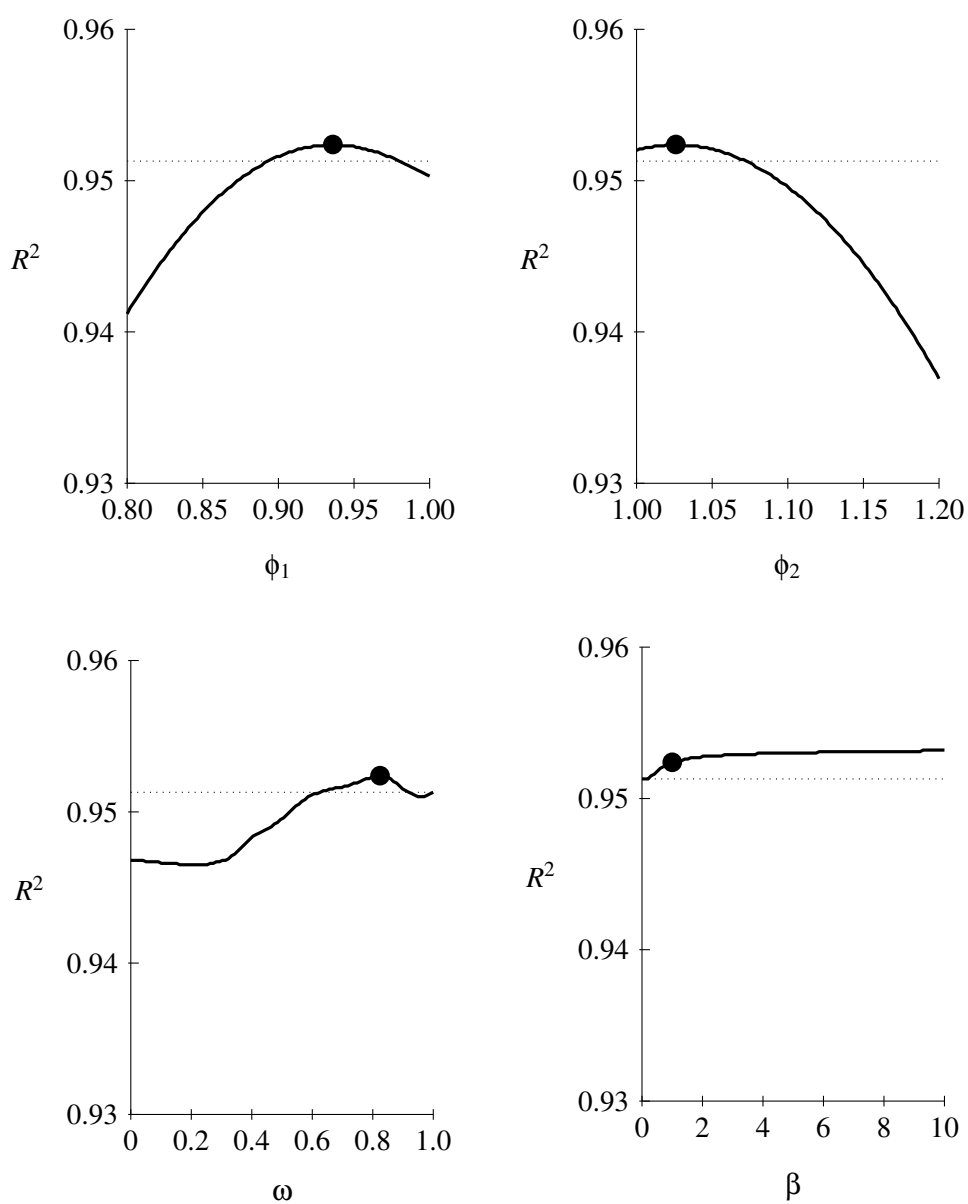


Figure 4.7: Sensitivity of  $R^2$  in the HAM. The dot denotes the choice of parameters as in model (A) giving  $R^2 = 0.952$ . In each of the diagrams, one of the four parameters ( $\phi_1$ ,  $\phi_2$ ,  $\omega$  or  $\beta$ ) is altered while retaining the values of the other three. The dotted line indicates the fit of the AR(1) model ( $R^2 = 0.951$ ).

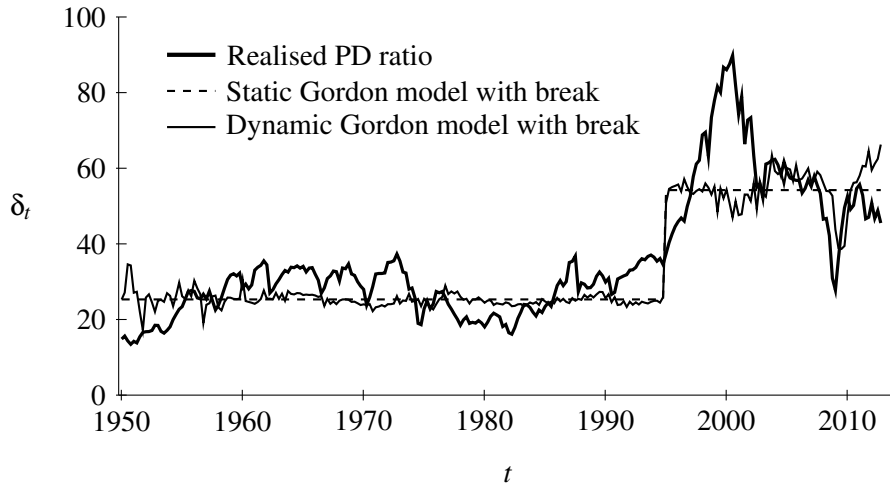


Figure 4.8: Realised PD ratio and its fundamental values based on the static and dynamic Gordon model with one structural break in 1995Q1.

	(A)	(B)	(D)
$\phi_1$	0.817*** (0.053)	0.827*** (0.048)	0.930*** (0.150)
$\phi_2$	1.019*** (0.036)	1.010*** (0.035)	0.945*** (0.164)
$\beta$	1	1.957 (3.106)	1
$\omega$	0.640** (0.276)	0.613** (0.271)	0
$\Delta\phi$	0.203*** (0.073)	0.184** (0.067)	0.016 (0.309)
$T$	252	252	252
$s^2$	10.90	10.88	11.39
$R^2$	0.955	0.955	0.953
$AIC$	2.525	2.531	2.556

The  $R^2$  denotes the proportion variation in  $\delta_t$  explained by the model.

Table 4.12: Estimation of the belief coefficients  $\phi_1$  and  $\phi_2$ , the intensity of choice  $\beta$  and the memory parameter  $\omega$  in the HAM  $x_t^{SG+break} = \frac{1}{R^*}(n_{1,t}\phi_1 + n_{2,t}\phi_2)x_{t-1}^{SG+break} + \varepsilon_t$ , with  $R^* = 1.008$  and the fractions  $n_{1,t}$  and  $n_{2,t}$  updated according to (4.18) and (4.19), under a static Gordon model with one structural break in 1995Q1. All specifications are estimated with nonlinear least squares.

	(A)	(B)	(D)
$\phi_1$	0.816*** (0.054)	0.829*** (0.052)	0.914 (-)
$\phi_2$	0.989*** (0.045)	0.977*** (0.043)	0.914 (-)
$\beta$	1	2.359 (5.080)	1
$\omega$	0.752*** (0.176)	0.759*** (0.123)	-
$\Delta\phi$	0.174** (0.082)	0.148** (0.075)	0 (-)
$T$	252	252	252
$s^2$	17.32	17.26	-
$R^2$	0.945	0.945	0.944
$AIC$	2.728	2.731	2.738

The  $R^2$  denotes the proportion variation in  $\delta_t$  explained by the model.

Table 4.13: Estimation of the belief coefficients  $\phi_1$  and  $\phi_2$ , the intensity of choice  $\beta$  and the memory parameter  $\omega$  in the HAM  $x_t^{DG+break} = \frac{1}{R^*}(n_{1,t}\phi_1 + n_{2,t}\phi_2)x_{t-1}^{DG+break} + \varepsilon_t$ , with  $R^* = 1.008$  and the fractions  $n_{1,t}$  and  $n_{2,t}$  updated according to (4.18) and (4.19), under a dynamic Gordon model with one structural break in 1995Q1. All specifications are estimated with nonlinear least squares except for specification (D), which is found by a grid search.



# Chapter 5

## Adverse effects of demand constraints in a financial market model with heterogeneous agents

### 5.1 Introduction

Should policy makers intervene in financial markets, and if so, how effective are market regulations? During times of large financial distress, the potentially amplifying role of excessive investment positions to asset price bubbles and crashes is heavily debated. It has been noted that leverage ratios increased to enormous levels, which allowed investors to buy assets with money they did not actually have. Short-selling is an inverse phenomenon, where investors sell assets they do not actually own. This chapter critically assesses the conjecture that constraining leverage and short-selling stabilises financial markets and restores prices to fundamentals.

The main message of this chapter is that the stabilising role of demand constraints – short-selling and leverage constraints – is severely limited if investors have heterogeneous beliefs about future prices. This result is shown to be robust for various forms of belief formation by investors. The intuition for these limitations follows logically from the idea that different beliefs

themselves can be stabilising or destabilising. As demand constraints are imposed on all market participants, there is no guarantee that destabilising beliefs become most restricted. In fact, this chapter shows examples in which demand constraints have adverse effects because the investors that should stabilise the market become restricted. It is possible that demand constraints lead to a new steady state, but then the market is dominated by the most optimistic investors and the steady state price lies always above the fundamental value.

To investigate leverage and short-selling constraints we use the well-known Brock and Hommes (1998) dynamic asset pricing framework with heterogeneous beliefs. The dynamical system arising in this model is called an *adaptive belief system* (ABS) and describes how prices and beliefs coevolve over time. Agents switch between beliefs if they observe that one of them has proven more profitable in the past. If agents observe that other beliefs are too biased and generate losses, a majority of agents switches endogenously to better performing strategies. In adaptive belief systems strategies producing consistent losses are driven out of the market, which reduces the degrees of freedom in modelling heterogeneous beliefs. An attractive feature of the ABSs in this chapter that they exhibit excess volatility (Shiller, 1981) that is typically observed in empirical financial time series. Adaptive belief systems thus provide a simple stylised framework to study the effects of leverage and short-selling constraints.

If investors take large positions in financial markets, inevitably they should be different in some way: investors are only able to buy large amounts of stock if other are willing to sell these amounts. Heterogeneity is thus a key feature to model leverage and short-selling in financial markets. A considerable literature of bubbles and crises in asset markets with heterogeneous investors exists and is surveyed in e.g. Brunnermeier and Oehmke (2012). The research that uses heterogeneity to evaluate leverage and short-selling constraints is reviewed below.

*Leverage* has mainly been analysed in equilibrium models with heterogeneous, yet fully rational agents. Gromb and Vayanos (2002) proposes a model with financially constrained arbitrageurs who can only invest one out of two risky assets. Longstaff (2008) considers two types of agents who differ in risk aversion. In a more recent stream of the literature, Geanakoplos

(2010) investigates the leverage cycle arising in a model with a continuum of investors who differ in their optimism, and shows how the margin changes over time. Simsek (2013) focuses on two belief types, optimists and pessimists, and models in greater detail the loan contract optimist use to leverage their investment. However, these papers evaluate how leverage affects equilibrium prices under rational agents with different characteristics, but say little about dynamic stability as discussed in this chapter.

The paper that comes closest to this chapter in investigating the dynamic effects of leverage is Thurner, Farmer, and Geanakoplos (2012). They develop a quantitative dynamics model where fundamentalist investors buy an underpriced asset. Underpricing occurs when noise traders drive the price below the fundamental value. The fundamentalists' optimal investment is proportional to their wealth, but leverage constraints bind the maximum amount they can invest. The results of the model show a downward spiral of margin calls: the natural buyers of the asset can turn into sellers when a price fall leads to a margin call. However, short-selling by investors is excluded, and therefore the model of Thurner, Farmer, and Geanakoplos (2012) can only be applied to situations where the asset is underpriced. This chapter combines possible effects of short-selling and leverage, and applies to both underpriced and overpriced assets.

*Short-selling* constraints have mostly been assessed as a policy tool on its own account and without considerations of leverage, for example in the classical paper of Miller (1977). In the seminal model of Harrison and Kreps (1978), risk-neutral investors with different valuations about the future asset price would take infinitely large positions in the absence of short-selling constraints. With a full ban on short-selling, speculative bubbles arise: the price can exceed even the expectation of the most optimistic agent because this agent believes he can resell the asset at an even higher price in the future. Anufriev and Tuinstra (2013) evaluate the impact of short-selling constraints in a dynamic model with heterogeneous, boundedly rational agents. They show that trading costs for selling an asset short can increase mispricing and price volatility when the asset is overvalued.

To our knowledge this is the first attempt to explicitly combine the effects of leverage and

short-selling constraints in a financial market model. Buss, Dumas, Uppal, and Vilkov (2013) compare different regulatory measures to control stock market volatility in a dynamic stochastic general equilibrium (DSGE) model with heterogeneous agents. The regulations that are considered are leverage constraints, short-selling constraints and the Tobin tax. Interestingly, although their methodology is very different from the one followed in this chapter, Buss, Dumas, Uppal, and Vilkov (2013) find similar adverse effects of market regulations in their model. Their elaborate model allows them to analyse indirect effects on real sectors, but cannot be used to impose different regulations simultaneously. The approach in the present chapter is more straightforward and focuses on adverse effects of demand constraints in adaptive belief systems.

This chapter fits in a growing literature on using heterogeneous beliefs models for policy evaluations; see e.g. Westerhoff (2008).<sup>1</sup> In particular, we extend the model of Anufriev and Tuinstra (2013) to allow for leverage constraints. Additional to the possibility of trading costs for short-selling, we introduce leverage constraints through a maximum investment level that applies to all investors in every period. In this general model we evaluate the effects of leverage constraints in comparison with trading costs for short-selling, and we consider a mixed policy that restricts demand on both long and short positions. Overall, the analysis points out that demand constraints do not necessarily stabilise financial markets and may have adverse effects.

To check the robustness of the results, we evaluate the asset pricing model for different stylised examples of adaptive belief systems. We consider not only a typical two-type model with fundamentalists versus chartists as in Anufriev and Tuinstra (2013), but also a three-type model with fundamentalists versus optimists and pessimists. Together these stylised adaptive belief systems capture a number of important cases. The time series generated by these ABSs exhibit two typical stylised facts observed in financial markets, namely bubbles and crashes, and oscillating behaviour around a fundamental price. By constraining investment positions along the fluctuations, the coevolution of prices and beliefs can be distorted.

---

<sup>1</sup> Brock, Hommes, and Wagener (2009) extend the Brock-Hommes asset pricing framework with heterogeneous beliefs by adding futures (Arrow securities) and show that in the presence of these additional hedging instruments the market destabilises because agents take excessively leveraged positions.

The chapter is organised as follows. Section 5.2 introduces the heterogeneous agents model with leverage and short-selling constraints. Section 5.3 presents the results for the typical two-type and three-type adaptive belief systems. In Section 5.4, we discuss the policy implications of the results for the stock market during the dot-com bubble and the financial crisis in 2008. The proofs can be found in the Appendix.

## 5.2 An asset pricing model with heterogeneous beliefs and demand constraints

### 5.2.1 Individual asset demand

In this section we present the asset pricing model with heterogeneous agents of Brock and Hommes (1998), extended with short-selling constraints by Anufriev and Tuinstra (2013), and introduce additional leverage constraints. Denote  $y_t$  as the dividend payoff,  $p_t$  as the asset price and  $z_{h,t}$  as the number of shares bought by investor type  $h$ , all at time  $t$ . Investors differ in their beliefs about what the price of the asset will be in the next time period: investor type  $h$  expects  $E_{h,t}[p_{t+1}]$ . Individual beliefs will influence demand  $z_{h,t}$  and thereby the equilibrium price  $p_t$ . Apart from their beliefs, investors are homogeneous: they know that dividends are identically and independently distributed with mean  $\bar{y}$ , and believe the variance of excess returns is  $\sigma^2$ .<sup>2</sup>

It is assumed that investors are mean-variance maximisers as in Brock and Hommes (1998):

$$\max_{z_{h,t}} \left( E_{h,t}[W_{h,t+1}] - \frac{a}{2} V_{h,t}[W_{h,t+1}] \right), \quad (5.1)$$

where  $a$  is the risk aversion, and  $E_{h,t}[\cdot]$  and  $V_{h,t}[\cdot]$  denote the beliefs of type  $h$  about the conditional expectation and variance. This maximisation problem can be derived from a standard utility function with constant absolute risk aversion (CARA).

---

<sup>2</sup> Gaunersdorfer (2000) extends the model of Brock and Hommes (1998) with homogeneous but time-varying beliefs  $\widehat{\sigma}_t^2$  about the variance of asset prices, and shows that for the model without demand constraints, the results are quite similar.

Wealth evolves as according to

$$\begin{aligned}
 W_{h,t+1} &= RW_{h,t} + \rho_{t+1}z_{h,t} - R\tau(z_{h,t}) \\
 &= RW_{h,t} + (p_{t+1} + \bar{y} - Rp_t)z_{h,t} - R\tau(z_{h,t}).
 \end{aligned} \tag{5.2}$$

The first term in equation (5.2) represents accumulation of wealth provided by a risk free asset under a fixed return  $R \equiv 1 + r_f$ , where  $r_f$  is the risk free rate. The second term denotes wealth effects from investing in the risky asset, depending on the number of shares  $z_{h,t}$  and the excess return  $\rho_{t+1}$  of the risky asset over the risk free asset. This excess return consists of the next period's price and dividend payoff less the opportunity costs of investing in the risky asset. The last term in the wealth equation (5.2),  $\tau(z_{h,t})$ , represents the trading costs for selling assets short, and has been introduced in Anufriev and Tuinstra (2013) as

$$\tau(z_{h,t}) = \begin{cases} 0 & \text{if } z_{h,t} \geq 0 \\ T|z_{h,t}| & \text{if } z_{h,t} < 0, \end{cases} \tag{5.3}$$

implying that for every unit of the asset sold short the investor has to pay a fixed tax of  $T$ .

The novel feature in this chapter is that traders face leverage constraints. Given their own limited amount of wealth, they have to attract loans if they want to buy large amounts of assets. Denoting  $\lambda^{max} \geq 1$  as the maximum leverage level, the leverage constraint can be written as

$$\begin{aligned}
 p_t z_{h,t} &\leq \lambda^{max} W_{h,t} \\
 &\leq I_{h,t}^{max}.
 \end{aligned} \tag{5.4}$$

The variable  $I_{h,t}^{max}$  represents the *maximum investment level* agent type  $h$  can invest as a monetary amount.

In the absence of trading costs and if the maximum investment level is sufficiently high, the optimal demand follows from the unconstrained maximisation problem (5.1) and is a decreasing

linear function in the price  $p_t$ :

$$\begin{aligned} z_{h,t}^*(p_t) &= \frac{E_{h,t}[\rho_{t+1}]}{aV_{h,t}[\rho_{t+1}]} \\ &= \frac{E_{h,t}[p_{t+1}] + \bar{y} - Rp_t}{a\sigma^2}. \end{aligned} \quad (5.5)$$

Note that, because the investment problem has been derived from a CARA utility function as in (5.1), demand is independent of the wealth level.

Along the same lines as the CARA utility function, we assume that the leverage constraint is independent of the wealth level, and is rather an absolute constraint. In other words, the maximum investment level is fixed over all agent types and periods:  $I_{h,t}^{max} = I^{max} > 0$  for all  $h, t$ . While this is clearly a simplifying assumption, it provides a natural starting point for evaluating the effects of leverage constraints on asset price dynamics.<sup>3</sup>

In general with possibly binding leverage and/or short-selling constraints, the relation between price and demand becomes nonlinear. Consider the possibility that trader type  $h$  is leverage constrained as in (5.4) because the optimal investment  $z_{h,t}^*$  is too high under price  $p_t$ . In this case, when the price  $p_t$  decreases further to 0, demand  $z_{h,t}^*$  will become even higher. At some point though the acquisition costs  $p_t z_{h,t}^*$  will start decreasing and the leverage constraint will stop being binding. In other words, demand is leverage constrained for a bounded range of strictly positive prices.<sup>4</sup> More precisely, the leverage constraint is binding when

$$z_{h,t}^* \geq I^{max} / p_t, \quad (5.6)$$

<sup>3</sup> A more developed model taking wealth effects into account could be based on a constant *relative* risk aversion (CRRA) utility function as in Chiarella and He (2001) or Anufriev and Bottazzi (2006). Wealth dynamics however could exhibit feedback of leverage constraints. When buyers become leverage constrained, prices drop, which decreases the buyers' profits and wealth accumulation. Less wealth make the buyers even more constrained, driving prices further down.

<sup>4</sup> It can be checked that  $p_t z_{h,t}^* = 0$  for  $p_t = 0$ , such that demand can only be leverage constrained for strictly positive prices.

or equivalently

$$p_t \in [p_t^{h-}, p_t^{h+}], p_t^{h\pm} \equiv \frac{1}{2R} \left( E_{h,t}[p_{t+1}] + \bar{y} \pm \sqrt{(E_{h,t}[p_{t+1}] + \bar{y})^2 - 4a\sigma^2 I^{max}R} \right). \quad (5.7)$$

Hence, there are two *cut-off values*  $p_t^{h\pm}$  that define the area in which agent type  $h$  is leverage constrained. Of course, if the determinant is negative the cut-off values do not exist, and the agent is never constrained. In the constrained area the agent demands his maximal number of shares  $I^{max}/p_t$ .

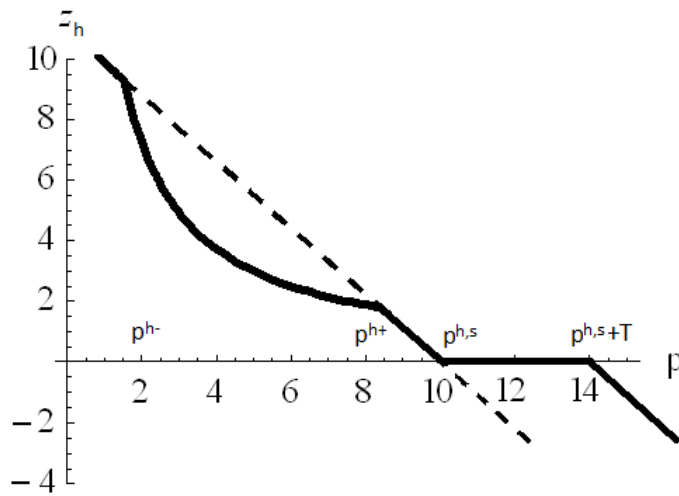


Figure 5.1: Demand  $z_h$  as a function of the price  $p$  (solid line) compared to linearly decreasing unconstrained demand (dashed line). For low prices  $p^{h-} < p < p^{h+}$ , demand is driven down by leverage constraints to  $I^{max}/p$ . For high prices  $p^{h,s} < p < p^{h,s} + T$ , demand is zero, while for even higher prices  $p > p^{h,s} + T$  the short-selling tax  $T$  is covered and demand is again linearly decreasing.

Anufriev and Tuinstra (2013) observed that similar cut-off values exist for the short-selling constraint. The short-selling tax has to be paid when  $z_{h,t}^* < 0$ , or

$$p_t > p_t^{h,s} = \frac{1}{R}(E_{h,t}[p_{t+1}] + \bar{y}). \quad (5.8)$$

When the price  $p_t$  increases above  $p_t^{h,s}$ , the demand function stays flat as both buying and short-selling is not beneficial. For even higher prices, the expected excess return can cover the additional cost of  $T$  per unit sold short that have to be paid. Both leverage and short-selling



constraints generate nonlinearities in the demand curve.

Taking the two constraints together gives a demand schedule that consists of five pieces:<sup>5</sup>

$$z_{h,t}(p_t) = \begin{cases} \frac{1}{a\sigma^2}(E_{h,t}[p_{t+1}] + \bar{y} - Rp_t) & \text{if } p_t < p_t^{h-} \\ I^{max}/p_t & \text{if } p_t \in [p_t^{h-}, p_t^{h+}] \\ \frac{1}{a\sigma^2}(E_{h,t}[p_{t+1}] + \bar{y} - Rp_t) & \text{if } p_t \in [p_t^{h+}, p_t^{h,s}] \\ 0 & \text{if } p_t \in [p_t^{h,s}, p_t^{h,s} + T] \\ \frac{1}{a\sigma^2}(E_{h,t}[p_{t+1}] + \bar{y} - R(p_t - T)) & \text{if } p_t > p_t^{h,s} + T. \end{cases} \quad (5.9)$$

Figure 5.1 illustrates the nonlinear individual demand for trader type  $h$  as a function of  $p_t$ , for given values of  $E_{h,t}[p_{t+1}]$  and parameters  $\bar{y}$ ,  $R$ ,  $a\sigma^2$ ,  $T$  and  $I^{max}$ . It shows how the demand function is affected by the two main parameters of interest:  $T$ , the trading costs for short-selling, and  $I^{max}$ , the maximal investment level. For  $T = 0$  and  $I^{max} = \infty$ , demand is unconstrained. The limit case  $T = \infty$  represents a full ban on short-selling. A policy maker can either increase  $T$ , decrease  $I^{max}$  or both, in order to influence demand of individual investors and thereby the stability of the market.

## 5.2.2 Market equilibrium

Consider  $H$  types of agents who differ in their beliefs  $E_{h,t}[p_{t+1}]$  for  $h \in 1, \dots, H$ . Denote the fractions of the different types at time  $t$  as  $n_{h,t}$ . The supply of the asset is fixed at  $z_s$  per trader. The intertemporal market equilibrium price  $p_t$  is implicitly defined by the market clearing equation,

<sup>5</sup> One can check that  $p_t^{h,s} > p_t^{h+}$ , such that the interval  $[p_t^{h+}, p_t^{h,s}]$  is non-empty.

given by

$$\sum_{h=1}^H n_{h,t} z_{h,t}(p_t) = z_s. \quad (5.10)$$

First consider the case that agents are rational. Then all investors can be described by a representative type ( $H = 1$ ) with perfect foresight:

$$E_t^*[p_{t+1}] = p_{t+1}. \quad (5.11)$$

In equilibrium this representative agent will hold  $z_s$  units of the asset, that is, if this investment level is allowed by the leverage constraint (5.4). In case the leverage constraint is not binding, the equilibrium price is the discounted sum of expected risk-adjusted dividends, given by

$$p_f = \frac{\bar{y} - a\sigma^2 z_s}{r_f}, \quad (5.12)$$

and denoted as the *fundamental price*  $p_f$ . In this chapter we assume that the representative rational agent is not constrained by the maximum investment level:

**Assumption 5.1.** *Leverage constraints are sufficiently loose for the fundamental price (5.12) to hold, i.e.*

$$\begin{aligned} I^{max} &> p_f z_s \\ &> \frac{\bar{y} - a\sigma^2 z_s}{r_f} z_s. \end{aligned} \quad (5.13)$$

Now consider the case  $H > 1$  with different beliefs  $E_{h,t}[p_{t+1}]$ . In the absence of trading costs and if the maximum investment level is sufficiently high, the equilibrium price can be derived from optimal unconstrained demand (5.5) and the market clearing equation (5.10), leading to

$$p_t = \frac{1}{R} \left( \sum_{h=1}^H n_{h,t} E_{h,t}[p_{t+1}] + \bar{y} - a\sigma^2 z_s \right). \quad (5.14)$$

In this market equilibrium equation the term  $a\sigma^2 z_s$  denotes a risk premium for investors holding the risky asset. Brock and Hommes (1998) focus on the special case of zero outside supply, i.e.  $z^s = 0$ , that ignores the risk premium.

In general under demand constraints, it becomes unfeasible to derive the equilibrium price analytically, because the individual demand schedules are highly nonlinear. The following proposition state that a unique equilibrium price exists.

**Proposition 5.1.** *Consider an asset market with heterogeneous beliefs and possible demand constraints. If  $z_s > 0$ , there exists a unique solution to  $\sum_{h=1}^H n_{h,t} z_{h,t}(p_t) = z_s$  providing the market clearing price  $p_t$  at time  $t$ .*

**Proof.** See Appendix 5.A □

The proof of Proposition 5.1 follows similar reasoning as in Anufriev and Tuinstra (2013). Individual demand is strictly decreasing when it is positive (see Figure 5.1), which ensures a unique solution for the market clearing price if outside supply  $z_s$  is strictly positive.<sup>6</sup>

The next section investigates the price dynamics in this model if agents can switch between strategies. Both the fraction  $n_{h,t}$  and the beliefs  $E_{h,t}[p_{t+1}]$  can change over time. Using the result of Proposition 5.1 that a unique price  $p_t$  exists in every period  $t$ , we analyse the effects of demand constraints on the dynamics.

### 5.3 Stylised adaptive belief systems

In this section we introduce the dynamics in the asset pricing market if agents can switch between different belief strategies. The fractions of agents with different types are updated using

---

<sup>6</sup> A prerequisite for the existence of a unique equilibrium is that leverage constraints are modelled as a maximum investment level  $I^{max}$  in a monetary amount. If instead a maximum were imposed on the amount of shares  $z_{h,t}$ , the demand would be constant in  $p_t$  for two ranges of prices. Then aggregate demand  $\sum_{h=1}^H n_{h,t} z_{h,t}$  may be constant for a range of prices and a continuum of equilibrium prices may exist.

the multinomial logit model as in Brock and Hommes (1997), i.e.

$$n_{h,t+1} = \frac{e^{\beta U_{h,t}}}{\sum_{j=1}^H e^{\beta U_{j,t}}}. \quad (5.15)$$

The parameter  $\beta$  is the *intensity of choice* with which agents react on differences in performance. If  $\beta = 0$ , agents do not switch and fractions are evenly distributed over the  $H$  rules:  $n_{h,t+1} = 1/H$  for all  $t$ . For increasing  $\beta$ , agents are more sensitive to performance in the last period and switch quicker to better performing rules.

For the performance  $U_{h,t}$  of the different rules, the main component is formed by the excess returns  $\rho_t$  times the asset holdings  $z_{h,t-1}$ . Apart from excess returns, agents consider two possible sources of costs: first, the trading costs  $R\tau(z_{h,t-1})$ , depending on whether the asset is sold short, and second, exogenous costs  $C_h$  associated with gathering information to use strategy  $h$ . The performance measure is therefore given by

$$U_{h,t} = \rho_t z_{h,t-1} - R\tau(z_{h,t-1}) - C_h. \quad (5.16)$$

In describing the expectation rule of the agents and price dynamics, we follow the convention to write prices in terms of the deviation from the fundamental price  $p_f$ ,

$$x_t = p_t - p_f. \quad (5.17)$$

In terms of deviations from the fundamental price, the excess return on the risky asset is

$$\begin{aligned} \rho_{t+1} &= p_{t+1} + \bar{y} - Rp_t \\ &= x_{t+1} - Rx_t + a\sigma^2 z_s, \end{aligned} \quad (5.18)$$

and the demand schedule becomes:

$$z_{h,t}(x_t) = \begin{cases} \frac{1}{a\sigma^2}(E_{h,t}[x_{t+1}] - Rx_t) + z_s & \text{if } x_t < x_t^{h-} \\ I^{max}/(x_t + p_f) & \text{if } x_t \in [x_t^{h-}, x_t^{h+}] \\ \frac{1}{a\sigma^2}(E_{h,t}[x_{t+1}] - Rx_t) + z_s & \text{if } x_t \in [x_t^{h+}, x_t^{h,s}] \\ 0 & \text{if } x_t \in [x_t^{h,s}, x_t^{h,s} + T] \\ \frac{1}{a\sigma^2}(E_{h,t}[x_{t+1}] - R(x_t - T)) + z_s & \text{if } x_t > x_t^{h,s} + T. \end{cases} \quad (5.19)$$

using appropriate definitions for  $x_t^{h\pm}$  and  $x_t^{h,s}$  corresponding with  $p_t^{h\pm}$  and  $p_t^{h,s}$ .<sup>7</sup>

As Brock and Hommes (1998), we use simple beliefs that are linear in the last observation, given by

$$E_{h,t}[x_{t+1}] = b_h + g_h x_{t-1}. \quad (5.20)$$

As will become clear, this simple structure of belief formation is sufficient to capture a number of important cases. Asset pricing models inhabited by agents with beliefs of the form (5.20) are able to generate important properties in medium-run asset pricing dynamics, namely recurrent bubbles and crashes, and oscillations around a fundamental value.

We use two examples of adaptive belief systems (ABSs). In Section 5.3.1, we follow Anufriev and Tuinstra (2013) in taking the ABS with fundamentalists versus chartists, where fundamentalists pay a cost for acquiring information about the fundamental value. In Section 5.3.2, we consider another example of an ABS from Brock and Hommes (1998) with three

<sup>7</sup> In contrast to Brock and Hommes (1998) and Anufriev and Tuinstra (2013), price dynamics are *not* independent from the value of the fundamental price. In the leverage constrained part of the demand function (5.9),  $I^{max}/p_t$ , the term  $p_f$  cannot be eliminated, because the constraint is expressed as a maximum monetary amount.

fixed beliefs ( $g_h = 0$  for all  $h$ ) and without information costs. Here the population of investors consists of fundamentalists, optimists and pessimists. In both examples, fundamental traders *cannot* drive out irrational traders even for increasingly large intensity of switching.

### 5.3.1 Fundamentalists versus chartists

Assume that there are two types of agents, the first *fundamentalists* and the second *chartists* with respective expectations

$$E_{1,t}[x_{t+1}] = 0, \quad (5.21)$$

$$E_{2,t}[x_{t+1}] = b + gx_{t-1}. \quad (5.22)$$

Fundamentalists pay a fixed cost  $C > 0$  to calculate the fundamental value. Chartists do not acquire information to find the fundamental value, but update their beliefs given the last observed price  $x_{t-1}$  to extrapolate trends with some possible bias  $b$ . For simplicity we set  $b = 0$ , such that the chartists extrapolate trends around the fundamental value.<sup>8</sup>

For the baseline parametrisation in the fundamentalist-chartist market (irrespective of demand constraints), we follow Anufriev and Tuinstra (2013):

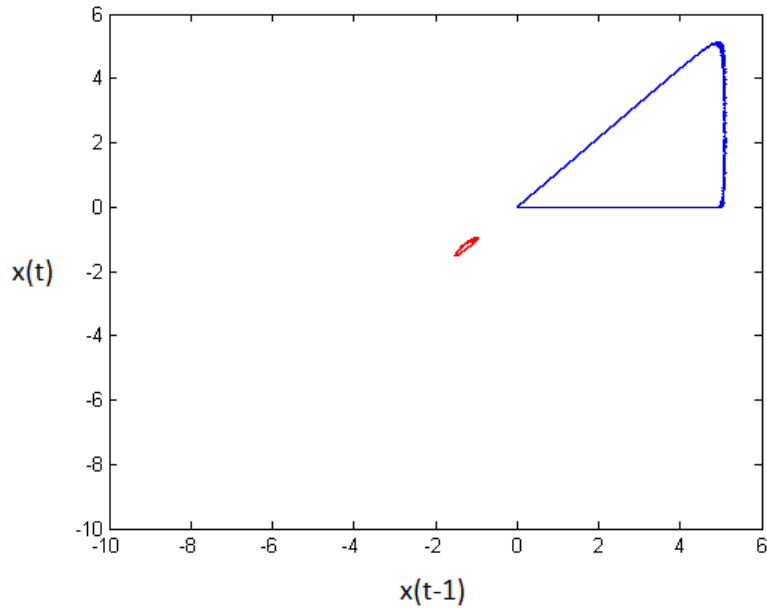
**Definition 5.1** (Baseline parametrisation under fundamentalists and chartists). *Parameters:*  $z_s = 0.1$ ,  $\alpha\sigma^2 = 1$ ,  $r_f = 0.1$ ,  $\bar{y} = 1.1$  ( $\rightarrow p_f = 10$ ),  $g = 1.2$ ,  $C = 1$  and  $\beta = 4$ .

Figure 5.2 shows the benchmark dynamics without constraints. Figure 5.2a shows two attractors in the  $(x_{t-1}, x_t)$ -plane. In Figure 5.2b, the left panels show dynamics on the upper attractor; the right panels on the lower attractor. The three panels in vertical order depict the evolution of prices, fractions of fundamentalists and positions, respectively.

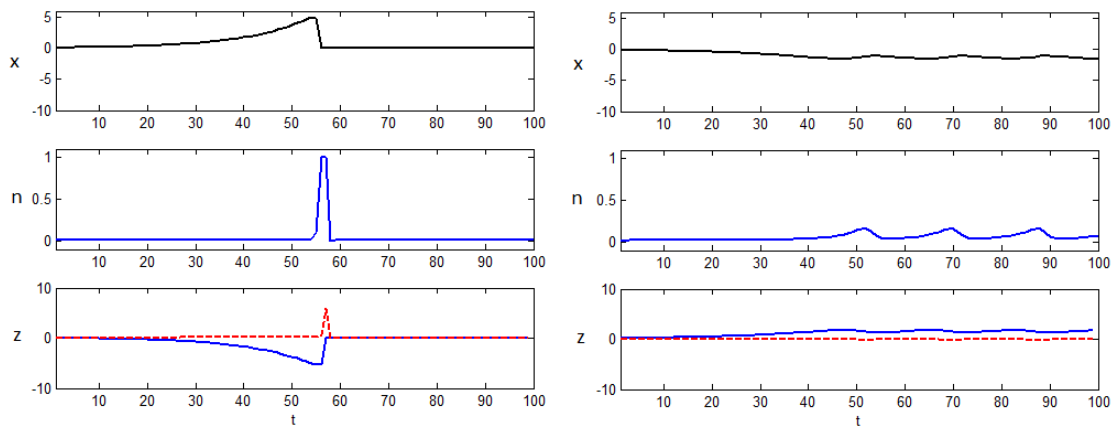
In this baseline model, the fundamental price is unstable for the relatively large extrapolation factor of chartists  $g > R$ . The reason is that around the fundamental steady state chartists earn

---

<sup>8</sup> One can argue that unbiased extrapolation requires  $p_f$  to be in the information set of the chartists, inconsistent with the information-gathering costs  $C$  of fundamentalists. However, for high values  $g > R$  the term  $gx_{t-1}$  generally dominates a relatively small bias  $b$ .



(a) Phase plot showing two quasi-periodic attractors, one above the fundamental value (blue) and one below the fundamental value (red)



(b) Time series. *Left panels:* dynamics on the upper attractor. *Right panels:* dynamics on the lower attractor. *Top panels:* prices. *Middle panels:* fraction of fundamentalists. *Bottom panels:* positions of fundamentalists (blue) and chartists (red, dashed).

Figure 5.2: Phase plot (a) and time series (b) for the benchmark ABS with fundamentalists versus chartists without constraints. The baseline parametrisation and  $T = 0$ ,  $I^{max} = \infty$ .

equally large excess returns as fundamentalists, while circumventing the information costs. As chartists attract more agents, market prices are influenced by trend-following behaviour, and price bubbles arise endogenously. Gradually, though, as chartists overshoot the growing prices  $x_t$  more and more by their extrapolative expectations  $gx_{t-1}$ , fundamentalists regain followers and eventually there is an abrupt crash. For this choice of parameters, there are also small recurrent negative ‘bubbles’ or troughs below the fundamental value.

Recurrent bubbles are a robust phenomenon in this model, and it is possible to give bounds for how large bubbles can grow. Lemma 4 of Brock and Hommes (1998) states that for  $z_s = 0$ ,  $R < g < R^2$  and  $\beta = \infty$  the dynamics are bounded. The following proposition specifies the bounds of price dynamics for  $\beta = \infty$  given any value of  $z_s > 0$ .

**Proposition 5.2.** *Consider the ABS with fundamentalists and chartists without demand constraints. Assume  $z_s > 0$ ,  $C > 0$ , and  $\beta = \infty$ . Then:*

- (a) *For  $0 < g < R$ , the fundamental steady state  $E_1 = (x^*, n_1^*) = (0, 0)$  is the unique, globally stable steady state.*
- (b) *For  $R < g < R^2$ , the fundamental steady state  $E_1 = (0, 0)$  is locally unstable. Price dynamics are bounded and converge to the (locally unstable) steady state  $E_1 = (0, 0)$ . For an initial price above the fundamental  $x_0 = \varepsilon > 0$ , the dynamics are bounded by some smallest interval  $[0, x^{max})$  with*

$$x^{max} \equiv \frac{g}{R} \frac{\sqrt{(a\sigma^2 z_s)^2 + 4(1 - g/R^2)a\sigma^2 C + a\sigma^2 z_s}}{2(R^2/g - 1)}. \quad (5.23)$$

*For an initial price below the fundamental  $x_0 = \varepsilon < 0$ , the dynamics are bounded by some smallest interval  $(x^{min}, 0]$  with*

$$x^{min} \equiv -\frac{g}{R} \frac{\sqrt{(a\sigma^2 z_s)^2 + 4(1 - g/R^2)a\sigma^2 C - a\sigma^2 z_s}}{2(R^2/g - 1)}. \quad (5.24)$$

- (c) *For  $R^2 < g < R^2(1 + \frac{a\sigma^2 z_s^2}{4C})$ , the fundamental steady state  $E_1 = (0, 0)$  is locally unstable.*



Price dynamics with an initial price above the fundamental  $x_0 = \varepsilon > 0$  diverge. For an initial price below the fundamental  $x_0 = \varepsilon < 0$ , price dynamics converge to the (locally unstable) steady state  $E_1 = (0,0)$  and are bounded by some smallest interval  $(x^{min}, 0]$  with

$$x^{min} \equiv -\frac{g}{R} \frac{\sqrt{(a\sigma^2 z_s)^2 - 4(g/R^2 - 1)a\sigma^2 C - a\sigma^2 z_s}}{2(1 - R^2/g)}. \quad (5.25)$$

(d) For  $g > R^2(1 + \frac{a\sigma^2 z_s^2}{4C})$ , the fundamental steady state  $E_1 = (0,0)$  is locally unstable and all price dynamics with an initial price  $x_0 = \varepsilon \neq 0$  diverge.

**Proof.** See Appendix 5.A □

Proposition 5.2 gives important insights about the magnitude of bubbles when the intensity of choice is large but finite. For  $R < g < R^2$ , prices move away from the fundamental with a rate close to  $g/R$  but remain within endogenous bounds. During the bubble chartists' forecasts overshoot the evolution of prices and, with finite  $\beta$ , fundamentalists start to attract more followers before  $x_t$  reaches  $x^{max}$  or  $x^{min}$ . After the bubble is being stalled, prices return very closely to the unstable steady state, and a new bubble appears.

For the special case  $z_s = 0$ , prices are bounded symmetrically by

$$|x_t| < \frac{g^2}{R^2} \sqrt{\frac{a\sigma^2 C}{R^2 - g}}, \quad (5.26)$$

a value increasing in  $g$ ,  $a\sigma^2$  and  $C$ , and decreasing in  $R$ . As already found by Anufriev and Tuinstra (2013), when there is positive supply  $z_s > 0$ , bubbles are larger above the fundamental value than below. Proposition 5.2 makes clear that for  $C \downarrow 0$ , bubbles below the fundamental value disappear ( $x^{min} = 0$ ) and the upperbound becomes:

$$x^{max} = \frac{g^2}{R} \frac{a\sigma^2 z_s}{R^2 - g}, \quad (5.27)$$

which is linearly increasing in  $z_s$ . The magnitude of bubbles depends intuitively on the pa-

rameters and can be evaluated analytically, which helps to investigate the effects of demand constraints in the next sections.

### Short-selling constraints in the two-type market

First we consider short-selling constraints in this market. Figure 5.3 gives the bifurcation diagrams for the short-selling tax  $T$  for the baseline parametrisation, given  $I^{max} = \infty$  to exclude the leverage constraint. The plot shows two attractors of time series for the baseline parameter values and different initial conditions. In the model with fundamentalists and chartists, prices can never cross the fundamental price  $p_f$  and the time series  $x_t$  keeps the same sign as the initial value  $x_0$ . So in this model (without noise) prices are either always above or always below the fundamental price.

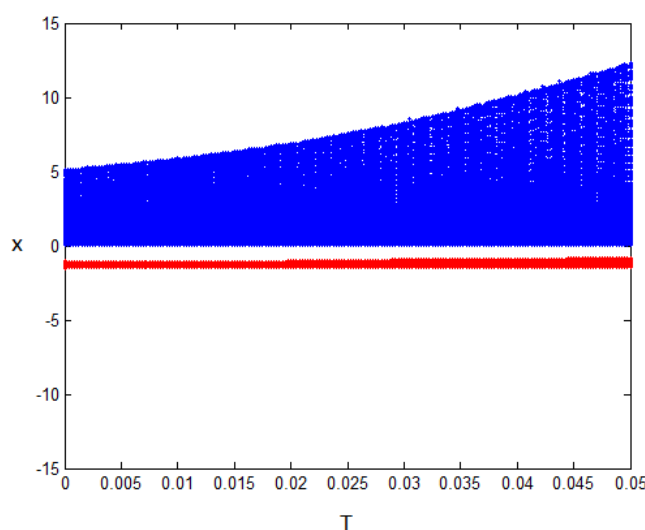


Figure 5.3: Bifurcation diagram for short-selling tax  $T$  given the baseline parametrisation and  $I^{max} = \infty$ . Small noise is added to the price  $x_t$  ( $\sigma = 10^{-3}$  with the same sign as  $x_t$ ) to avoid that the numerical simulation gets stuck in the locally unstable steady state  $x_t = 0$ . Two attractors are shown, one above the fundamental value (blue) and one below the fundamental value (red).

Above the fundamental value an increase of the tax  $T$  on short selling increases average mispricing and volatility in the market. The reason is that as the bubble builds up above the fundamental value fundamentalists want to sell large amounts of the asset short: see the bottom left panel of Figure 5.2b. Because the large number of short-sold shares carry high trading

costs, the fundamentalist strategy is less profitable relative to the case without a short-selling tax. Chartists remain the dominant agent type and continue to ‘ride’ the bubble by extrapolating the price for a longer time. As prices are driven up further, at some point fundamentalists’ excess returns  $\rho_{t,z_{1,t}}$  of betting on a price drop will become so large to cover even the additional short-selling tax  $R\tau(z_{h,t})$ , and prices crash back close to the fundamental price.

On the other hand, below the fundamental value, the effects of a trading tax are negligible, as chartists short sell the assets but in relatively small amounts. This can be checked in the bottom right panel of Figure 5.2b. The bifurcation diagram for  $T$  in Figure 5.3 confirms the main result of Anufriev and Tuinstra (2013), namely that costs on short-selling increase mispricing and price volatility in the two-type market when the asset is overvalued.

### **Leverage constraints in the two-type market**

To allow for an unconstrained fundamental value as in Assumption 5.1, the baseline parameter choices in Definition 1 require  $I^{max} > 1$ . The effects of the leverage constraints on the dynamics are shown in Figure 5.4. On the lower attractor, bubbles become larger as the constraint is tightened (i.e.  $I^{max}$  decreases), an effect that is similar for short-selling constraints on the upper attractor. The reason is that the fundamentalist strategy, which should drive back prices to the fundamental, requires much more extreme positions and soon face the leverage constraint (see Figure 5.2b, bottom right panel).

On the upper attractor, there are no visible effects for large values of  $I^{max}$ . The leverage constraint should restrict the price to rise far above the fundamental value, as chartists become constrained by the maximum investment level  $I^{max}$ . Because the chartists have relatively low demand levels along the bubble (see Figure 5.2b, bottom left panel), the leverage constraint only becomes binding after fundamentalists start have a higher performance and the crash already started.

However, in contrast to short-selling constraints, tight leverage constraints can have a stabilising effect, if prices are above the fundamental value. As seen in the upper part of Figure 5.4,

leverage constraints do affect the dynamics for  $I^{max}$  smaller 3. In this case chartists become constrained during a rising market, making the bubble burst faster.

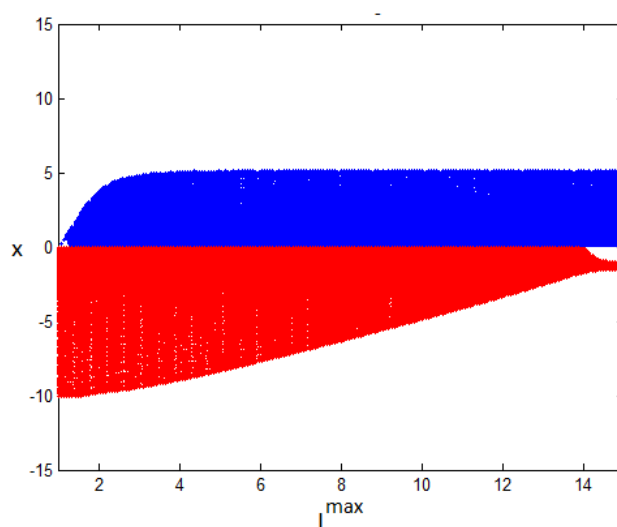


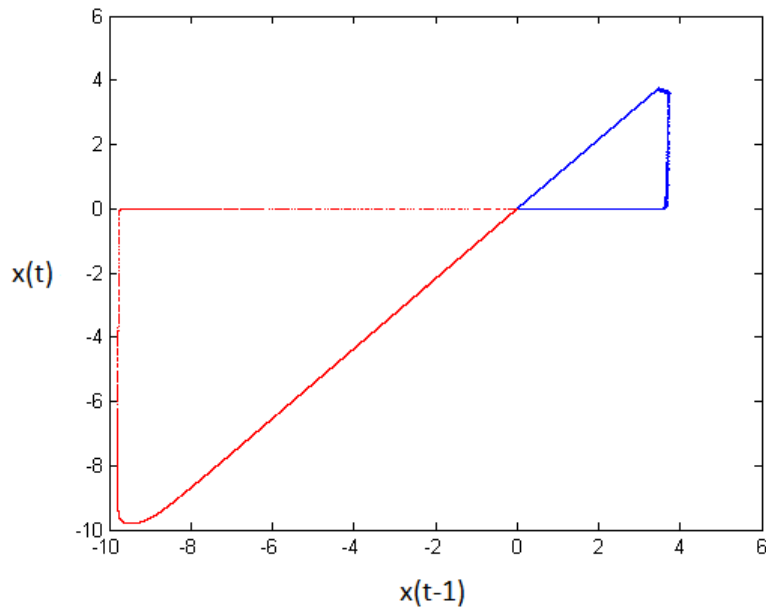
Figure 5.4: Bifurcation diagram for leverage constraint  $I^{max}$  given the baseline parametrisation and  $T = 0$ . Small noise is added to the price  $x_t$  ( $\sigma = 10^{-3}$  with the same sign as  $x_t$ ) to avoid that the numerical simulation gets stuck in the locally unstable steady state  $x_t = 0$ . Two attractors are shown, one above the fundamental value (blue) and one below the fundamental value (red).

Figure 5.5 shows the phase plots and time series for a relatively tight leverage constraint  $I^{max} = 2$  (without short-selling tax,  $T = 0$ ). In Figure 5.5b, the evolution of prices, fractions and demand levels are presented on both attractors. On the upper attractor (left panels), chartists have small yet positive positions  $z_{2,t}$  during the complete time series, through which the leverage constraint invokes a slightly earlier switch to the fundamentalist strategy and therefore a smaller bubble. On the lower attractor (right panels), fundamentalists can buy only small amounts leading to a large drop of the price  $x_t$  to almost  $-10$ .

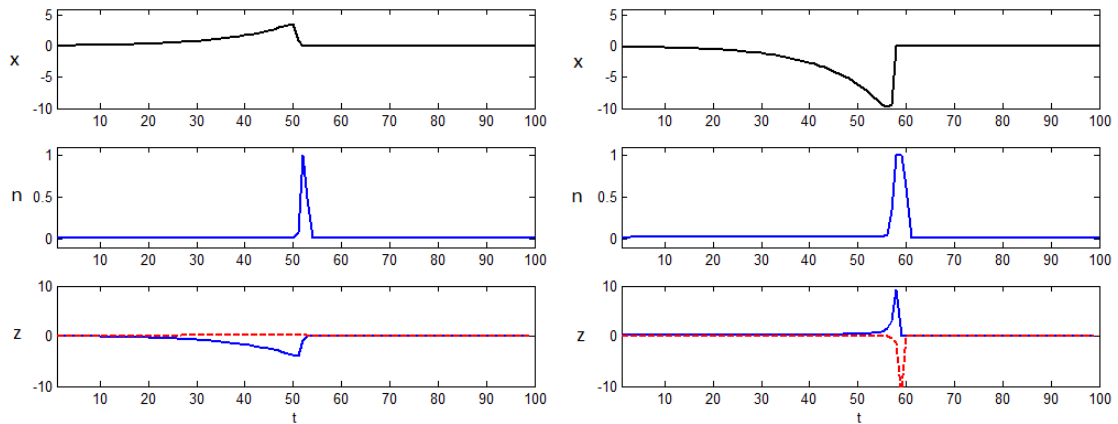
### Combining the two demand constraints in the two-type market

Taking the results in the two-type ABS together, a recurring point is that constraints have the adverse effect to restricting the stabilising fundamentalist strategy, that requires larger positions. As fundamentalists are the stabilising type in the market, both constraints can increase price volatility and mispricing, but on different sides of the fundamental price. Short-selling

constraints have destabilising effects above the fundamental price, when the fundamentalist strategy requires selling the asset; leverage constraints destabilise below the fundamental when it requires buying the asset. In both cases it takes a longer bubble before traders switch back to the fundamentalist belief and correct the destabilising trend-following behaviour of chartists.



(a) Phase plot showing two quasi-periodic attractors, one above the fundamental value (blue) and one below the fundamental value (red).



(b) Time series. *Left panels:* dynamics on the upper attractor. *Right panels:* dynamics on the lower attractor. *Top panels:* prices. *Middle panels:* fraction of fundamentalists. *Bottom panels:* positions of fundamentalists (blue) and chartists (red, dashed).

Figure 5.5: Phase plot (a) and time series (b) for the ABS with fundamentalists versus chartists and a leverage constraint. The baseline parametrisation and  $T = 0, I^{max} = 2$ .

Only by using tight leverage constraints for prices above the fundamental price, the destabilising effect of chartists can be limited.

The following proposition shows that the effects found under the baseline parametrisation hold in general for  $g > R$  in the limit case  $\beta = \infty$ .

**Proposition 5.3.** *Consider the ABS with fundamentalists and chartists under possible demand constraints. Assume  $z_s > 0$ ,  $C > 0$  and  $\beta = \infty$ . Then if  $g > R$ , the fundamental steady state  $E_1 = (x^*, n_1^*) = (0, 0)$  is locally unstable for all  $T \geq 0$  and  $I^{max} > p_f z_s$ . Furthermore, if the price dynamics are for all initial prices  $x_0 = \varepsilon$  bounded by some smallest interval  $(x^{min}, x^{max})$  with  $x^{min} < 0 < x^{max}$ , the following effects hold:*

- (a) *an increase in  $T$  (i.e. a tighter short-selling constraint) has no effect on  $x^{min}$ ;*
- (b) *an increase in  $T$  (i.e. a tighter short-selling constraint) has a (weakly) increasing effect on  $x^{max}$ ;*
- (c) *a decrease in  $I^{max}$  (i.e. a tighter leverage constraint) has a (weakly) decreasing effect on  $x^{min}$ ;*
- (d) *a decrease in  $I^{max}$  (i.e. a tighter leverage constraint) has a (weakly) decreasing effect on  $x^{max}$ . There is a price ceiling*

$$x^{L+} \equiv I^{max}/z_s - p_f > 0 \tag{5.28}$$

*on the dynamics implying  $x^{max} \leq x^{L+}$ . For sufficiently low  $I^{max}$  there exists a locally stable, non-fundamental steady state  $E_2 = (x^{L+}, 0)$  at the pricing ceiling.*

**Proof.** See Appendix 5.A □

In Proposition 5.2, it was shown that for  $\beta = \infty$ ,  $R < g < R^2$  and in the absence of demand constraints, prices are bounded by  $x^{max}$  and  $x^{min}$ . Proposition 5.3 now indicates for  $g > R$  how these price bounds – if they exist – are affected by changes in the demand constraints. Parts (b)

and (c) show the adverse effects of respectively short-selling and leverage constraints on price stability.

Proposition 5.3 further states in part (d) that a new, non-fundamental steady state can arise under a strong leverage constraint. In this equilibrium all agents are chartists ( $n_{2,t} = 1$ ) and buy the entire stock of assets  $z_s$  at price  $p_t = p_f + x^{L+} = I^{max}/z_s$ . Note that the chartist investors would like to buy more for the given price, as their forecasts are high given the extrapolation factor  $g > R$ . In the absence of the leverage constraint, chartists indeed buy more (from fundamentalists) and prices increase until the chartists' forecasts overshoot the realised prices, leading to a price crash. But if chartists are leverage constrained, they cannot buy more than the maximum monetary amount  $I^{max}$  permits, and the asset price remains in this non-fundamental equilibrium.

The optimistic steady state occurs because chartists become the *natural buyers* of the asset when prices are above the fundamental value. This term was coined by Geanakoplos (2010) to denote the proportion of a heterogeneous population that want the asset the most (for example because they have the most optimistic expectations). Natural buyers can push up asset prices when they are not counterbalanced by natural sellers. This is a well-known finding when heterogeneous beliefs are combined with short-selling constraints. Miller (1977) and Harrison and Kreps (1978) argued that restricting short-selling can lead to overpricing because the natural sellers of the asset with more pessimistic beliefs become constrained.

In our model with changing fractions, natural buyers can dominate the market even without short-selling constraints, if natural buyers receive consistently higher payoffs than natural sellers. This is exactly what happens in part (d) of Proposition 5.3. The payoffs of leverage-constrained chartists are higher than those of the fundamentalists, leading to a non-fundamental steady state all investors are chartists. Importantly, the steady state price  $x^{L+}$  is an increasing function of  $I^{max}$ . This underlines the earlier finding that leverage constraints can have a stabilising effect if prices are above the fundamental value.

The results suggest that combining the market regulations cannot help to mitigate their po-

tential adverse effects, as long as the constraints affect the stabilising forces in the market. Above the fundamental value leverage constraints reduce mispricing and price volatility, but additional short-selling constraints can undo this effect and increase financial instability. Below the fundamental value leverage constraints have a similar adverse effect, and additional short-selling taxes are irrelevant. The example of fundamentalists versus chartists thus points to a serious limitation of demand constraints in stabilising financial markets.

### 5.3.2 Fundamentalists versus optimists and pessimists

We saw that under the adaptive belief system (ABS) with costly fundamentalists versus chartist strategies, demand constraints may increase mispricing and price volatility. The question arises how robust these results are with respect to the set of boundedly rational strategies agents choose from. In this section we provide another simple and typical example of an adaptive belief system with three agent types (see also Brock and Hommes, 1998; Hommes, 2013). The forecasting rules are

$$E_{1,t}[x_{t+1}] = 0, \tag{5.29}$$

$$E_{2,t}[x_{t+1}] = +b, \tag{5.30}$$

$$E_{3,t}[x_{t+1}] = -b, \tag{5.31}$$

where  $b > 0$  is a constant. The first rule is again the (stabilising) fundamentalist belief, this time without information-gathering costs. The two alternative strategies in this ABS are optimists with a fixed biased belief above fundamental value, and pessimists with a similarly fixed biased belief below the fundamental value. These strategies do not require information costs either.

First we introduce a positive outside supply  $z_s \geq 0$  into this adaptive belief system as in the two-type ABS. This extension is motivated by the aim of the chapter to investigate demand constraints. A strictly positive supply is required to ensure that a market clearing price in general always exists (see Proposition 5.1). In economic terms, an important motivation to allow for



strictly positive supply is that traders who sell the asset, not necessarily sell the asset short (Anufriev and Tuinstra, 2013, p. 1526). Allowing for positive outside supply also introduces a risk premium into the model, see equation (5.14).

The following proposition specifies the nature of price dynamics in the limit case  $\beta = \infty$ . It generalises Lemma 9 of Brock and Hommes (1998) to the case of positive supply  $z_s > 0$ .

**Proposition 5.4.** *Consider the ABS with fundamentalists, optimists and pessimists without demand constraints. Assume  $z_s > 0$ ,  $b > 0$ , and  $\beta = \infty$ . Then:*

- (a) *For  $0 < b < \frac{r_f}{R} a \sigma^2 z_s$ , a non-fundamental steady state  $E = (x^*, n_1^*, n_2^*, n_3^*) = (b/R, 0, 1, 0)$  exists and is the unique, globally stable steady state.*
- (b) *For  $b > \frac{r_f}{R} a \sigma^2 z_s$ , the system has a globally stable 4-cycle with prices  $(b/R)\{1, 1, -1, -1\}$  and price volatility  $\sigma = b/R$ .*

**Proof.** See Appendix 5.A □

A difference with a market inhabited by chartists is that the range of feasible prices is always limited, as the beliefs of all three strategies are fixed and bounded. More precisely, the price  $x_t$  can only fluctuate in the interval  $[-b/R, b/R]$ . Proposition 5.4 shows that fundamentalist traders cannot drive out simple rule-of-thumb strategies even though biased expectations  $b$  and  $-b$  lie outside this interval, when agents have a high intensity of switching to better performing strategies. For sufficiently large biases, cycles occur and the market is alternately dominated by optimists and pessimists.

For positive outside supply and small biases, the market converges to a non-fundamental steady state even if  $\beta = \infty$ . In this steady state occurs optimists – the natural buyers of the asset – become the dominating agent type in the market (cf. part (d) of Proposition 5.2 in which chartists were the natural buyers possibly dominating the market). Notice that optimists always buy more of the asset than other agent types as they expect a higher price in the next period. Hence optimists receive the highest payoffs if excess returns of the risky asset  $\rho_t$ . Positive

outside supply increases the excess return in (5.18) via the risk premium  $a\sigma^2 z_s$ . A series of high prices  $x_{t-1} = x_t = b/R$  leads to positive excess returns if

$$\begin{aligned} \rho_t &= b/R - R(b/R) + a\sigma^2 z_s > 0 \\ \Leftrightarrow \quad b &< \frac{r_f}{R} a\sigma^2 z_s, \end{aligned} \quad (5.32)$$

showing the dominance of the optimistic natural buyers, and explaining the non-fundamental steady state with  $x_t = b/R$ .

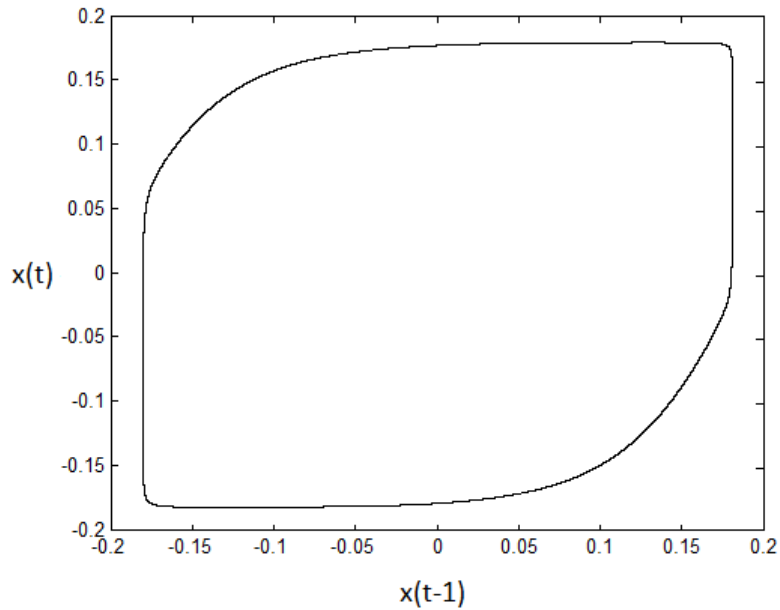
For the baseline parametrisation without constraints, we follow Hommes (2013, p. 175), with a small value of supply  $z_s = 0.01$ . Under this parameters the bias  $b$  is sufficiently large to investigate the effect of demand constraints on cycles around the fundamental value.

**Definition 5.2** (Baseline parametrisation under fundamentalists and biases). *Parameters:*  $z_s = 0.01$ ,  $r_f = 0.1$ ,  $a\sigma^2 = 1$ ,  $\bar{y} = 1.01$  ( $\rightarrow p_f = 10$ ),  $b = 0.2$  and  $\beta = 100$ .

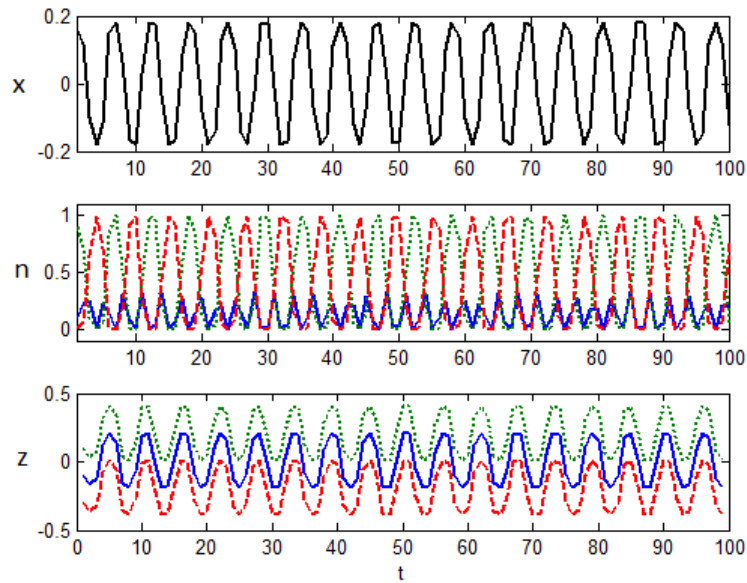
Given the finite value of  $\beta$ , we observe quasi-periodic cycles of prices around the fundamental value (Figure 5.6). The middle and lower time series in Figure 5.6b show the fractions and positions of all three agent types. Fundamentalists have a cyclical demand following the price fluctuations, while the demand of optimists and pessimists equals fundamentalists' demand shifted vertically by  $\frac{1}{a\sigma^2}b$  and  $-\frac{1}{a\sigma^2}b$  (see the optimal unconstrained investment level (5.5)). As in the tops and troughs of the price cycle optimists and pessimists receive larger profits, the market is alternately dominated by either optimists or pessimists, reinforcing the cycle. Under the baseline parameters the three-type ABS thus exhibits characteristic feature of prices oscillating around the fundamental price.

### Short-selling constraints in the three-type market

We introduce short-selling constraints in the same way as we did in the two-type ABS. Figure 5.7 presents the bifurcation diagram for the short-selling tax  $T$  given the baseline parametrisation. In this three-type ABS, short-selling constraints lead to lower price volatility, and to a



(a) Phase plot showing a quasi-periodic attractor.

(b) Time series. *Top panel:* prices. *Middle panel:* fraction of fundamentalists (blue), optimists (green, dotted) and pessimists (red, dashed). *Bottom panel:* positions of fundamentalists (blue), optimists (green, dotted) and pessimists (red, dashed).Figure 5.6: Phase plot (a) and time series (b) for the benchmark ABS with fundamentalists versus opposite biases without constraints. The baseline parametrisation and  $T = 0$ ,  $I^{max} = \infty$ .

stable equilibrium price if  $T$  is larger than 0.06.

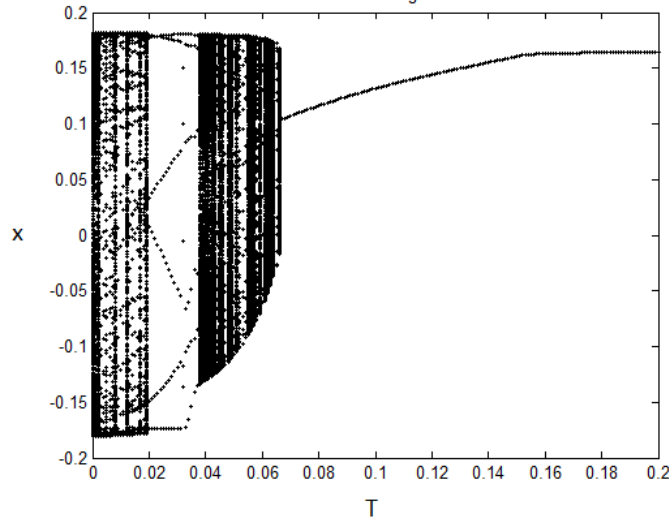
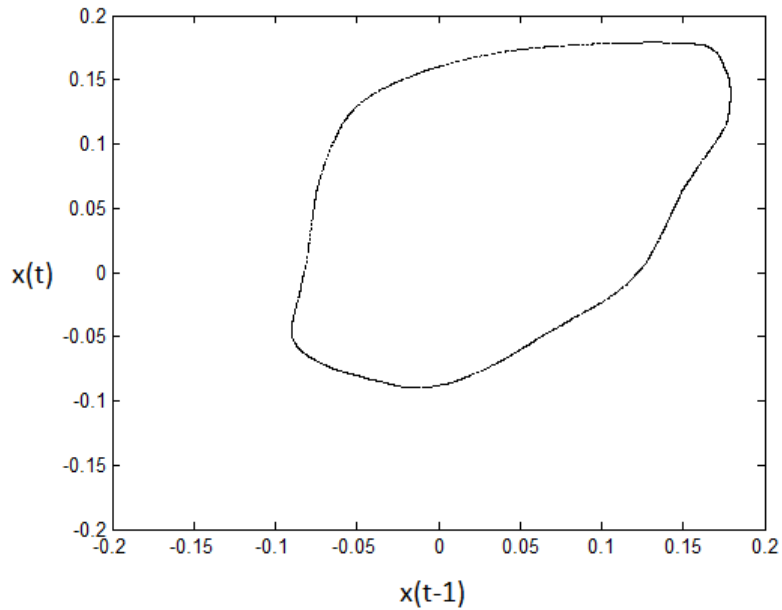


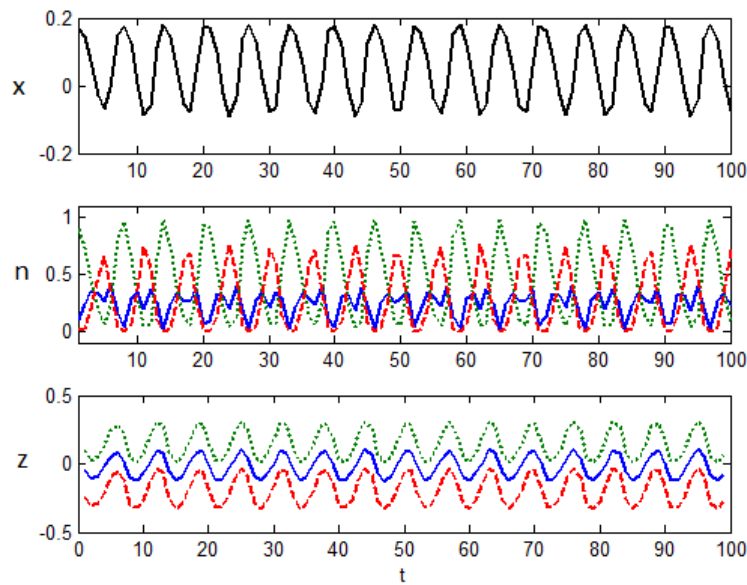
Figure 5.7: Bifurcation diagram for short-selling tax  $T$  given the baseline parametrisation and  $I^{max} = \infty$ .

The stabilising effect of strict short-selling restrictions under fixed expectation rules can be understood intuitively as follows. Consider first a relatively small short-selling tax of  $T = 0.055$  as in Figure 5.8. The positions of agents (depicted in the lower panel Figure 5.8b) remain almost unchanged, but the trading costs  $R\tau(z_{h,t})$  for short-selling reduce the fractions of pessimists at the price troughs, which are lower than the fractions of optimists at the price tops (see the middle panel). This in turn brings the troughs closer to the fundamental value. As the minimum price during the cycle is shifted upwards, the amplitude below the fundamental decreases.

For a larger tax  $T > 0.06$ , the fraction of pessimists is further limited, and prices do not follow a cycle any longer, but converge to a steady state value  $x = \frac{1}{R}(n_2 - n_3)b > 0$ . At this price the demand of fundamentalists and optimists clears the market  $n_1 z_1 + n_2 z_2 = z_s$  with  $z_1 \approx 0$  because fundamentalists are also short-selling constrained. The difference in performance between fundamentalists and optimists, who demand  $z_2 > 0$ , determine the fractions  $n_1$  and  $n_2$ , depending on the values of  $z_s$  and  $\beta$ . In either case there are very few pessimists ( $n_3 \ll n_1, n_2$ ). The steady state in deviation from the fundamental value is positive. Tight short-selling constraint drive out the most pessimistic traders, increasing asset prices.



(a) Phase plot showing a quasi-periodic attractor.

(b) Time series. *Top panel:* prices. *Middle panel:* fraction of fundamentalists (blue), optimists (green, dotted) and pessimists (red, dashed). *Bottom panel:* positions of fundamentalists (blue), optimists (green, dotted) and pessimists (red, dashed).Figure 5.8: Phase plot (a) and time series (b) for the ABS with fundamentalists versus opposite biases and a short-selling tax. The baseline parametrisation and  $T = 0.055$ ,  $I^{max} = \infty$ .

### Leverage constraints in the three-type market

Next, we investigate the effect of leverage constraints in the three-type ABS. The parameter choices in Definition 2 require  $I^{max} > 0.1$  to have an unconstrained fundamental value as in Assumption 5.1. The bifurcation diagram of the price dynamics (see Figure 5.9) show an opposite effect to short-selling constraints. For a leverage constraint  $I^{max}$  lower than 2.5, the maximum price during the cycle is shifted downwards and the amplitude above the fundamental decreases. Leverage constraints lead to a stable steady state below the fundamental value if they are sufficiently strict: in this case if  $I^{max}$  is smaller than 0.5.

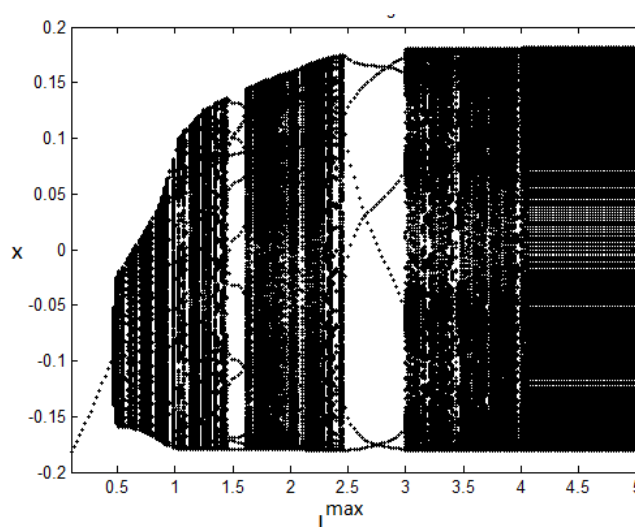


Figure 5.9: Bifurcation diagram for leverage constraint  $I^{max}$  given the baseline parametrisation and  $T = 0$ .

The intuition why leverage constraints stabilise the dynamics is closely related to the discussion of short-selling policies. For a sufficiently strict constraint on buying assets, the optimistic and fundamental type have a demand equal to the maximal investment level, which they maintain under small deviations in the price. The market price is only influenced by the pessimistic agent type that is able to adjust his investment. The most optimistic types are affected most by leverage constraints and are therefore dominated by more pessimistic traders, leading to lower prices.

Figure 5.10 shows the attractor and dynamics for a maximal investment level of  $I^{max} = 0.75$ .

Optimists are constrained at the maximal investment level (see the bottom panel of Figure 5.10) and therefore the maximum level along the price fluctuations decreases. Also fundamentalists can be seen to be leverage constrained for part of the cycle. Agents with the pessimistic belief are never restricted in their demand and drive down prices during the cycle.

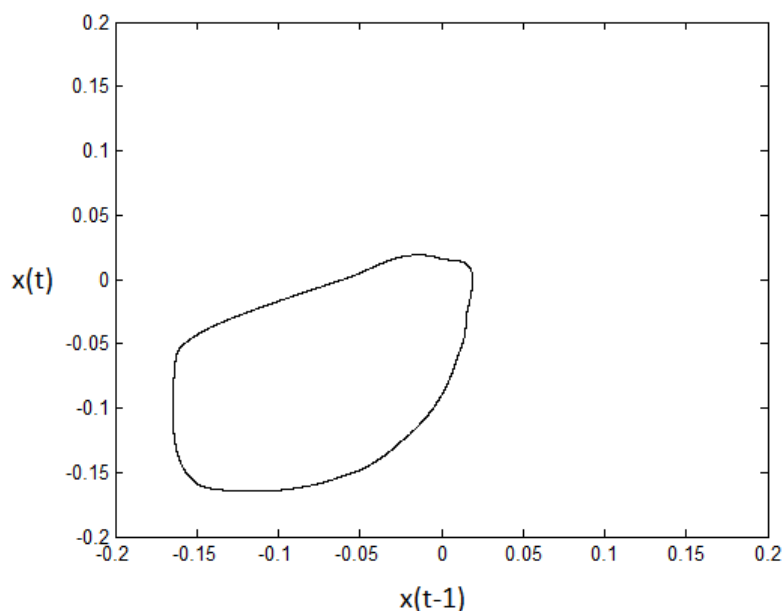
We have seen that when agents can choose from three fixed belief types, demand constraints reduce the amplitude of cycles and thus the price volatility. A sufficiently tight constraint assures that either optimistic or pessimistic beliefs have limited impact on the market price. For short-selling constraints, the impact of the pessimist belief is limited by the trading costs making it less attractive to traders. Under leverage constraints, optimists (and fundamentalists as well) can take only limited positions. One important difference with the market inhabited by fundamentalists and chartists is that the remaining agent type(s) cannot generate an almost self-fulfilling bubble when beliefs are fixed and bounded.

### **Combining the two demand constraints in the three-type market**

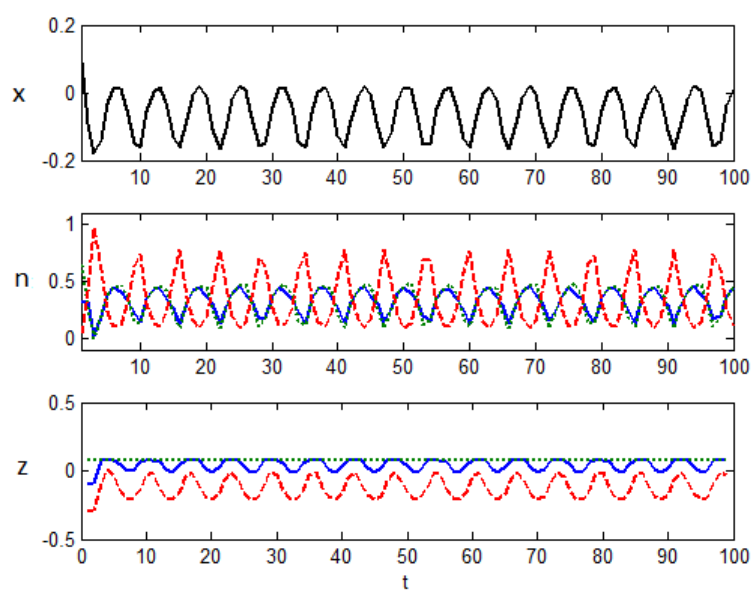
Although demand constraints can drive down price volatility in the three-type ABS, mispricing remains when a single constraint is imposed. The reason is that although the fundamentalist belief will attract more followers in general, one type of biased traders will survive when only one constraint is imposed. With only short-selling constraints prices converge to a level above the fundamental value, and with only leverage constraints under the fundamental value. A natural next step is therefore to investigate combined effect of short-selling and leverage constraints.

In Figure 5.11a, the basin of attraction of a stable steady state is presented given the baseline parametrisation and different values of  $T$  and  $I^{max}$ . Figure 5.11b shows a similar basin of attraction with a higher intensity of choice,  $\beta = 500$ , for reasons explained below. Each small square indicates parameter values for which the dynamics converge to a stable steady state. As expected, for loose constraints in the left upper corner of the plot the dynamics do not converge, but lead to (quasi-)periodic cycles as in Figures 5.6, 5.8 and 5.10. Imposing a single constraint by increasing  $T$  (with  $I^{max} = \infty$ ) or lowering  $I^{max}$  (with  $T = 0$ ) leads to convergence to

a non-fundamental steady state This steady state lies above the fundamental for the short-selling constraint and below the fundamental for the leverage constraint.



(a) Phase plot showing a quasi-periodic attractor.



(b) Time series. *Top panel:* prices. *Middle panel:* fraction of fundamentalists (blue), optimists (green, dotted) and pessimists (red, dashed). *Bottom panel:* positions of fundamentalists (blue), optimists (green, dotted) and pessimists (red, dashed).

Figure 5.10: Phase plot (a) and time series (b) for the ABS with fundamentalists versus opposite biases and a leverage constraint. The baseline parametrisation and  $T = 0$ ,  $I^{max} = 0.75$ .



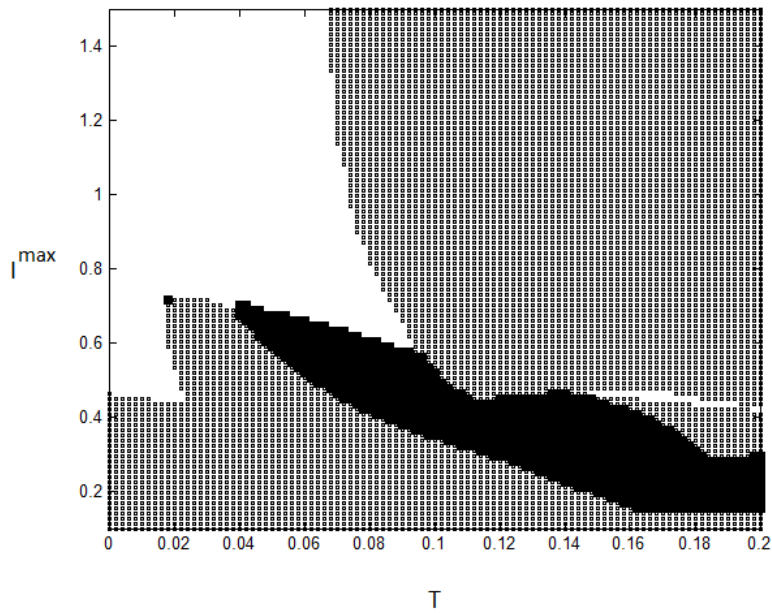
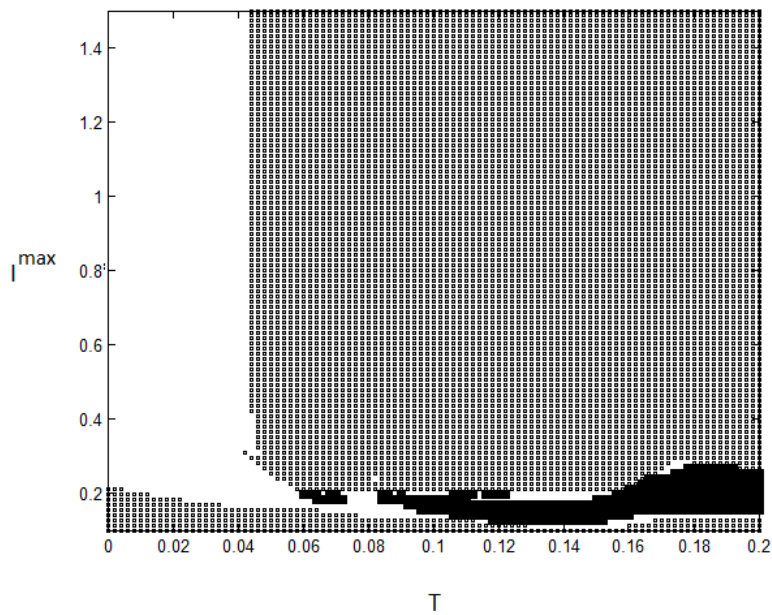
(a) The baseline intensity of choice:  $\beta = 100$ .(b) A higher intensity of choice:  $\beta = 500$ .

Figure 5.11: Basin of attraction of a stable steady state for different values of  $T$  and  $I^{\max}$  given the baseline parametrisation with  $\beta = 100$  (a) and  $\beta = 500$  (b). Small squares indicate parameter values for which the dynamics has a stable steady state; in the filled regions the price converges to a steady state value  $x^*$  with  $|x^*| < 0.01$  close to the fundamental solution.

To stabilise the dynamics close to the fundamental value, both constraints should be combined. The filled regions in Figure 5.11a show combinations of  $T$  and  $I^{max}$  that stabilise prices to a steady state value  $x^*$  with  $|x| < 0.01$ , close to the fundamental. A combination of a high short-selling tax and a tight leverage constraint (i.e. low  $I^{max}$ ) stabilises asset prices close to the fundamental value.

The possibility to stabilise the market close the fundamental price depends, however, on the relative low value chosen for the intensity of choice  $\beta$ . In Figure 5.11b, for which  $\beta$  has been increased form 100 to 500, the filled region of a steady state values close to the fundamental value has shifted down and is much smaller. The following proposition makes clear that no steady state with  $x^* = 0$  exists for any choice of the other parameters if  $\beta = \infty$ .

**Proposition 5.5.** *Consider the ABS with fundamentalists, optimists and pessimists under possible demand constraints. Assume  $z_s > 0$ ,  $b > 0$  and  $\beta = \infty$ . Then price dynamics are bounded by  $[-b/R, b/R]$  and price volatility is bounded by  $\sigma \leq b/R$ . If a steady state  $E = (x^*, n_1^*, n_2^*, n_3^*)$  exists, it must have a non-fundamental price  $x^* > 0$  for all  $T \geq 0$  and  $I^{max} > p_f z_s$ .*

**Proof.** See Appendix 5.A □

Proposition 5.5 shows that demand constraints weakly decrease price volatility below the level  $\sigma = b/R$  that arises in the unconstrained 4-cycle for  $b > \frac{r_f}{R} a \sigma^2 z_s$ . On the other hand, mispricing cannot be completely removed by demand constraints if investors are highly sensitive to realised profits and switch rapidly between beliefs. If constraints lead to a steady state, the price must be above the fundamental value.

## 5.4 Discussion

The above analysis suggests that demand constraints do not necessarily stabilise asset markets with heterogeneous agents and may have adverse effects. In this section we discuss policy implications of these results for the U.S. stock market. As a starting point for policy intervention, policy makers must have some estimate of a fundamental stock price. In reality, there is a lot

of debate about the underlying fundamental value of stock prices, and it is not self-evident that policy makers have a fully satisfactory understanding about underlying fundamentals. Nevertheless, policy makers could have concerns about large mispricing and market volatility even if the fundamental is not fully known to them, and could consider imposing demand constraints.<sup>9</sup>

Figure 5.12 plots the S&P500 price-dividend (PD) ratios from 1995Q1 to 2012Q4, as well as the average PD ratio over this period equal to 57.7. As an example, the fundamental value based on the consumption-habit model of Campbell and Cochrane (1999) is added (the dashed line). This model has been named one of the most successful rational-agent approaches in explaining asset price behaviour by the Nobel Prize Committee (2013) in their motivation of the 2013 Nobel Prize Laureates.<sup>10</sup>

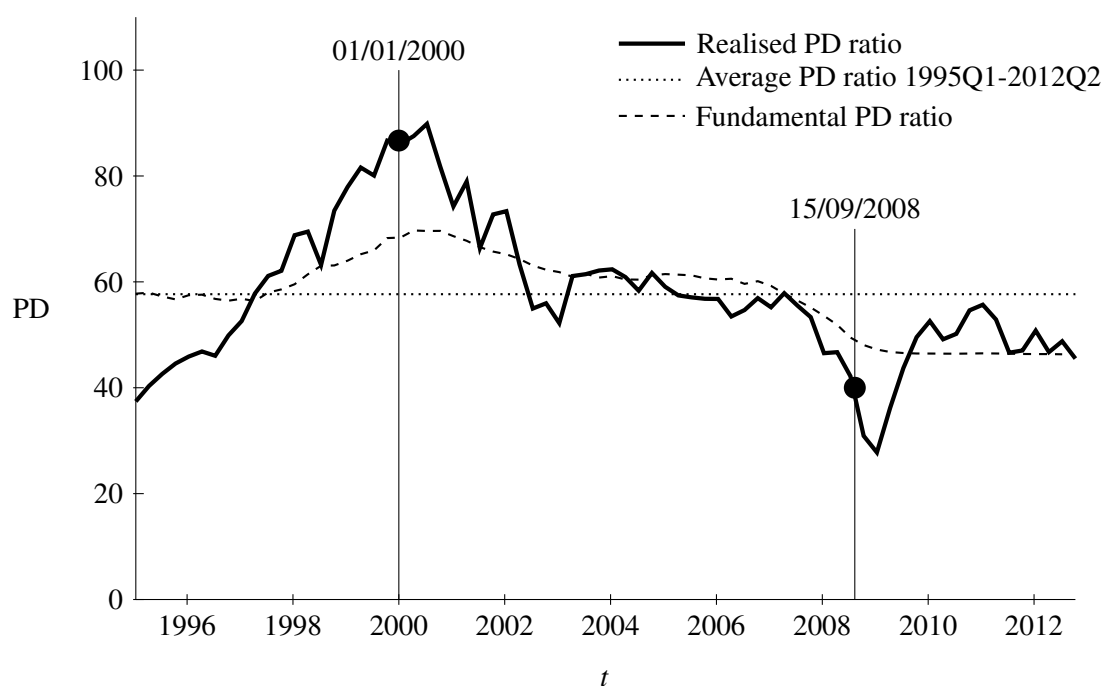


Figure 5.12: Realised PD ratio of the S&P500, its average and the fundamental value based on the Campbell-Cochrane consumption-habit model, 1995Q1-2012Q2.

<sup>9</sup> Naturally, as for any model, one has to take caution in applying the theoretical results in this chapter directly to economic reality. Besides the fundamental value, another crucial question is what kind of rules actual investors use to forecast prices. See Hommes (2011) for a survey on asset market experiments with human subjects, which provides evidence for the use of simple rules in making investment decisions similar as in our stylised model.

<sup>10</sup> See Chapter 4 for the details of the calculation of the Campbell-Cochrane fundamental value and other choices for the fundamental value such as the Gordon model.

Although many other possible fundamental values can be chosen, a robust finding in the literature is *excess volatility*, as found in the seminal paper of Shiller (1981) and subsequently confirmed by many others. Excess volatility means that prices fluctuate much more than can be justified by fundamentals. The Campbell-Cochrane fundamental value plotted in Figure 5.12 reproduces some general patterns in asset prices, but does not fully explain the asset price movements. The most notable deviations are around 2000, the top of the dot-com bubble, and after the bankruptcy of Lehman Brothers on 15 September 2008, the height of the financial crisis.

The PD ratios of the S&P500 show the typical behaviour generated by the stylised heterogeneous agents models in this chapter: recurrent bubbles and crashes and oscillating fluctuations. In Chapter 4 we have estimated a two-type heterogeneous agents model – closely related to model with fundamentalists and chartists in this chapter – on the S&P500 PD ratios in deviations from the Campbell-Cochrane fundamental. There it was found that the switching of the agents between the rules has significant effects on asset prices and strongly amplifies booms and busts. Here, we focus on the effects of demand constraints in the U.S. stock market, taking the heterogeneity of agents as given.

During the financial crisis in 2008 there have been recurrent calls for bans on short-selling, and on 19 September 2008 the Securities and Exchange Commission (SEC) indeed prohibited short-selling for 299 financial stocks. The motivation was to prevent investors from betting on a fall in stock prices, which would supposedly “restore equilibrium” to the market. On 2 October 2008 the short-selling bans ended (see Anufriev and Tuinstra, 2013, for an account of the major events from July to September 2008).

The present analysis of short-selling constraints suggests that the SEC ban is unlikely to have helped stabilising financial markets, as prices seem to have been below the fundamental value in 2008. In the two-type market with fundamentalists and chartists, prices have an endogenous lower bound of prices below the fundamental that depends on many parameters, but not on the short-selling tax  $T$ . In the three-type market with fundamentalists, optimists and chartists, price

volatility can be reduced by short-selling constraints, but steady states (if they occur) lie above the fundamental. So in both examples of adaptive belief systems in the asset pricing model with heterogeneous agents, short-selling constraints can *not* restore the fundamental equilibrium.

At the same time, if the price was indeed below the fundamental in the crisis, the short-selling ban most likely has not increased mispricing either. Such an adverse effect of short-selling constraints can happen in markets with fundamentalists and chartists, but only above the fundamental value. It is therefore consistent with our stylised heterogeneous agent model that the short-selling ban of the SEC had a negligible effect. This is supported by the observation that stock prices were falling before, during and after the short-sell ban in 2008. Our model suggests that the observed limited effect of the SEC ban in stabilising the financial markets can be expected to hold for demand constraints under fairly general conditions.<sup>11</sup>

Despite the potential adverse effects of demand constraints in general, leverage constraints can reduce mispricing and price volatility when prices are above the fundamental. This was for example the case during the dot-com bubble in 2000 (see Figure 5.12). Overpricing of a financial asset can occur when the natural buyers with the most optimistic beliefs are not counterbalanced by natural sellers and dominate the market. If these dominating natural buyers become constrained in borrowing money to buy stocks, prices decrease. The heterogeneous agents model thus suggests that only if policy makers are confident that asset prices are too high compared to fundamentals, they should consider constraining leverage ratios in order to deflate financial bubbles like the dot-com bubble.

---

<sup>11</sup> It is not clear what the effect of leverage constraints is on prices below the fundamental value. In the two-type market, price volatility increases, but in the three-type market price volatility decreases.

## Appendix 5.A Proofs

### *Proof of Proposition 5.1.*

Individual market demand given by (5.9) is continuous and decreasing. Aggregate demand  $\sum_{h=1}^H n_{h,t} z_{h,t}(p_t)$  is also continuous and decreasing. It will be *strictly* decreasing as long as aggregate demand is strictly positive. Conversely, if and only if for some price  $p'_t$  every agent type  $h$  demands  $z_{h,t}(p'_t) = 0$  there exists an interval of prices  $I_t \subset \mathbb{R}^+$  (with  $p'_t \in I_t$ ) such that  $\sum_{h=1}^H n_{h,t} z_{h,t}(p_t)$  is constant (and equal to 0) for all  $p_t \in I_t$ . As market clearing by (5.10) requires  $z_{h,t}(p_t) > 0$  for some  $h \in H$  if  $z_s > 0$ , a unique equilibrium price  $p_t$  exists.  $\square$

### *Proof of Proposition 5.2.*

From the definition of multinomial logit fractions (5.15) and  $\beta = \infty$ , it follows that at every  $t$  the market is dominated by either fundamentalists ( $n_{1,t} = 1$ ) or chartists ( $n_{1,t} = 0$ ). As soon as  $n_{1,t} = 1$ , we get  $x_{t+1} = 0$ . In this fundamental state, both agent types believe  $E_{h,t+j}[x_{t+j+1}] = 0$  and demand  $z_{h,t+j} = z_s$ , but fundamentalist pay additional costs  $C > 0$ . As this implies  $U_{1,t} < U_{2,t}$  and thus  $n_{1,t+1} = 0$ , it follows that  $E_1 = (0, 0)$  is a steady state.

Consider an initial state  $n_{1,t} = 0$  for  $t = 0, 1$  and  $x_0 = \varepsilon$  with  $\varepsilon \neq 0$  small. The price in period  $t = 1$  is determined by clearing of the market dominated by chartists, i.e. the price that makes sure that chartists buy all outside supply:

$$z_{2,t}(x_t) = z_s. \quad (5.33)$$

We get  $x_1 = \varepsilon(g/R)$ , and  $x_t = \varepsilon(g/R)^t$  as long as  $n_{1,t} = 0$ . For  $0 < g < R$ , prices converge to  $E_1 = (0, 0)$ , and  $E$  must be globally stable.

For  $g > R$ , the fundamental steady state  $E_1 = (0, 0)$  is locally unstable. Chartists dominate the market for subsequent periods ( $n_{1,t} = 0$  for  $t = 1, 2, \dots$ ) if their performance exceeds the

performance of fundamentalists:

$$\begin{aligned}
 n_{1,t+1} = 0 &\Leftrightarrow \\
 U_{1,t} &< U_{2,t} \\
 \rho_t z_{1,t-1} - C &< \rho_t z_{2,t-1}.
 \end{aligned} \tag{5.34}$$

and, using the definition of  $\rho_t$  (5.18) and that  $z_{2,t-1} = z_s$ ,

$$(x_t - Rx_{t-1} + a\sigma^2 z_s)(z_{1,t-1}(x_{t-1}) - z_s) - C < 0. \tag{5.35}$$

To find the magnitude of prices bubbles, we solve for the price  $x_t$  at which fundamentalists and chartists have an equal performance, with  $x_{t-1} = (R/g)x_t$  and  $x_{t-2} = (R/g)^2 x_t$ .

$$\begin{aligned}
 (x_t - Rx_{t-1} + a\sigma^2 z_s)\left(-\frac{R}{a\sigma^2}x_{1,t-1} + z_s - z_s\right) - C &= 0 \\
 (x_t - (R^2/g)x_t + a\sigma^2 z_s)\left(-\frac{R^2}{a\sigma^2 g}x_t\right) - C &= 0 \\
 (R^2/g - 1)x_t^2 - a\sigma^2 z_s x_t - a\sigma^2 (g/R^2)C &= 0.
 \end{aligned} \tag{5.36}$$

If it exists, the solution to this quadratic equation is

$$x_t = x^{b\pm} \equiv \frac{a\sigma^2 z_s \pm \sqrt{(a\sigma^2 z_s)^2 + 4(1 - g/R^2)a\sigma^2 C}}{2(R^2/g - 1)}. \tag{5.37}$$

For all  $0 < x_t < x^{b+}$ , the bubble continues for at least one more period in  $t + 1$ . The bubble ends in period  $t + 1$  if  $x_{t+1} = (g/R)x_t > x^{b+}$ . The upper bound on prices is therefore

$$x^{max} = \frac{g}{R}x^{b+} = \frac{g}{R} \frac{a\sigma^2 z_s + \sqrt{(a\sigma^2 z_s)^2 + 4(1 - g/R^2)a\sigma^2 C}}{2(R^2/g - 1)}. \tag{5.38}$$

Similarly the lower bound on prices is

$$x^{min} = \frac{g}{R}x^{b-} = \frac{g}{R} \frac{a\sigma^2 z_s - \sqrt{(a\sigma^2 z_s)^2 + 4(1 - g/R^2)a\sigma^2 C}}{2(R^2/g - 1)}. \quad (5.39)$$

Finally, if no solution to (5.36) exists, the performance of chartists exceeds the performance of fundamentalists for all  $t$ , and the dynamics diverge with rate  $g/R$ . Notice that the first term in (5.36),  $(R^2/g - 1)x_t^2$ , is negative if  $g > R^2$ . Therefore no positive solution to this equation exists, proving that dynamics with positive initial prices  $x_0 = \varepsilon > 0$  diverge if  $g > R^2$ . For a negative initial state  $x_0 = \varepsilon < 0$ , the existence of  $x^{b\pm}$  requires

$$\begin{aligned} (a\sigma^2 z_s)^2 + 4(1 - g/R^2)a\sigma^2 C &> 0 \\ R^2(1 + \frac{a\sigma^2 z_s^2}{4C}) &> g, \end{aligned} \quad (5.40)$$

completing the proof. □

***Proof of Proposition 5.3.***

Observe that demand constraints cannot have an effect on the steady state  $E_1 = (0, 0)$ . This follows because the beliefs for both types coincide with the belief of a rational representative agent ( $E_{1,t}[x_{t+1}] = E_{2,t}[x_{t+1}] = 0$  for  $x_t = 0$ ), and the rational representative agent is unconstrained by Assumption 5.1. Also, given any  $T$  and  $I^{max} > p_f z_s$ , both types are unconstrained for  $x_t = \varepsilon$  with a sufficiently small  $\varepsilon \neq 0$ . The steady state  $E_1 = (0, 0)$  is therefore locally stable if and only if  $g < R$  (as in Proposition 5.2) with or without demand constraints.

Consider an initial state  $n_{1,t} = 0$  for  $t = 0, 1$  and  $x_0 = \varepsilon$  with  $\varepsilon \neq 0$  small. We get  $z_{1,1} = z_s$ . Chartists dominate the market for subsequent periods ( $n_{1,t} = 0$  for  $t = 1, 2, \dots$ ) if their perfor-



mance exceeds the performance of fundamentalists:

$$\begin{aligned}
n_{1,t+1} = 0 &\Leftrightarrow \\
U_{1,t} &< U_{2,t} \\
\rho_t z_{1,t-1} - R\tau(z_{1,t-1}) - C &< \rho_t z_{2,t-1} - R\tau(z_{2,t-1}) \\
(x_t - Rx_{t-1} + a\sigma^2 z_s)(z_{1,t-1}(x_{t-1}) - z_s) - R\tau(z_{1,t-1}) - C &< 0, \tag{5.41}
\end{aligned}$$

using the fact that chartists never pay the short-selling tax, as  $z_{2,t-1} = z_s > 0$ .

Consider first the case that chartists are not leverage constrained. Then we get  $x_t = \varepsilon(g/R)^t$  as long as  $n_{2,t} = 1$ , irrespective of possible demand constraints operative on fundamentalists. By the assumption in the proposition that price dynamics are bounded, there must exist a  $x_t$  with  $x_{t-1} = (R/g)x_t$  and  $x_{t-2} = (R/g)^2 x_t$  at which fundamentalists and chartists have an equal performance:

$$\begin{aligned}
0 &= (x_t - Rx_{t-1} + a\sigma^2 z_s)(z_{1,t-1}(x_{t-1}) - z_s) - R\tau(z_{1,t-1}(x_{t-1})) - C \\
0 &= (x_t - \frac{R^2}{g}x_t + a\sigma^2 z_s)(z_{1,t-1}(\frac{R}{g}x_t) - z_s) - R\tau(z_{1,t-1}(\frac{R}{g}x_t)) - C, \tag{5.42}
\end{aligned}$$

which implicitly defines two solutions  $x^{b\pm}$ . For all  $x^{b-} < x_t < x^{b+}$ , we have  $n_{1,t+1} = 0$  and the bubble continues for at least one more period in  $t+1$ . The bubble ends in period  $t+1$  if  $x_{t+1} = (g/R)x_t \notin [x^{b-}, x^{b+}]$ . The upper and lower bounds on prices are therefore  $x^{max} = (g/R)x^{b+}$  and  $x^{min} = (g/R)x^{b-}$ .

Three effects of demand constraints follow from condition (5.42). First, at price  $x_t < 0$  below the fundamental value, fundamentalists buy  $z_{1,t}(x_t) > 0$ , so fundamentalists pay no short-selling tax. An increase in  $T$  has thus no effect on  $x^{min}$ .

Second, the short-selling tax  $T$  has an effect on  $x^{max}$  through two channels. Notice that for  $x_t > 0$  fundamentalists have a strictly lower demand  $z_{1,t} < z_{2,t}$ , although  $z_{1,t}$  can be both positive and negative depending on the parameters. The first channel runs through the short-selling tax,

which weakly increases the costs for fundamentalists to short-sell in the rising market, and reduces the performance of fundamentalists. The second channel runs through demand. A tighter short-selling constraint weakly increases fundamentalists' demand  $z_{1,t}$  and reduces the difference  $z_{1,t-1}(x_{1,t-1}) - z_s$  in (5.42). Fundamentalists require a larger positive excess return  $\rho_t$  and thus a more positive  $x_t$  to overcome costs. An increase in  $T$  strictly decreases  $x^{max}$  if  $z_{1,t}(x^{max}) < 0$ , and in general weakly decreases  $x^{max}$ .

Third, a decrease in  $I^{max}$  can have a decreasing effect on  $x^{min}$  for the following reason. Notice that for  $x_t < 0$  fundamentalists have a strictly higher demand  $z_{1,t} > z_{2,t}$ . A tighter leverage constraint weakly decreases fundamentalists' demand  $z_{1,t}$ , and as  $z_{1,t} > z_s$  it reduces the difference  $z_{1,t-1}(x_{1,t-1}) - z_s$  in (5.42). Fundamentalists require a larger positive excess return  $\rho_t$  and thus a more negative  $x_t$  to overcome the information costs  $C$ . The lower bound on prices  $x_{min}$  is weakly decreasing if  $I^{max}$  is decreased.

Finally, consider the case that chartists are leverage constrained in buying  $z_s > 0$ . The leverage constraint implies

$$z_s = z_{2,t} \leq \frac{I^{max}}{p_f + x_t} \quad (5.43)$$

$$\Rightarrow x_t \leq x^{L+} \equiv \frac{I^{max}}{z_s} - p_f, \quad (5.44)$$

which imposes a *price ceiling*  $x^{L+} > 0$  on the price dynamics. Consider that at  $t - 1$  chartists are leverage constrained leading to  $x_{t-1} = x^{L+}$ . Then the condition (5.41) determines whether in period  $t$  chartists dominate the market, i.e.  $n_{1,t} = 0$ . If  $n_{1,t} = 0$  then  $x_t = x^{L+}$ , and it is possible to end up in the steady state  $E_2 = (x^{L+}, 0)$ . For  $E_2 = (x^{L+}, 0)$  to be a steady state, the performance of chartists should be higher for  $x_t = x_{t-1} = x^{L+}$

$$(x^{L+} - Rx^{L+} + a\sigma^2 z_s)(z_{1,t-1}(x^{L+}) - z_s) - R\tau(z_{1,t-1}(x^{L+})) - C < 0. \quad (5.45)$$

Condition (5.45) holds for  $x^{L+} = 0$  (but not for  $x^{L+} = \infty$ ) so it must be fulfilled for sufficiently low  $x^{L+}$ . As  $x^{L+} = I^{max}/z_s - p_f$ , this implies that  $E_2 = (x^{L+}, 0)$  is a steady state for  $I^{max}$  close

enough to  $p_f z_s$ , i.e. for a sufficiently low  $I^{max}$ . □

***Proof of Proposition 5.4.***

In the absence of demand constraints, the demand for each type is given by

$$z_{1,t} = \frac{1}{a\sigma^2}(-Rx_t) + z_s, \quad (5.46)$$

$$z_{2,t} = \frac{1}{a\sigma^2}(b - Rx_t) + z_s, \quad (5.47)$$

$$z_{3,t} = \frac{1}{a\sigma^2}(-b - Rx_t) + z_s. \quad (5.48)$$

In every period  $t$  we have  $z_2 > z_1 > z_3$ . Realised profits are  $U_{h,t} = \rho_t z_{h,t}$ . This implies either  $U_{2,t} > U_{1,t} > U_{3,t}$  or  $U_{3,t} > U_{1,t} > U_{2,t}$  for all  $\rho_t \neq 0$ . If  $\rho_t = 0$  for  $t = 0$ , we have  $U_{h,0} = 0$  for all  $h$  leading to  $n_{h,1} = 1/3$  and  $x_1 = 0$ . However, the fundamental state cannot be a steady state, as  $\rho_t > 0$  for  $x_t = x_{t-1} = 0$ . So for  $\beta = \infty$  the market is always dominated at  $t = 1$  by either optimists or pessimists.

Consider an initial state  $n_{3,0} = 1$  and  $x_0 = -b/R$  with  $\rho_{t=0} < 0$ . Then in period  $t = 1$  pessimists dominate the market, implying  $n_{3,1} = 1$  and  $x_1 = -b/R$ . This pessimistic state can not be a steady state, as excess returns at  $t = 1$ , because

$$\rho_1 = -b/R + b + a\sigma^2 z_s > 0, \quad (5.49)$$

are strictly positive, implying  $U_{2,1} > U_{1,1} > U_{3,1}$ . This leads to an optimistic state  $n_{2,2} = 1$  and  $x_2 = b/R$ . The excess returns at  $t = 2$  are

$$\rho_2 = b/R + b + a\sigma^2 z_s > 0, \quad (5.50)$$

and strictly positive, maintaining the optimistic state  $n_{2,3} = 1$  and  $x_3 = b/R$ . At  $t = 3$  the excess

returns are

$$\rho_3 = b/R - b + a\sigma^2 z_s. \quad (5.51)$$

As this expression can be either positive or negative, two possibilities arise for the dynamics depending on  $b$ . For low biases  $0 < b < \frac{r_f}{R} a\sigma^2 z_s$ , excess returns remain positive in the optimistic state, and  $E = (b/R, 0, 1, 0)$  is the globally stable steady state.

For large biases supply  $b > \frac{r_f}{R} a\sigma^2 z_s$ , excess returns  $\rho_3$  are strictly negative, returning the price to the negative state  $n_{3,4} = 1$  and  $x_4 = -b/R$ . Under this low level of outside supply,

$$\rho_4 = -b/R - b + a\sigma^2 z_s < 0, \quad (5.52)$$

and  $t = 4$  we are in back in the initial pessimistic state started in  $t = 0$ . The result is a stable 4-cycle with prices  $\frac{1}{R}\{b, b, -b, -b\}$ . The average price is  $\bar{x} = 0$  and the price volatility (the square root of the variance) along the cycle is  $b/R$ .  $\square$

***Proof of Proposition 5.5.***

Under general demand constraints, we have  $z_2 \geq z_1 \geq z_3$ , where the inequality is strict unless some types are both leverage constrained at  $z_{h,t} = I^{max}/(p_f + x_t)$ , or short-selling constrained at  $z_{h,t} = 0$ . As aggregate demand is decreasing in the price, market clearing implies that prices are maximal when optimists dominate the market ( $z_{2,t}(x_t) = z_s$ ), and minimal when pessimists dominate the market ( $z_{3,t}(x_t) = z_s$ ). As demand constraints pushes large investment positions in the direction of  $z_{h,t} = 0$ , price dynamics must be bounded by the interval  $[-b/R, b/R]$ .

From these price bounds it is straightforward to derive the maximal price volatility  $\sigma$ . By definition the price variance is given by

$$\sigma^2 = \int_{-b/R}^{b/R} (x_t - \bar{x})^2 f(x_t) dx_t, \quad (5.53)$$

where  $\bar{x}$  is the average price and  $f(x_t)$  is the probability density function for the price  $x_t$ . Because of convexity of the function  $(x_t - \bar{x})^2$ , the price variance is maximised when there are only two price states at the bounds

$$f(x_t) = \begin{cases} \frac{1}{2} & \text{if } x_t = -b/R \\ 0 & \text{if } x_t \in (-b/R, b/R) \\ \frac{1}{2} & \text{if } x_t = b/R, \end{cases} \quad (5.54)$$

which gives  $\sigma^2 = (b/R)^2$  and price volatility  $\sigma = b/R$ . For any other distribution of prices  $f(x_t)$ , price volatility is strictly lower  $\sigma < b/R$ .

Now consider a possible steady state  $E = (x^*, n_1^*, n_2^*, n_3^*)$ . It will be shown that  $x^* > 0$  must hold using a proof by contradiction. Assume there is a steady state with  $x^* \leq 0$ . In this steady state, excess returns are strictly positive:  $\rho^* = x^* - Rx^* + a\sigma^2 z_s = -r_f x^* + a\sigma^2 z_s > 0$ . Because  $z_2 \geq z_1 \geq z_3$ , optimists have the highest performance  $U_2 \geq U_1 \geq U_3$ . The steady state has three possible distributions of agents:  $n_2^* = 1$ ,  $n_2^* = n_1^* = \frac{1}{2}$  or  $n_2^* = n_1^* = n_3^* = \frac{1}{3}$ . If optimists coexist with other types ( $n_2^* = n_1^* = \frac{1}{2}$  or  $n_2^* = n_1^* = n_3^* = \frac{1}{3}$ ), optimists must have equal performance, and also equal demand to the other existing types:  $z_2 = z_1$  or  $z_2 = z_1 = z_3$ . In all three cases, market clearing implies  $z_2(x^*) = z_s$ . However, Assumption 5.1 assures that  $z_2(0) > z_s$  by bounding leverage constraints  $I^{max}$  from below, and demand  $z_2(x_t)$  is decreasing in  $x_t$ . This implies  $x^* > 0$ , contradicting the assumption of a steady state price equal to or below the fundamental price.

□



# Chapter 6

## Summary

This thesis investigates complex systems in financial economics and applies the theory of complex systems to interbank and stock markets. The questions that this thesis intends to answer originate mainly from the financial crisis of 2007-2008. In retrospect, the single most important event of this crisis was the bankruptcy of the bank Lehman Brothers on 15 September 2008. The central underlying question in this thesis is how the relatively minor problems in the U.S. mortgage market could grow out to an endogenously emerging and global financial-economic crisis.

Economists have found it difficult to find a satisfactory explanation for the emergence of the crisis, because, in essence, many traditional economic models relate to simple systems. Large groups of decision-making agents such as consumers, producers or investors are preferably forced into one single rational representative agent. Even if the model allows for explicit interaction between agents, the focus is often on a small number of homogeneous agents. In contrast, complex systems typically exhibit strong interaction at the micro level that can induce large changes at the macro level – so-called critical transitions. The far-reaching events during the financial crisis bear close resemblance to these critical transitions. This thesis therefore models financial markets as complex systems.

Models of complex systems are not by definition exceptionally complicated. On the contrary: the complex systems models in this thesis evolve around a relatively small number of

key economic concepts. A small number of economic ideas, such as trading relationships between banks and market sentiment of heterogeneous investors, will be used to investigate the emergence of the financial turmoil.

A first methodological step in applying the theory of complex systems is to establish empirical relevance. Even as many existing economic frameworks turned out to be of limited use to policy makers during the financial crisis, the burden of proof lies on alternative systems to improve empirical performance. This thesis establishes the empirical relevance for two applications of complex systems: firstly, it presents evidence for a non-random network structure in the interbank market, and secondly, it finds distinct temporal regimes in asset price dynamics that may shift to one another through a critical transition. The empirical relevance of complex systems leads to new questions for financial market regulation which are examined subsequently.

The thesis contains four research chapters organised in two parts, preceded by a general introduction. The first part is concerned with the interbank market, and the second with the stock market – markets that have had an decidedly amplifying effect on the financial crisis. As such, the analysis helps to understand the mechanisms behind the emergence of the crisis.

Part I (chapters 2 and 3) considers interbank markets, taking particular notice of the interbank market in the Netherlands. Mainly through its tight connections to American banks suffering from the subprime mortgage crisis, the open Dutch banking system came under serious pressure. In October 2008, ABN AMRO bank was nationalised by the Dutch government, and also ING bank received a large capital injection. An important motivation for these government interventions was reduce the risk of further financial contagion.

Chapter 2 analyses data on interbank exposures in the Netherlands for the period from 1998 up to 2008. The focus is on the *core periphery structure*, which is found to give a better fit to the data than other network models proposed in the literature. The good fit of the core periphery structure suggests the presence of a stable set of densely connected core banks that intermediate between periphery banks. Monitoring core banks, which are systemically more important in passing through shocks, can help reduce systemic risk.



---

Chapter 3 aims at giving an explanation for the core periphery structure in a network formation model of a financial market with an explicit role for intermediation by middlemen. In a core periphery network, core banks receive intermediation benefits for supporting indirect connections between periphery banks. The core periphery network is found to be strategically unstable if agents are homogeneous: there is always at least one agent who has an incentive to change its links. After allowing for heterogeneity among agents, a core periphery network can arise endogenously, with the big banks forming the core. This is an interesting finding especially in combination with the estimated core periphery network in the Netherlands. The Dutch core banks have a higher average size than periphery banks, even though some unexpected medium-sized banks are part of the core as well.

Part II of the thesis (Chapters 4 and 5) is concerned with stock markets. In Chapter 4, a complex systems model for asset pricing is estimated using U.S. stock prices from 1950 up to 2012. Market sentiment is modelled by *behavioural heterogeneity*: traders in the stock market are boundedly rational and disagree about how quickly stock prices return to the fundamental value. Fundamentalist traders believe in mean-reversion, while chartists extrapolate trends. Investors gradually switch between the two rules, based upon their relative performance, leading to self-reinforcing regimes of mean-reversion and trend-following. The estimated behavioural heterogeneity is both statistically and economically significant. The analysis in this chapter suggests that market sentiment switches between different behavioural regimes, which amplify fundamental shocks such as technology shocks and the subsequent dot-com bubble, and the Lehman bankruptcy in 2008 and the associated financial crisis.

Chapter 5 investigates the consequences of bounded rationality and heterogeneous beliefs for financial market regulations constraining large investment positions. It has been noted that the crisis in stock markets occurred at a time when *leverage* ratios had increased to enormous amounts and when simultaneously investors engaged massively in *short-selling*. This chapter shows that leverage and short-selling constraints can have large adverse effects on financial stability if investors have heterogeneous beliefs about future prices. The intuition for these

adverse effects is that different beliefs themselves can be stabilising or destabilising. If the market regulations (unintentionally) restrict the stabilising fundamentalist strategy, mispricing and price volatility may in fact increase.

The applications in this thesis relate to the debate on policy measures aiming to prevent a future financial crisis of the kind witnessed in 2007 and 2008. For the relatively unexplored interbank market, policy measures aiming to create a more robust structure require knowledge of why individual optimisation of banks leads to the observed network structure. Market regulation for stock markets is already much more common. However, in studying the effects of different measures policy makers should realise that the price formation on a financial market is part of a complex system in which investors make use of simple, boundedly rational strategies. Policy analysis based on realistic models of complex systems can contribute to an improved understanding and a better regulation of the financial economy.

# Bibliography

- ACEMOGLU, D., A. OZDAGLAR, AND A. TAHBAZ-SALEHI (2013): “Systemic risk and stability in financial networks,” NBER Working Paper 18727, National Bureau of Economic Research.
- ADAM, K., J. BEUTEL, AND A. MARCET (2013): “Stock price booms and expected capital gains,” mimeo, University of Mannheim.
- ADAM, K., AND A. MARCET (2011): “Internal rationality, imperfect market knowledge and asset prices,” *Journal of Economic Theory*, 146(3), 1224–1252.
- AFONSO, G., AND R. LAGOS (2012): “An empirical study of trade dynamics in the Fed funds market,” FRB of New York staff report 550.
- AKRAM, Q. F., AND C. CHRISTOPHERSEN (2010): “Interbank overnight interest rates – gains from systemic importance,” Norges Bank Working Paper 2010-11.
- ALFARANO, S., T. LUX, AND F. WAGNER (2005): “Estimation of agent-based models: The case of an asymmetric herding model,” *Computational Economics*, 26(1), 19–49.
- ALLEN, F., AND D. GALE (2000): “Financial contagion,” *Journal of Political Economy*, 108(1), 1–33.
- AMINI, H., R. CONT, AND A. MINCA (2012): “Stress testing the resilience of financial networks,” *International Journal of Theoretical and Applied Finance*, 15(01).
- ANUFRIEV, M., AND G. BOTTAZZI (2006): “Price and wealth dynamics in a speculative market with generic procedurally rational traders,” Warwick Business School Working Paper 06-02.
- ANUFRIEV, M., AND C. HOMMES (2012): “Evolution of market heuristics,” *The Knowledge Engineering Review*, 27(2), 255–271.
- ANUFRIEV, M., AND J. TUINSTRA (2013): “The impact of short-selling constraints on financial market stability in a heterogeneous agents model,” *Journal of Economic Dynamics and Control*, 37(8), 1523–1543.
- AOKI, M. (2002): “Open models of share markets with two dominant types of participants,” *Journal of Economic Behavior & Organization*, 49(2), 199–216.
- BABUS, A. (2011): “Endogenous intermediation in over-the-counter markets,” mimeo, Imperial College London.

## BIBLIOGRAPHY

---

- BANSAL, R., AND A. YARON (2004): “Risks for the long run: A potential resolution of asset pricing puzzles,” *The Journal of Finance*, 59(4), 1481–1509.
- BARABÁSI, A.-L., AND R. ALBERT (1999): “Emergence of scaling in random networks,” *Science*, 286, 509–512.
- BARBERIS, N., AND R. THALER (2003): “A survey of behavioral finance,” in *Handbook of the Economics of Finance*, ed. by G. Constantinides, M. Harris, and R. M. Stulz, vol. 1, pp. 1053–1128. North-Holland: Elsevier.
- BATTISTON, S., D. DELLI GATTI, M. GALLEGATI, B. GREENWALD, AND J. E. STIGLITZ (2012): “Liaisons dangereuses: Increasing connectivity, risk sharing, and systemic risk,” *Journal of Economic Dynamics and Control*, 36(8), 1121–1141.
- BEDAYO, M., A. MAULEON, AND V. VANNETELBOSCH (2013): “Bargaining and delay in trading networks,” FEEM Working Paper 752, Fondazione Eni Enrico Mattei.
- BERNANKE, B. S. (2004): “The Great Moderation,” Speech delivered at the meetings of the Eastern Economic Association, Washington, DC (February).
- BHATTACHARYA, S., AND D. GALE (1987): “Preference shocks, liquidity, and central bank policy,” in *New Approaches to Monetary Theory*, Cambridge University Press, ed. by W. Singleton, and J. Barnett, pp. 69–88. New York.
- BLANCHARD, O. J. (2008): “The crisis: Basic mechanisms, and appropriate policies,” MIT Working Paper 09-01, Massachusetts Institute of Technology.
- BLUME, L. E., D. EASLEY, J. KLEINBERG, AND E. TARDOS (2009): “Trading networks with price-setting agents,” *Games and Economic Behavior*, 67(1), 36–50.
- BOLLERSLEV, T. (1986): “Generalized autoregressive conditional heteroskedasticity,” *Journal of Econometrics*, 31(3), 307–327.
- BORGATTI, S. P., AND M. G. EVERETT (1999): “Models of core/periphery structures,” *Social Networks*, 21, 375–395.
- BOSS, M., H. ELSINGER, M. SUMMER, AND S. THURNER (2004): “The network topology of the interbank market,” *Quantitative Finance*, 4(6), 677–684.
- BOSWIJK, H. P., C. H. HOMMES, AND S. MANZAN (2007): “Behavioral heterogeneity in stock prices,” *Journal of Economic Dynamics and Control*, 31(6), 1938–1970.
- BOYD, J., W. FITZGERALD, AND R. BECK (2006): “Computing core/periphery structures and permutation tests for social relations data,” *Social Networks*, 28(2), 165–178.
- BRANCH, W., AND G. EVANS (2010): “Asset return dynamics and learning,” *Review of Financial Studies*, 23(4), 1651–1680.
- BRÄUNING, F., AND F. FECHT (2012): “Relationship lending in the interbank market and the price of liquidity,” Deutsche Bundesbank Discussion Paper 2012-22.

- BROCK, W. A., AND C. H. HOMMES (1997): "A rational route to randomness," *Econometrica*, 65(5), 1059–1096.
- (1998): "Heterogeneous beliefs and routes to chaos in a simple asset pricing model," *Journal of Economic Dynamics and Control*, 22(8-9), 1235–1274.
- BROCK, W. A., C. H. HOMMES, AND F. O. O. WAGENER (2009): "More hedging instruments may destabilize markets," *Journal of Economic Dynamics and Control*, 33(11), 1912–1928.
- BRUNNERMEIER, M., AND M. OEHMKE (2012): "Bubbles, financial crises, and systemic risk," NBER Working Paper 18398, National Bureau of Economic Research.
- BRUNNERMEIER, M. K. (2009): "Deciphering the liquidity and credit crunch 2007-2008," *Journal of Economic Perspectives*, 23(1), 77–100.
- BUSKENS, V., AND A. VAN DE RIJT (2008): "Strategic network formation with structural holes," *American Journal of Sociology*, 114, 37–407.
- BUSS, A., B. DUMAS, R. UPPAL, AND G. VILKOV (2013): "The equilibrium dynamics of liquidity and illiquid asset prices," mimeo, Banque de France.
- CAMPBELL, J. Y., AND J. H. COCHRANE (1999): "By force of habit: A consumption-based explanation of aggregate stock market behavior," *The Journal of Political Economy*, 107(2), 205–251.
- CAMPBELL, J. Y., AND R. J. SHILLER (1988a): "The dividend-price ratio and expectations of future dividends and discount factors," *Review of Financial Studies*, 1(3), 195–228.
- (1988b): "Stock prices, earnings, and expected dividends," *Journal of Finance*, 43(3), 661–76.
- CAMPBELL, J. Y., AND R. J. SHILLER (2001): "Valuation ratios and the long-run stock market outlook: An update," NBER Working Paper 8221, National Bureau of Economic Research, Inc.
- CASTIGLIONESI, F. (2007): "Financial contagion and the role of the central bank," *Journal of Banking & Finance*, 31(1), 81–101.
- CHAN, S. (2010): "Is Ben Bernanke having fun yet?," *New York Times*, (15 May).
- CHEN, X., AND S. C. LUDVIGSON (2009): "Land of addicts? An empirical investigation of habit-based asset pricing models," *Journal of Applied Econometrics*, 24(7), 1057–1093.
- CHENG, S.-H., C.-L. CHANG, AND Y.-R. DU (2012): "Agent-based economic models and econometrics," *The Knowledge Engineering Review*, 27(2), 187–219.
- CHEVALIER, J., AND G. ELLISON (1997): "Risk taking by mutual funds as a response to incentives," *Journal of Political Economy*, 105(6), 1167–1200.
- CHIARELLA, C., AND X.-Z. HE (2001): "Asset price and wealth dynamics under heterogeneous expectations," *Quantitative Finance*, 1(5), 509–526.

## BIBLIOGRAPHY

---

- CHIARELLA, C., X.-Z. HE, AND R. C. ZWINKELS (2014): “Heterogeneous expectations in asset pricing: Empirical evidence from the S&P500,” *Journal of Economic Behavior and Organization*, (forthcoming).
- CHINAZZI, M., AND G. FAGIOLO (2013): “Systemic risk, contagion, and financial networks: A survey,” LEM Papers 2013-08, Laboratory of Economics and Management, Sant’Anna School of Advanced Studies, Pisa.
- CHOI, S., A. GALEOTTI, AND S. GOYAL (2013): “Trading in networks: Theory and experiment,” mimeo.
- COCCO, J. F., F. J. GOMES, AND N. C. MARTINS (2009): “Lending relationships in the interbank market,” *Journal of Financial Intermediation*, 18(1), 24–48.
- COCHRANE, J. H. (2001): *Asset Pricing*. Princeton University Press.
- (2011): “Presidential address: Discount rates,” *The Journal of Finance*, 66(4), 1047–1108.
- COHEN-COLE, E., E. PATACCHINI, AND Y. ZENOU (2011): “Systemic risk and network formation in the interbank market,” CEPR discussion paper 8332, Centre for Economic Policy Research.
- CONT, R., A. MOUSA, AND E. BASTOS E SANTOS (2013): “Network structure and systemic risk in banking systems,” in *Handbook of Systemic Risk*, ed. by J. Fouque, and J. Langsam. Cambridge University Press.
- COROMINAS-BOSCH, M. (2004): “Bargaining in a network of buyers and sellers,” *Journal of Economic Theory*, 115(1), 35–77.
- CRAIG, B., AND G. VON PETER (2014): “Interbank tiering and money center banks,” *Journal of Financial Intermediation*, (forthcoming).
- DE JONG, E., W. VERSCHOOR, AND R. ZWINKELS (2010): “Heterogeneity of agents and exchange rate dynamics: Evidence from the EMS,” *Journal of International Money and Finance*, 29(8), 1652–1669.
- DE NOOY, W., A. MRVAR, AND V. BATAGELJ (2011): *Exploratory Social Network Analysis with Pajek*. Cambridge University Press, 2 edn.
- DEGRYSE, H., AND G. NGUYEN (2007): “Interbank exposures: An empirical examination of contagion risk in the Belgian banking system,” *International Journal of Central Banking*, 3(2), 123–171.
- DEL GUERCIO, D., AND P. A. TKAC (2002): “The determinants of the flow of funds of managed portfolios: Mutual funds vs. pension funds,” *Journal of Financial and Quantitative Analysis*, 37(4), 523–558.

- DELONG, J., A. SHLEIFER, L. SUMMERS, AND R. WALDMANN (1990a): "Positive feedback investment strategies and destabilizing rational speculation," *Journal of Finance*, 45(2), 379–395.
- (1990b): "Noise trader risk in financial markets," *Journal of Political Economy*, 98(4), 703–738.
- ERDŐS, P., AND A. RÉNYI (1959): "On random graphs," *Publicationes Mathematicae*, 6, 290–297.
- FRANKEL, J. A., AND K. A. FROOT (1987): "Using survey data to test standard propositions regarding exchange rate expectations," *American Economic Review*, 77(1), 133–53.
- FREIXAS, X., B. PARIGI, AND J. C. ROCHET (2000): "Systemic risk, interbank relations and liquidity provision by the central bank," *Journal of Money, Credit and Banking*, 3(3), 611–638.
- FRICKE, D., AND T. LUX (2012): "Core-periphery structure in the overnight money market: Evidence from the e-MID trading platform," Kiel Working Paper 1759.
- FRIJNS, B., T. LEHNERT, AND R. C. ZWINKELS (2010): "Behavioral heterogeneity in the option market," *Journal of Economic Dynamics and Control*, 34(11), 2273–2287.
- FURFINE, C. H. (2003): "Interbank exposures: Quantifying the risk of contagion," *Journal of Money, Credit and Banking*, 35(1), 111–128.
- GAI, P., A. HALDANE, AND S. KAPADIA (2011): "Complexity, concentration and contagion," *Journal of Monetary Economics*, 58(5), 453–470.
- GAI, P., AND S. KAPADIA (2010): "Contagion in financial networks," Bank of England Working Paper 383.
- GALE, D. M., AND S. KARIV (2007): "Financial networks," *American Economic Review*, 97(2), 99–103.
- GALEOTTI, A., AND S. GOYAL (2010): "The law of the few," *American Economic Review*, 100(4), 1468–92.
- GALEOTTI, A., S. GOYAL, AND J. KAMPHORST (2006): "Network formation with heterogeneous players," *Games and Economic Behavior*, 54(2), 353–372.
- GAUNERSDORFER, A. (2000): "Endogenous fluctuations in a simple asset pricing model with heterogeneous agents," *Journal of Economic Dynamics and Control*, 24(5), 799–831.
- GEANAKOPOLOS, J. (2010): "The leverage cycle," in *NBER Macroeconomics Annual 2009, Volume 24*, ed. by D. Acemoglu, K. Rogoff, and M. Woodford, pp. 1–65. University of Chicago Press.
- GEORG, C.-P. (2013): "The effect of the interbank network structure on contagion and common shocks," *Journal of Banking & Finance*, 37(7), 2216–2228.

## BIBLIOGRAPHY

---

- GOFMAN, M. (2011): “A network-based analysis of over-the-counter markets,” Chicago Meetings Paper, pp. 1–71.
- GOH, K.-I., B. KAHNG, AND D. KIM (2001): “Universal behavior of load distribution in scale-free networks,” *Physical Review Letters*, 87(27), 278701.
- GORDON, M. J. (1962): *The Investment, Financing, and Valuation of the Corporation*. Irwin.
- GOYAL, S. (2009): *Connections: An Introduction to the Economics of Networks*. Princeton University Press.
- GOYAL, S., AND F. VEGA-REDONDO (2007): “Structural holes in social networks,” *Journal of Economic Theory*, 137(1), 460–492.
- GREENWOOD, R., AND A. SHLEIFER (2013): “Expectations of returns and expected returns,” NBER Working Paper 18686, National Bureau of Economic Research.
- GROMB, D., AND D. VAYANOS (2002): “Equilibrium and welfare in markets with financially constrained arbitrageurs,” *Journal of Financial Economics*, 66(2), 361–407.
- HALDANE, A. G. (2009): “Rethinking the financial network,” Speech delivered at the Financial Student Association in Amsterdam (April).
- HALDANE, A. G., AND V. MADOUROS (2012): “The dog and the frisbee,” Speech delivered at Jackson Hole conference (August).
- HARRISON, J. M., AND D. M. KREPS (1978): “Speculative investor behavior in a stock market with heterogeneous expectations,” *The Quarterly Journal of Economics*, 92(2), 323–336.
- HAYEK, F. A. (1964): “The theory of complex phenomena,” in *The Critical Approach to Science and Philosophy*, pp. 332–349. London: McMillan.
- HEEMEIJER, P., C. HOMMES, J. SONNEMANS, AND J. TUINSTRA (2009): “Price stability and volatility in markets with positive and negative expectations feedback: An experimental investigation,” *Journal of Economic Dynamics and Control*, 33(5), 1052–1072.
- HILLER, T. (2012): “Peer effects in endogenous networks,” Bristol Economics Working Paper 12-633.
- HILSCHER, J., AND M. WILSON (2011): “Credit ratings and credit risk,” Brandeis University Working Paper 31.
- HIRSHLEIFER, D. (2001): “Investor psychology and asset pricing,” *The Journal of Finance*, 56(4), 1533–1597.
- HOMMES, C. (2006): “Heterogeneous agent models in economics and finance,” in *Handbook of Computational Economics*, ed. by L. Tesfatsion, and K. L. Judd, vol. 2, pp. 1109–1186. North-Holland: Elsevier.
- (2011): “The heterogeneous expectations hypothesis: Some evidence from the lab,” *Journal of Economic Dynamics and Control*, 35(1), 1–24.



- (2013): *Behavioral Rationality and Heterogeneous Expectations in Complex Economic Systems*. Cambridge University Press.
- HOMMES, C., 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.
- HOMMES, C., J. SONNEMANS, J. TUINSTRA, AND H. VAN DE VELDEN (2008): “Expectations and bubbles in asset pricing experiments,” *Journal of Economic Behavior & Organization*, 67(1), 116–133.
- HONG, H., AND J. C. STEIN (1999): “A unified theory of underreaction, momentum trading, and overreaction in asset markets,” *The Journal of Finance*, 54(6), 2143–2184.
- HOUY, N. (2009): “Structural holes in social networks: A remark,” *Journal of Economic Theory*, 144(1), 422–431.
- IMAKUBO, K., AND Y. SOEJIMA (2010): “The transaction network in Japan’s interbank money markets,” *Monetary and Economic Studies*, 28, 107–150.
- INTERNATIONAL MONETARY FUND (2007): “Global Financial Stability Report,” (October).
- IPPOLITO, R. A. (1989): “Efficiency with costly information: A study of mutual fund performance, 1965–1984,” *The Quarterly Journal of Economics*, 104(1), 1–23.
- JACKSON, M., AND A. WOLINSKY (1996): “A strategic model of social and economic networks,” *Journal of Economic Theory*, 71(1), 44–74.
- JACKSON, M. O. (2005): “A Survey of Network Formation Models: Stability and Efficiency,” in *Group Formation in Economics: Networks, Clubs and Coalitions*, ed. by G. Demange, and M. Wooders, pp. 11–57. Cambridge University Press.
- JACKSON, M. O. (2008): *Social and Economic Networks*. Princeton University Press.
- JACKSON, M. O., AND A. WATTS (2002): “The evolution of social and economic networks,” *Journal of Economic Theory*, 106(2), 265–295.
- KARCESKI, J. (2002): “Returns-chasing behavior, mutual funds, and beta’s death,” *Journal of Financial and Quantitative Analysis*, 37(4), 559–594.
- KEYNES, J. M. (1936): *The General Theory of Interest, Employment and Money*. London: Macmillan.
- KIRMAN, A. (2010): *Complex Economics: Individual and Collective Rationality*. London: Routledge.
- KIRMAN, A., AND G. TEYSSIÈRE (2002): “Microeconomic models for long memory in the volatility of financial time series,” *Studies in Nonlinear Dynamics & Econometrics*, 5(4), 3.

## BIBLIOGRAPHY

---

- KLEINBERG, J., S. SURI, E. TARDOS, AND T. WEXLER (2008): “Strategic network formation with structural holes,” *Proceedings 9th ACM Conference on Electronic Commerce*, pp. 480–489.
- KÖNIG, M. D., C. J. TESSONE, AND Y. ZENOU (2014): “Nestedness in networks: A theoretical model and some applications,” *Economic Theory*, (forthcoming).
- KRANTON, R. E., AND D. F. MINEHART (2001): “A theory of buyer-seller networks,” *American Economic Review*, 91(3), 485–508.
- KUBELEC, C., AND F. SÁ (2010): “The geographical composition of national external balance sheets: 1980-2005,” Bank of England Working Paper 384.
- LANGFIELD, S., Z. LIU, AND T. OTA (2012): “Mapping the UK interbank system,” mimeo.
- LEBARON, B. (2006): “Agent-based computational finance,” in *Handbook of Computational Economics*, ed. by L. Tesfatsion, and K. L. Judd, vol. 2, pp. 1187–1233. North-Holland: Elsevier.
- LIEDORP, F., L. MEDEMA, M. KOETTER, R. H. KONING, AND I. VAN LELYVELD (2010): “Peer monitoring or contagion? Interbank market exposure and bank risk,” DNB Working Paper 248.
- LOF, M. (2012): “Heterogeneity in stock prices: A STAR model with multivariate transition function,” *Journal of Economic Dynamics and Control*, 36(12), 1845–1854.
- (2014): “Rational speculators, contrarians and excess volatility,” *Management Science*, (forthcoming).
- LONGSTAFF, F. A. (2008): “Flight-from-leverage in distressed financial markets,” mimeo, UCLA Anderson School of Management, University of California, Los Angeles.
- LUX, T. (2009): “Stochastic behavioral asset pricing models and the stylized facts,” in *Handbook of Financial Markets: Dynamics and Evolution*, ed. by T. Hens IV, and K. R. Schenk-Hoppé, vol. 1, p. 161. North-Holland Amsterdam.
- MANNILA, H., AND E. TERZI (2007): “Nestedness and segmented nestedness,” *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 480–489.
- MARTÍNEZ-JARAMILLO, S., B. ALEXANDROVA-KABADJOVA, B. BRAVO-BENÍTEZ, AND J. P. SOLÓRZANO-MARGAIN (2014): “An empirical study of the Mexican banking system’s network and its implications for systemic risk,” *Journal of Economic Dynamics and Control*, 40, 242–265.
- MERLO, A., AND C. A. WILSON (1995): “A stochastic model of sequential bargaining with complete information,” *Econometrica*, 63(2), 371–99.
- MILLER, E. M. (1977): “Risk, uncertainty, and divergence of opinion,” *The Journal of Finance*, 32(4), 1151–1168.

- MINSKY, H. P. (1986): *Stabilizing an Unstable Economy*. Yale University Press.
- MISTRULLI, P. E. (2011): “Assessing financial contagion in the interbank market: Maximum entropy versus observed interbank lending patterns,” *Journal of Banking & Finance*, 35(5), 1114–1127.
- NAKAMOTO, M., AND D. WIGHTON (2007): “Citigroup chief stays bullish on buyouts,” *Financial Times*, (9 July).
- NEWMAN, M. E. J. (2010): *Networks: An Introduction*. Oxford University Press.
- NIER, E., J. YANG, T. YORULMAZER, AND A. ALENTORN (2007): “Network models and financial stability,” *Journal of Economic Dynamics and Control*, 31(6), 2033–2060.
- NOBEL PRIZE COMMITTEE (2013): “Understanding asset prices,” Scientific background on the Sveriges Riksbank prize in economic sciences in memory of Alfred Nobel.
- OFEK, E., AND M. RICHARDSON (2003): “Dotcom mania: The rise and fall of internet stock prices,” *The Journal of Finance*, 58(3), 1113–1138.
- PÁSTOR, L., AND R. F. STAMBAUGH (2001): “The equity premium and structural breaks,” *The Journal of Finance*, 56(4), 1207–1239.
- PÁSTOR, L., AND P. VERONESI (2006): “Was there a Nasdaq bubble in the late 1990s?,” *Journal of Financial Economics*, 81(1), 61–100.
- PERSITZ, D. (2009): “Power and core-periphery networks,” FEEM Working Paper 42, Fondazione Eni Enrico Mattei.
- POTERBA, J. M., AND L. H. SUMMERS (1988): “Mean reversion in stock prices: Evidence and implications,” *Journal of Financial Economics*, 22(1), 27–59.
- ROCHET, J. C., AND J. TIROLE (1996): “Interbank lending and systemic risk,” *Journal of Money, Credit, and Banking*, 28(4), 733–762.
- SAMUELSON, P. (1947): *Foundations of Economic Analysis*. Harvard University Press.
- SHILLER, R. J. (1981): “Do stock prices move too much to be justified by subsequent changes in dividends?,” *American Economic Review*, 71(3), 421–36.
- (1987): “Investor behavior in the October 1987 stock market crash: Survey evidence,” NBER Working Paper 2446, National Bureau of Economic Research.
- (2000): “Measuring bubble expectations and investor confidence,” *The Journal of Psychology and Financial Markets*, 1(1), 49–60.
- (2005): *Irrational Exuberance*. Princeton University Press, 2 edn.
- SIEDLAREK, J.-P. (2012a): “Intermediation in networks,” FEEM Working Paper 693, Fondazione Eni Enrico Mattei.

## BIBLIOGRAPHY

---

- (2012b): “Making friends meet: A model of mediated network formation,” mimeo, European University Florence.
- SIMON, H. A. (1962): “The architecture of complexity,” *Proceedings of the American Philosophical Society*, 106(6), 467–482.
- SIMSEK, A. (2013): “Belief disagreements and collateral constraints,” *Econometrica*, 81(1), 1–53.
- SIRRI, E. R., AND P. TUFANO (1998): “Costly search and mutual fund flows,” *The Journal of Finance*, 53(5), 1589–1622.
- SOLIS-MONTES, M. P. (2013): “The structure of the Mexican interbank market,” mimeo, Banco de Mexico.
- SORAMÄKI, K., M. L. BECH, J. ARNOLD, R. J. GLASS, AND W. E. BEYELER (2007): “The topology of interbank payment flows,” *Physica A*, 379, 317–333.
- SQUARTINI, T., AND D. GARLASCHELLI (2011): “Analytical maximum-likelihood method to detect patterns in real networks,” *New Journal of Physics*, 13(8), 083001.
- SQUARTINI, T., I. VAN LELYVELD, AND D. GARLASCHELLI (2013): “Early-warning signals of topological collapse in interbank networks,” *Scientific Reports*, 3, 3357.
- STUMPF, P. H., AND PORTER M. A (2012): “Critical truths about power laws,” *Science*, 335, 665–666.
- TER ELLEN, S., AND R. C. ZWINKELS (2010): “Oil price dynamics: A behavioral finance approach with heterogeneous agents,” *Energy Economics*, 32(6), 1427–1434.
- TERÄSVIRTA, T. (1994): “Specification, estimation and evaluation of smooth transition autoregressive models,” *Journal of the American Statistical Association*, 89(425), 208–218.
- THURNER, S., J. D. FARMER, AND J. GEANAKOPOLOS (2012): “Leverage causes fat tails and clustered volatility,” *Quantitative Finance*, 12(5), 695–707.
- TIROLE, J. (2011): “Illiquidity and all its friends,” *Journal of Economic Literature*, 49(2), 287–325.
- TRICHET, J.-C. (2010): “Reflections on the nature of monetary policy non-standard measures and finance theory,” Speech at the ECB Central Banking Conference, Frankfurt (November).
- UPPER, C. (2011): “Simulation methods to assess the danger of contagion in interbank markets,” *Journal of Financial Stability*, 7(3), 111–125.
- UPPER, C., AND A. WORMS (2004): “Estimating bilateral exposures in the German interbank market: Is there a danger of contagion?,” *European Economic Review*, 48(4), 827–849.
- VAN LELYVELD, I., AND F. LIEDORP (2006): “Interbank contagion in the Dutch banking sector: A sensitivity analysis,” *International Journal of Central Banking*, 31, 99–133.

- VAN LELYVELD, I., F. LIEDORP, AND M. KAMPMAN (2011): “An empirical assessment of reinsurance risk,” *Journal of Financial Stability*, 7(4), 191–203.
- VISSING-JORGENSEN, A. (2004): “Perspectives on behavioral finance: Does “irrationality” disappear with wealth? Evidence from wxpectations and actions,” in *NBER Macroeconomics Annual 2003, Volume 18*, pp. 139–208. The MIT Press.
- WEAVER, W. (1948): “Science and complexity,” *American Scientist*, 36(4), 536–544.
- WESTERHOFF, F. H. (2008): “The use of agent-based financial market models to test the effectiveness of regulatory policies,” *Jahrbucher fur Nationalokonomie und Statistik*, 228(2), 195.
- WESTERHOFF, F. H., AND S. REITZ (2003): “Nonlinearities and cyclical behavior: The role of chartists and fundamentalists,” *Studies in Nonlinear Dynamics & Econometrics*, 7(4).
- XIE, Y.-B., T. ZHOU, AND B.-H. WANG (2008): “Scale-free networks without growth,” *Physica A*, 387(7), 1683–1688.
- YELLEN, J. L. (2013): “Interconnectedness and systemic risk: Lessons from the financial crisis and policy implications remarks,” Remarks at the American Economic Association/American Finance Association Joint Luncheon San Diego, California (January).
- ZHONG, W., AND J. LIU (2012): “Comments on “Scale-free networks without growth”,” *Physica A*, 391, 263–265.



# Краткое содержание (Summary in Russian)

Данная докторская диссертация исследует сложные системы (complex systems) в финансовой экономике и применяет теорию сложных систем к межбанковскому и фондовому рынкам. Вопросы, на которые мы пытаемся найти ответы в данной диссертации, вызваны, во многом, финансовым кризисом 2007-2008 годов. В ретроспективе самым важным событием этого кризиса стало банкротство банка Lehman Brothers 15 сентября 2008 года. Центральный вопрос, лежащий в основе диссертации, таков: как относительно второстепенные проблемы ипотечного рынка США переросли во внутрисистемный, глобальный финансово-экономический кризис.

Экономисты столкнулись со сложностями в поисках удовлетворительного объяснения, откуда появился экономический кризис и как он развивался и распространялся. Это связано с тем, что традиционные экономические модели относятся, в сущности, к простым системам. В таких системах экономические решения о потреблении, производстве и инвестировании, принимаемые в реальности большой группой агентов, моделируются как рациональные решения одного, “репрезентативного” агента. Даже если модель позволяет агентам взаимодействовать, непосредственно влияя на решения друг друга, внимание уделяется лишь небольшому числу однородных агентов. Сложные системы, в свою очередь, наоборот сосредоточены на сильных взаимодействиях на микроуровне, последствием которых могут стать так называемые

критические переходы (**critical transitions**), то есть несоизмеримо большие изменения на макроуровне.

Модели сложных систем не обязательно сложные по концепции. Наоборот, модели в данной диссертации строятся на относительно малом количестве исходных экономических идей. С помощью таких ключевых понятий, как кредитно-заемные отношения между банками и «настроение» фондового рынка неоднородных инвесторов, исследуется распространение финансовых потрясений.

Первый методологический шаг в применении теории сложных систем – это установка эмпирической значимости данной теории. Хотя кризис показал несостоятельность многих существующих экономических моделей, это еще не означает, что альтернативные модели смогут предложить эмпирически лучшие прогнозы и рекомендации людям, принимающим решения. Данная диссертация устанавливает эмпирическую значимость двух приложений теории сложных систем. Во-первых, в ней, в результате исследования структуры связей на межбанковском рынке, показано, что сеть межбанковских взаимодействий не описывается моделью случайного графа, но согласуется с моделью сложных сетей. Во-вторых, исследование динамики фондового рынка выявляет четко различающиеся временные состояния динамики цен на акции, которые посредством критических переходов могут внезапно сменять друг друга. Эмпирическая значимость сложных систем ведёт к новым вопросам регулирования финансовых рынков, которые также рассматриваются в данной диссертации.

Диссертация состоит из четырех глав и разделена на две части, которым предшествует общее введение. В первой части рассматривается межбанковский рынок, а во второй – фондовый рынок, так что обе части относятся к определенным рынкам, которые очевидно создавали усиливающий эффект на финансовый кризис. Таким образом, анализ помогает лучше понять механизмы возникновения и распространения кризиса.

Часть I (главы 2 и 3) охватывает межбанковские рынки и в частности рассмат-



ривает межбанковский рынок Нидерландов. С 2008 года банковский сектор Нидерландов столкнулся с трудностями, в основном из-за тесных связей с американскими банками, которые несли большие убытки в результате ипотечного кризиса. Такие крупные игроки на рынке, как ABN AMRO и ING, были спасены правительством Нидерландов, чтобы как можно сильнее уменьшить финансовые потери.

Глава 2 анализирует данные о кредитах между нидерландскими банками в период 1998-2008 гг. Акцент лежит на центрo-периферической структуре (*core periphery structure*), которая, как показывает анализ, описывает данные лучше, чем другие известные модели. Центрo-периферическая структура предполагает, что существует стабильное ядро тесно связанных между собой банков и что эти банки являются посредниками между остальными, периферийными банками, если те хотят торговать на рынке. Как следствие, мониторинг банков ядра, играющих ключевую роль в передаче шоков другим банкам, может помочь в решении задачи по снижению систематических рисков.

Цель главы 3 – поиск экономических причин для возникновения центрo-периферической структуры через построение модели формирования сети на финансовом рынке с акцентом на межбанковское посредничество. В центрo-периферической модели банки ядра получают финансовое вознаграждение за роль посредников для создания косвенных связей между периферийными банками. Если банки однородные, то центрo-периферическая сеть является стратегически нестабильной: это означает, что всегда есть хотя бы один банк, у которого есть стимул изменить свои связи. Однако в том случае, если банки неоднородные, возникает эндогенная стабильная центрo-периферическая сеть, в которых большие банки формируют центральное ядро. Эта теоретическая модель довольно хорошо описывает центрo-периферическую сеть банков Нидерландов. Ключевые банки Нидерландов, принадлежащие к ядру, в среднем больше, чем периферийные банки, хотя некоторые банки среднего размера также входят в ядро.

Часть II диссертации (главы 4 и 5) направлена на изучение фондового рынка, где ценообразование акций моделируется внутри сложной системы. У инвесторов нет всеохватывающей, рациональной картины ценообразования: при принятии инвестиционных решений они используют простые стратегии основываясь на принципах ограниченной рациональности (**bounded rationality**). В частности, на инвесторов влияет «настроение рынка» (**market sentiment**): они стремятся переключиться на другую, в текущий момент более прибыльную стратегию, даже если фундаментальный анализ не выявляет причин для этого. Важно подчеркнуть, что такое поведение наблюдалось у индивидуальных инвесторов в ходе предыдущих экспериментальных и эмпирических исследований.

В главе 4 рассматривается подобная модель так называемой неоднородности поведения (**behavioural heterogeneity**) на примере фондового рынка США в 1950-2012 гг. Предполагается что на рынке есть два типа инвесторов, имеющих разные представления о динамике цен и скорости, с которой цены возвращаются к фундаментальной стоимости. Фундаменталисты верят в то, что этот процесс происходит быстро, в то время как чартисты (инвесторы, использующие технический анализ) считают, что тренды на фондовом рынке могут превалировать, по крайней мере в коротком периоде. Инвесторы могут переходить от фундаментального анализа к техническому, или наоборот, если какая-то из этих стратегий оказывается более успешной. В результате анализа данных фондового рынка США показывается, что неоднородность поведения является как статистически, так и экономически значимой. Вывод этой главы заключается в том, что настроение рынка может неожиданно меняться, из-за чего эффект фундаментальных шоков усиливается, как случилось, например, с инновациями в информационных и коммуникационных технологиях и последующим «пузырем доткомов», а также с банкротством банка Lehman Brothers и связанным с этим финансовым кризисом.

Глава 5 исследует последствия неоднородности поведения и неоднородных ожи-

даний (*heterogeneous expectations*) для предложенных мер по регуляции фондового рынка, заключающихся в противостоянии крупным позициям инвесторов. Эти предложения появились снова, потому что кризис на фондовом рынке шел в ногу с массовым использованием с одной стороны финансового рычага (т.е. покупок акций на заемный капитал), а с другой, коротких продаж (т.е. продаж акций без покрытия). Из этой главы следует, что если у инвесторов наблюдаются неоднородные ожидания по поводу будущих цен на акции, то ограничения использования финансового рычага и/или коротких продаж может произвести значительный неблагоприятный эффект. Интуитивное объяснение такого неблагоприятного эффекта – это тот факт, что различные ожидания инвесторов могут иметь стабилизирующий или дестабилизирующий характер. Если эти меры (непроизвольно) ограничивают стабилизирующую фундаментальную стратегию, то оценочные ошибки и нестабильность цен на акции могут возрасти.

Приложения в данной диссертации соприкасаются с обсуждениями о том, какие меры должны быть приняты, чтобы предотвратить будущий финансовый кризис, сопоставимый с кризисом 2007-2008 годов. Важная для этих обсуждений сетевая структура межбанковского рынка является мало исследованной сферой. Перед тем, чтобы определить, какие меры повышают стабильность финансовой системы, людям, принимающим решения, нужно понять, как наблюдаемая сетевая структура возникает из индивидуальной оптимизации банков в системе. На фондовом рынке некоторые меры урегулирования уже рассматривались. Однако в изучении эффектов разных мер нужно принимать в расчёт, что ценообразование на финансовом рынке составляет часть сложной системы, в которой инвесторы зачастую используют простые, ограниченно рациональные стратегии. Анализ принимаемых мер, основанных на реалистических моделях сложных систем, способствует лучшему пониманию и более успешному регулированию финансовой экономики.



# Samenvatting (Summary in Dutch)

Dit proefschrift onderzoekt complexe systemen in de financiële economie en past de theorie van complexe systemen toe op de interbancaire markt en de aandelenmarkt. De onderzoeksvragen waarop een antwoord wordt gezocht vinden hoofdzakelijk hun oorsprong in de financiële crisis van 2007 en 2008. Het brandpunt van deze crisis was, achteraf gezien, het faillissement van de Lehman Brothers bank op 15 september 2008. De centrale onderliggende vraag in het onderzoek is hoe de financiële crisis, in beweging gezet door relatief beperkte problemen in de hypotheekmarkt in de VS, uit kon groeien tot een endogeen voortgedreven economische crisis van wereldwijde omvang.

De economische wetenschap heeft moeite een goede verklaring te vinden voor het ontstaan en de uitbreiding van de wereldwijde crisis. Dit komt doordat veel traditionele economische modellen in wezen betrekking hebben op simpele systemen. Grote groepen economische agenten, die keuzes maken over bijvoorbeeld consumptie, productie of investeringen, worden bij voorkeur in één rationele representatieve agent samengegoten. Als het model al expliciete interactie tussen agenten omvat, dan gaat het vaak om interactie tussen een klein aantal homogene agenten. Complexe systemen daarentegen worden juist gekenmerkt door sterke interacties op microniveau, die grote veranderingen teweeg kunnen brengen op macroniveau – zogeheten kritische transities. De ingrijpende gebeurtenissen tijdens de financiële crisis vertonen grote gelijkenis met deze kritische transities. Financiële markten worden in dit proefschrift daarom gemodelleerd als complexe systemen.

Modellen van complexe systemen zijn niet per definitie uitzonderlijk gecompliceerd. Inte-

gendeel, in dit proefschrift draait het juist om een relatief klein aantal economische concepten. Aan de hand van centrale begrippen zoals handelsrelaties tussen banken en marktsentiment van heterogene beleggers wordt de verspreiding van de financiële onrust onderzocht.

Een eerste methodologische stap in het toepassen van de theorie van complexe systemen is het aantonen van empirische relevantie. Veel bestaande economische raamwerken bleken tijdens de crisis weinig verklarende kracht te hebben en waren van weinig waarde voor beleidsmakers, maar het staat niet *a priori* vast dat alternatieve modellen tot een betere empirische fit zullen leiden. Voor de toepassingen in dit proefschrift wordt de empirische relevantie wel aangetoond: ten eerste wordt een niet-triviale netwerkstructuur in de interbancaire markt in kaart gebracht, en ten tweede blijken er duidelijk verschillende toestanden in de dynamica van aandelenprijzen te bestaan die door middel van een kritische transitie abrupt in elkaar kunnen overgaan. Na het vaststellen van de empirische relevantie wordt ingegaan op de nieuwe vragen die complexe systemen oproepen met betrekking tot de regulering van financiële markten.

Het proefschrift bevat vier hoofdstukken onderverdeeld in twee delen, voorafgegaan door een algemene inleiding. De beide delen hebben betrekking op twee specifieke markten die een aantoonbaar versterkend effect hebben gehad op de financiële crisis, namelijk de interbancaire markt en de aandelenmarkt. De analyse draagt daarmee bij aan het begrip van het ontstaan en de uitbreiding van de financiële crisis.

Deel I (hoofdstukken 2 en 3) behandelt interbancaire markten en heeft speciale aandacht voor de interbancaire markt in Nederland. De Nederlandse bankwereld kwam vanaf 2008 in de problemen doordat ze in sterke mate waren blootgesteld aan Amerikaanse banken, die grote verliezen maakten door de beruchte ‘rommelhypotheken’. Belangrijke instellingen zoals ABN AMRO en ING moesten worden gered door de Nederlandse overheid om verdere verspreiding van financiële verliezen zoveel mogelijk te beperken.

Hoofdstuk 2 analyseert data van uitstaande leningen tussen Nederlandse banken van 1998 tot en met 2008. De nadruk ligt op de *kern-periferiestructuur*, die beter op de data blijkt te passen dan andere netwerkmodellen uit de literatuur. De goede fit van de kern-periferiestructuur

suggereert dat er een stabiele kern van sterk vervlochten banken bestaat, en dat de kernbanken bemiddelen tussen periferiebanken als deze willen handelen in de interbancaire markt. Aangezien kernbanken een systeemfunctie vervullen in het opvangen van schokken, kan speciaal toezicht op deze banken de stabiliteit van het systeem als geheel mogelijk vergroten.

In hoofdstuk 3 wordt een verklaring gezocht voor de kern-periferiestructuur in een netwerkformatiemodel voor een financiële markt waarin een expliciete rol is weggelegd voor bemiddeling door tussenhandelaren. In het kern-periferienetwerk ontvangen kernbanken een bemiddelingsprovisie voor het creëren van indirecte verbindingen tussen periferiebanken. Het kern-periferienetwerk blijkt strategisch instabiel indien de banken homogeen zijn: er is altijd ten minste één bank die een prikkel heeft zijn verbindingen aan te passen. Als er rekening gehouden wordt met de heterogeniteit van banken wordt wel een endogeen kern-periferienetwerk gevormd, waarbij de grote banken de kern vormen. Dit theoretische model sluit goed aan bij het geschatte kern-periferienetwerk in de Nederlandse markt. De Nederlandse kernbanken zijn over het algemeen groter dan de periferiebanken, maar ook enkele banken van gemiddelde grootte behoren tot de kern.

Deel II van het proefschrift (hoofdstukken 4 en 5) richt zich op de aandelenmarkt, waarin de prijsbepaling van een aandeel wordt gemodelleerd binnen een complex systeem. Beleggers hebben geen allesomvattend, perfect rationeel beeld van deze prijsbepaling, maar gebruiken eenvoudige strategieën om investeringsbeslissingen te maken. Daarbij zijn beleggers gevoelig voor *marktsentiment*: ze hebben de neiging over te stappen naar een andere, beter presterende strategie, zelfs als er op basis van fundamentele analyse van de aandelen daartoe weinig reden is. Deze neiging om te switchen naar beter presterende beleggingsstrategieën is al eerder aangetoond bij individuele beleggers in experimenteel en empirisch onderzoek.

In hoofdstuk 4 wordt een dergelijk model met zogeheten gedragsheterogeniteit geschat op aandelenprijzen in de VS tussen 1950 en 2012. Er wordt onderscheid gemaakt tussen twee typen beleggers op de markt, die het oneens zijn over de snelheid waarmee prijzen terugkeren naar de fundamentele waarde. Fundamentalisten vertrouwen erop dat dit proces snel verloopt, terwijl

zogeheten chartisten geloven dat een eventuele trend in aandelenprijzen op korte termijn doorgetrokken zal worden. De populatie van beleggers kan geleidelijk verschuiven in de richting van fundamentele analyse of technische analyse (chartisme) als met een van deze strategieën betere prestaties worden behaald. De geschatte gedragsheterogeniteit blijkt zowel statistisch als economisch significant. De resultaten in dit hoofdstuk suggereren dat marktsentiment kan overslaan van de ene gedragstoestand naar de andere waardoor fundamentele schokken kunnen worden opgeblazen, zoals gebeurde met innovaties in informatie- en communicatietechnologie en de daaropvolgende internetzeepbel, en met het faillissement van Lehman Brothers en de daarmee samenhangende financiële crisis.

Hoofdstuk 5 onderzoekt de gevolgen van gedragsheterogeniteit en heterogene verwachtingen voor voorgestelde beleidsmaatregelen die buitengewoon grote beleggingsposities op financiële markten proberen tegen te gaan. Deze voorstellen zijn opnieuw opgekomen omdat de crisis op de aandelenmarkten hand in hand ging met het massale gebruik van enerzijds *leverage* (het kopen van aandelen met geleend geld) en anderzijds *short-selling* (het verkopen van aandelen die men niet zelf bezit). Dit hoofdstuk toont aan dat, als beleggers heterogene verwachtingen hebben van toekomstige aandelenprijzen, een beperking van leverage en/of short-selling praktijken een aanzienlijk averechts effect kan hebben. De intuïtieve verklaring voor dit averechtse effect is dat de verschillende verwachtingen van beleggers zelf ook stabiliserend of destabiliserend voor de markt kunnen zijn. Als de maatregelen (onbedoeld) de stabiliserende fundamentele strategie inperken, kunnen de waarderingsfouten en volatiliteit in aandelenprijzen juist toenemen.

De toepassingen in dit proefschrift sluiten aan bij het debat over beleidsmaatregelen die beogen een toekomstige financiële crisis van het formaat van de crisis in 2007 en 2008 te voorkomen. De netwerkstructuur in de interbancaire markt is hierin relatief onontgonnen gebied. Voordat nieuwe regels kunnen worden overwogen die de stabiliteit van het financiële netwerk moeten verhogen, dienen beleidsmakers grondige kennis te vergaren over hoe de geobserveerde netwerkstructuur ontstaat uit de individuele optimalisatie van banken. Voor de aandelenmarkt is



al veel meer ervaring opgedaan met verschillende vormen van regulering. Bij het onderzoeken van de effecten van beleidsmaatregelen moet er echter rekening mee worden gehouden dat de prijsbepaling op een financiële markt onderdeel uitmaakt van een complex systeem waarin beleggers gebruik maken van simpele, begreepd rationele strategieën. Beleidsanalyse gebaseerd op realistische modellen van complexe systemen kan bijdragen aan een beter begrip en betere regulering van de financiële economie.



The Tinbergen Institute is the Institute for Economic Research, which was founded in 1987 by the Faculties of Economics and Econometrics of the Erasmus University Rotterdam, University of Amsterdam and VU University Amsterdam. The Institute is named after the late Professor Jan Tinbergen, Dutch Nobel Prize laureate in economics in 1969. The Tinbergen Institute is located in Amsterdam and Rotterdam. The following books recently appeared in the Tinbergen Institute Research Series:

- 542 T. BAO, *Experiments on Heterogeneous Expectations and Switching Behavior*
- 543 S.D. LANSDORP, *On Risks and Opportunities in Financial Markets*
- 544 N. MOES, *Cooperative decision making in river water allocation problems*
- 545 P. STAKENAS, *Fractional integration and cointegration in financial time series*
- 546 M. SCHARTH, *Essays on Monte Carlo Methods for State Space Models*
- 547 J. ZENHORST, *Macroeconomic Perspectives on the Equity Premium Puzzle*
- 548 B. PELLOUX, *The Role of Emotions and Social Ties in Public Good Games: Behavioral and Neuroeconomic Studies*
- 549 N. YANG, *Markov-Perfect Industry Dynamics: Theory, Computation, and Applications*
- 550 R.R. VAN VELDHUIZEN, *Essays in Experimental Economics*
- 551 X. ZHANG, *Modeling Time Variation in Systemic Risk*
- 552 H.R.A. KOSTER, *The internal structure of cities: the economics of agglomeration, amenities and accessibility.*
- 553 S.P.T. GROOT, *Agglomeration, globalization and regional labor markets: micro evidence for the Netherlands.*
- 554 J.L. MÖHLMANN, *Globalization and Productivity Micro-Evidence on Heterogeneous Firms, Workers and Products*
- 555 S.M. HOOGENDOORN, *Diversity and Team Performance: A Series of Field Experiments*
- 556 C.L. BEHRENS, *Product differentiation in aviation passenger markets: The impact of demand heterogeneity on competition*
- 557 G. SMRKOLJ, *Dynamic Models of Research and Development*
- 558 S. PEER, *The economics of trip scheduling, travel time variability and traffic information*
- 559 V. SPINU, *Nonadditive Beliefs: From Measurement to Extensions*
- 560 S.P. KASTORYANO, *Essays in Applied Dynamic Microeconomics*
- 561 M. VAN DUIJN, *Location, choice, cultural heritage and house prices*
- 562 T. SALIMANS, *Essays in Likelihood-Based Computational Econometrics*
- 563 P. SUN, *Tail Risk of Equity Returns*

- 564 C.G.J. KARSTEN, *The Law and Finance of M&A Contracts*
- 565 C. OZGEN, *Impacts of Immigration and Cultural Diversity on Innovation and Economic Growth*
- 566 R.S. SCHOLTE, *The interplay between early-life conditions, major events and health later in life*
- 567 B.N.KRAMER, *Why don't they take a card? Essays on the demand for micro health insurance*
- 568 M. KILIÇ, *Fundamental Insights in Power Futures Prices*
- 569 A.G.B. DE VRIES, *Venture Capital: Relations with the Economy and Intellectual Property*
- 570 E.M.F. VAN DEN BROEK, *Keeping up Appearances*
- 571 K.T. MOORE, *A Tale of Risk: Essays on Financial Extremes*
- 572 F.T. ZOUTMAN, *A Symphony of Redistributive Instruments*
- 573 M.J. GERRITSE, *Policy Competition and the Spatial Economy*
- 574 A. OPSCHOOR, *Understanding Financial Market Volatility*
- 575 R.R. VAN LOON, *Tourism and the Economic Valuation of Cultural Heritage*
- 576 I.L. LYUBIMOV, *Essays on Political Economy and Economic Development*
- 577 A.A.F. GERRITSEN, *Essays in Optimal Government Policy*
- 578 M.L. SCHOLTUS, *The Impact of High-Frequency Trading on Financial Markets*
- 579 E. RAVIV, *Forecasting Financial and Macroeconomic Variables: Shrinkage, Dimension reduction, and Aggregation*
- 580 J. TICHEM, *Altruism, Conformism, and Incentives in the Workplace*
- 581 E.S. HENDRIKS, *Essays in Law and Economics*
- 582 X. SHEN, *Essays on Empirical Asset Pricing*
- 583 L.T. GATAREK, *Econometric Contributions to Financial Trading, Hedging and Risk Measurement*
- 584 X. LI, *Temporary Price Deviation, Limited Attention and Information Acquisition in the Stock Market*
- 585 Y. DAI, *Efficiency in Corporate Takeovers*
- 586 S.L. VAN DER STER, *Approximate feasibility in real-time scheduling: Speeding up in order to meet deadlines*
- 587 A. SELIM, *An Examination of Uncertainty from a Psychological and Economic Viewpoint*
- 588 B.Z. YUESHEN, *Frictions in Modern Financial Markets and the Implications for Market Quality*
- 589 D. VAN DOLDER, *Game Shows, Gambles, and Economic Behavior*
- 590 S.P. CEYHAN, *Essays on Bayesian Analysis of Time Varying Economic Patterns*
- 591 S. RENES, *Never the Single Measure*