

Gedruckt mit Genehmigung der Wirtschafts- und Sozialwissenschaftlichen
Fakultät der Christian-Albrechts-Universität zu Kiel
Dekan: Professor Dr. Achim Walter
Erstberichterstattender: Professor Dr. Thomas Lux
Zweitberichterstattender: Professor Dr. Lukas Menkhoff
Tag der Abgabe der Arbeit: 30.06.2014
Tag der mündlichen Prüfung: 18.11.2014

Complex Interactions in Financial Markets

Inauguraldissertation zur Erlangung des akademischen Grades
eines Doktors der Wirtschafts- und Sozialwissenschaften
der Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von
Diplom-Volkswirt Karl Finger
aus Erlangen
Kiel, Juni 2014

To Vicky and Katrin

Acknowledgements

Over the past four years, I have received support, encouragement and more from various people and without them it would not have been possible for me to complete my thesis. First, my gratitude goes to my supervisor Prof. Thomas Lux. I am very fortunate that he encouraged me to work on (for me) completely new economic models and significantly broadened my horizon and knowledge. Moreover, he offered help and guidance, in his calm and supportive manner, whenever necessary.

Second, but with equal gratitude, I want to thank my co-advisor Prof. Stefan Reitz for being a wonderful boss and supporting my time-consuming research projects during the last years.

I also want to thank my colleagues at the QBER Institute Simon, Alexander, Alexander, Ethel, Maria, Markus and Sebastian for their help and support.

Similarly, I am indebted to participants of conferences and workshops, for their valuable comments and suggestions. I am especially thankful to Prof. Tony He, who arranged a fantastic research stay at the University of Technology in Sydney, Australia.

Of course, special thanks go to my family and friends, who always supported me. Special thanks go to Daniel, Martini, Fenz, Sebi, Felix, Ulli, Hannes, Amelie, Anna, Tillman, Lorena, Esther, Giovanni, Marlene, and Otto.

CONTENTS

| | |
|---|----|
| <i>Contents</i> | i |
| <i>List of Figures</i> | iv |
| <i>List of Tables</i> | vi |
| 1. <i>General Introduction</i> | 1 |
| 2. <i>Financial Transaction Tax</i> | 8 |
| 2.1 Introduction and Existing Literature | 9 |
| 2.2 Model | 11 |
| 2.2.1 CDA and Information | 11 |
| 2.2.2 Trader Types | 12 |
| 2.2.3 Implementing the Tax Instruments | 14 |
| 2.2.4 Asset Demand and Cancellation | 15 |
| 2.2.5 Trading Process and Time | 16 |
| 2.3 Pseudo-Empirical Results | 16 |
| 2.3.1 Baseline Scenario with Two Markets | 17 |
| 2.3.2 Effects of a FTT in the Two-Market Case | 22 |
| 2.3.3 Effects of FAT | 24 |
| 2.4 Discussion and Conclusion | 25 |
| 3. <i>Network Analysis of the e-MID</i> | 30 |
| 3.1 Introduction and Existing Literature | 31 |
| 3.2 Networks | 32 |
| 3.3 The Italian Interbank Market e-MID | 33 |
| 3.4 Results | 34 |
| 3.4.1 General Features | 35 |
| 3.4.2 Density | 39 |
| 3.4.3 What is a Sensible Aggregation Period? | 42 |
| 3.4.4 Transitivity | 47 |
| 3.4.5 Small-World Property | 50 |
| 3.4.6 Effects of the Global Financial Crisis | 52 |

| | | |
|-------|--|-----|
| 3.5 | Discussion | 54 |
| 3.6 | Conclusions and Outlook | 57 |
| 4. | <i>Assortativity in Scale-Free Interbank Credit Networks</i> | 59 |
| 4.1 | Introduction and Existing Literature | 60 |
| 4.2 | Networks and Assortativity | 62 |
| 4.3 | The Italian Interbank Market (e-MID) | 64 |
| 4.4 | Assortativity Patterns in the Italian Interbank Market | 65 |
| 4.5 | Assortativity in Finite-Size Scale-Free Networks | 71 |
| 4.6 | Conclusions | 76 |
| 4.7 | Appendix | 78 |
| 4.7.1 | Robustness Checks for the Empirical Results | 78 |
| 4.7.2 | Assortativity in Rewired Scale-Free Networks | 79 |
| 5. | <i>Actor-Oriented Relations in the Interbank Market</i> | 83 |
| 5.1 | Introduction | 84 |
| 5.2 | Data Set and Network Construction | 86 |
| 5.3 | The Stochastic Actor Oriented Model of Network Formation | 89 |
| 5.3.1 | Model Specification | 90 |
| 5.3.2 | Network Effects | 93 |
| 5.3.3 | Model Estimation | 101 |
| 5.4 | Estimation Results | 104 |
| 5.4.1 | Full Sample Estimation | 104 |
| 5.4.2 | Before and After the Financial Crisis | 110 |
| 5.5 | Concluding Remarks | 115 |
| 5.6 | Appendix | 118 |
| 5.6.1 | Composition Change | 118 |
| 5.6.2 | Quarterly Results | 118 |
| 6. | <i>Speculation on Agricultural Commodity Futures Markets</i> | 159 |
| 6.1 | Introduction | 160 |
| 6.2 | Data | 162 |
| 6.3 | Trader Groups | 164 |
| 6.4 | Empirical Analysis | 168 |
| 6.5 | Conclusion | 173 |
| 7. | <i>Speculation on Energy Commodities and Precious Metals</i> | 174 |
| 7.1 | Introduction | 175 |
| 7.2 | Data | 176 |
| 7.3 | Trader Groups | 179 |
| 7.4 | Empirical Analysis | 182 |

| | | |
|-------|-------------------------------------|-----|
| 7.5 | Conclusion | 189 |
| 7.6 | Appendix | 191 |
| 7.6.1 | Correlation | 191 |
| 7.6.2 | Complementary Results | 191 |
| 7.6.3 | Structural Breaks | 191 |
| 7.6.4 | Robustness Check | 192 |
| 7.6.5 | Other Trader Groups | 192 |
| 8. | <i>Concluding Remarks</i> | 203 |
| | <i>References</i> | 206 |
| | <i>Bibliography</i> | 206 |

LIST OF FIGURES

| | | |
|------|---|----|
| 2.1 | Exemplary simulation results for a single run without tax . . . | 18 |
| 2.2 | Activity of informed traders for 200 runs without tax | 19 |
| 2.3 | Activity of informed traders in one exemplary run without tax | 20 |
| 2.4 | Crossroads of switching process | 21 |
| 2.5 | Monte-Carlo simulation results with dependence on χ_{FTT} . . . | 28 |
| 2.6 | Monte-Carlo simulation results with dependence on χ_{FAT} . . . | 29 |
| | | |
| 3.1 | Number of active banks, traded volume and number of trades over time | 36 |
| 3.2 | Fraction of trades and traded volume between banks from dif- ferent countries | 37 |
| 3.3 | The banking network in the 4th quarter of 2010 | 38 |
| 3.4 | Observed density | 40 |
| 3.5 | Density for aggregated networks | 41 |
| 3.6 | Jaccard Index | 44 |
| 3.7 | Reciprocity for aggregated networks | 46 |
| 3.8 | Transitivity. | 48 |
| 3.9 | Clustering coefficients. | 49 |
| 3.10 | Clustering coefficients for aggregated networks | 50 |
| | | |
| 4.1 | Assortativity coefficient for the undirected version of the net- work over time | 66 |
| 4.2 | The Italian interbank network in 1999Q1 and in 2010Q4 . . . | 67 |
| 4.3 | Assortativity coefficients for the directed version of the Italian network | 69 |
| 4.4 | Assortativity coefficient for different aggregation periods . . . | 70 |
| 4.5 | Assortativity coefficient for daily, quarterly and sampled net- works | 70 |
| 4.6 | Assortativity coefficient in scale-free networks (Goh <i>et al.</i> , 2001) with $N = 100$ | 73 |
| 4.7 | Assortativity coefficient in scale-free networks (fitness-based algorithm) with $N = 100$ | 74 |

| | | |
|------|---|-----|
| 4.8 | Assortative mixing in scale-free networks (Goh <i>et al.</i> , 2001) with $N = 100$ | 75 |
| 4.9 | Assortativity coefficient of the complete network for different aggregation periods | 79 |
| 4.10 | Assortativity coefficient for the valued network (transaction volumes) and different aggregation periods | 80 |
| 4.11 | Assortativity coefficient for the valued network (number of trades) and different aggregation periods | 80 |
| 4.12 | Network matrix for scale-free networks with $N = 100$, $M = 1980$, $\alpha = 2.0$, and different levels of assortativity | 81 |
| 4.13 | Network matrix for scale-free networks with $N = 100$, $M = 1980$, $\alpha = 3.0$, and different levels of assortativity | 82 |
| 5.1 | The two directed clustering coefficients and their corresponding effects | 97 |
| 5.2 | The evolution of the basic rate parameter and the number of banks from 2002 to 2010 | 107 |
| 6.1 | Time series of the S&P GSCI spot price index and net-long positions for soybeans and cocoa | 168 |
| 7.1 | The S&P GSCI spot price index for crude oil and silver | 183 |
| 7.2 | Statistics of the cumulative OLS residuals | 202 |

LIST OF TABLES

| | | |
|------|---|-----|
| 2.1 | Baseline parameter setting for the simulations. | 17 |
| 3.1 | Basic network statistics for different aggregation periods | 43 |
| 3.2 | Jaccard Index for daily, monthly, quarterly and yearly networks. | 45 |
| 3.3 | Reciprocity of the Italian Banking network | 45 |
| 3.4 | The average CC and ASPL for the observed and random networks | 51 |
| 3.5 | Summary statistics for quarters 36-45 | 53 |
| 5.1 | Descriptive statistics calculated for the 36 quarters from 2002-2010 | 93 |
| 5.2 | The effects of the objective function | 95 |
| 5.3 | The quarterly estimation results from 2002 until 2010 | 106 |
| 5.4 | The quarterly estimation results from 2002 until the 2nd quarter of 2008 | 111 |
| 5.5 | The quarterly estimation results from the 2nd quarter of 2008 until the 4th quarter of 2010 | 112 |
| 5.6 | Two sample t-test comparing the estimation results for the two subperiods | 114 |
| 5.7 | The quarterly estimation results from 2002 until 2010 with composition change | 119 |
| 5.8 | The quarterly estimation results from 2002 until the 2nd quarter of 2008 with composition change | 120 |
| 5.9 | The quarterly estimation results from the 2nd quarter of 2008 until the 4th quarter of 2010 with composition change | 121 |
| 5.10 | The estimation results for each quarter from 2002Q1 until 2002Q4 | 122 |
| 5.11 | Joint significance of the actor covariate related effects for each covariate | 158 |
| 6.1 | Summary of the 10 commodities | 164 |
| 6.2 | Trader classifications of the CFTC | 165 |

| | | |
|------|--|-----|
| 6.3 | Average correlation coefficients between the net-long positions with each other and the S&P GSCI Spot Price Index | 166 |
| 6.4 | Test statistics for the Augmented Dickey Fuller and the Kwiatkowski-Phillips-Schmidt-Shin test in levels and first differences | 169 |
| 6.5 | Results for Granger non causality between the net-long positions and the S&P GSCI spot price index | 171 |
| 7.1 | Basic information about the six commodities | 178 |
| 7.2 | Trader classifications of the CFTC | 179 |
| 7.3 | Correlations between the net-long positions and the S&P GSCI Spot Price Indices | 180 |
| 7.4 | Unit root tests for prices and net-long positions | 184 |
| 7.5 | Test of Granger non-causality for the whole sample | 187 |
| 7.6 | Test of Granger non-causality for the subperiods | 189 |
| 7.7 | Correlations between the net-long positions and the S&P GSCI Spot Price Indices for the 3 energy commodities | 191 |
| 7.8 | Correlations between the net-long positions and the S&P GSCI Spot Price Indices for the 3 precious metals | 192 |
| 7.9 | Test of Granger non-causality for the pre-GFC period | 193 |
| 7.10 | Test of Granger non-causality for the GFC period | 194 |
| 7.11 | Test of Granger non-causality for after the GFC | 195 |
| 7.12 | Test of Granger non-causality for gold if the price series is considered to be I(2) | 196 |
| 7.13 | Unit-root tests for Prod and Other | 197 |
| 7.14 | Granger causality test for Prod and Other | 198 |
| 7.15 | Granger causality test for Prod and Other before GFC | 199 |
| 7.16 | Granger causality test for Prod and Other during GFC | 200 |
| 7.17 | Granger causality test for Prod and Other after GFC | 201 |

1. GENERAL INTRODUCTION

This thesis consists of the following six papers:¹

Finger, K., Fricke, D. (2014), Tax Evasion in an Artificial Financial Market, QBER Discussion Papers 1-2014, Institute for Quantitative Business and Economics Research

Finger, K., Fricke, D., Lux, T. (2013), Network Analysis of the e-MID Overnight Money Market: The Informational Value of Different Aggregation Levels for Intrinsic Dynamic Processes, *Computational Management Science*, 10 (2-3), pp. 187-211

Fricke, D., Finger, K., Lux, T. (2013), On Assortative and Disassortative Mixing in Scale-Free Networks: The Case of Interbank Credit Networks, Kiel Working Paper 1830, Kiel Institute for the World Economy

Finger, K., Lux, T. (2014), Friendship between Banks: An Application of the Actor-Oriented Model of Network Formation to Interbank Credit Relations, Kiel Working Paper 1916, Kiel Institute for the World Economy

Finger, K., Haas, M., Klos, A., Reitz, S. (2013), Who are the Speculators on Commodity Markets?, QBER Discussion Papers 9-2013, Institute for Quantitative Business and Economics Research

Finger, K. (2014), Speculative Activity on Futures Markets for Energy Commodities and Precious Metals, QBER Discussion Papers 4-2014, Institute for Quantitative Business and Economics Research

The global financial crises (GFC) led to the deepest worldwide recession since the great depression in 1929. It also provoked a big controversy regarding the failure of the economic discipline which steered my interest in this direction.

A financial crisis is typically primarily associated with a specific financial market or market segment, from which the crisis potentially evolves and adversely affects other financial markets and finally the real economy. In the beginning some traders detect outstanding profit opportunities in this market and the corresponding high demand increases the asset price. To become problematic, however, the initial increase has to trigger some kind of reinforcing dynamic which leads to a further increase not backed by fundamentals. The reason may be that (a) traders influence the perception of

¹ At the end of this introduction I describe my contribution to each of them.

each other (e.g. copying successful strategies) and/or (b) technical traders rely on a continuation of the price trend (Lux, 1998).

The eventual burst of the price bubble might again produce reinforcing dynamics this time driving the price below the fundamental value (an overvaluation is followed by an undervaluation). This is because many financial institutions have to repeatedly adjust their portfolios after the initial price drop due to similar asset holdings, mark to market evaluation and high leverage close to the regulatory constraints. Throughout this process the need for liquidity is becoming more desperate because, e.g., credit lines will not be extended. Hence, the urgent need of liquidity makes it necessary to sell other, more liquid assets, and creates contagion.

In the GFC the worldwide trading of asset backed securities (ABS) of US subprime loans made the crisis a global problem.² In general, the financial world has grown and is growing together ever faster due to technical innovations and deregulation during the last decades. The ever increasing demand for ABS was also driven by the wrong judgement of the rating agencies giving even the junior tranches very solid ratings (Coval *et al.*, 2009). This enabled pension funds, insurance companies and state owned banks to participate in this market in the first place.

This short and simplistic summary already shows that not all economic agents share the same expectations and that the action of one agent often triggers feedback loops that might lead to quite different results compared to the initial effect. Therefore it seems natural to see (financial) markets as a complex system in which heterogeneous agents (e.g. private investors, financial institutions, and firms) interact and produce complicated dynamics on the macro-level (Lux, 2011). In sharp contrast mainstream macroeconomics is still relying on the rational agent framework and often these models are only slightly adjusted to cope with the observations made during the GFC.³

In my dissertation I tried, together with my co-authors, to better understand the interaction of heterogeneous agents and the associated consequences on different financial markets. The first paper extends the artificial financial market of Fricke and Lux (2013a) twofold. Firstly, a second asset market is introduced and, secondly, not only the effects of a financial transaction tax (a small transaction fee on each financial transaction) but also that of a financial activity tax (taxing the speculative profits of financial institutions) are considered. In this model the so called informed traders, each following their own strategy, produce rich price dynamics which are able to replicate some of the ‘stylized facts’ of financial markets. These informed traders show very

² See Brunnermeier (2009) for a more elaborate description of the GFC.

³ See Colander *et al.* (2009) for a detailed discussion of this topic.

strong herding behavior (without any tax) in one of the two markets, while the ‘herd’ potentially switches between the two markets at certain crossroads. In respect to the two tax instruments the effects are quite diverse. The financial transaction tax (FTT) is able to guard the taxed market against potentially destabilizing short-term ‘chartist’ traders (believing in the continuation of price trends), because the tax is a fixed value irrespective of the intended investment period. On that account the misalignment from the fundamental value is reduced for very small tax rates. The financial activity tax (FAT) affects the traders more equally and, therefore, the switching process between the two markets continues. Hence, the FTT should be better suited to target short-term traders, while the FAT might affect the market structure less. However, the model shows that the tax rate is absolutely crucial in both cases. In this respect policy makers got (potentially) ahead of our research project and eleven EU countries agreed upon implementing a common FTT, however, neither the date nor a precise agenda when this tax will be implemented exist.⁴ Moreover, France (August 2012) and Italy (March 2013) already introduced an FTT and the effects (in case of France) are according to Colliard and Hoffmann (2013) far from devastating.

The next three papers are concerned with the Italian banking network using the e-MID data (the only electronic platform for interbank loan) for overnight loans from 2002 until 2010. The GFC showed the importance of linkages between banks (default of Lehman) and the potentially devastating effects of default cascades transmitting through the financial system.⁵ In this respect the network perspective widely used in many scientific disciplines is able to shed new insights and improve our understanding of the banking system.

A generally interesting aspect of the complex system literature and especially the network section is that it serves as a melting pot for various scientific disciplines. It has been applied in many of the natural sciences (technical or biological networks), but also in the social sciences (social networks) for decades and economic networks seem to lie in between those kind of networks in some aspects. However, the network approach (as a complex system) is relatively new to economics and in case of banking networks is dominated by studies undergone by researchers and models related to the natural sciences. They mainly focus on investigating the overall topology of the network to find universal ‘stylized facts’ and analyze the resilience of the network structure, which is of course very important. Nevertheless, the

⁴ On January 22, 2013 the Council of the European Union allowed Belgium, Germany, Estonia, Greece, Spain, France, Italy, Austria, Portugal, Slovenia and Slovakia to introduce the tax.

⁵ See e.g. Haldane and May (2011) and Battiston *et al.* (2012).

actions of banks or other economic agents are always the result of humans (individual or groups) making a conscious decision and therefore economists should more openly look to models and ideas developed in the social sciences.

The first of these three papers predominantly questions whether the use of overnight loans as an indicator for a link automatically justifies to restrict our analysis to daily networks as it is mainly done in the literature. However, we are able to show that much of the randomness of the daily networks disappears, if we prolong the aggregation period and this finding is applied in the other two papers by using quarterly networks.

The second paper focuses on the compatibility of the use of scale-free generating mechanisms to replicate banking networks and the finding that real world banking networks (also the Italian) are disassortative in respect to their degree, i.e. banks with many links tend to be linked to banks with few links. The starting point was that the most prominent generating mechanisms for scale-free networks called preferential attachment from Barabasi and Albert (1999) is under some assumptions (e.g. infinite network size) unassortative. However, our investigation of several generating mechanisms showed that scale-free networks can be disassortative, but the disassortativity is never of the degree observed in the Italian banking network. This especially true for the observed values of density and scaling exponent.⁶

The third paper adapts the stochastic actor-oriented model developed by Snijders (1996) to investigate the dynamics of the interbank network. It enables us to provide insights on the driving forces behind banks' decisions about their counterparties in the interbank market by taking an actor point of view. The behavior is astonishingly robust in almost all respects over the complete sample period. This stability should have played an important role in the resilience of the interbank market during the GFC. One important factor for this finding is that banks heavily rely on lasting relations, although the electronic nature of the e-MID platform reduces the costs to trade with 'new' counterparties practically to zero. In spite of the general stability, we observe two changes after the beginning of the GFC: firstly, large banks and 'core' banks (identified by Fricke and Lux, 2014) become more important borrowers. Secondly, banks avoid intermediate relations more rigorously than before. Together, these findings indicate that counterparty risk played a more important role in the lending decisions of banks after the GFC. Moreover, the first aspect supports the 'too big to fail' and 'too interconnected to fail' argument for the size and core variables, respectively.

In response to the GFC and the observed vulnerability of the financial

⁶ See Fricke and Lux (2013b) for a discussion on whether a power-law is at all appropriate to describe the degree distribution of banking networks.

system the Basel III agreement by the Basel Committee on Banking Supervision claimed more restrictive regulation, e.g. capital requirements or leverage ratios. For instance, the new leverage ratios even entail a stricter rule for systemically important financial institution to account for the ‘too big to fail’ and ‘too interconnected to fail’ issues. However, in my view, a financial system (regardless of other issues like too high complexity) can not be considered ‘safe’ as long as financial institutions know that governments would have to bail them out.

The next two papers are concerned with the potentially destabilizing effects of speculation on the prices of commodities between 2006 and 2013. The first concentrates on ten agricultural commodities. The identification of ‘speculators’ is in itself a difficult and crucial task, for which we rely on the weekly ‘Commitments of Traders’ reports of the Commodity Futures Trading Commission (CFTC). They classify the biggest traders in three different ways and thereby distinguish nine groups in total. To study the relation between these groups we proxy their strategy by their net-long (long minus short) positions, since this measures the aggregated expected price change of the respective group.⁷ By looking at the correlation coefficients among the net-long positions we observe three distinct trading strategies: (a) commodity index traders, (b) commercials and (c) non-commercials. The commodity index traders have increasingly attracted scholarly attention, since their investment strategy exploits the weak correlation of commodity, bonds and stocks returns. Therefore, their trading decisions seem not to depend on the fundamentals of commodity markets.⁸ The second ‘harmful’ speculators candidates are the commercials which also participate in the market for profit reasons. However, an investigation of Granger (non) causality between the price of a commodity and the corresponding net-long position of the respective trader group shows that neither of them influences the price. Yet the commercials seem to react to the price. Hence, we find no evidence that excessive speculative activity significantly contributed to the price spikes during our sample period. Of course, this does not prove that there is no causal relationship and especially the data leaves room for improvements. The aggregation period (weekly) might be too long and the level of aggregation (traders are at most arranged into four groups) most likely blends traders with still relatively diverse strategies into one group.

The last paper applies a similar methodology to investigate the effect of speculation on the prices of energy commodities and precious metals. Since

⁷ A positive net-long position indicates that they profit from an increasing price.

⁸ See e.g. Gorton and Rouwenhorst (2006), Erb and Harvey (2006) and Conover *et al.* (2010).

none of the six commodities under scrutiny are listed in the ‘Supplemental’ report of the CFTC, it is not possible to examine commodity index traders in this respect. Money Managers and Swap Dealers are identified as potential candidates for excessive speculative activity in this study. The only robust finding for the complete sample is that Swap Dealers often tend to adjust their trading strategy with respect to the price. According to an OLS-CUSUM test for structural breaks, the sample is split into three subperiods. Analyzing the individual subperiods, I find that Money Managers seem to at least partially Granger cause the strongest price increases of crude oil and silver. Hence, there is some evidence that excessive speculation might have driven the prices of commodities above their fundamental values.

Here, I briefly describe my contribution to the studies presented in chapters 2 – 7 and finally chapter 8 concludes.

- Chapter 2: This paper is joint work with Daniel Fricke. The extension of the technical implementation was done by myself, analyzing, writing and editing was shared equally.
- Chapter 3 and 4: These papers are joint work with Daniel Fricke and Thomas Lux. The technical implementation and the writing, usually resulting from extensive discussions with Professor Lux, were principally done by Daniel and myself.
- Chapter 5: This paper is joint work with Thomas Lux. It was Professor Lux’ idea to adapt the Stochastic-Actor-Oriented Model from Snijders (1996) to the interbank loan market. The technical implementation, estimation, analysis, and most of the writing, were done by myself.
- Chapter 6: This paper is joint work with Stefan Reitz, Markus Haas and Alexander Klos. Prof. Reitz provided the idea to analyse the impact of different trader groups on commodity prices using their net-long position on futures markets. The technical implementation, estimation, analysis, and most of the writing were done by myself. During the whole process extensive discussions with Prof Haas, Prof. Klos and Prof Reitz greatly contributed to the final outcome.
- Chapter 7: The complete work was done by myself.

2. TAX EVASION IN AN ARTIFICIAL FINANCIAL MARKET

2.1 *Introduction and Existing Literature*

The International Monetary Fund was instructed by the G-20 to prepare a range of options that make the financial sector contribute to the costs induced by the global financial crisis. Three main ideas were presented (cf. IMF, 2010):

1. a financial stability contribution also called ‘bank tax’ or ‘bank levy’ taxing balance sheet items, where the individual tax rates are proportional to the systemic importance of each financial institution,
2. a financial activity tax (FAT), which aims at raising revenues by taxing the (speculative) profits of financial institutions,
3. a financial transaction tax (FTT), which puts a small transaction fee on the notional value of each financial transaction.

This paper is motivated by the fact that surprisingly few attempts have been made to compare the effects of the different tax instruments quantitatively. Here we compare the effects of the latter two instruments.

The IMF states that the FAT should be better suited to generate tax revenues and stabilize financial markets. Shaviro (2012) compares non-technically both taxes and comes to the same conclusion, because the FAT relies on a broad ‘net’ measure. Relatively recently, a set of 11 EU countries agreed on an ‘enhanced cooperation’ in implementing a common FTT in 2014, given that that no consensus was reached on the G-20 or EU level.

One of the major arguments against the FTT is that it can be strongly destabilizing due to capital flights from a taxed towards untaxed markets. Somewhat surprisingly, at least under the hypothesis of a linear response to the FTT, empirical results on the effects of the FTT are far from pointing in the same direction.¹ Therefore, we simulate an artificial financial market to non-invasively evaluate the effects of the different regulatory measures.

Several authors have investigated the effects of the FTT in such systems, whereas, to the best of our knowledge only Lengnick and Wohltmann (2012) compared the effects of the FTT and the FAT in an artificial market setting. Their main finding is that the FTT tends to be preferable for reasonably small tax rates. since it reduces volatility and generates substantial tax revenues. In contrast, while the FAT is less able to stabilize markets and also generates

¹ See, e.g. Jones and Seguin (1997), Hau (2006) and Roll (1989), respectively. These papers find very different effects of FTTs. The diversity of these empirical findings is an indicator for nonlinear effects of the tax - something that we can replicate in our analysis.

less tax revenues, it has the advantage that it does not create distortions for very large tax rates.

However, the results of Lengnick and Wohltmann (2012) should be treated with caution since they are based on a market-maker setting, where one liquidity provider is assumed to be exempt of the tax. A priori, there is no reason why certain market participants do not have to pay the tax. Hence, we rely on the model of Fricke and Lux (2013a) which takes the endogenous nature of liquidity (in part driven by behavioral changes of traders) into account by simulating a more realistic order-driven continuous double auction (CDA) setting. The model differentiates between two groups of traders, namely noise traders and informed traders. While noise traders submit random orders, informed traders take both fundamental values and past prices into account when forming their price expectations. Depending on their time horizons, informed traders' strategies are therefore chartist, fundamentalist or something in between. The model replicates some of the stylized facts of financial time-series for certain parameter combinations. Within a single asset market, Fricke and Lux (2013a) found the FTT to be more attractive for very small tax rates, since it takes only marginally profitable orders out of the system and at the same time generates substantial tax revenues.

Here we aim at extending their framework in two ways: first, we implement the FAT in their framework and compare it to the FTT. Second, we introduce another asset market, since the potential of drying up of liquidity due to tax evasion is the most prominent argument of single countries to not implement the tax individually. Altering the taxation of the two assets allows us to analyze the effects of a tax haven on the taxed market and vice versa. Only few studies have dealt with this issue in detail and if so, only for the FTT. For example, Westerhoff and Dieci (2006) find that the taxed market will become more stable, while the untaxed market becomes less stable.² Imposing a uniform tax stabilizes both markets. However, these results were also obtained under the assumption that a market-maker is exempt of the tax.

Our main findings can be summarized as follows: the introduction of a second market produces alternating 'herding' behavior of the informed traders in being active in either of the two markets. Interestingly, this result does not rely on the interaction or coordination among the informed traders, but rather results from simple individual profit considerations. In the current setting, noise traders respond similarly to the two tax instruments, so the interesting results are driven by the behavioral changes of informed traders.

² Hanke *et al.* (2010) find comparable results in an experimental setup show additionally that the relative market size may be an important determinant for the effects of the FTT.

As in the one market case, the FTT harms short-term oriented traders more. As a result, these traders tend to prefer trading in the untaxed market already for small tax rates, which reduces the likelihood of strongly cyclical assets prices in the taxed market at the expense of the untaxed market. Accordingly, asset prices tend to be closer to their fundamentals for very small tax rates in the taxed market. The FAT affects all trading horizons equally and therefore has a smaller impact on the trading process. However, the decreasing liquidity with corresponding higher spreads in the taxed market even tend to attract ‘chartist’ traders and destabilize the market in terms of higher distortions and volatility. Hence, the FTT might stabilize markets but has to be adjusted cautiously for different asset classes, while the FAT has less ramifications on the trading process for small tax rates and can easily be used to tax several asset classes.

The remainder of this paper is structured as follows: Section 2.2 introduces the structure of the model. Section 2.3 presents pseudo-empirical results and Section 2.4 concludes.

2.2 Model

Here we briefly summarize the model setup. More details on the original model can be found in Fricke and Lux (2013a).

2.2.1 CDA and Information

The financial market consists of N heterogenous agents trading two assets (dividendless, fixed supply) against cash. For simplicity, we only use one currency in the following, i.e. cash can be used to buy either asset, earning zero interest. The market is order-driven and the quoted prices p_t^m (midprice), for assets $m = 1, 2$, are the average of the best asks (a_1^m) and best bids (b_1^m) in the respective limit-order book (LOB), while $a_1^m - b_1^m > 0$ is the bid-ask spread. In case there are no orders in the LOB, the quoted price is simply the last quoted price. Prices are discrete and can only be submitted on a specified grid, defined by the tick-size Δ .

In the CDA market, agents can submit either limit or market orders. A market order specifies the amount to buy or sell at the current best price, whereas a limit order puts a limit price at which the agent is still willing to trade. Price dynamics are therefore driven both by the arrival of new orders and the cancellation of existing limit orders. Limit orders cannot be immediately executed and are stored in the LOB according to price-time priority, i.e. orders with better prices have a higher chance of execution

(price priority). If two limit orders have the same price, the order that was posted earlier then enjoys (time) priority.

Despite being dividendless, we assume that the two assets have constant positive fundamental values, which we for simplicity fix at the same level ($p_f^1 = p_f^2 = p_f$). Only informed agents know the fundamental values, whereas the state of the LOB and the history of the quoted prices are public information. Each agent i is initially endowed with an amount of $S_0^{i,1} = S_0^{i,2} = N_s$ of the two assets, and $C_0^i = 2N_s p_0$ units of cash. We impose short-selling and capital constraints, such that $S_t^{i,m}, C_t^i \geq 0, \forall t, i$.

2.2.2 Trader Types

Two groups of traders compete in the system: a fraction $\theta \in [0, 1]$ of informed traders and a fraction $(1 - \theta)$ of noise traders. Each agent acts once every $H^{w,i}$ time-steps on average. We model strategic order placement, since agents create two orders: one unconditional order, which will be sent to the LOB in case of positive expected profits, and a second conditional (take-profit) order, which will only be submitted to the LOB, if the first order was executed.

Noise Traders

As noted in Fricke and Lux (2013a), noise traders are needed to ensure a steady inflow of limit and market orders. Noise traders choose random limit prices for both markets and post their orders in the market, where they expect the largest returns. For simplicity, their price expectations are randomly chosen around the current quoted prices according to

$$p^{(m,l)} = p_t^m e(\epsilon), \quad (2.1)$$

for the two assets and the two limit prices $l, m = 1, 2$, respectively, with ϵ denoting $\text{iiN}(0, \sigma_\epsilon)$ random numbers. With this approach, we directly obtain four limit prices, two for each asset: for example for asset 1, $p^{(1,1)}$ for the first (unconditional) and $p^{(1,2)}$ for the second (conditional) take-profit order. By comparing the percentage differences between the first and second limit price in the two markets, the agent decides in which market to act. Based on these limit prices, we can identify the market side that the trader acts upon. This can be done by comparing the two limit prices, as we impose that their orders should not create a sure loss, i.e. the agent buys first and sells later if $p^{(m,2)} > p^{(m,1)}$, or the agent sells first and buys later if $p^{(m,2)} < p^{(m,1)}$. The expected return then equals $r^e = |\ln(p^{(2)}/p^{(1)})|$. All traders will only post orders with positive expected returns (after tax).

Informed Traders: Chartists and Fundamentalists

Price Expectations Whether an informed agent buys or sells the asset depends on his expectation of the asset's future price at the end of his investment horizon. When forming price expectations, informed traders use information about past prices and the fundamental value. For simplicity we drop the superscripts for the two assets in most of the following.

For each asset, price expectations evolve according to

$$\frac{d\hat{p}_{t+\tau}^i}{d\tau} = -\frac{\hat{p}_{t+\tau-1}^i - p_t^f}{H^{t,i}} + \left(T_t^i + \frac{p_t - p_t^f}{H^{t,i}} \right) e\left(-\frac{\tau}{H^{t,i}}\right), \quad (2.2)$$

where \hat{p}_t^i is agent i 's the expected price at t , T_t^i is the calculated trend and $H^{t,i}$ is the forward trend horizon over which agent i expects the trend to last (cf. Youssefmir *et al.*, 1998). The trend itself is an exponentially smoothed average rate of price changes over a backward horizon $H^{b,i}$, i.e.

$$\frac{dT_t^i}{dt} = \frac{1}{H^{b,i}} \left(\frac{dp}{dt} - T_t^i \right), \quad (2.3)$$

where dp is the price change between $t - \tau$ and t and $H^{b,i}$ is the backward trend horizon of i . Thus, the trend measures the deviation from the moving average of prices.

Subject to the boundary condition $\hat{p}_t^i = p_t$, each agent formulates his expected price development for both assets over the next $H^{w,i}$ time-steps via Eq. (2.2) using the calculated trend from Eq. (2.3). This system incorporates, depending on the relative magnitude of $H^{t,i}$ to $H^{w,i}$, chartist and fundamentalist behavior: agents with a small value of $H^{t,i}/H^{w,i}$ use a fundamentalist strategy, while a larger value indicates a more chartist strategy. Intermediate values are a combination of the two.

Limit Prices Based on the expected prices, the determination of limit prices of informed agents can be split into three parts: In the first part the agent uses Eqs. (2.2) and (2.3) to independently forecast the evolution of the two midprices between t and $t + H^{w,i}$.³ When the agent expects the price to be higher (lower) than the current price he will plan to submit a buy (sell) order. The limit prices of this order and the corresponding conditional order depend on the expected development of the price between t and $t + H^{w,i}$.

In the second step, the agent uses information about the midprices and a proxy for the expected price volatility to form more detailed expectations

³ We could also think of forecasting the joint development of assets 1 and 2, e.g. incorporating dependency between the expected asset price dynamics.

about the best bid and ask over time. As proxy for the expected price volatility, we use the average distance between the current best and second best prices on the two market sides, see Fricke and Lux (2013a). This is calculated as

$$\hat{\sigma}_t^i = \frac{((a_2 - a_1) + (b_1 - b_2))}{2}, \quad (2.4)$$

where a_2 and b_2 denote the second best ask and bid prices of the asset under study, respectively.

Equipped with expectations about the midprices, the corresponding price volatility, and the decision to buy or sell from the first step, the agent then chooses his limit prices as follows: Defining

$$\begin{aligned} \hat{p}_{\max}^i &= \max\{\hat{p}_{t:\tau:t+H^w,i}\} \\ \hat{p}_{\min}^i &= \min\{\hat{p}_{t:\tau:t+H^w,i}\} \end{aligned} \quad (2.5)$$

as the maximum and minimum of the expected midprice over the investment horizon of agent i , informed traders pick their limit prices as:

Definition 1. *Buy limit prices:* $\hat{p}_{t+H^w,i} > p_t$.

$$\begin{aligned} p^{(1)} &= b^i = \min(a_1, \hat{p}_{\min}^i + 0.5\hat{\sigma}_t^i) \\ p^{(2)} &= a^i = \max(\hat{p}_{\max}^i - 0.5\hat{\sigma}_t^i, \hat{p}_{t+H^w,i} - 0.5\hat{\sigma}_t^i). \end{aligned} \quad (2.6)$$

Definition 2. *Sell limit prices:* $\hat{p}_{t+H^w,i} < p_t$.

$$\begin{aligned} p^{(1)} &= a^i = \max(b_1, \hat{p}_{\max}^i - 0.5\hat{\sigma}_t^i) \\ p^{(2)} &= b^i = \min(\hat{p}_{\min}^i + 0.5\hat{\sigma}_t^i, \hat{p}_{t+H^w,i} + 0.5\hat{\sigma}_t^i). \end{aligned} \quad (2.7)$$

Informed traders calculate limit prices for both markets. They will post orders in the market, where the expected return corresponding to the limit prices is largest.

2.2.3 Implementing the Tax Instruments

Here we briefly discuss the implementation of the different tax instruments. For the case of the FTT this is very easy: The expected return after tax equals $\tilde{r}^e = r^e - \chi_{FTT}$, with χ_{FTT} denoting the two-sided tax rate. Thus, the FTT has the effect of reducing the expected return of individual trades. In case that the expected return becomes negative, agents do not submit their order.⁴ In this way, trades with expected negative return are not submitted

⁴ Note that this implies risk neutrality, since agents only look at the expected returns. One could also incorporate a measure of risk (volatility) in order to incorporate risk aversion in the order submission process.

to the LOB, thereby reducing the amount of liquidity available. Note that the two counterparties involved in the trade share the tax burden equally among themselves.

For the FAT, the setting becomes slightly more complicated: in this case not a (gross) transaction value is taxed, but rather a (net) measure of trading profits. A flat tax is meant to take a certain fraction, denoted as χ_{FAT} , of the realized profits. Thus, the expected return (ignoring the order size) equals $\tilde{r}^e = (1 - \chi_{FAT})r^e$. By reducing the expected profit of individual orders, the FAT also has a direct effect on the order submission process. However, given that the sign of the expected return after tax for the FAT cannot change, the FAT has no effect on a single asset market. Therefore, we explicitly focus on the two-market case in the following, which is likely to yield much more interesting dynamics through tax evasion. Note that, in contrast to the FTT, the tax only applies to realized profits, i.e. only realized profits through successful trading are being taxed. In this way, defining the tax base is substantially more difficult as compared to the FTT case, since we need to track the profits of individual traders. This is easy in a simulated financial market, but likely to be far more complicated in reality.

In the following, the FAT is imposed only in cases when a trader is able to completely execute his two orders, i.e. made a quantifiable profit. Even though this definition may miss important cases,⁵ we still believe that this approach has a rough similarity to the proposed designs of FATs. In particular, many agents might have positive asset holdings, but we are reluctant to constantly put a tax on potential balance sheet profits.

Eventually, this brief discussion shows that it is much easier to define the tax base for the FTT as compared to the FAT. Policymakers should keep this in mind. In reality, the tax would be unlikely to be imposed on every single transaction, but rather on the total profits at the end of, say, each quarter. We should also stress that the values used for the FTT in Fricke and Lux (2013a) were typically very small percentage values (i.e. $<0.005\%$). Since the FAT is meant to take a certain fraction of the agents' profits, χ_{FAT} can take values between 0 (no tax) and 1 (all profits are taken away). For these extreme values, there will be no tax revenues whatsoever. For values in between, however, there will be positive tax revenues.

⁵ For example, consider a trader who bought assets, but is unable to execute his conditional take-profit order afterwards. The trader might still generate profits by submitting a new unconditional sell order, but that trade would not count as a profit realization as we define it.

2.2.4 Asset Demand and Cancellation

Without repeating the discussion in Fricke and Lux (2013a), we follow their approach in choosing the size of their limit orders and use the following demand function:

$$d_t^i = \begin{cases} \left[\bar{\alpha} \frac{C_t^i}{b^i} \right] & \text{if buy order} \\ \left[\bar{\alpha} S_t^i \right] & \text{if sell order,} \end{cases} \quad (2.8)$$

where $[x]$ denotes rounding towards minus infinity, and $\bar{\alpha} \in (0, 1]$ is an aggressiveness parameter determining the proportion of cash/assets an agent actually wants to use for investment. For simplicity we use the same parameter for all agents. When submitting market orders, agents will at most trade the amount available at the opposite best price, denoted as d^q , i.e.

$$\bar{d}_t^i = \min(d^q, d_t^i). \quad (2.9)$$

The results of Farmer *et al.* (2004) and Mike and Farmer (2008) suggest that the cancellation process is complex. Nonetheless, in order to keep the model simple, we assume a Poisson process for order cancellation and the exogenous cancellation probability π_{canc} is the same for all noise traders. For informed traders, we neglect this channel of order cancellation, since this would inject a significant amount of (additional) randomness to their strategies.

2.2.5 Trading Process and Time

We model the LOB at the highest possible frequency, but our basic analysis is concerned with the daily frequency. A few remarks on the timing of the model is in order here: The agents' time horizons are measured in terms of days, such that, for example, each agent approaches the market every $H^{w,i}$ days on average. At the beginning of each day t , we randomly reshuffle the list of agents willing to approach the market on this given day (active agents). Within the day, each active agent approaches the market and potentially submits new orders. Informed traders calculate their price expectations over the next $H^{w,i}$ days.

In everything that follows, we restrict ourselves to analyzing daily data only. A detailed description of the algorithm can be found in Fricke and Lux (2013a).

2.3 Pseudo-Empirical Results

In this section we present pseudo-empirical results from the model simulations. If not stated otherwise, the reported results are the average outcome

| Parameter | Value | Description |
|---------------------|----------------------------------|--|
| H^w | $\in \{20, 40, 60, \dots, 640\}$ | Admissible range, H^w |
| N | = 500 | Total number of agents (one market) |
| N_2 | = 900 | Total number of agents (two markets) |
| N_s | = 100 | Parameter for initial endowment |
| p_0, p_0^f | = 100 | Starting value: price/fundamental value |
| $\bar{\alpha}$ | = .10 | Order aggressiveness |
| Δ | = 10^{-3} | Tick size for the price |
| θ | = .20 | Fraction of informed traders (one market) |
| θ_2 | = .10 | Fraction of informed traders (two markets) |
| π_{canc} | = .01 | Cancellation probability (noise traders) |
| σ_ϵ | = 10^{-3} | Volatility for noise traders' expectations |
| χ_{FTT} | = $\{0, .0001, \dots, .0016\}$ | Tax rate of the FTT |
| χ_{FAT} | = $\{0, .05, \dots, .80\}$ | Tax rate of the FAT |

Tab. 2.1: Baseline parameter setting for the simulations. The subscript 2 indicates the adjusted parameter for the two market case.

of Monte-Carlo simulations of 22,500 days (time-steps), where we disregard an initial period of 2,500 days, each of which repeated 20 times with different random seeds. We also plot error bars, showing one standard deviation. Before investigating the effects of the different tax instruments, we start out by introducing the two market scenario in order to illustrate the changes inflicted by the second asset market.

The basic parameter values used in our simulations are shown in Table 2.1, which we borrow from Fricke and Lux (2013a). In this baseline scenario, agents' investment horizons are picked uniformly from the admissible range.

2.3.1 Baseline Scenario with Two Markets

In order to closely replicate the results of the one market case it is necessary to adjust the number of agents (N) and the fraction of informed traders (IT) θ . The reason for additionally adjusting θ is that the IT concentrate (most of the time) strongly in one of the two markets. Therefore we approximately doubled the number of noise traders (NT) to 810 (400 before) to preserve the level of liquidity and cut the fraction of IT $\theta_2 = .10$ in half. That makes in total $N_2 = 900$ agents compared to 500 before. Both trader groups face no restriction whatsoever in respect to submitting orders in any of the two markets.

Figure 2.1 shows the exemplary results for a single simulation run. The top left panel shows that both prices fluctuate around the fundamental value.

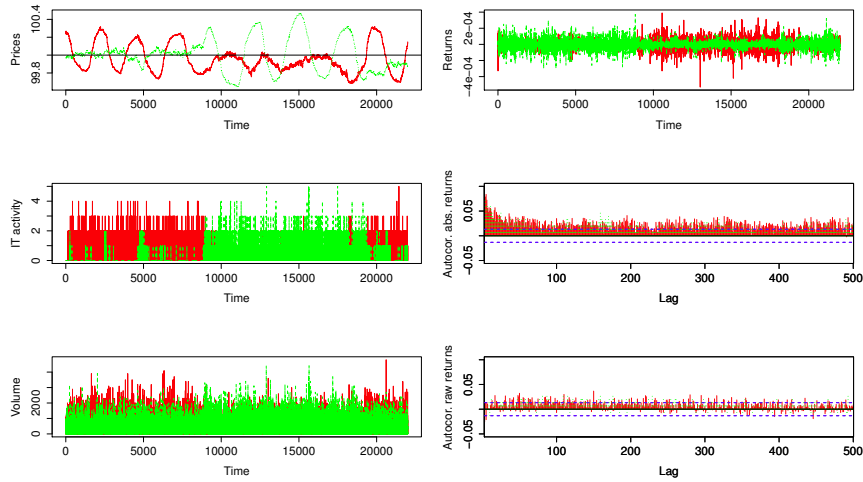


Fig. 2.1: Simulation results for a single run without tax. Plotted are market 1 (green) and market 2 (red). Top left: prices and fundamental value (black). Top right: log-returns. Center left: orders of IT. Center right: autocorrelation of absolute returns and 95% confidence interval. Bottom left: transaction volume. Bottom right: autocorrelation of absolute returns and 95% confidence interval.

Additionally, we alternatingly observe the building up and bursting of bubbles in one market and rather random price movements in the other market. This phenomenon is due to the concentration of ITs in the trending market (cf. center left panel for the number of informed traders active in market 1 and 2, respectively), which switch infrequently to the other market. In the top right panel we see the log-returns for the two assets, with a kurtosis of 11.576 and 18.088 and a skewness of -0.083 and $.237$. Hence, the return distribution is fat-tailed as observed on financial markets. Interestingly, the returns are seemingly higher in the market in which the ITs are predominantly inactive. However, this is only true in case of the extreme values and the mechanics leading to this are: the additional orders of ITs reduce the spread and accordingly more trades are executed in the market. So, there are fewer days without any trade in the trending market, but if a trade is executed in the noisy market it is more likely to result in a relatively high return. The center right panel shows that in both markets the autocorrelation of the absolute returns is significantly positive up to a lag order of around 40 and hovers around the significance threshold for higher lags. Thus, the

model is able to replicate to some extent the ‘volatility clustering’ observed on financial markets, yet the empirical findings indicate stronger autocorrelation and longer memory.⁶ In contrast, the autocorrelation of the raw returns in the bottom right panel is insignificant for most of the lags showing no clear pattern, except of the negative autocorrelation for small lags which is an indicator for the mean-reversion in asset prices. The bottom left panel shows that the market in which more orders arrive at the time due to the IT has higher transaction volume, while the NT secure that both markets are liquid at all time.

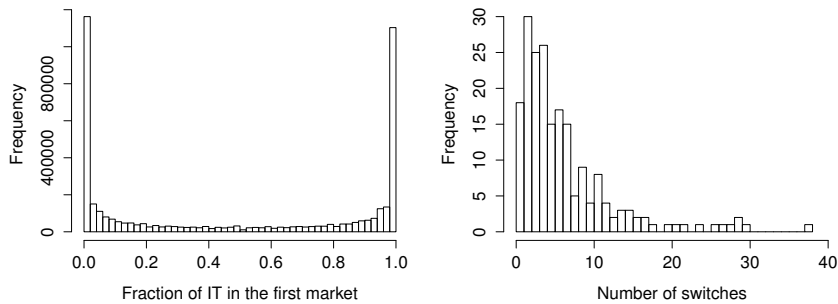


Fig. 2.2: Simulation results for 200 runs without tax. Left: histogram of the fraction of IT in the 1st market (aggregated over the last 50 time steps) per time step. Right: histogram of the number of switches per run.

In order to analyze the switching behavior of the ITs more thoroughly, we aggregate the IT activity (cf. center left panel 2.1) over the last 50 days. This is necessary, since on each day very few IT are active producing a noisy signal. Out of these values we calculate the fraction of IT which have been active in the first market to obtain an easy to interpret measure for concentration. In the left panel of Figure 2.2 the histogram over 200 simulation runs (more than 4 million observations) displays that the distribution is bimodal and symmetric around the mean (.5). The two very distinct modes are at the edges of the distribution with 27.1% (25.8%) of all observations having values below .02 (above .98) and can be seen as the majority states of the system. As we will discuss below, these findings bear some similarity to Kirman (1993).⁷ The bimodal distribution of IT activity in either market is due to

⁶ Fricke and Lux (2013a) also report that the model has problems in replicating realistic levels of volatility clustering.

⁷ Note that the equilibrium distribution of the process can, in contrast to Kirman (1993),

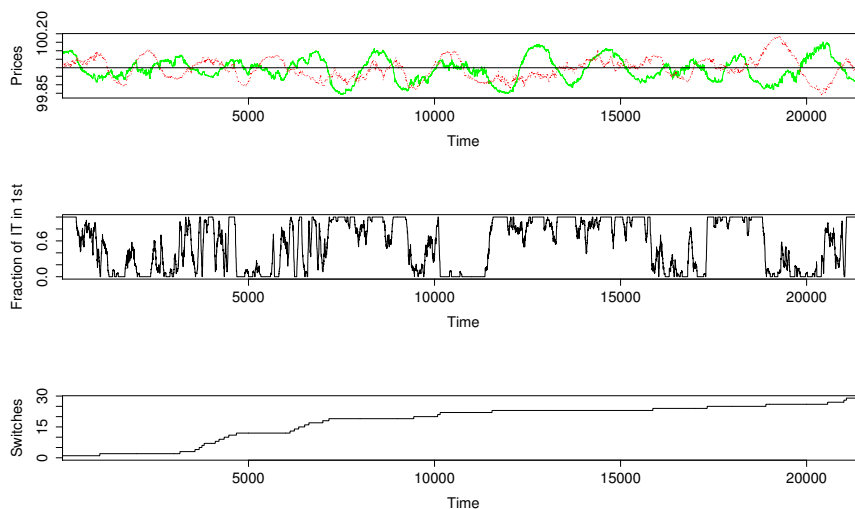


Fig. 2.3: Simulation results for a single run without tax. Top: prices market 1 (green), market 2 (red) and fundamental value (black). Middle: the fraction of IT in the 1st market (aggregated over the last 50 time steps). Bottom: the number of cumulated switches of IT between the two markets.

the reinforcing activity ‘creating’ the cyclical price behaviour in the first place and attracting, via the corresponding strong trends (chartist component) and/or high deviations from the FV (fundamentalist component) around the turning points, ever more ITs to the same market.

Defining the majority states as $.9 (.1)$, we count switches as follows: suppose we are in a majority state, and 95% of the ITs trade in market one (in this sense, the ITs are understood to be ‘locked into’ the first market). A switch involves that the fraction of ITs in market 1 falls below 10%. The histogram of the number of switches over the 200 runs in the right panel of Figure 2.2 indicates that up to 38 switches can occur in a single run, yet still in 18 runs (9%) no switch and in 146 runs (73%) less than seven switches take place. To better understand the mechanics leading to the relatively high number of switches, Figure 2.3 shows the prices, the fraction of IT and the accumulated switches over time (top to bottom) of an exemplary run with 28 switches. The middle panel illustrates that the fraction of IT is often either close to one or zero, yet these majority states are fragile. In the

not be solved analytically but only be approximated by the histogram in the left panel of Figure 2.2.

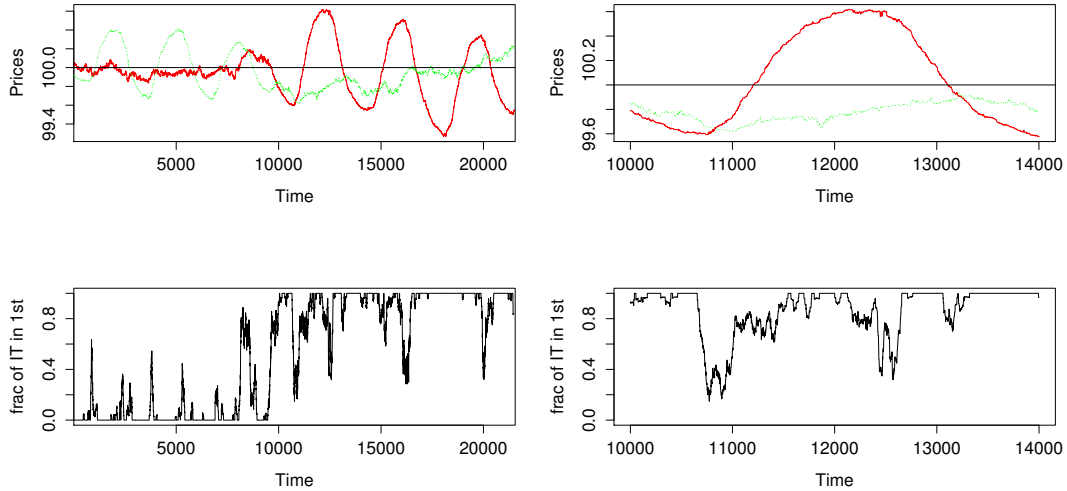


Fig. 2.4: Simulation results for a single run without tax. Top: prices market 1 (green), market 2 (red) and fundamental value (black). Bottom: the fraction of IT in the 1st market (aggregated over the last 50 days). Left: the complete run. Right: from day 10,000 to 14,000.

period between time step 3,500 and 4,700, a total number of nine switches occur while both price series are accordingly in a mixed state between cyclical boom/bust and random walk (cf. top panel). Hence, the turning away from a majority state seems to destabilize the system and so additional switches are becoming more likely until a new stable regime is reached.

Figure 2.4 illustrates the seemingly coordinated switches of ITs: the top row shows the price series of the two markets, while the bottom row shows the fraction of IT in the first market. The complete simulation run is shown on the left, on the right we show the period between day 10,000 and 14,000. In this run a single switch took place somewhere around day 9,500 in which the majority of the ITs moved their trading from the second to the first market. A closer look at the time series slightly after the switch reveals two conditions under which ITs are more likely to switch markets: first, at the end of the turning points of the cyclical price process when the price stops diverting and ultimately returns to the FV, which happens twice around day 10,800 and day 12,500. The reason is that the ‘trend component’ in Equation (2.3) is very small and so it is more likely that especially short-term oriented traders (chartists) expect a higher return from the second market. Second, the price

series of the first market intersects with the FV which it does around days 11, 100 and 13, 100. This time the ‘fundamental component’ in Equation (2.3) is close to zero and therefore longer-term oriented traders (fundamentalist) are more likely to switch markets. Hence, at these two points the positive feedback effects of IT to stay in the same market are weakened and switches might occur. A strong negative correlation of $-.607$ between the number of switches and the average distortion over both markets in a single run indicates that big swings make the ongoing regime more persistent. This stabilizing effect of big bubbles should have two reasons. On the one hand, yet the according component is still weakened as described above the probability for a switch at each crossroad is lowered, because the other component is strengthened in both cases. The trend is stronger if the price intersects with the FV and at the turning points the misalignments are larger. On the other hand the swings need on average more time and so the number of crossroads per run is reduced.⁸Note that due to path-dependence it is impossible to tell why the cyclical bubbles are larger in some simulation runs.

In many respects, our findings are similar to the findings of Kirman (1993). In both models no exogenous shock is necessary to alter the behaviour of the majority of traders and flip between the two majority (quasi-equilibrium) states, but if one of this states is reached it is likely to be stable for a certain time. However, in contrast to Kirman’s ants, we do not require direct interactions between the individual traders. Hence, in this framework the observed ‘herding’ behaviour is achieved without any kind of coordination among the agents, since they solely rely on the expected profits of their fixed strategy.

To sum up, the model output for the two markets can replicate some of the stylized facts of financial markets and is in this respect comparable to the one market case in Fricke and Lux (2013a). At the same time the second market produces rich dynamics among the IT being highly concentrated in one most of the time, but sometimes abandoning this equilibrium kind of state at certain crossroads until a new one is reached.

2.3.2 *Effects of a FTT in the Two-Market Case*

This section analyses the effects of the FTT in the two-market setting, for varying tax rates and compares to results to the one market case analysed in Fricke and Lux (2013a). We restrict ourselves to the scenario in which the tax is solely introduced in one market and so all traders have the possibility to trade in an untaxed market (tax haven). The reason for this is that

⁸ The effects are visible if we look at the two example runs in Figure 2.3 and Figure 2.4.

the introduction of the tax in both markets leads to very similar outcomes as in the one market scenario, while the volatility is increased due to the switches of the majority of the IT between the markets. Most interesting to us is clearly the behavior of ITs, as NTs are simply posting random orders. Already at this point, we should stress that the effects of the FTT with a tax haven are for the taxed market similar to those with only one market, because ‘chartist’ traders are reluctant to trade in the taxed market even for very small tax rates.

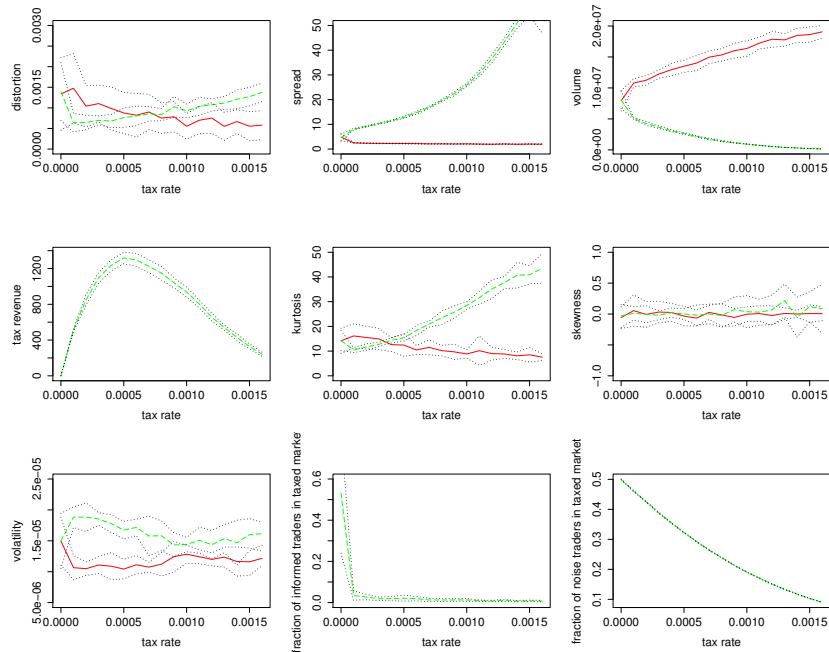


Fig. 2.5: Monte-Carlo simulation results with dependence on χ_{FTT} . Plotted are mean values for 20 runs (solid lines) taxed market (green) tax haven (red), plus and minus one standard deviation (dashed lines). Top left: distortion from FV. Top middle: spread. Top right: transaction volume. Center left: tax revenue. Center middle: kurtosis. Center right: skewness. Bottom left: volatility (absolute price change). Bottom middle: fraction of orders by IT in the taxed market. Bottom right: fraction of orders by NT in the taxed market.

The results of our Monte-Carlo simulations are summarized in Figure 2.5, where the green lines correspond to the taxed market and the red lines to

the tax haven. The top left panel shows that the average distortion in the taxed market is being reduced for small tax rates, but tends to increase for larger values and exceeds the initial value for large tax rates. In contrast, the average distortion in the untaxed market is clearly negatively related to the tax rate. Note however, that the error bands for both series are rather large, making them indistinguishable from a statistical perspective for intermediate tax rates. The top middle panel shows that liquidity decreases (increases) significantly in the taxed (untaxed) market. Similarly, market activity is being reduced significantly in the taxed market, at the same time increasing roughly proportionally in the untaxed market, cf. top right panel of Figure 2.5. Interestingly, as in the one-market case we find an inverted U-shaped relationship between tax revenues and the tax rate (Laffer curve). The center middle panel shows that the average kurtosis for the return distribution strongly increases in the taxed market, while it decreases in the untaxed market. Thus, extreme price changes become more (less) likely in the taxed (untaxed) market.

The volatility in the bottom left panel and the distortion in the top left panel appear to be minimally affected by the FTT. However, the high kurtosis implies not only many extreme returns, but also more values close to the center of the distribution (zero). Hence, the relatively low volatility for higher tax rates is a fragile and somewhat misleading result, since the market price will strongly react when a trade occurs. The bottom center panel shows that the fraction of orders submitted from ITs in the taxed market is already very small for a very low tax rate. Furthermore, the investigation of single simulation runs displays that the periods of relatively high activity of IT are accompanied by large distortions from the FV and can be attributed to ‘fundamentalists’ traders anticipating the price to return to the FV. After this drop the decline of the order activity of IT continues until the tax rate reaches .07%. For even higher tax rates up to .16% the order activity of IT fluctuates around 1%, since the distortions created by the remaining NT are at points still high enough to attract ‘fundamentalists’, albeit the high tax rate. The absence of ‘chartists’ stops the aforementioned switching process and we therefore observe the boom and bust cycle of the price solely in the tax haven. Moreover, since all IT have both components (trend and reversion to FV) the absence of ‘chartist’ traders and the corresponding smaller trends let the ‘fundamentalists’ react earlier to misalignments. This partially explains that not only the average distortion but also the maximum and highest 10% of the distortions shrink for very small tax rates in the taxed markets. Note that the trend component in the expected returns for ‘fundamentalists’ could be seen as the fear of the trend (away from the FV) to continue, while the absence of trends strengthens its confidence in the fundamental analysis.

Overall the properties for the taxed market are similar to the one market case in Fricke and Lux (2013a), where the FTT was found to be somewhat stabilizing for small tax rates, but less so for larger ones. However, we should stress that our findings for the tax haven strongly differ from those in the literature, as basically all measures of market quality improve.

2.3.3 *Effects of FAT*

Here we move to the analysis of the two market case for the FAT. We should stress that the one market case is trivial for the FAT, since the behavior of both trader groups is (almost) unaffected by the tax (as long as it is smaller than 100%), because the expected return stays always positive. Therefore, the results are basically identical to the scenario without the tax (unreported result). Apart from generating tax revenues, the FAT has no effect on market quality in the one market case. Hence, we turn to the two market case in order to capture evasive effects of the traders in response to the FAT.⁹

Figure 2.6 shows that the effects of the FAT differ substantially from those of the FTT. The distortion in the taxed market visible in the top left panel hovers close to the initial level for low to medium tax rates, but in contrast to the FTT it never falls below its starting value. The spread in the top middle panel is for the taxed market very close to the one of the tax haven up to a medium tax rate and increases steadily for higher values. Additionally it gets highly volatile for high tax rates above. The transaction volume in the top right panel is despite the retreat of the NT similar for both markets up to medium tax rates, because it is compensated by the slightly higher participation of IT. For higher values the gap between the transaction volume of both markets increases constantly. The tax revenue in the center left panel is far from a Laffer curve relationship, but rather tends to slowly increase in the tax rate. Overall, the tax revenues are much lower compared to the FTT. The kurtosis in the center middle panel increases with the tax rate and its value as well as its volatility explodes for very high tax rates. The volatility in the bottom left panel fluctuates around the starting values for low to medium tax rates in both markets, yet unlike to the FTT the volatility increases in the taxed market for very high tax rates.

The activity of IT in the taxed market visible in the bottom middle panel slightly increases, yet the error bars display that the variation in between

⁹ Moreover, the scenario in which both markets are taxed equally gives qualitatively the same results as the one market case and is practically identical to the two market case without tax except for the tax revenues. Again, the volatility is higher than in the one market case due to changing concentration of IT.

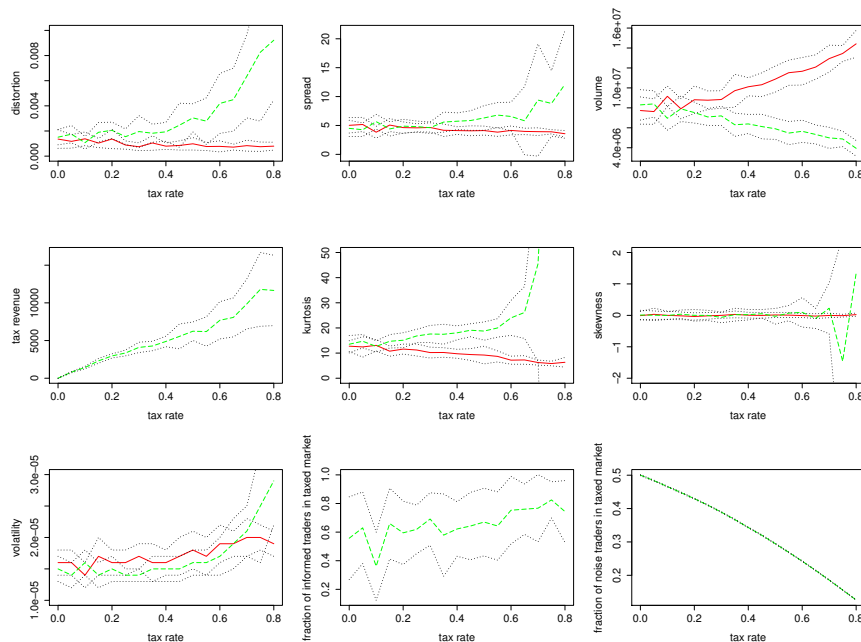


Fig. 2.6: Monte-Carlo simulation results with dependence on χ_{FAT} . Plotted are mean values for 20 runs (solid lines) taxed market (green) tax haven (red), plus and minus one standard deviation (dashed lines). Top left: distortion from FV. Top middle: spread. Top right: transaction volume. Center left: tax revenue. Center middle: kurtosis. Center right: skewness. Bottom left: volatility (absolute price change). Bottom middle: fraction of orders by IT in the taxed market. Bottom right: fraction of orders by NT in the taxed market.

the single runs is high. In case of the NT the limit prices are again not affected by any information obtained during the trading process, except for the current price. The bottom right panel shows that their participation in the taxed market, accordingly declines in a similar fashion to the FTT. This is important since the similar behaviour of the NT relative to the FTT is necessary to meaningfully compare both tax instruments. In case of the tax haven the increasing number of NT and relatively stable activity in respect to different tax rates of the IT leads to a decrease of the kurtosis and the return distribution is getting closer to the gaussian case.

The main reason for the differences compared to the FTT is the much higher activity of the IT, since the FAT does not adversely affect short term

investors. The continuing trading of IT in combination with the orderly retreat of the NT creates stronger trends and this (over)compensates the negative effect of a higher tax rate for IT. Hence, the rare switches of the majority of IT and the corresponding alternating bubble process between the two markets are conserved. Again, depending on the timing of these switches the activity of IT varies in between single runs dramatically and this leads to the higher volatility of almost all properties compared to the FTT. The high distortion and the in general extreme results in the taxed market for high tax rates are the result of persistent trading of IT in a market environment with decreasing liquidity (fewer NT) and accordingly increasing spreads. This implies that the lower level of liquidity and the accompanied higher spreads strengthen the ‘chartist’ trader in observing and exploiting stronger trends until the ‘fundamentalist’ are driving the price back to its FV. Hence, the most dangerous situation seems to be a market environment with low liquidity and still high participation of (short-term) ‘chartist’ traders

Somewhat surprisingly, given the absence of pronounced effects in the one market case, the market quality of the taxed market tends to decrease in the tax rate. In contrast, apart from larger trading volumes, there is basically no effect on the tax haven. In the end, it appears to be more lucrative to not introduce the tax.

2.4 *Discussion and Conclusion*

In this paper, we extended the framework of Fricke and Lux (2013a) by incorporating a second asset market. The introduction of a second market leads to interesting dynamics among the ITs, because they interchangeably agglomerate in one market. Hence, the cyclical bubble process known from the one market model is limited to this market, while the other market shows seemingly random price movements. The switches of the majority between the two markets are more likely to occur at certain points. These crossroads are if the price intersects with the FV and the ‘fundamentalist’ component is close to zero or when the overshooting/undershooting of the price stops and it turns back to the FV and the trend component is close to zero. That the participation of IT in one market reinforces the attractiveness of this market for other ITs in the future can be interpreted as ‘herding’ behavior without the need for coordination between the agents or exogenous effects. On the other hand the concentration of the ITs in one market leads to returns showing no correlation whatsoever between the markets and hence the model is unable to produce any kind of ‘contagion’ effects. At the same time, the

two market model is able to replicate some of the well-known stylized facts, such as the martingale property, fat tails and to some extent the long memory process of absolute returns.

The most striking in respect to the tax instruments is that the FAT and FTT produce different results due to the behavioural changes of the IT, while the NT in both cases orderly retreat from the taxed market and therefore are not the main drivers of the results. The FTT adversely affects short-term oriented traders. These drop out of the taxed market already for very small tax rates and so the building up and bursting of speculative bubbles is restricted to the tax haven. The ‘fundamentalist’ traders still participate, if the misalignments are big enough, and the NT have almost no preference for small tax rates and ensure a high level of liquidity. This leads to smaller average and extreme distortions, because the ‘fundamentalists’ expect the price to return back to the FV for smaller misalignments, since their own trend following component produces smaller values. Furthermore, there is no systematic pressure against the ‘fundamentalist’ strategy without ‘chartists’. The corresponding evaporation of liquidity leads to more extreme returns. The relatively low level of volatility for this situation is misleading, since it is driven by many days without a single trade. Overall, the qualitative properties of the taxed market are remain surprisingly robust in the presence of a tax haven.

On the other hand the FAT does almost not affect the activity of the IT in the taxed market even for high tax rates for two reasons: first, all traders independent of their investment horizon are equally penalized by paying a fraction of their (expected) profits. Second, the trends and expected profits for ‘chartist’ are potentially higher when the liquidity is reduced for high tax rates. Hence, the alternating cyclical bubble process in one of the markets and a ‘seemingly’ random walk with boundaries set by the ‘fundamentalist’ traders in the other with random switches between the two markets continues. This leads in general to more volatile results, because the single runs differ in respect to the time the IT are concentrated in the two markets. Overall, the properties of the taxed market are quite persistent for small tax rates, yet the FAT appears to show no clear stabilizing impact for any tax rate. In fact, for very high tax rates the taxed market displays high levels of both distortion and volatility plus many extreme returns. Therefore, the worst situation for the taxed market is if the IT, especially the ‘chartist’ contribute a large fraction of the orders in the taxed market, because most of the NT already switched to the tax haven.

The FTT affects short-term ‘chartist’ traders worst and therefore has the potential to stabilize the taxed market by reducing the distortions. This potentially beneficial effect should not be underrated since misalignments

cause wrong allocation of goods and negatively influence the real sectors of the economy. A look at financial data of the last years stresses this, since even aggregated stock indices like the German DAX recently reached 7500 points the third time since 2000 while falling below 4000 points in between.

However, the issue of financial taxes is not limited to stock and bonds markets, since the global trend towards more financialization included more and more resources into financial markets in the last decades. An example are commodities which became a new asset class and are largely affected by speculative behaviour nowadays. This is important, since there are important differences in the market microstructure of various financial markets and it is necessary to take this into account, when we want to address the question which tax instrument should be used. In LOB settings, the FTT should be potentially advantageous at least for small tax rates. However, in a market-maker setting even a very small FTT might erase their profitability and let the market break down. This seems to be a big advantage of the FAT, because it has less impact on specific groups and should work more equally for different market structures.

To sum up, both tax instruments tend to create substantial tax revenues, yet decrease the liquidity and become destabilizing for high tax rates. The FTT might be able to even stabilize specific financial markets by reducing the misalignments, yet has to be specifically adjusted for different financial markets. The coming FTT of the EU is in line with this suggestion by setting different rates for bond and stock transaction. The FAT on the other hand should be implemented on all speculative profits made by financial institutions and therefore could additionally be an incentive for banks to focus more on credit intermediation rather than speculative activity. Hence, both tax instruments would make the financial sector contribute to the large cost induced by the global financial crisis, while the governments should be cautious in setting the tax rates. In the end, policymakers might consider the implementation of a general FAT and an additional FTT for specific markets.

3. NETWORK ANALYSIS OF THE E-MID OVERNIGHT
MONEY MARKET: THE INFORMATIONAL VALUE OF
DIFFERENT AGGREGATION LEVELS FOR INTRINSIC
DYNAMIC PROCESSES

3.1 Introduction and Existing Literature

Interbank markets are crucial for the functioning of the economy. The exchange of money on a very short-term basis allows banks to manage liquidity risks, cf. Allen and Gale (2000). However, as painfully illustrated by the global financial crisis (GFC) in 2007/08, creating links at the micro-level may generate systemic risk at the macro-level.¹ Thus, the interbank network, based on connections in terms of credit relationships, is a complete representation of the liquidity management decisions of financial institutions. Clearly the structure of such a network is important for its stability. The economy depends on stable interbank markets, since short-term money market rates affect those of longer maturities and thus the real economy. From this viewpoint, it appears quite surprising that the economic profession has not been concerned much with the functioning of interbank markets until recently. The usual focus is on the overnight segment of the interbank deposit market, since it tends to be the largest spot segment of money markets. In this paper, we analyze the network properties of the Italian e-MID (electronic market for interbank deposits) data based on overnight loans during the period 1999-2010.

Most existing studies on the structure of real interbank markets have been conducted by physicists trying to get an idea of the topology of different interbank markets. Examples include Boss *et al.* (2004) for the Austrian interbank market, Inaoka *et al.* (2004) for the Japanese BOJ-Net, Soramäki *et al.* (2007) for the US Fedwire network, Bech and Atalay (2010) for the US Federal funds market, and De Masi *et al.* (2006) and Iori *et al.* (2008) for the Italian e-MID (electronic market for interbank deposits). The most important findings reported in this literature are: (1) most interbank networks are quite large (e.g. more than 5000 banks in the Fedwire network), (2) interbank networks are sparse, meaning that only a minority of all possible links do actually exist, (3) degree distributions appear to be scale-free (with coefficients between 2-3), (4) transaction volumes appear to follow scale-free distributions as well, (5) clustering coefficients are usually quite small, (6) interbank networks are small worlds and (7) the networks show disassortative mixing with respect to the bank size, so small banks tend to trade with large banks and vice versa.

In this paper we also analyze the Italian interbank market. We should

¹ In the literature, the traditional approach is to focus on the optimal liquidity management of a particular financial institution, see e.g. Brodt (1978) and Ho and Saunders (1985). Recently, Brunnermeier and Pedersen (2009) introduced the distinction between funding liquidity and market liquidity in a theoretical model, showing that their relationship may give rise to systemic crises.

stress that the e-MID data are the only interbank data which can be purchased freely without any restrictions. In contrast, getting access to similar datasets for other markets is usually far more complicated. Most relevant for our study are the two papers on the e-MID. The authors analyze daily networks from 1999-2002, finding intradaily and intramonthly seasonalities. The most important conclusion is that the networks appear to be random at the daily level. This finding is in stark contrast with the findings of preferential lending relationships in the Portuguese interbank market by Cocco *et al.* (2009). In the present paper, we are mostly concerned with matching these seemingly incompatible findings, by showing that the aggregation period has an effect on the informational value of the underlying networks. The main result is that daily networks indeed feature a substantial amount of randomness and cannot be considered as being representative for the underlying ‘latent’ network. This is illustrated on the basis of a number of network statistics, which are compared to those of random networks. Furthermore, we find a substantial amount of asymmetry in the network. Last but not least, we find that the GFC can be identified as a significant structural break for many network measures. While our findings do not solve the issue of what is the ‘optimal’ structure of the interbank network (whatever optimal means in this context), our analysis is a first, albeit very important step, in understanding the structure and dynamics of the banking system. Equipped with this knowledge, we can hope to build more realistic models in the future, which could be used to provide policy recommendations for the optimal management of interbank networks.

The remainder of this paper is structured as follows: section 3.2 gives a brief introduction into (interbank) networks, section 3.3 introduces the Italian e-MID trading system and gives an overview of the data set we have access to. Section 3.4 describes our findings and section 3.6 concludes and discusses the relevance of these findings for future research.

3.2 Networks

A network consists of a set of N nodes that are connected by M edges (links). Taking each bank as a node and the interbank positions between them as links, the interbank network can be represented as a square matrix of dimension $N \times N$ (data matrix, denoted \mathbf{D}).² An element d_{ij} of this matrix represents a gross interbank claim, the total value of credit extended by bank i to bank j within a certain period. The size of d_{ij} can thus be seen as a

² In the following, matrices will be written in bold, capital letters. Vectors and scalars will be written as lower-case letters.

measure of link intensity. Row (column) i shows bank i 's interbank claims (liabilities) towards all other banks. The diagonal elements d_{ii} are zero, since a bank will not trade with itself.³ Off-diagonal elements are positive in the presence of a link and zero otherwise.

Interbank data usually give rise to directed, sparse and valued networks.⁴ However, much of the extant network research ignores the last aspect by focusing on binary adjacency matrices only. An adjacency matrix \mathbf{A} contains elements a_{ij} equal to 1, if there is a directed link from bank i to j and 0 otherwise. Since the network is directed, both \mathbf{A} and \mathbf{D} are asymmetric in general. In this paper, we also take into account valued information by using both the raw data matrix as well as a matrix containing the number of trades between banks, denoted as \mathbf{T} . In some cases it is also useful to work with the undirected version of the adjacency matrices, \mathbf{A}^u , where $a_{ij}^u = \max(a_{ij}, a_{ji})$.

As usual, some data aggregation is necessary to represent the system as a network. In this paper we compare different aggregation periods.

3.3 The Italian Interbank Market e-MID

The Italian electronic market for interbank deposits (e-MID) is a screen-based platform for trading of unsecured money-market deposits in Euros, US-Dollars, Pound Sterling, and Zloty operating in Milan through e-MID SpA.⁵ The market is fully centralized and very liquid; in 2006 e-MID accounted for 17% of total turnover in the unsecured money market in the Euro area. Average daily trading volumes were 24.2 bn Euro in 2006, 22.4 bn Euro in 2007 and only 14 bn Euro in 2008. We should mention that researchers from the European Central Bank have repeatedly stated that the e-MID data is representative for the interbank overnight activity, cf. Beaupain and Durré (2012).

Available maturities range from overnight up to one year. Most of the transactions are overnight. While the fraction was roughly 80% of all trades in 1999, this figure has been continuously increasing over time with a value of more than 90% in 2010.⁶ As of August 2011, e-MID had 192 members from

³ This is true when we think of individual banks as consolidated entities.

⁴ Directed means that $d_{i,j} \neq d_{j,i}$ in general. Sparse means that at any point in time the number of links is only a small fraction of the $N(N-1)$ possible links. Valued means that interbank claims are reported in monetary values as opposed to 1 or 0 in the presence or absence of a claim, respectively.

⁵ The vast majority of trades (roughly 95%) is conducted in Euro.

⁶ This development is driven by the fact that the market is unsecured. The recent financial crisis made unsecured loans in general less attractive, with stronger impact for longer maturities. See below. It should be noted, that there is also a market for secured

EU countries and the US. Members were 29 central banks acting as market observers, 1 ministry of finance, 101 domestic banks and 61 international banks. We will see below that the composition of the active market participants has been changing substantially over time. Trades are bilateral and are executed within the limits of the credit lines agreed upon directly between participants. Contracts are automatically settled through the TARGET2 system.

The trading mechanism follows a quote-driven market and is similar to a limit-order-book in a stock market, but without consolidation. The market is transparent in the sense that the quoting banks' IDs are visible to all other banks. Quotes contain the market side (buy or sell money), the volume, the interest rate and the maturity. Trades are registered when a bank (aggressor) actively chooses a quoted order. The platform allows for credit line checking before a transaction will be carried out, so trades have to be confirmed by both counterparties. The market also allows direct bilateral trades between counterparties.

The minimum quote size is 1.5 million Euros, whereas the minimum trade size is only 50,000 Euros. Thus, aggressors do not have to trade the entire amount quoted.⁷ Additional participant requirements, for example a certain amount of total assets, may pose an upward bias on the size of the participating banks. In any case, e-MID covers essentially the entire domestic overnight deposit market in Italy.⁸

We have access to all registered trades in Euro in the period from January 1999 to December 2010. For each trade we know the two banks' ID numbers (not the names), their relative position (aggressor and quoter), the maturity and the transaction type (buy or sell). As mentioned above, the majority of trades is conducted overnight and due to the GFC markets for longer maturities essentially dried up. We will focus on all overnight trades conducted on the platform, leaving a total number of 1,317,679 trades. The vast majority of trades (1,215,759 in total) involved two Italian counterparties. The large sample size of 12 years allows us to analyze the network evolution over time. For example, at the quarterly level we end up with 48 snapshots of the network.

loans called e-MIDER.

⁷ The minimum quote size could pose an upward bias for participating banks. It would be interesting to check who are the quoting banks and who are the aggressors. Furthermore it would be interesting to look at quote data, as we only have access to actual trades.

⁸ More details can be found on the e-MID website, see <http://www.e-mid.it/>.

3.4 Results

In this section, we look at the network structures formed by interbank lending over various horizons of time aggregation of the underlying data. We will see that comparing various network measures at different levels of time aggregation reveals interesting features suggestive of underlying behavioral regularities. Given that most studies focus on overnight data, it has become quite standard to focus on networks constructed from daily data. Here we find that, at least for the Italian interbank network, it may be more sensible to focus on longer aggregation periods, namely monthly or quarterly data. We also discuss in how far the network structure has changed (and, in how far it has remained intact) after the default of Lehman Brothers in September 2008.

3.4.1 General Features

In total, 350 banks (255 Italian and 95 foreign) were active at least once during the sample period. However, the number of active banks changes substantially over time as can be seen from the left panel of Figure 3.1.⁹ We see a clear downward trend in the number of active Italian banks over time, whereas the additional large drop after the onset of the GFC is mainly due to the exit of foreign banks. The center panel shows that the decline of the number of active Italian banks went along with a relatively constant trading volume in this segment until 2008. This suggests that the decline of active Italian banks was mainly due to mergers and acquisitions within the Italian banking sector. Given the anonymity of the data set, it is impossible to shed more light on this interesting issue. The overall upward trend of trading volumes was due to the increase of active foreign banks until 2008, while their activities in this market virtually faded away after the onset of the crisis. Interestingly, the average volume per trade tends to increase over time, as can be seen from the strong negative trend in the total number of trades (right panel in Figure 3.1), at least for the Italian banks.

An interesting question in this regard is, who trades with whom. Figure 3.2 illustrates this for the number of trades (top) and the transacted volume (bottom) by country. For example, the green lines show the total number of trades (traded volumes) of foreign banks lending money to Italian banks, relative to all outgoing trades of foreign banks. Similarly, the blue lines show the total number of trades (traded volumes) of money flowing between Italian banks, as a fraction of all outgoing trades of Italian banks. The general

⁹ Similar developments are reported by Bech and Atalay (2010) for the federal funds market.

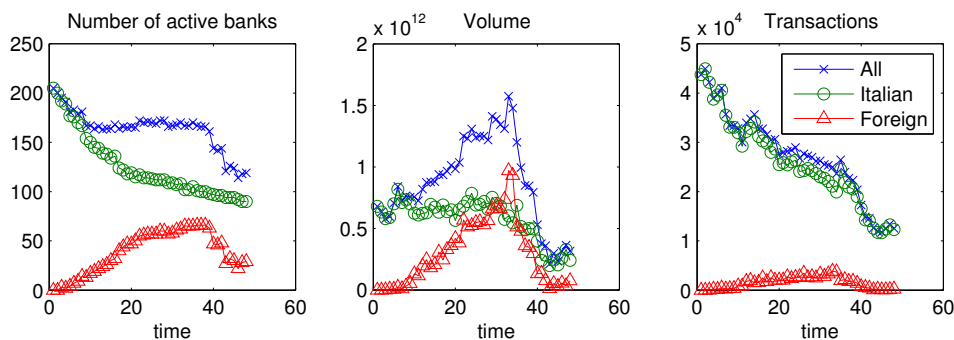


Fig. 3.1: Number of active banks (left), traded volume (center) and number of trades (right) over time (quarterly). We also split the traded volume into money lent by Italian and foreign banks, respectively. The same holds for the number of trades.

patterns are the same for both Figures: Italian banks lend most of the time to other Italian banks (99.31% on average) and only a negligible amount to foreign banks (0.61% on average). This pattern is remarkably stable over time. In contrast, at the beginning of the sample period, foreign banks mostly used the market in the absence of (many) other foreign counterparties to lend money to Italian banks. This has changed over time and foreign banks later on mostly used the platform to trade with other foreign banks. It is not quite clear why this is the case, the underlying trend seems to point towards structural changes altering the (foreign) banks' behavior. For many research questions, one should therefore only use the subsample of Italian banks. In most of what follows, we stick to this choice.

This leads us to a first glance at the network structure. Figure 3.3 shows the banking network formed by the 119 active banks (89 Italian) in the last quarter of 2010.¹⁰ The network mainly consists of two components: The very dense part formed by the Italian banks (circles) on the right-hand side and the far less interconnected foreign banks (triangles) on the left-hand side. The higher activity of the Italian banks is not represented in terms of the volume traded. We use total outgoing volume as a proxy for banks size and group the banks into 4 classes according to which percentile (30th, 60th, 90th or above) they belong to. This attribute is shown in the Figure as the size and the brightness of the nodes. We should note that 3 out of 12 banks of

¹⁰ The Figure was produced using visone, <http://www.visone.info/>, by Brandes and Wagner (2004).

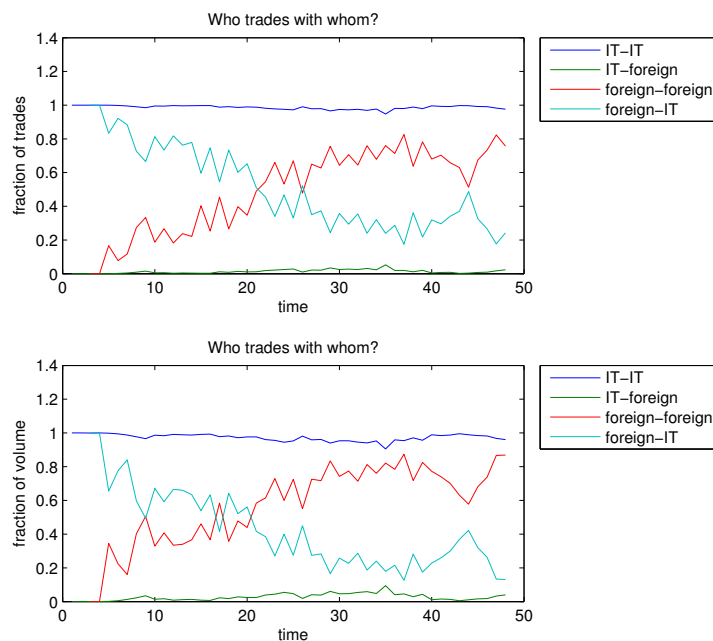


Fig. 3.2: Fraction of trades (top) and traded volume (bottom) between banks from different countries.

group 4 are foreign banks which is in line with their fraction of the total banks (30 out of 119). Hence, foreign banks trade less on average (both in terms of volume and number of trades), however, the volume per trade is higher.

It is also interesting to highlight some specific features of the trading behavior of individual banks in this particular quarter, since not all banks use the market in the same way:¹¹ There are 14 banks with zero in-degree and 29 banks with zero out-degree.¹² Surprisingly these banks are quite heterogeneous and not, as one might expect, just small banks. As an example the highest overall transaction volume of 58.6 bn Euro for a single bank, and therefore roughly 9.3% of the total trading volume, was traded by a German bank borrowing this sum in 90 trades from 8 counterparties. Another interesting case is an Italian bank trading only with one counterparty, lending this other bank 5.02 billion Euro in 76 trades, whereas borrowing just 0.03 billion in 3 trades. Even though these special relationships are quite interesting,

¹¹ For a detailed analysis of the trading strategies in the e-MID, see Fricke (2012).

¹² In a companion paper, we focus explicitly on fitting the degree distribution, see Fricke and Lux (2013b). The main findings are: (1) the degree distributions are unlikely to follow power-laws, and (2) the in- and out-degrees do not follow the same distribution.

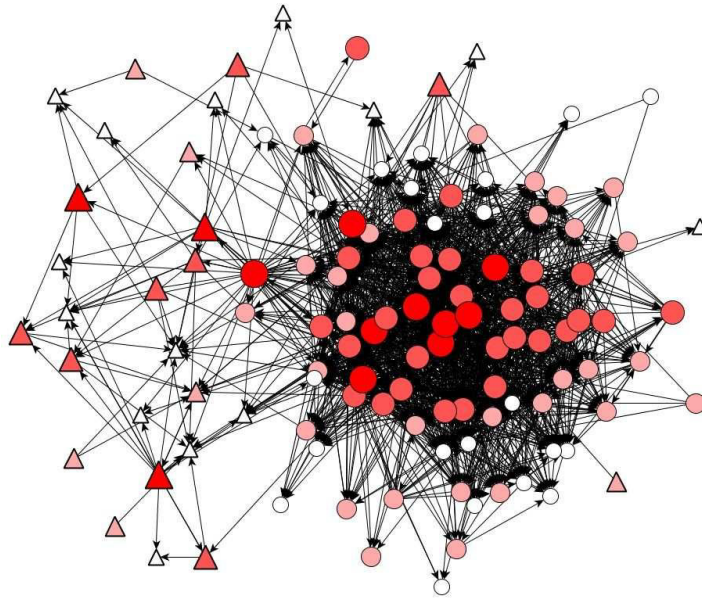


Fig. 3.3: The banking network in the 4th quarter of 2010: triangles are foreign banks. The size of the node as well as the brightness of the red color indicate the size in terms of volume lent.

the anonymity of the data set makes it impossible for us to say more on the particular relationships that might lead to these interesting outcomes.

As should be clear from the discussion above, we will mostly focus on the (sub)network formed by Italian banks only. Table 3.1 summarizes, for this subset, some basic descriptive statistics for the four different aggregation periods under investigation to get a first impression of the influence of the aggregation level.¹³ Note that the number of active banks, trading volume, trades and links has to increase with the aggregation period. The difference is, however, that volume and trades are simply multiples of the daily values, whereas the number of links grow at a much lower rate. This already indicates that banks prefer to trade with specific counterparty. The trades per link show that even on the daily level on average every fourth link already represents two trades rather than only a single one. After this broad overview of the market and the ongoing interactions, we turn to the question of a sensible aggregation period for the network.

3.4.2 Density

The density ρ of a network is defined as the number of existing links (M) relative to the maximum possible number of links. It can be calculated as

$$\rho = \frac{M}{N(N-1)}. \quad (3.1)$$

Figure 3.4 illustrates the evolution of the density for four different aggregation periods (day, month, quarter, year). Except for the daily networks the density is quite stable over time and slightly increases until the GFC, which was a significant structural break for the monthly and quarterly network. We should note that the breakpoint (quarter 39), coincides with the quarter during which Lehman Brothers collapsed. The daily density fluctuates much more strongly, but overall increases throughout the sample.¹⁴

We should stress that, compared to the findings for other interbank networks, the density of the Italian interbank network is quite high. For example, Bech and Atalay (2010) and Soramäki *et al.* (2007) report an average density of below 1% in daily interbank networks, compared to an average density of roughly 3.1% for the daily e-MID networks. The main reasons for

¹³ Note that mostly because of the shrinking number of active banks all statistics decrease over time, irrespective of the aggregation period.

¹⁴ The density for the total network, including the foreign banks, seems to steadily decline over the sample period. This illustrates the fact that the increasing fraction of foreign banks are less interconnected with the (smaller) Italian banks.

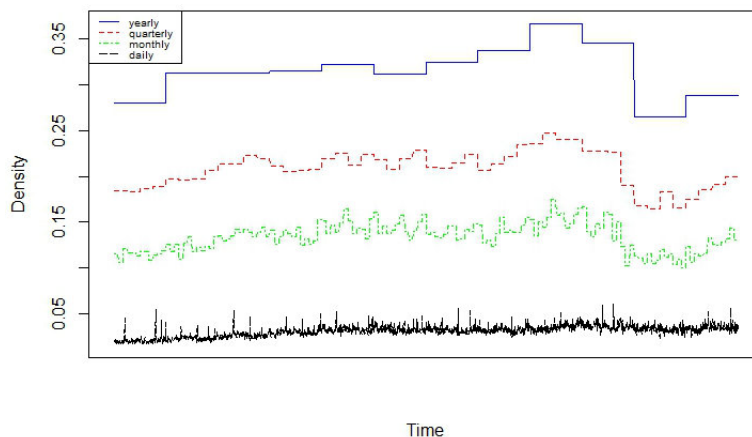


Fig. 3.4: The density for yearly (blue), quarterly (red), monthly (green) and daily (black) aggregated networks. A Chow-test and an additional CUSUM-test indicate a structural break for quarter 39 (month 117) at the 1% significance level, but not for the yearly or daily networks.

the higher density are most likely the relatively small number of participating banks in the market and the transparent market structure which easily allows each bank to trade with any other bank in the market. For comparison, the Fedwire network investigated by Soramäki *et al.* (2007) contains 5,086 institutions.¹⁵ The means of 20.8% for quarterly aggregated networks (13.4% monthly) reveal much higher densities for longer aggregation levels. Obviously the network density is positively related to the aggregation period, but to our knowledge the structure of this relation has not been investigated for interbank networks so far.

For this reason, we compare the aggregation properties of the empirical networks with those of random networks. Here we use Erdős-Renyi networks, i.e. completely random networks, and random scale-free networks, where the out-degrees follow a power-law distribution with scaling parameter 2.3.¹⁶ The

¹⁵ Additionally, the electronic nature of the trading platform simplifies the trading between any two counterparties, i.e. increases the probability of observing any link.

¹⁶ The power-law distribution with tail exponent 2.3 is a common finding in many interbank markets, see e.g. Boss *et al.* (2004). The resulting sequences of the out-degrees, i.e. the simulated quasi daily networks, are attributed to the nodes by ranking those according to the observed out-degrees, considering only active banks during the particular day. Note that if we did not account for the ordering of the observed degree sequences, we would

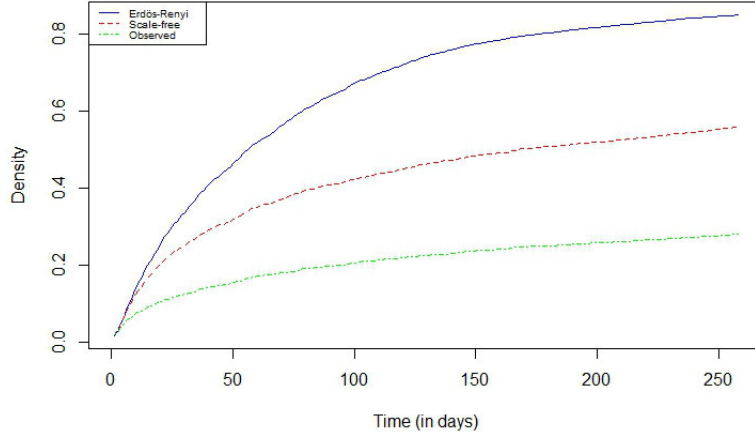


Fig. 3.5: Data for 1999. Density for the aggregated Erdős-Renyi (blue), Scale-Free (red) with $\alpha = 2.3$, and observed networks (green). Aggregation period in days. Note: we do not plot standard deviations, since these are negligible.

experiments work as follows: For each year, we aggregate the daily networks and plot the resulting density in dependence of the aggregation period, from one day up to one year (roughly 250 days). For the random networks, we aggregate artificial Erdős-Renyi and scale-free networks for each day with the same number of active banks and density as the observed daily network. The results are the average values for 100 runs for the Erdős-Renyi and the scale-free networks.¹⁷

As an example, Figure 3.5 illustrates the results for 1999, but we find very similar qualitative results for all 12 sample years. For all three networks, there appears to be a saturation level for the density, however at different levels.

end up with very similar aggregation properties as the Erdős-Renyi case. The in-degrees are distributed in a random uniform way, ruling out self-links and counting each link at most once.

¹⁷ Note that the density of aggregated Erdős-Renyi networks can be written as

$$\rho_T^r = 1 - \prod_{t=1}^T (1 - \rho_t^r),$$

where $(1 - \rho_t^r)$ is the probability of observing no links in the network at time t whatsoever, but since we adjust the number of active banks on a daily basis these probabilities would not exactly coincide with the simulated results.

The Erdős-Renyi networks always show the highest density (up to .851), followed by the Scale-Free networks (up to .559) and the observed networks (only up to .280). Apparently, it is much more likely for the empirical data that the same link gets activated several times than for the randomized data of Erdős-Renyi and scale-free networks (where the overall number of links is the same by construction). This is supported by the fact that in the observed networks just 2,757 links are observed exactly once in this year, while for the scale-free and random networks these values are on average 5,040 and 5,746, respectively. Hence, these results indicate the existence of lasting (preferential) lending relationships in the actual banking network.

3.4.3 What is a Sensible Aggregation Period?

After showing that longer than daily aggregation tends to reveal non-random structures for the Italian banking network, we are concerned with determining the ‘correct’ aggregation period in more detail in this section. This question is crucial for extracting relevant information, since the banking network typically cannot be observed at a given point in time, but has to be approximated by aggregating trades over a certain period.¹⁸ Most studies (including the present one) tend to analyze overnight loans only, in which case daily networks give an accurate picture of the actual linkages between financial institutions at a certain point in time. Most of the (apparently universal) properties of interbank networks, as mentioned in the literature overview, were derived based on such daily networks. From the perspective of a policymaker, real-time information on the structure of the interbank network is important as well, e.g. in order to assess the risk of systemic crises.

However, it should be clear that focusing on daily networks can also be potentially (dangerously) misleading, given that the low level of network connectivity suggests a negligible degree of systemic risk in general. Aggregating over a longer period is preferable, if it can reveal a non-random structure of the banking network. The existence of preferential relationships would imply that daily transactions are not determined myopically, but that a virtual network of longer lasting relationships exists. Daily transactions would then be akin to random draws from this underlying network with the realizations depending on current liquidity needs and liquidity overhang. Aggregation over a sufficiently long time horizon might reveal more and more of the hidden links, rather than adding up purely random draws from all possible links.

¹⁸ The literature on interbank networks is surprisingly silent about the choice of the aggregation period. We are aware of only one paper, namely Kyriakopoulos *et al.*, 2009, investigating this issue.

| Statistic | Year | Quarter | Month | Day |
|-----------------|-----------|---------|---------|--------|
| Banks | 131 | 125 | 123 | 107 |
| Volume | 3,505,195 | 876,299 | 292,100 | 13,688 |
| Trades | 109,807 | 27,452 | 9,151 | 429 |
| Links | 5,670 | 3,384 | 2,088 | 344 |
| Trades per Link | 19.37 | 8.11 | 4.38 | 1.25 |

Tab. 3.1: The average values for the number of active banks, total volume (million Euros), number of trades, number of links, average degree and trades per link for daily, monthly, quarterly and yearly networks.

Thus, for the purpose of identifying community structures and systemically important institutions, longer aggregation periods may be more sensible even if this comes at the cost of losing information.¹⁹ Moreover, from an economic viewpoint, overnight loans can be seen as longer-term loans, where the lender can decide every day whether to prolong the loan or not. In this way, aggregating overnight credit relationships over longer frequencies could provide a more accurate picture of the interbank network.

In the following, we compare several statistical properties of the e-MID networks based on yearly, quarterly, monthly and daily aggregated networks. Table 3.1 briefly summarizes a number of summary statistics for the e-MID networks in dependence of the aggregation period. The use of a ‘sensible’ aggregation period should ensure that we extract stable features (if they exist) of the banking network rather than noisy trading patterns at different points in time. In this regard, it is important to investigate the stability of the link structure in order to assess whether subsequent occurrences of the network share many common links. In order to do this, we rely on the Jaccard Index (JI),²⁰ which can be used to quantify the similarity of two sample sets in general. Here it is defined as

$$J = \frac{S_{11}}{S_{01} + S_{10} + S_{11}}, \quad (3.2)$$

where S_{xy} counts the number of relations having status $a_{ij} = x$ at the first

¹⁹ As an example, Craig and von Peter (2010) and Fricke and Lux (2014) identify core-periphery structures in the German and Italian interbank network, respectively, using quarterly networks. Fricke and Lux (2014) also discuss the results for other aggregation periods.

²⁰ The so-called graph correlation, see e.g. Butts and Carley (2001), shows qualitatively very similar results, but is not able to cope with banks entering or exiting market. The correlation of both measures is always above .9 irrespective of the aggregation period.

instance and $a_{ij} = y$ at the second. The JI measures links which survive as a fraction of links which are established at any of the two points in time. Hence, it also takes into account those banks which are active in only one of the two periods.

Figure 3.6 shows that the JI is very stable over time for longer aggregation periods, but not for the daily level. As expected, the JI tends to be higher for longer aggregation intervals. The daily measures are much more unstable and increase substantially until the GFC. More problematic than the smaller average level are however the extreme outliers on the downside. As a rule-of-thumb, in social network analysis one considers networks with JI values above .3 as substantially stable.²¹

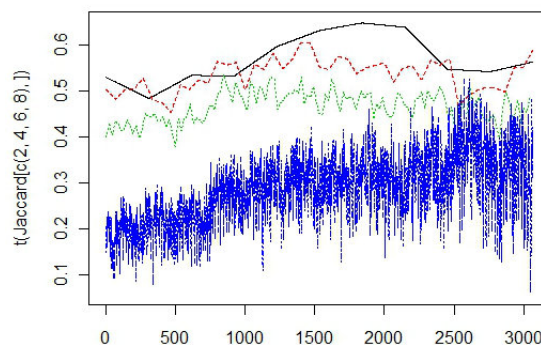


Fig. 3.6: Jaccard Index for daily (black), monthly (green), quarterly (red) and yearly (blue) networks.

Table 3.2 shows the mean, minimum, 10th percentile and standard deviation of the JIs for different aggregation periods. Again, the most evident observation is that the daily networks are rather special: the minimum and the 10th percentile of the JI are significantly smaller, indicating that we observe values below .2 in at least 10% of the sample, which is not a rare event.²² These results suggest a high degree of randomness in the daily networks.

Obviously, higher values of the JI are no guarantee that we are closer to the ‘real’ network per se. Note that in a network with randomly drawn

²¹ See Snijders *et al.* (2010).

²² We should note that, as apparent from Figure 3.6, the reason for the 10th percentile to be below the .2 threshold is not the GFC.

| Jaccard Index | Year | Quarter | Month | Day |
|--------------------|-------|---------|-------|-------|
| mean | .5543 | .5302 | .4638 | .2861 |
| standard deviation | .0535 | .0368 | .0333 | .0740 |
| min | .4652 | .4479 | .3735 | .0603 |
| 10th percentile | .5072 | .4835 | .4183 | .1904 |

Tab. 3.2: Jaccard Index for daily, monthly, quarterly and yearly networks. Calculations were carried out for all subsequent networks at the different aggregation periods. Standard deviations based on all observations.

| Reciprocity | Year | Quarter | Month | Day |
|--------------------|-------|---------|-------|-------|
| mean | .4264 | .2085 | .0829 | .0042 |
| standard deviation | .0580 | .0423 | .0244 | .0060 |

Tab. 3.3: Reciprocity of the Italian Banking network. Calculations were carried out for all networks at the different aggregation periods. Standard deviations based on all observations.

connections, the index should be positively related to the length of the aggregation period. Thus, it is important to show that other network measures also take on values significantly different from random networks for longer aggregation periods. In the following, we will therefore have a closer look at the reciprocity of the network.

Reciprocity is a global concept for directed networks that measures how many of the existing links are mutual. It can be calculated by adding up all loops of length two, i.e. reciprocal links, and dividing them by the total number of links.

Table 3.3 shows higher levels of reciprocity for longer aggregation periods.²³ In the case of daily networks we observe very few mutual links. As Iori *et al.* (2008) stated this is a very plausible finding, since banks rarely borrow and lend money from the same bank within a particular day. However, the values for longer aggregation periods show that the banking network is not one-sided, supporting the evidence on the inability of daily networks to represent the ‘true’ underlying (directed) banking network. The left panel of

²³ A structural break (after the GFC) is detected by a Chow-test as well as an additional CUSUM test for the 10th year, the 39th quarter and the 117th month respectively, but not for daily networks. For the yearly networks only the Chow-test indicates a structural break. Note that the yearly analysis involves only 12 data points.

Figure 3.7 illustrates the results for 1999, where we perform a similar analysis as for the density above, by comparing the observed network reciprocity to those of Erdős-Renyi and scale-free random networks.²⁴ The actual values are, again, always the lowest. The right panel of Figure 3.7 shows that the reciprocity (after 19 days for the real network) exceeds the density for all three networks. Note that for random networks one would expect the two measures to be almost identical. However, different banks are active at different days and if many banks are often simultaneously active the chance of forming a reciprocal link is higher (remember that we used the actually active banks of each day in the Monte-Carlo exercise). However, more important for our analysis is that for the real network the difference between reciprocity and density increases steadily and exceeds the difference of the random networks. Hence, using longer than daily aggregation is not only capable of taking mutual credit relationships into account, but even indicates a preference of banks to form them. On the other hand, the saturation of the reciprocity indicates that banks do not have mutual credit relationships with most of their counterparties.²⁵ Thus, the noise level of networks with longer aggregation periods is smaller and the directed version of the networks contains a substantial amount of information.

On the base of the Jaccard index, monthly and quarterly networks appear most stable as they have a high index with a very low standard deviation, i.e., the highest degree of structural stability of lending relationships over time. Yearly aggregation levels, in contrast, have somewhat higher variation in their JI and might be problematic, because the banking network (in particular during unstable times) is likely to evolve much faster. Somewhat related is the change in the composition of banks, i.e. banks leaving and entering the market, since we consider a bank as active for the whole year even if it leaves the market after the first trading day.²⁶ Concerning the monthly level, Iori *et al.* (2008) discovered intradaily and -monthly seasonalities which may affect our results. In most of the following, we will therefore focus on the quarterly networks.

We should emphasize that providing detailed guidelines on the optimal aggregation period of interbank network data is impossible. Clearly, it depends on the research question at hand what the most sensible aggregation period is. Our findings for the Italian interbank network suggest that ag-

²⁴ Again the results are qualitatively very similar for the other years as well.

²⁵ For instance, Finger and Lux (2013) show that the possibility to complete a mutual relation makes the link more likely, yet the existence of a mutual relationship increases the probability of these two links to be broken in the future.

²⁶ This problem occurs for each aggregation period, but is likely to become more pronounced for longer frequencies.

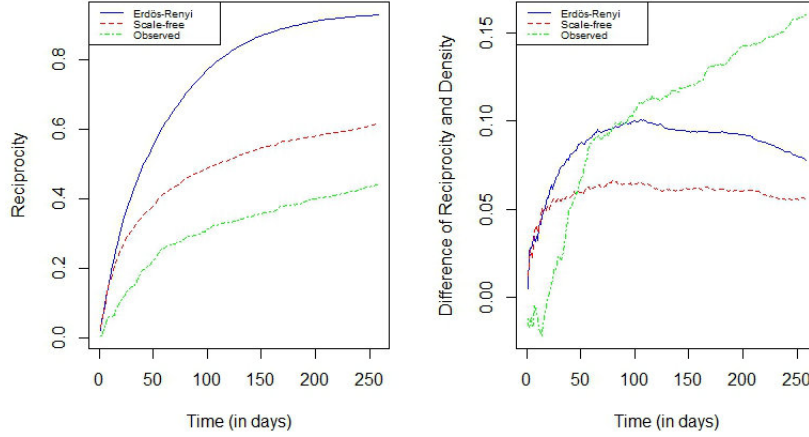


Fig. 3.7: Data for 1999. Left: reciprocity for the Erdős-Renyi (blue), Scale-Free (red) with $\alpha = 2.3$, and observed networks (green). Right: difference between reciprocity and density for the respective networks. Aggregation period in days.

gregation over a sufficiently long time horizon indeed reveals more and more of the hidden links, rather than adding up purely random draws from all possible links. The findings for the JI and the reciprocity, combined with the relatively fast saturation of the empirical density observed in Figure 3.5, is consistent with this interpretation. However, for policymakers, who need to monitor the network and act on real-time data, monthly or quarterly aggregation periods may be too long. For other purposes, e.g. identifying community structure, longer aggregation is definitely necessary. Similar to our approach, we recommend checking the saturation of certain network statistics in order to quantify the structural stability of the network at hand.

3.4.4 Transitivity

Here we are interested in transitive relations between three banks. The concept of transitivity states that a specific relationship \diamond is transitive if from $i \diamond j$ and $j \diamond k$ it follows that $i \diamond k$ holds. Equality is a transitive relation, but inequality is not. From $i = j = k$ follows $i = k$, yet $i \neq j \neq k$ does not imply $i \neq k$.

The relation we are interested in is i has a link to j or $a_{ij} = 1$. The relationship is obviously not transitive, since i has a link to j and j has a

link to k does not strictly imply that k also has a link to i . However, it is interesting to investigate how many such closed triplets occur. More generally speaking, transitivity measures whether the existence of certain links depend both on the relation between the two counterparties and on the existence of links with a third party. The measure most prominently used for this

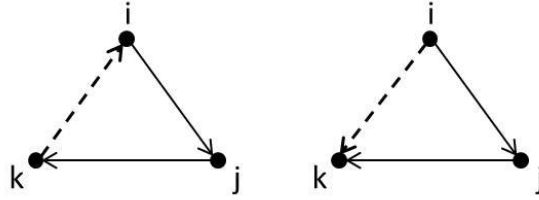


Fig. 3.8: The two possibilities how the directed path of length two (solid lines) between i and k can be closed. On the left hand side the path is closed into a loop of length three (CC^1). On the right hand side the triplet is interconnected but not in the same single direction (CC^2).

purpose is the (directed) clustering coefficient²⁷, which, despite its name, has no relation to cluster identification whatsoever.²⁸ It measures the number of (directed) paths of length two in the network and takes the fraction of those that are closed.²⁹ Figure 3.8 illustrates two ways to close the triplet i, j, k in a directed network. First, the directed path may be closed into a loop as shown on the left hand side of the figure. The function of such closure is given by the coefficient CC^1 :

$$CC^1 = \frac{\sum_{j \neq i \neq k} a_{ij} a_{jk} a_{ki}}{\sum_{j \neq i \neq k} a_{ij} a_{jk}} \quad (3.3)$$

Second, the link from i to k may be reversed. The function of such closure is given by the coefficient:

$$CC^2 = \frac{\sum_{j \neq i \neq k} a_{ij} a_{jk} a_{ik}}{\sum_{j \neq i \neq k} a_{ij} a_{jk}} \quad (3.4)$$

²⁷ For more detailed definitions of clustering coefficients see Zhou (2002).

²⁸ See Fricke and Lux (2014) and Fricke (2012) for detailed approaches of cluster identification in the e-MID market.

²⁹ Any connection along directed links between two nodes i and j is called a path and the length of the path is defined as the number of edges crossed. There are no restrictions on visiting a node or link more than once alongside a path.

An important distinction is that the nodes in the case of CC^1 , as apparent from Figure 3.8, have each one in- and one outgoing link and therefore show no hierarchical ordering. Figure 3.9 shows that the results for the two coeffi-

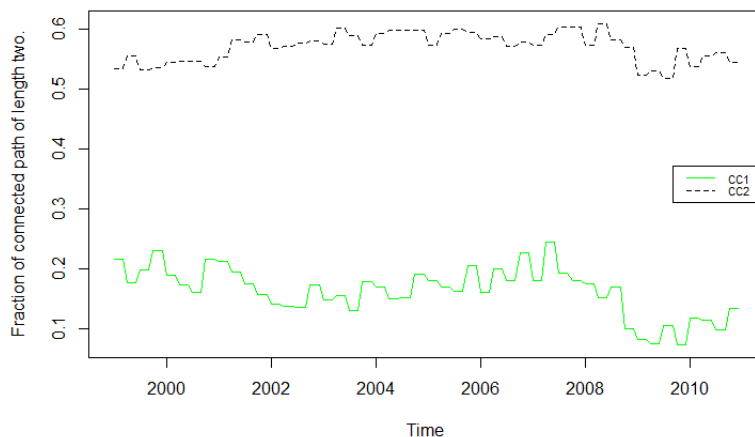


Fig. 3.9: The two clustering coefficients of quarterly networks: CC^1 (green) is the fraction of path of length two which are closed into a loop by a third link. CC^2 characterizes links in which the triangle is closed in a hierarchical way.

icients are very different. The mean for CC^1 is .164 and .571 for CC^2 . This is further evidence for the non-random character of the banking network, since the probability of an ‘average’ link to exist is just equal to the density of .208. Hence, the existence of a path of length two between i and k via j makes it 2.75 times more likely that the link from i to k exists compared to a random link³⁰, but reduces the probability that there is a link from k to i by 21%. The huge difference indicates that the banking network has a hierarchical ordering on the triadic level.

Figure 3.10 illustrates that for the Erdős-Renyi networks the evolution of the clustering coefficients is almost identical (correlation above .999). The exact numbers of the clustering coefficients for the observed networks change with the aggregation, but CC^1 is always much smaller than CC^2 for all aggregation levels as shown by figure 3.10. CC^2 is at the beginning even higher for the observed network than for the random networks, but saturates after a steep increase relatively quickly on a much lower level (up to .624).

³⁰ The probability of random link is exactly the density of .208.

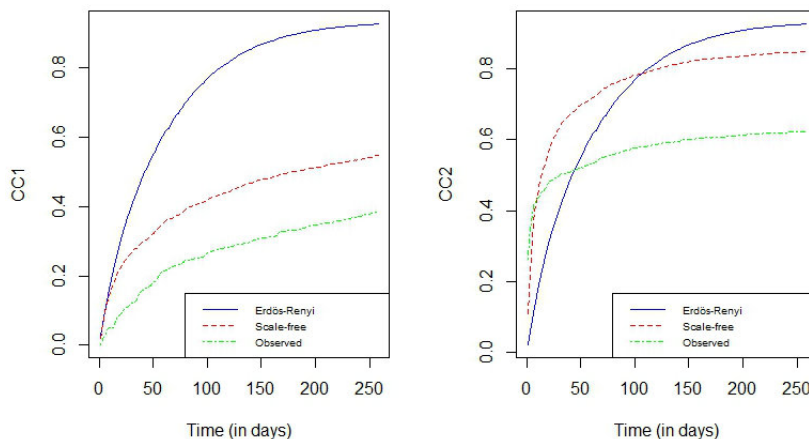


Fig. 3.10: Left: CC^1 for the Erdős-Renyi (blue), Scale-free (red) with $\alpha = 2.3$, and observed networks (green), aggregated on a daily basis up to the horizon of one year(1999). Right: CC^2 .

CC^1 on the other hand is almost zero at the beginning. Note that a loop of length three at a single day implies that each involved bank would get back some of its own lending via an intermediate bank, which appears very unlikely. However, CC^1 increases (up to .386) for longer aggregation periods showing that such relations do exist. Note that the daily or undirected networks are not capable of taking this into account.

3.4.5 Small-World Property

Another very prominent measure in the network literature is the average shortest path length (ASPL). Interest in this measure stems from the remarkable finding that in many ‘real world’ networks the ASPL is quite small, also known as the small-world phenomenon.³¹ Here we focus on the undirected version of the network.

Watts and Strogatz (1998) show that completely random networks already have a very small ASPL, but at the same time a relatively low (undi-

³¹ Small ASPLs have been detected for social, information, technological and biological networks. The first empirical finding dates back to the chain letter experiments conducted by Milgram (1967). His finding that on average only six acquaintances are needed to form a link between two random selected persons led to the famous phrase of ‘six degrees of separation’.

| | CC | ASPL |
|-------------|------------------|-------------------|
| Erdős-Renyi | .3736 (.0058) | 1.6470 (.0073) |
| Scale-free | .4131 (.0103) | 1.6259 (.0134) |
| Observed | .5422 (.0434) | 1.6486 (.0373) |

Tab. 3.4: The average CC and ASPL for the observed network and for 100 Erdős-Renyi and scale-free (with $\alpha = 2.3$) random networks. Standard deviations in brackets.

rected) clustering coefficient (CC) equal to their density. On the other hand, a regular network, where all nodes have connections to their s nearest neighbours, has a CC of 1, but a very high ASPL. Interestingly, already the introduction of a few ‘random’ links reduces the ASPL significantly, because these ‘long range’ links connect different clusters of the network. The authors argue that the interesting ‘real world’ networks are neither purely regular nor random. Therefore, Watts and Strogatz (1998) define small-world networks to additionally have a higher CC than random networks. Hence, to further investigate if the banking network exhibits the small-world property, we calculate the ASPL and the CC and compare the values to random networks.

Let the (symmetric) matrix \mathbf{G} of dimension $N \times N$ have elements g_{ij} equal to the geodesic distances between any two nodes, i.e. each element g_{ij} gives the length of the geodesic path between node i and j . The ASPL is calculated by dividing the sum of all (existing) geodesic path lengths by the total number of (existing) geodesic paths.³² The CC is calculated similar to the directed case (CC^1 and CC^2), but ignores the directedness of the links.

In this case a comparison of the networks aggregated day by day is not suitable, since the density of both random networks and scale-free networks becomes much higher than the density of the empirical network (cf. Figure 3.5). Therefore, we simulate random and scale-free(undirected) networks corresponding to the aggregated quarterly networks with respect to their density (and their degree distribution for scale-free networks). Table 3.4 summarizes the results: The ASPL for all networks is small and almost

³² It is not necessary for two nodes to have a shortest path, since there might be no link leading from one to the other. In this case the two nodes lie in different components of the network and by convention the length of these non-existing geodesic paths are set to infinity. The undirected banking network consists always of only one component and the same is true for the investigated random networks, because of the high density.

identical, in which the very low value is due to the high density. The CC for the Erdős-Renyi networks is by construction close to the density, since all links have per quarter exactly this probability to occur. The CC for the scale-free is higher (.413), but is exceeded by the statistics of the observed network (.542). This indicates higher regularity in the link structure. Hence, the banking network lies midway between regular and completely random graphs. As can be seen from the last column of Table 3.4, the ASPL does not provide much scope for distinguishing between the benchmark Erdős-Renyi and scale-free networks and the empirical ones. The reason might be that the relatively high density leads to relatively short path lengths anyway.³³

3.4.6 Effects of the Global Financial Crisis

Finally, we take a closer look at the effects of the GFC on the banking network.³⁴ The start of the GFC is not easy to determine, but we have seen that the collapse of Lehman Brothers in quarter 39 (2008 Q4) was a major shock for the global financial markets in general and the Italian interbank market as well.³⁵ The effects of this event were twofold: first, the counterparties of Lehman Brothers realized huge losses. Second, the perception of risks changed, since Lehman had been considered to be ‘too-big-too-fail’ before. The resulting dramatic increase of perceived counterparty risk reduced the willingness of banks to lend to each other, which ultimately affected the real economy due to tighter lending restrictions. The monetary authorities and governments around the world injected substantial amounts of capital into the financial system to prevent interbank markets from freezing in the following weeks. We have seen, that important network measures such as density and reciprocity, were significantly affected by these events as well.

Here we investigate the change in banks’ behavior during and after the breakdown of Lehman Brothers. To begin with, Table 3.5 contains several basic network statistics for ten quarters around the breakpoint (quarter 39). Interestingly, the number of active Italian banks remained relatively stable during this period, and in fact 86 banks were active in all of the ten quarters. The stability of this composition is important, since under these circumstances changes in the behavior of the banks on the aggregated or individual

³³ The results are qualitatively the same for the directed version of the network. The ASPL for the quarterly networks is 1.912, while 1.7765 for the scale-free and 1.802 for the Erdős-Renyi networks. The CC^2 (.571) is significantly higher than for the scale free with .342 and Erdős-Renyi networks with .208, whereas CC^1 (.164) is smaller. However, in this case the network consists not only of one giant component, but might contain separate clusters, which makes the interpretation of the ASPL more difficult.

³⁴ See also Fricke and Lux (2014).

³⁵ See Brunnermeier (2009).

| | | | | | |
|-----------------|---------|---------|---------|---------|---------|
| Quarter | 36 | 37 | 38 | 39 | 40 |
| Banks | 101 | 100 | 100 | 98 | 97 |
| Volume | 445,991 | 409,340 | 435,338 | 404,353 | 385,819 |
| Trades | 20,984 | 20,078 | 19,963 | 18,160 | 16,477 |
| Links | 2,425 | 2,253 | 2,249 | 2,153 | 1,768 |
| Trades per Link | 8.65 | 8.91 | 8.88 | 8.43 | 9.32 |
| Quarter | 41 | 42 | 43 | 44 | 45 |
| Banks | 96 | 96 | 94 | 94 | 94 |
| Volume | 234,102 | 267,057 | 197,021 | 227,076 | 196,503 |
| Trades | 14,184 | 13,981 | 12,525 | 11,636 | 11,577 |
| Links | 1,533 | 1,506 | 1,599 | 1,449 | 1,530 |
| Trades per Link | 9.25 | 9.28 | 7.83 | 8.03 | 7.57 |

Tab. 3.5: The table summarizes the number of Italian banks, total volume (million Euros), number of trades, number of links and links per trade for the quarters 36-45, while the Lehman collapse has been in quarter 39.

level should be mainly driven by their response to this exogenous shocks.

We also see that the total trading volume, the number of trades and the number of links are all decreasing over this period, but the exact patterns are distinct. Surprisingly, the volume is quite stable until the 40th quarter, but drops by 39.3 percent in the 41st quarter. In contrast, the number of trades starts to fall in the 39th quarter and decreases further until quarter 44. The total number of links already starts to decrease in quarter 37, but overall tends to develop in a very similar way to the number of trades.³⁶ In the end, the most immediate reaction to the crisis was that banks traded similar total volumes but in fewer trades and with a smaller number of counterparties in order to minimize their (perceived) counterparty risks.³⁷

Interestingly, the trades per link are the highest for the three quarters after the Lehman collapse. This indicates that the banks relied stronger on their preferred counterparties. Preferred in this context might simply mean that the banks had more reliable information about these banks, which, however, should coincide with the existence of a previous trading relationship. Eventually, we conclude that the breakdown of Lehman Brothers significantly affected the behavior of individual banks and thus had a clear impact on the

³⁶ The correlation between both is .963 for this period and .957 over the complete sample.

³⁷ Fricke and Lux (2014) show that most of these changes were in fact driven by behavioral changes of core banks.

structure of the network. Quite surprisingly, we find that the link structure of subsequent networks remained rather stable during this period, since no structural break is detected for the Jaccard Index. Furthermore, despite the significant impact of the GFC, we do not find evidence for a complete drying up of the e-MID market, even at the daily level.³⁸

3.5 Discussion

What do we learn from this analysis of the network properties of the interbank market? Network analysis is one way to cope with the complexity of a huge system of dispersed activity and extract some hopefully important general characteristics from such a system. Without such an analysis, the raw data of more than one million trades between financial institutions is relatively uninformative for the researcher. We have, however, demonstrated that different ways to approach these data (different aggregation levels) might reveal different pieces of information on the underlying structure, with more or less robust outcomes. If we think about the data *revealing* an underlying network structure through random sampling from this net of connections (as the activation of a link depends on the random realization of liquidity needs or liquidity overhangs), it might seem sensible to allow for a longer observation (aggregation) period to catch more of the existing but often dormant links. As it appears, practically all network statistics are more robust over longer horizons, while the daily horizon might just provide information on a small random part of the ‘true’ network. Under this perspective, daily data might not be a useful format for network inference at all - despite the fact that the underlying data are for daily maturities of interbank lending. This important distinction between the time horizon of the legal contract, and the appropriate time horizon for a meaningful network analysis seems to have been largely ignored in extant literature on interbank networks.

Besides this basic message, another interesting result is that the network is asymmetrical in many respects. For the quarterly network the fraction of reciprocal links is very similar to the density and therefore mutual relationships tend to be neither especially attractive nor repellent. Furthermore, the two directed clustering coefficients are very different. The probability for a path of length two to be closed into a loop is 3.48 times smaller than the alternative closure. Note that this speaks in favor of a hierarchical structure: if bank A receives credit from bank B and extends credit to bank C, it is much more likely that C also receives credit from B rather than being

³⁸ As noted above, this is not true for loans longer than overnight. These markets essentially collapsed completely, which is not surprising given that the loans are unsecured.

its creditor itself. Additionally, the correlation between in- and out-degree is merely .12 for the complete sample. Therefore, the information that a bank has a large number of incoming links is a surprisingly poor indicator of how many outgoing links the bank has. The high level of asymmetry in the network indicates that different banks might play different roles in the network. Fricke (2012) and Fricke and Lux (2014) provide evidence that this is indeed the case. Moreover, for many measures the GFC could be identified as a structural break and also the decreasing number of volume, trades and links support that the GFC heavily affected the Italian interbank market. However, the network overall remained surprisingly stable and despite the decrease of its volume (in the beginning of 2009) the e-MID market was never close to drying up completely.

In terms of managerial insights, our paper sheds light on the practical implementation of liquidity management at financial institutions. Indeed, while liquidity management is of outermost importance for the stability of both individual financial institutions and the financial system in its entirety³⁹, relatively little is known about the empirical conduct of liquidity management. What we can infer from our present findings is that a purely neoclassical approach to liquidity management (as detailed, for example, in Brodt, 1978) would probably not provide a very accurate characterization of empirical practices. Individual banks (or their liquidity managers) would pursue an intertemporal linear programming strategy in order to maximize the profits under given constraints on reserve requirements and capital adequacy regulations. Thus, we would expect an atomistic market without any stable link formation between participants. Indeed, the very evidence of an interesting (i.e. not completely random) network structure speaks against a purely competitive market. While the present paper sheds light on the non-random structure of links and their stability over time on a purely phenomenological level, our model-based analyses in two companion papers very well underscores the non-random structure of link formation: Fricke and Lux (2014) show that participating banks in the e-MID platform can be very consistently categorized as either core or periphery banks. Typically, core banks borrow from a large pool of periphery banks, and lend to other core or periphery banks thus creating indirect links particularly between members of the ‘periphery’. Many of these links are very stable over time, as also demonstrated by Finger and Lux (2013) in a so-called ‘stochastic actor-oriented’ model of network formation (borrowed from sociology). Here it turns out that solely characteristics of the specific counterparty, most prominently past

³⁹ See e.g. IIF Special Committee on Liquidity Risk (2007), Basle Committee on Banking Supervision (2008) and 2010, Préfontaine *et al.*

trades with the same counterparty, and not higher level relations⁴⁰ are the strongest explanatory factors for link activation/termination in the interbank market. Surprisingly, interest rates have only a minor and non systematic effect. Further evidence for the existence of stable long-run relationships in the interbank market is also provided by Cocco *et al.* (2009) for a different segment (Portuguese banks) with different econometric methodology.

Overall, it seems that market participants focus strongly on lasting relationships rather than shopping for the best deals in terms of interest rates offered. It is particularly remarkable that these findings are based on data for an electronic platform that easily provides traders access to a long list of publicly quoted bids and asks. Obviously, trusted working relationships for liquidity adjustment are of value to the managers-in-charge who are often prepared to forego more advantageous offers from other market participants in favor of their existing working relationships. Formation of stable links might also work as a way to reduce the complexity of the day-to-day task of balancing liquidity needs and excess liquidity (cf. Wilkinson and Young, 2002).⁴¹ The overall non-random network is the ‘emergent’ macro structure from such behavioral strategies of individual market participants.

The importance of stable links also becomes apparent during the financial crisis: while we saw a dramatic decrease of overall activity in the network, no structural break in the Jaccard index is found. This means that the structure of link formation and continuation has remained very much intact, but trading activity has typically been concentrated on fewer counterparties than before. From this perspective, the non-random character of link formation might have prevented an even more detrimental collapse of network activity.⁴² While the trigger for the crisis was an event outside the network under investigation (Lehman Brothers did not participate in the Euro denominated e-MID market we are investigating.)⁴³, our analysis sheds some light on how a network of lasting business relationships reacts on such a

⁴⁰ This means that banks do not consider the complete network, e.g. indirect relations, during the network formation process, although they have via the e-MID platform all the necessary information.

⁴¹ With respect to the stability of relationships, the working of the interbank market seems not too different from that of the often-studied fish market. Here as well, stable links might underscore the importance of trust between buyer and seller as well as the element of complexity reduction in a day-to-day task where excessive shopping for good prices seems much less important than trust in the reliability of the counterpart, cf. Kirman and Vriend, 2001, Gallegati *et al.*, 2011.

⁴² Fricke and Lux (2014) show that mostly core banks reduced their lending activity but at the same time increased their borrowing from their counterparties in the periphery.

⁴³ No bank from outside Europe traded in the Euro segment of the e-MID interbank market in the whole sample period.

disturbance (which might entail a loss of trust in some counterparties depending on experience and interaction or their exposure towards the source of the disruption). Both phenomenological and model-based analyses should, therefore, contribute essentially to understanding the vulnerability of the interbank network which is of crucial importance for the smooth functioning of the financial sector. Knowledge of its intrinsic structure would also be instrumental in identifying feedback effects from the interbank market in liquidity stress tests for single financial institutions, as to start conducting system-wide stress tests (cf. Elsinger *et al.*, 2006 and Montagna and Lux (2013)).

3.6 Conclusions and Outlook

In this paper, we have investigated the interbank lending activity as documented in the e-MID data from 1999 until the end of 2010 from a network perspective. Our main finding is that daily networks feature too much randomness to be considered a representative statistic of some underlying latent network. The Jaccard index shows the higher consistency over time for longer aggregation periods and the very low density compared to (aggregated) random networks indicates the existence of preferred trading relations. In general the evolution of all global network measures for longer aggregation periods (month, quarter, year) is very similar in their deviation from the Erdős-Renyi and scale-free benchmarks. Moreover, the monthly and quarterly networks are characterized by a significantly higher than random clustering coefficient, and thus reveal some regularity in the link structure. The (almost) zero reciprocity and CC^1 of daily networks proves the inability of this aggregation level to reveal information on such structural elements. However, quarterly networks consistently exhibit a non-random structure and allow us to consider the mutuality of the relations and are therefore a preferable subject of study, especially if one is interested in the evolution of the network over time.

Essentially, these results show that it is far from trivial to map a given data structure into a ‘network’. While daily records of the interbank trading system can be arranged in an adjacency matrix and treated with all types of network statistics, they provide probably only a very small sample of realizations from a richer structure of relationships. Just like daily contacts of humans provide very incomplete information of networks of friendship and acquaintances, the daily interbank data might only provide a small selection of existing, dormant established trading channels. Hence, inference based on such high-frequency data may be misleading while a higher level of time

aggregation might provide a more complete view on the interbank market. What level of aggregation is sufficient for certain purposes is an empirical question depending on the research questions at hand. However, saturation of certain measures may be a good indicator that most dormant links have been activated at least once over a certain time horizon. At the same time, such dependence of statistics on the time horizon serves to sort out a number of simple generating mechanisms (i.e. completely randomly determined networks in every period) and reveal interesting dynamic structure.

In the future, more attention should be given to the analysis of directed banking networks using longer aggregation periods to identify structural commonalities. From a management perspective, our findings carry important implications for the regulation of credit institutions, since the low level of connectivity in the noisy daily networks suggests a low degree of systemic risk at any point in time. While real-time information about the structure of the interbank network is important for policymakers having to take immediate action, regulators should therefore rather concentrate on longer network aggregates. Getting a better idea on the wider pool of counterparties of all credit institutions would allow the identification of possible behavioral changes among the set of relatively active banks. Such changes might then serve as an indicator for funding problems of individual institutions, cf. Fricke and Lux (2014). In the end, it would be important to extend our phenomenological analysis in order to test hypotheses about the behavior of banks at the micro-level that drives the system's properties. We should also stress that, while we did not tackle the question of the optimal regulation of interbank networks, our results help in understanding the relationship between the structure of interbank networks at different aggregation periods. These results can be used to feed simulation models of interbank contagion with more realistic network structures.

4. ON ASSORTATIVE AND DISASSORTATIVE MIXING IN
SCALE-FREE NETWORKS: THE CASE OF INTERBANK
CREDIT NETWORKS

4.1 Introduction and Existing Literature

The global financial crisis (GFC) in 2007/08 moved interbank networks into the focus of academia, since the linkages created by liabilities among banks and other financial institutions played a crucial role and yet are poorly understood. An intriguing example is the bankruptcy of Lehman in September 2008, where the immense effects on the system as a whole were dramatically underestimated by the authorities. As a consequence, investigations of complex systems in terms of their network properties gain more and more attention in economics following the lead of other disciplines in which network analyses have already a long tradition. One remarkable finding is that many complex¹ real world networks share some apparently universal features.² A prominent example is a highly skewed degree distribution, where the degree of a node is the number of (incoming/outgoing) links.³

This paper focuses on the relationship between the degree distribution and another aspect of the network topology, namely the concept of degree assortativity, which measures whether nodes prefer links to other nodes with similar degree. The main finding in the network literature in this respect is that social networks tend to be assortative, meaning that low (high) degree nodes form links to other low (high) degree nodes (Newman, 2002, 2003). König *et al.* (2010) argue that capacity constraints in social networks lead to assortative networks and introduce a network formation algorithm taking this into account. A counterexample to this rule of assortativity of social networks is reported by Holme *et al.* (2004). They find disassortativity together with a skewed degree distribution in an internet dating community which, thus, appears to share some characteristics with networks constructed from interbank credit relations. On the other hand, technological and biological networks tend to be disassortative. Economic networks can be seen as a hybrid case sharing features of both depending on the context (Jackson, 2008).

The question to which class banking networks belong is crucial, since the different types of networks have very different degrees of resilience and vulnerability to shocks. For instance, Newman (2003) found that assortative networks are more resilient, whereas disassortative networks are particularly vulnerable against directed attacks on high degree nodes (hubs). In the existing empirical literature Soramäki *et al.* (2007) for the Fedwire payments network, Bech and Atalay (2010) for the Federal funds network and Iori *et*

¹ Complex in this regard refers to the fact that the networks are neither purely random nor regular.

² See e.g. Reka and Barabási (2002).

³ See e.g. Newman (2010), ch. 8.

al. (2008) for the Italian overnight money market find banking networks to be disassortative.

This paper takes as its empirical starting point interbank networks derived from the credit extended via the e-MID (electronic market for interbank deposits) trading platform for overnight loans from 1999 to 2010, which is maintained by a privately owned Italian company and currently the only electronic broker market for interbank deposits. The focus of the analysis will be on Italian banks, since the participation of foreign banks in the e-MID had been very volatile and faded away almost completely after 2008. Our results support the earlier findings that banking networks are disassortative and this result holds for both the directed and the undirected case.⁴ The strongest result for the directed case is that banks with a high in-degree tend to have links to banks with a small out-degree, which holds also for the out-in case but to a lesser degree. This robust finding of interbank networks to be disassortative emphasizes the importance to identify the so-called ‘systemically important’ banks, which would adversely affect large parts of the network in case of their failure. For instance, Craig and von Peter (2010) and Fricke and Lux (2014) estimated a core-periphery model for the German and the Italian interbank network to identify these important components, while Sachs (2013) shows that such a banking system becomes less stable the more concentrated the assets are.

However, the finding of interbank networks to be scale-free⁵ (SF) and disassortative at the same time are hard to combine with the analytical results of Newman (2002), showing that SF networks (based on the preferential attachment, or PA, algorithm from Barabasi and Albert, 1999) with a typical scaling exponent of $\alpha = 3$ are unassortative, i.e. neither assortative nor disassortative, at least in the limit of very large networks. Therefore, the evidence for a SF distribution together with disassortativity motivates our investigation of whether SF networks are unassortative in general or only for specific parametrizations, and whether the results are different for a finite size of the network. Monte-Carlo simulations show that Erdős-Renyi (ER) (1959) random networks are unassortative, independent of the system parameters. In contrast, SF random networks with relatively small scaling exponents tend to be disassortative. However, the degree of disassortativity is still stronger for the observed networks than for SF networks with the scaling exponent of 2.3 reported in the literature. Another finding is that even small changes in the scaling exponent can lead to substantially different mixing patterns.

⁴ Directed networks take the direction of the liability into account and therefore links do not have to be mutual.

⁵ See e.g. Boss *et al.* (2004) for the Austrian interbank market and Iori *et al.* (2008) for the Italian e-MID trading platform.

If scaling exponents are even higher (as the ones estimated by Fricke and Lux, 2013b, for the e-MID network lending data) a variety of well-known generating mechanisms for SF networks would be inconsistent with the observed combination of density, degree distribution and (dis)assortativity in the data. Of course, SF networks would also be typically misspecified for such data, if a power-law distribution for the degree could be rejected. An interesting new approach to model network formation is presented in König *et al.* 2009 applying a two-stage game in which links are formed as a best response to the present network. The generating mechanism is able to obtain a degree distribution with a specific power-law and produces always disassortative networks. Cohen-Cole *et al.* 2011 propose a related mechanism for interbank markets in which the link formation is microfounded on the banks profitability.

The remaining part of this paper is organized as follows. Section 4.2 gives a brief introduction into interbank networks. Section 4.3 introduces the data set obtained from the e-MID trading system. Section 4.4 presents the empirical analysis of the observed networks, with a number of robustness checks to be found in the Appendix. Section 4.5 contrasts this with Monte-Carlo simulations of various generating mechanisms for SF networks and section 4.6 concludes.

4.2 Networks and Assortativity

A network consists of a set of N nodes that are connected by M directed edges (links). Taking each bank as a node and the interbank positions between them as links, the interbank network can be represented as a square matrix of dimension $N \times N$ (data matrix, denoted \mathbf{D}). An element d_{ij} of this matrix represents a gross interbank claim, the total value of credit extended by bank i to bank j within a certain period. The size of d_{ij} can thus be seen as a measure of link intensity. Row (column) i shows bank i 's interbank claims (liabilities) towards all other banks. The diagonal elements d_{ii} are zero, since a bank will not trade with itself. Off-diagonal elements are positive in the presence of a link and zero otherwise.

Interbank data usually give rise to directed, sparse and valued networks.⁶ However, much of the extant network research ignores the last aspect by focusing on binary adjacency matrices only. An adjacency matrix \mathbf{A} contains

⁶ Directed means that $d_{ij} \neq d_{ji}$ in general. Sparse means that at any point in time the number of links is only a small fraction of the $N(N-1)$ possible links. Valued means that interbank claims are reported in monetary values as opposed to 1 or 0 in the presence or absence of a claim, respectively.

elements a_{ij} equal to 1, if there is a directed link from bank i to j and 0 otherwise. Since the network is directed, both \mathbf{A} and \mathbf{D} are asymmetric in general.

The degree of a node gives the total number of links that a bank has with all other banks and can thus be seen as a measure for the importance of individual nodes. Undirected networks imply symmetric adjacency matrices (with elements denoted by a_{ij}^u). In this case bank i 's total degree k_i is simply the total number of relationships bank i has with other banks, i.e.

$$k_i = \sum_{j \neq i} a_{ij}^u. \quad (4.1)$$

For directed networks, we differentiate between incoming links (bank i borrows money from other banks) and outgoing links (i lends money to other banks), and define the in- and out-degree of i (k_i^{in} and k_i^{out}) as the number of pertinent directed links

$$\begin{aligned} k_i^{in} &= \sum_{j \neq i} a_{ji} \\ k_i^{out} &= \sum_{j \neq i} a_{ij}, \end{aligned} \quad (4.2)$$

respectively. In our case the diagonal elements are zero, so in- and out-degree can be calculated by simply taking column- and row-sums of the adjacency matrix, respectively.

In this paper, we are mainly interested in the assortative mixing patterns of interbank networks. In general, the concept of assortativity is concerned with the similarity, in terms of some attribute, of connected nodes. Here we are interested in assortative mixing by degree⁷, i.e. how similar the degrees of connected nodes are. A network shows assortative mixing, if high-degree nodes tend to have many connections with other high-degree nodes. There exist several measures of assortativity in the literature. In the following we define the assortativity coefficient r as the Pearson correlation coefficient of degree between pairs of linked nodes, see Newman (2002). Hence, positive values of r indicate a correlation between nodes of similar degree, while negative values indicate relationships between nodes of different degrees. Thus, r lies between -1 and 1, with $r = 1$ ($r = -1$) corresponding to perfectly assortative (disassortative) mixing patterns. The assortativity coefficient is most often reported for the undirected version of a network, a convention that we will stick to in most of the following.

⁷ The Appendix contains a number of robustness checks, where we also focus on valued networks.

Empirically, social networks tend to display assortative mixing patterns, while technological and biological networks are usually characterized by disassortative mixing patterns. Interbank networks have been reported to display disassortative mixing patterns as well, so in this respect they tend to be closer to most technological and biological networks than to most social networks.⁸

Note that the assortativity is a global property of a network. However, it is not necessarily the case that all parts of a network display the same mixing patterns, cf. Fricke and Lux (2014). Therefore, it is often useful to plot the average neighbor degree of a network, which is formally defined as

$$\langle k_{nn} \rangle = \sum_{k'} k' P(k'|k), \quad (4.3)$$

where $P(k'|k)$ is the conditional probability that a link belonging to node with degree k points to a node with degree k' , cf. Pastor-Satorras *et al.* (2001). If $\langle k_{nn} \rangle$ is an increasing (decreasing) function of k , the network is assortative (disassortative), but this function may display different behavior for low and high degree nodes, respectively.

4.3 The Italian Interbank Market (e-MID)

The Italian electronic market for interbank deposits (e-MID) is a screen-based platform for trading of unsecured money-market deposits in Euros, US-Dollars, Pound Sterling, and Zloty operating in Milan through e-MID SpA.⁹ The market is fully centralized and very liquid; in 2006 e-MID accounted for 17% of total turnover in the unsecured money market in the Euro area. Average daily trading volumes were 24.2 bn Euro in 2006, 22.4 bn Euro in 2007 and dropped to only 14 bn Euro in 2008 as a consequence of the financial crisis. We should mention that researchers from the European Central Bank have repeatedly stated that the e-MID data is representative for the interbank overnight activity, cf. Beaupain and Durré (2012).

Detailed descriptions of the market and the corresponding network properties can be found in Finger *et al.* (2013).¹⁰ In this paper we used all registered trades in Euro in the period from January 1999 to December 2010. For each trade we know the banks' ID numbers (not the names), their relative position (aggressor and quoter), the maturity and the transaction type (buy or sell). The majority of trades is conducted overnight and due to the

⁸ See Soramäki *et al.* (2007), Bech and Atalay (2010), and Iori *et al.* (2008).

⁹ The vast majority of trades (roughly 95%) is conducted in Euro.

¹⁰ See also the e-MID website <http://www.e-mid.it/>.

GFC markets for longer maturities essentially dried up. We will focus on the overnight trades conducted on the platform. From these we take the trades conducted within the subset of Italian banks as foreign banks withdrew more or less completely from this market after 2008. Another reason to focus on the subnetwork formed by Italian banks only is that they are likely to use the e-MID market as their main source of funding over the entire period and are in general more homogenous. This leaves a total number of 1,215,759 trades.

As usual, some data aggregation is necessary to represent the system as a network. In the following, we define the network as the binary adjacency matrix obtained from quarterly aggregated data, i.e. a link exists if at least one transaction has been taking place between the banks i and j within a quarter (the number of transaction or their volume would define the entries of a valued network data matrix). Aggregation of quarterly data had been chosen because many existing links might be dormant at higher frequencies, while for even higher levels of time aggregation structural changes of the network (entry and exit of participants) might interfere with the requirement of a stable structure. Finger *et al.* (2013) indeed show that a certain saturation of the density and stabilization of the structure happens at our chosen level of aggregation. We should stress, however, that most of our findings remain largely unaffected by using different aggregation levels, see the Appendix for details.

4.4 Assortativity Patterns in the Italian Interbank Market

Figure 4.1 indicates that the Italian interbank network indeed displays disassortative mixing patterns, since we find negative values for r over the complete sample period (blue line). Note that the positive trend in the coefficient over time indicates that the network evolves into a more unassortative state over time. To what extent this is driven by behavioral changes of banks is not quite clear.

Figure 4.2 illustrates the increase in the assortativity coefficient over time. It shows, for the first quarter of 1999 (top) and the last quarter of 2010 (bottom), respectively, the directed network matrix (left), the corresponding undirected network matrix (center), and the average neighbor degree (left) as defined in Eq. (4.3). The first quarter of 1999 is the most disassortative network of our sample ($r = -0.38$), while the last quarter of 2010 is the least disassortative ($r = -0.21$). Links are indicated by black dots in the network matrices, where we sorted the nodes (rows/columns) according to their total degree. Focusing on the directed networks first, we see a substantial amount

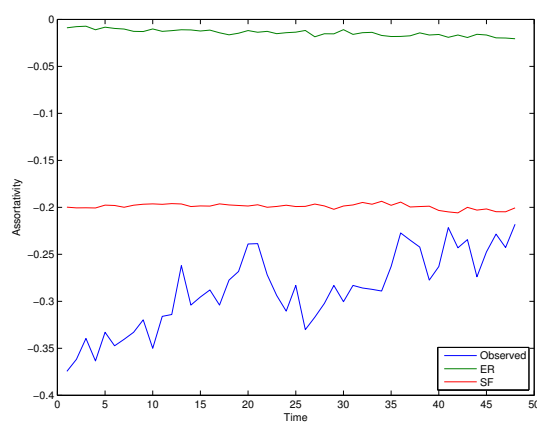


Fig. 4.1: Assortativity coefficient r of observed (blue) and random networks for ER (green), and SF networks (red) for the undirected version of the network over time. The densities in the SF and ER networks were chosen to match the observed ones (quarter by quarter) for the network of Italian banks. A scaling exponent of 2.3 was used for the in- and out-degrees in the SF networks. Results for random and SF networks are the average of 100 simulations, with dotted lines showing one standard deviation.

of asymmetry for both quarters, since high degree nodes tend to create more outgoing links to low-degree nodes than vice versa. The corresponding undirected network matrices are mainly plotted for the sake of comparison with scale-free networks in the next section, but also illustrate the effect of symmetrizing strongly asymmetric networks (see below). The average neighbor degree displays a quite pronounced downward slope in 1999Q1, whereas the relationship is much noisier in 2010Q4. Note that this noisiness is partly due to a significantly smaller network size in the latter case ($N = 90$ as compared to $N = 205$). Moreover, very high degree nodes are largely absent in 2010Q4, whereas several of such nodes are present in 1999Q1. Thus, it seems that a significant fraction of the increase in r is due to behavioral changes of large banks, who were found to cut down the number of outgoing links during the global financial crisis, cf. Fricke and Lux (2014). At the same time, mergers and acquisitions in the Italian banking sector have led to a decreasing number of Italian banks and, therefore, of participants within the e-MID platform. The tendency towards less pronounced disassortativity might be the consequence of this consolidation and the reduction of small banks (with presumably a small number of links) that could appear as counterparts in trades with banks with high degrees.

In order to check whether the above findings are affected by focusing on undirected interbank networks, we also calculated the assortativity coefficient for the directed version of the Italian interbank network.¹¹ The calculation works as follows: consider for example, the case where we want to calculate the assortativity coefficient for the combination of in- and out-degree of connected nodes (In-Out). For each link we collect the in-degree for the source node and the out-degree of the target node, and then calculate the correlation between the two vectors. Figure 4.3 shows that the findings are comparable to the undirected case. For example, for the assortativity coefficient for the In-Out case, we see negative values for the complete sample period. Thus, it is likely for a bank with high in-degree to have outgoing links to nodes with small out-degree. Similarly, the values for the Out-In case are also negative over the complete sample period, albeit on a significantly smaller absolute level closer to the unassortative case. Thus, a bank with high out-degree may have banks of any kind of in-degree as counterparty. These findings are in line with the results in Fricke and Lux (2014), where the Italian interbank market is shown to display a structure close to a hierarchical core-periphery network. The set of highly connected core banks tends to lend money to other core banks and a large number of loosely connected periphery banks, which in turn tend to lend money to a small number of selected core banks,

¹¹ See Newman (2002) and Piraveenan *et al.* (2012).

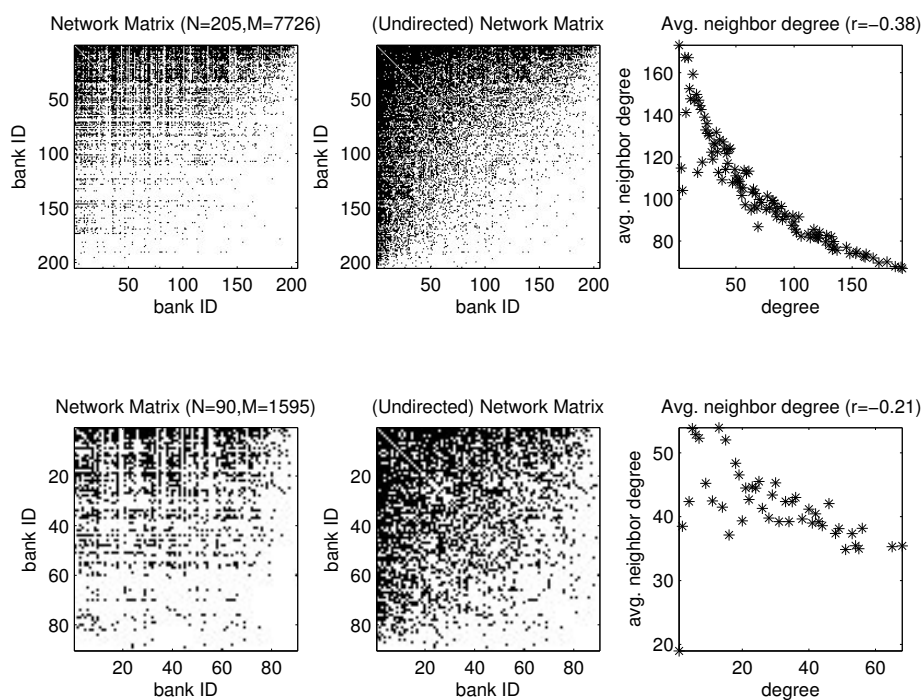


Fig. 4.2: The Italian interbank network in the first quarter of 1999 (top) and in the last quarter of 2010 (bottom). Left: directed network matrix. Black dots indicate links. Center: corresponding undirected network matrix. Right: average neighbor degree as a function of the degree.

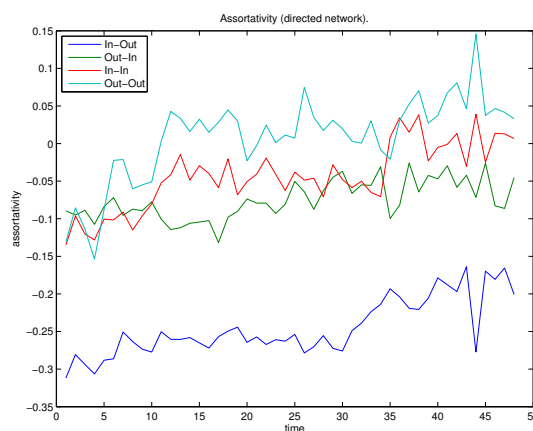


Fig. 4.3: Assortativity coefficients over time for the directed version of the network using Italian banks only.

but appear to trade relatively scarcely among themselves. The large (absolute) values of the In-Out combination are thus an indicator of core banks (who have high in-degrees in general), to lend money to a large number of periphery banks (who have small out-degrees in general). The same is true for the Out-In combination, but on a smaller scale, given that core banks' out-degrees tend to exceed their in-degrees. The results for the In-In and Out-Out combinations are less clear, since both coefficients are close to zero over the complete sample period. The Out-Out combination even shows positive values for most of the sample period. In this respect, it seems important to note that there is very little (at times even negative) correlation between single banks' in- and out-degrees. Thus, a bank with a high in-degree does not necessarily have a high out-degree, and vice versa.

To further investigate the universality of the finding that the Italian banking network is disassortative Figure 4.4 compares the assortativity coefficient r for different aggregation periods, namely daily, 2 weeks, monthly, quarterly, and yearly. We see that the daily networks are typically disassortative, but there is a high level of variation in the calculated coefficients. Interestingly, for longer aggregation periods, the coefficients are largely comparable. Using an aggregation period of 2 weeks already reduces a lot of the noisiness of the daily networks. Note that these results are in line with the findings of Finger *et al.* (2013), who show that the randomness of the daily networks is strongly reduced under sufficient time-aggregation of the networks.

Fricke and Lux (2014) have shown that the Italian interbank network displays a core-periphery structure, where core (periphery) banks are maxi-

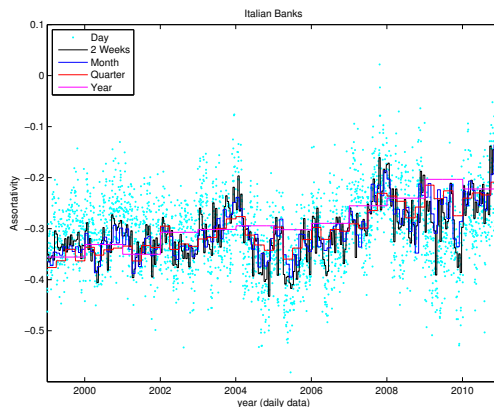


Fig. 4.4: Assortativity coefficient r for different aggregation periods, based on Italian banks only.

mally (minimally) connected with each other. This structure was shown to be rather stable over time, at least up to the global financial crisis, when the fit of the core-periphery model became significantly worse. One important question concerning the aggregation properties is how the daily networks, which appear to be quite noisy and rather random, are related to the quarterly ones. In order to investigate this issue, we performed a sampling exercise: during each day of the sample period, we randomly sample a subset of the links observed at the quarterly level and calculate the corresponding assortativity.

The assortativity values for the daily and quarterly networks are displayed in Figure 4.5. Quite interestingly, the assortativity coefficient of the sampled daily networks is always larger than the corresponding value for the quarterly networks. Similarly, in most cases the level of disassortativity in the sampled daily networks is typically smaller compared to the values for the observed daily networks. This indicates that the daily networks are unlikely to be simple random realizations of the underlying quarterly network. We performed a similar sampling exercise in a variety of synthetic networks with varying levels of degree disassortativity (unreported result), including stylized core-periphery and scale-free networks, cf. Stumpf *et al.* (2005). Quite interestingly, in all cases, the level of disassortativity is always smaller in random subsamples of the actual networks. We leave a detailed analysis of this issue for future research.

Overall, these findings indicate that high-degree banks tend to connect to low-degree banks and vice versa (albeit with a certain tendency towards less disassortative mixing over time). This finding can be seen as an indi-

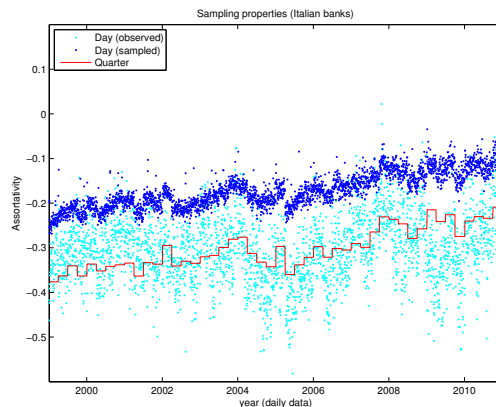


Fig. 4.5: Assortativity coefficient r for the Italian interbank networks (daily and quarterly) and sampled daily networks from the quarterly observations.

cator of the tiered structure of the interbank market¹² and is also useful for understanding the spread of contagion in the interbank network.¹³

4.5 Assortativity in Finite-Size Scale-Free Networks

Interestingly, Newman (2002) shows that both ER networks and SF networks (based on the Barabasi-Albert, 1999, PA mechanism, i.e. with tail parameter of 3) are unassortative for large network sizes, i.e. they display no correlation between the degrees of connected nodes. Since interbank networks have been reported to display smaller scaling parameters (typically around 2.3), we checked the assortativity of synthetic SF networks in a Monte-Carlo exercise. For this purpose, we generated (directed) SF random networks using the algorithm of Goh *et al.* (2001), which allows to choose N , M , and the scaling exponent(s) α freely. Moreover, Lee *et al.* (2006) analytically showed that the average neighbor degree is for this generating mechanism only independent of k (unassortative) if the scaling exponent $\alpha \geq 3$. In case of $2 < \alpha < 3$ the average neighbor degree decreases for high degree nodes indicating disassortativity in line with our results for e-MID network.¹⁴ We

¹² See Fricke and Lux (2014).

¹³ In assortative networks, diseases targeting high-degree individuals are likely to spread to other high-degree nodes. In disassortative networks, specifically targeted vaccination strategies for high-degree nodes can quickly stop the epidemic in the network.

¹⁴ However, the average neighbor degree for the Italian banking network visible in Figure 4.2 displays this kind of behavior also for low degree nodes.

should note that this algorithm belongs to the class of static SF networks, working as follows: each of the N vertices is indexed by an integer value i ($i = 1, \dots, N$). For each node we assign two weights, $p_i = i^{-\gamma_{\text{out}}}$ and $q_i = i^{-\gamma_{\text{in}}}$ to each vertex for outgoing and incoming edges, respectively. The control parameters γ_{out} and γ_{in} are in the interval $[0, 1)$. At each leg, we select two vertices with probability $p_i / \sum_k p_k$ and $q_j / \sum_k q_k$, and a link from i to j is added. This repeats until the network has the desired number of links. The absence of self-links and multiple links in this process explains the disassortativity. Goh *et al.* (2001) show that the out- and in-degrees are power-law distributed with parameters $\alpha_{\text{out}} = (1 + \gamma_{\text{out}}) / \gamma_{\text{out}}$ and $\alpha_{\text{in}} = (1 + \gamma_{\text{in}}) / \gamma_{\text{in}}$. In the following, we only show results using the same tail exponent for both in- and out-degrees, i.e. assuming symmetric in- and out-degree distributions such that $\alpha_{\text{in}} = \alpha_{\text{out}} = \alpha$. The analytical results are only valid in the limit of $N \rightarrow \infty$ and Lee *et al.* (2006) also show that the distortion of the numerical simulations increases if N decreases. Therefore, since our main goal is to test the ability of generating mechanism to replicate observed networks we rely in the remaining part of this paper on Monte-Carlo experiments. In a first step we compare the disassortativity for the observed network of Italian banks only with synthetic ER and (static) SF networks. For each quarter, we take the observed values of N and M , and for the SF networks we additionally assume a constant α of 2.3 both for the in- and out-degree distributions.

Figure 4.1 shows the average results for 100 random networks in each quarter. As expected the ER networks are unassortative, even though the average r is not exactly equal to zero. The static SF networks yield relatively constant values for r around -.2 for all quarters in line with the results from Lee *et al.* (2006). Thus, SF networks with $\alpha = 2.3$ are not too different from the observed ones in terms of disassortativity, even though the absolute level of the observed values for r is smaller throughout.¹⁵

To further investigate the appropriateness of the static SF networks we performed an additional Monte-Carlo experiment with varying density and tail exponent (while fixing $N = 100$), cf. Figure 4.6. According to the results in Lee *et al.* 2006 only the tail exponent should effect the assortativity, while it should be independent of the density. Quite interestingly, we find a U-shaped relationship between the assortativity coefficient and the density, i.e. r is smallest for intermediate densities. Note that the minimum is usually

¹⁵ The number of undirected links M in the synthetic networks matches the number of observed links. Since we symmetrize the networks in order to calculate the assortativity coefficient, the number of links will not necessarily be the same in the two cases. We checked that the results remain largely unaffected as compared to the reported results, when constructing undirected SF networks with exactly the same number of links as the symmetrized empirical networks.

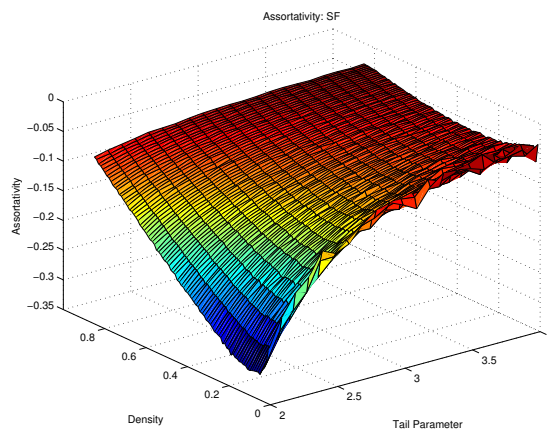


Fig. 4.6: Assortativity coefficient r in scale-free networks with $N = 100$, varying both the density and the tail exponent. The plot shows average results from a Monte-Carlo simulation with 100 runs for each parameter setting. SF networks were generated using the algorithm of Goh *et al.* (2001).

found for densities around .2, which is close to the observed density of our quarterly networks. Clearly, a higher density makes interesting mixing patterns less likely, since most nodes are connected. Similarly, for low densities only few nodes are connected, resulting in small values for r . Additionally, we find that r tends to be positively related to the scaling exponent, approaching the unassortative state for larger values. In any case, we see that the assortativity coefficient of the synthetic SF networks with $\alpha = 3$ is not zero in general, indicating that the analytical result of Lee *et al.* (2006) is only valid in the limit of very large networks (in particular with densities around 0.2). We also varied the size of the network, but the findings are remarkably robust (unreported result).¹⁶ This shows that (static) SF networks display disassortative mixing patterns and that the extent of their disassortativity depends on the density as well and not the tail exponent alone. Moreover, we observe disassortative mixing patterns for $\alpha \geq 3$, which is again contrary to the asymptotical findings for $N \rightarrow \infty$.

To see how general these findings are, we performed a number of robustness checks using alternative algorithms for generating SF networks. From the class of static SF networks, we used the fitness-based algorithm of De Masi *et al.* (2006) and found qualitatively very similar results, cf. Figure

¹⁶ We found very similar results using networks with up to 2,500 nodes. The calculations are computationally too demanding for even larger networks.

4.7.¹⁷ The major difference lies in a somewhat higher level of noise, but the general shape is very similar. We also checked the results using standard PA approaches. Typically, PA networks are constructed on the basis of growing networks, since nodes are added sequentially to the network until it has the desired size. Each newly added node creates a certain (fixed) number of links to existing nodes, where high-degree nodes have a higher probability of attracting additional links. Here we used the approaches of Barabasi and Albert (1999) and Xie *et al.* (2008) for constructing the synthetic networks.¹⁸ We should stress the difficulty in comparing the results to those from the static SF networks, as the tail exponent typically cannot be varied in the PA approaches. Still, we find that also the PA networks tend to be disassortative, with r approaching zero for larger densities. In contrast to the static approaches, we also found the size of the network to play a role for the structural properties of the growing network models, since larger PA networks are closer to the unassortative state for any density (unreported).

In order to illustrate the underlying structure of the SF network, Figure 4.8 shows the directed network matrices (left), with the nodes sorted according to their degree, and the corresponding average neighbor degree as a function of degree (right) for the case of $\alpha = 2.0$ (top) and 3.0 (bottom), respectively. The networks are again based on the Goh *et al.* (2001) algorithm and the density is .2 which is in line with the observed quarterly networks from the e-MID platform. Note that we restrict ourselves to plotting only the directed version of the network here, since the level of asymmetry in the SF networks is quite small in general, in particular compared to the empirical networks (cf. Figure 4.2). This is due to the fact that we always use the same tail exponent for the in- and out-degree distributions. Note also that the SF networks, while typically displaying a tightly connected core (top left corner of the network matrices), the structure of the remaining network is quite distinct from the observed networks both for the case of $\alpha = 2$ and 3. While the average neighbor degree is strongly downward sloping for small tail exponents, this is not necessarily the case for larger values. In fact, based on a rewiring exercise, we found that SF networks with small scaling exponents

¹⁷ To be precise, we used the probability function $P(f_i, f_j) = \left(\frac{f_i}{f_{\max}}\right)^{\beta_1} \left(\frac{f_j}{f_{\max}}\right)^{\beta_2}$, where f is a ‘fitness’ measure for the individual banks, and f_{\max} is the maximum fitness. We employed the basic parametrization of Montagna and Lux (2013), that had been calibrated to roughly fit the link distribution of the e-MID data. Hence, f follows a power-law distribution on the interval $[5, 100]$ with scaling exponent 2. The β parameters are related to the scaling exponents of the degree distributions as follows: $\alpha_{\text{out}} = (1 + \beta_1)/\beta_1$ and $\alpha_{\text{in}} = (1 + \beta_2)/\beta_2$.

¹⁸ Strictly speaking, the approach of Xie *et al.* (2008) is not based on growing networks, but rewires existing links in accordance with PA.

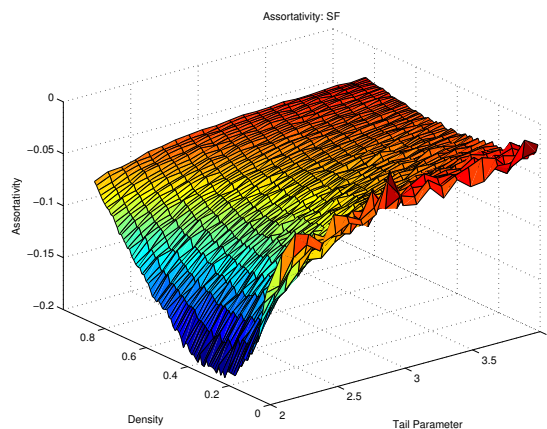


Fig. 4.7: Assortativity coefficient r in scale-free networks with $N = 100$, varying both the density and the tail exponent. The plot shows average results from a Monte-Carlo simulation with 100 runs for each parameter setting. SF networks were generated using the fitness-based algorithm, with parameters taken from Montagna and Lux (2013). See footnote 17 for details.

will always be disassortative, whereas this is not necessarily true for larger tail exponents. See the Appendix for details.

Our above experiments have shown that small changes in the tail exponent can lead to substantially different mixing patterns, at least in networks of finite size. Moreover, the assortativity is also strongly dependent on the network density. We should stress that Fricke and Lux (2013b) show that the scaling exponents of the degree distribution for the Italian interbank network tend to be substantially larger than those considered in this exercise. The data analysed by Fricke and Lux (2013b) partially overlaps with the data set investigated by Iori *et al.* (2008) whose claim of a SF degree distribution with low power-law exponent could not be continued under a variety of perspectives on the distribution of degrees and the aggregation level. With these values ($\alpha > 5$), however, it will be impossible to replicate disassortative mixing patterns. In the end, our findings show hardly any evidence in favor of practically all generating mechanisms for scale-free networks, including preferential attachment, as possible underlying models for the Italian interbank network. As this is the most popular generating mechanism for scale-free networks with power-law degree distributions, this inability to explain the financial network formation with a single mechanism should also warrant a closer look at the distributional features of the data.

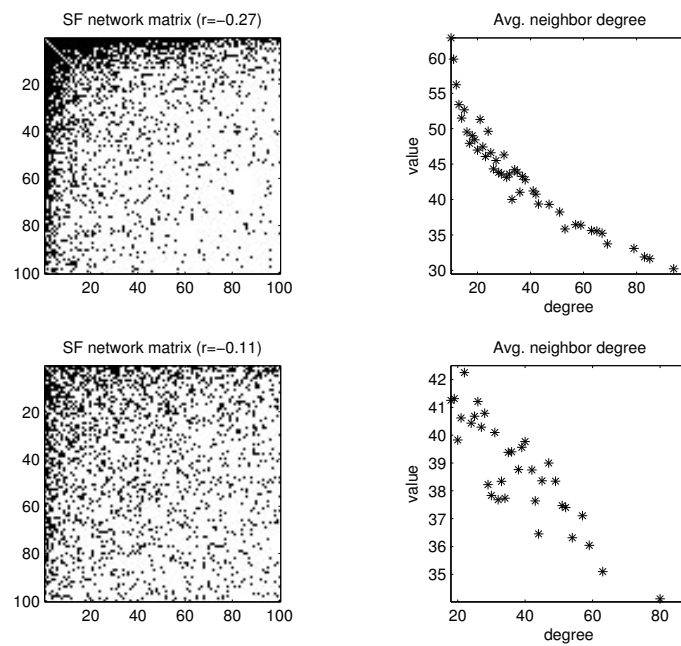


Fig. 4.8: Assortative mixing in scale-free networks with $N = 100$, $M = 1980$. The tail exponent is set to $\alpha = 2$ (top) and 3 (bottom), respectively. Left: directed network matrices. Black dots indicate links. Right: average neighbor degree as a function of degree.

4.6 Conclusions

This paper presents evidence for interbank networks having disassortative mixing patterns with respect to their degree, i.e. high-degree nodes tend to connect to low-degree nodes and vice versa. In the case of the Italian interbank market this result is true for directed networks and undirected networks over the whole sample period from 1999-2010. The well-known vulnerability of disassortative networks to the failure of central nodes makes a deeper analysis of the network formation process in the interbank market particularly worthwhile. This motivates us to search for network generating mechanism that could provide an explanation for the particular combination of features we encounter in the interbank data.¹⁹ Such generating mechanisms could then be used for conducting realistic simulations of contagion effects via credit relationships between banks, cf. Upper (2011).

Our finite size Monte-Carlo simulations show that for tail exponents smaller than $\alpha = 3$, SF networks become disassortative. The level of disassortativity is in general highly sensitive to changes of the scaling exponent. In addition, the density affects the degree of assortativity, with the strongest level of disassortativity for densities around .2, which is close to the observed values of the Italian interbank network. Hence, the frequent finding of a power-law distribution of degrees and disassortativity of interbank networks would be in line with our results. However, it is not clear whether interbank networks are in fact always characterized by a power-law distribution with low α . While disassortativity and a low density seem to be uniform features of all data investigated so far, the power-law distribution of degrees seems less clear-cut. Findings of high α or inappropriateness of a power-law to describe the degree distribution (Fricke and Lux, 2013b) would both imply that SF generating mechanisms are not consistent with the structure of interbank data. For a high scaling coefficient, known SF generating algorithms would not be consistent with disassortative behavior, while rejection of a scale-free distribution of the degrees would invalidate the very statistical foundation of SF networks. Therefore, our results indicate that new models of interbank networks formation should be developed, but for that purpose a more solid behavioral foundation of the link formation between financial institutions is needed. A more behaviorally motivated model of interbank lending has been proposed recently by Cohen-Cole *et al.* 2011.

¹⁹ See Basle Committee on Banking Supervision (2011) for more details.

4.7 Appendix

4.7.1 Robustness Checks for the Empirical Results

This Appendix investigates the robustness of the empirical results in section 4.4 of the main text. As a first step, we show that the results are broadly similar when including foreign banks in the analysis, while the results presented in the main text are for the subnetwork of Italian banks only. Secondly, we focused on binary networks. Here we investigate the assortativity in valued networks, i.e. networks of transaction volumes and number of trades, as well. The results are in line with those from the binary networks, however, the average level of the disassortativity is negatively related to the length of the aggregation period in the valued network.

Including Foreign Banks

So far, we have only focused on the subnetwork formed by Italian banks. Finger *et al.* (2013) have shown that both Italian and foreign banks mostly tend to trade among themselves on the e-MID market, such that the complete network can be decomposed into two subnetworks. Nevertheless, the assortativity of the complete networks might differ from the reported values for the Italian banks. In the light of our comments on the sampling properties of disassortative networks, we should stress that focusing on Italian banks only is therefore not equivalent to focusing on a random subnetwork. Hence, as long as the linking behavior among foreign banks is not too different from that of the Italian banks, we would not expect pronounced differences in the mixing patterns.

Figure 4.9 shows the assortativity coefficient r for the complete interbank network, containing both Italian and foreign banks, for different aggregation periods, namely daily, 2 weeks, monthly, quarterly, and yearly. As before, the daily networks are typically disassortative, but there is a high level of variation in the calculated coefficients. The values are again remarkably stable under different aggregation levels, showing disassortative mixing patterns in general. The most striking difference, compared to the case of Italian banks only, is that the increase in r over the sample period is much more pronounced and somewhat cyclical. Also, the average level of r appears to be somewhat larger, which is mainly due to the fact that foreign banks tend to connect mostly with other foreign banks, i.e. more unassortative behavior.

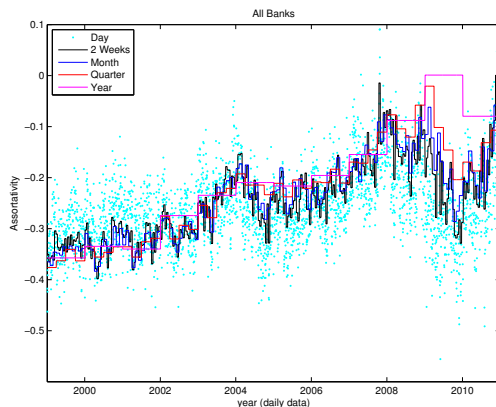


Fig. 4.9: Assortativity coefficient r for different aggregation periods, based on Italian and foreign banks.

Assortativity in Valued Networks

So far, we have treated interbank networks as binary adjacency matrices. Here we check the assortativity coefficient using valued networks, namely the transaction volumes and the number of trades, respectively.

Figures 4.10 and 4.11 show the results for the network matrices based on the transaction volumes and the number of transactions, respectively, again for a variety of aggregation periods. As before, the networks are typically disassortative on all frequencies, consistent with the findings for the binary networks. However, when aggregating valued networks, the coefficients r tend to increase with the aggregation period. It is not quite clear why this is the case.

4.7.2 Assortativity in Rewired Scale-Free Networks

To shed light on the underlying structure of the SF networks, Figures 4.12 and 4.13 (left) show examples for adjacency matrices of scale-free networks with $\alpha = 2.0$ and 3.0 , respectively, again based on the Goh *et al.* (2001) algorithm. In addition, we show their most assortative (center) and disassortative (right) counterparts, which were obtained using an adjusted version of the rewiring algorithm of Xulvi-Brunet and Sokolov (2004) for the case of directed networks.²⁰ We should stress that the rewiring algorithm

²⁰ The basic idea of the algorithm is as follows: we can vary the level of assortativity in the network by rewiring observed links. At each iteration, we choose two random links connecting four distinct nodes. With a certain probability we change the ends of the nodes, without changing the sources, to generate assortative or disassortative mixing patterns. In

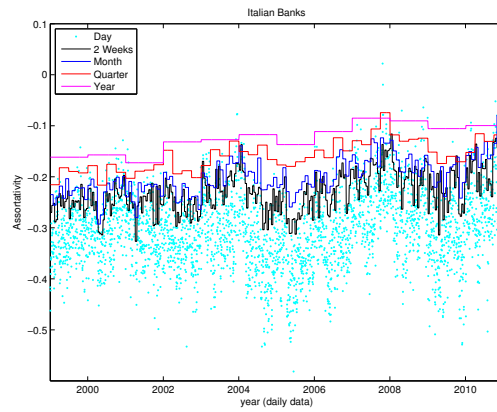


Fig. 4.10: Assortativity coefficient r in the valued network (transaction volumes) for different aggregation periods, based on Italian banks only.

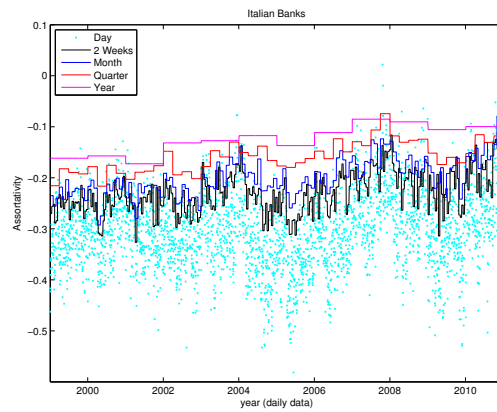


Fig. 4.11: Assortativity coefficient r in the valued network (number of trades) for different aggregation periods, based on Italian banks only.

ensures that, besides keeping the total number of links constant, the original in-degree and out-degree sequences remain unaffected. For $\alpha = 2.0$, the left panel of Figure 4.12 shows a relatively dense core on the top left, which contains the most highly connected nodes, whereas moving to the bottom right shows a large number of loosely connected periphery nodes. In contrast, the center panel shows the most assortative network (after rewiring), where nodes with similar degree tend to form clusters along the main diagonal. Note that this network displays disassortative mixing patterns due to the high degree nodes that are connected with practically all other nodes, i.e. also with the low degree nodes on the bottom right. The right panel shows the most disassortative case. Again, the highly connected core persists, but the remaining links do not appear in clusters along the main diagonal, but rather form a concave shape starting from the bottom left to the bottom right. Given that the nodes are sorted by their degrees, this indicates that highly dissimilar nodes are connected to each other. In contrast, Figure 4.13 shows the results for $\alpha = 3.0$. In this case, there are fewer highly connected nodes, which allows to rewire the links into a truly assortative state (center). The right panel shows the most disassortative case, leaving only relatively few links among the most highly connected nodes, but displaying a similar value of r as for $\alpha = 2.0$. Thus, SF networks with small scaling exponents will always be disassortative, whereas this is not necessarily true for larger values of α .

the former case, we connect the higher degree nodes with each other and similarly for the lower degree nodes. In the latter case, the node with the two high-degree nodes connect to the two lower degree nodes. For a sufficient number of iterations and different rewiring probabilities, this algorithm generates networks with different levels of r .

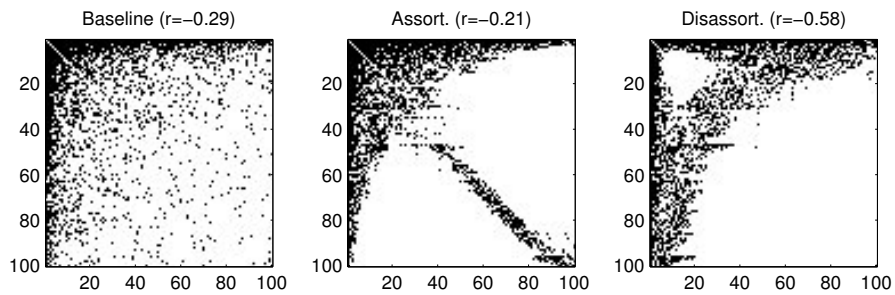


Fig. 4.12: Network matrix for scale-free networks with $N = 100$, $M = 1980$, $\alpha = 2.0$, and different levels of assortativity. Left: scale-free network, center: most assortative case, right: most unassortative case. Black dots indicate links. Assortativity coefficient in brackets.

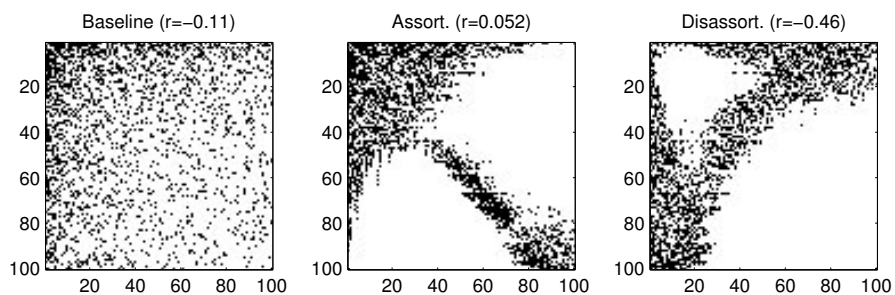


Fig. 4.13: Network matrix for scale-free networks with $N = 100$, $M = 1980$, $\alpha = 3.0$, and different levels of assortativity. Left: scale-free network, center: most assortative case, right: most unassortative case. Black dots indicate links. Assortativity coefficient in brackets.

5. FRIENDSHIP BETWEEN BANKS: AN APPLICATION OF
THE ACTOR-ORIENTED MODEL OF NETWORK
FORMATION TO INTERBANK CREDIT RELATIONS

5.1 Introduction

In recent years the focus of banking regulation as well as the academic banking literature has started to shift from the analysis of the optimizing behavior of single banks to what is called a systemic perspective. Pertinent literature has started to study default patterns and the possibility of a system-wide breakdown (as it seemed immanent in 2007/08) from the viewpoint of a network approach to the structure of the financial system.¹

The literature analysing the interbank market as a network has taken advantage of the fact that network related research has experienced a surge of interest recently in various scientific disciplines. The internet, friendship relations, cellular networks and ecosystems are just a few examples of complex systems investigated in terms of their network properties. The first applications of network theory to the banking system have been focusing on measures from the natural sciences describing the topology of the banking network to determine its general resilience or vulnerability in the presence of shocks. Examples are Inaoka *et al.* (2004) for the Japanese interbank market, Boss *et al.* (2004) for the Austrian banking sector, Soramäki *et al.* (2007) for the US Fedwire network, Bech *et al.* (2010) for the US Federal funds market and De Masi *et al.* (2006) and Iori *et al.* (2008) for the Italian interbank market. The most prominent findings of these studies are:(a) that degree distributions seem to follow a power law with coefficients between 2 – 3, (b) the density of the network is relatively low,² (c) the networks show disassortative mixing with respect to their degree, (d) the average shortest path is very small (i.e. the networks exhibit a small world structure).

More recently, Craig and von Peter (2010) and Fricke and Lux (2014) show that the interbank network is close to a core-periphery structure, in the spirit of Borgatti and Everett (2000), for both German and Italian data of interbank credit relationships. The basic idea is that a few money-center banks (the core) are highly interconnected among themselves and act as intermediaries for the banks in the periphery. Cocco *et al.* (2009) and Affinito (2012) have also provided evidence for relational trading in interbank markets. Cocco *et al.* (2009) show for the Portuguese interbank market that banks pay lower interest rates on loans from banks they had frequently traded with in the past. Affinito (2012) also emphasizes long lasting trading relations and their positive effect during the global financial crisis (GFC) in the Italian interbank market. Hence, their results support the prediction of Carlin *et al.* (2007) that cooperation among banks (traders) is an equilibrium outcome

¹ See e.g. Iori *et al.* (2006), Nier *et al.* (2007) and Haldane and May (2011).

² The density of a network is defined as the percentage of existing links to all possible links of a fully connected system.

under repeated interaction. Iori *et al.* (2014) model an agent-based interbank market with relational trading and compare the generated networks with real world networks aggregated from transactions of the e-MID platform.

In this paper we approach the same topic from a complementary perspective: we use the stochastic actor-oriented model (SAOM) developed by Snijders (1996) to analyse the dynamics of the interbank network using longitudinal panel data. While the above studies use traditional statistical methods, the SAOM analyzes the evolution of the network from an actor point of view and so provides an avenue for identification of the driving forces behind banks decisions about their counterparties in the interbank market. In the last few years this model has been heavily used in sociology, but also in neighbouring fields.³ However, to our best knowledge the SAOM has never been used in the economics literature.

The data set used comprises all Euro denominated overnight transactions by Italian banks in the electronic market for interbank deposits (e-MID) from 2001 to 2010 aggregated into quarterly networks. Since the test of Lospinoso *et. al* (2011) detects time heterogeneity for almost all effects the model will first be estimated separately for each quarterly aggregate. Afterwards the two-step multilevel approach proposed by Snijders and Baerveldt (2003) is applied to assess the overall significance of the hypothetical effects for the complete sample and the subperiods before and after the GFC.

In our version of the SAOM model we apply a large number of network-related effects, actor-specific and dyadic variables as potential determinants of agents' link formation. Overall the behaviour of banks appears relatively robust in almost all respects over the complete sample period and this stability should have contributed to the relative resilience of the interbank market during the GFC. The most salient result of our analysis is that banks heavily rely on lasting relations, although an electronic platform like the e-MID reduces the direct costs to trade with 'new' counterparties practically to zero. Only before the GFC (where the spreads were very small⁴) do we find that the behavior of banks was slightly influenced by the average interest rates of their counterparty. Two additional changes are observed after the onset of the GFC: (1) larger banks and 'core' (as identified by Fricke and Lux, 2014) banks becoming more popular and (2) banks avoiding intermediate relations more than before. Both aspects indicate that counterparty risk has become a more important factor. The first aspect might support the 'to big to fail' and 'to interconnected to fail' argument in case of the size and core variables,

³ See for example Checkley and Steglich (2007) for the impact of the mobility of managers on venture capital firms and Johnson *et al.* (2009) for an analysis of seasonal changes of an ecosystem.

⁴ See Raddant (2012).

respectively.

In general, banks' choices of counterparties in the interbank credit market seem to favour hierarchical ordering on the local (no cycles) and global level (degree distributions). Moreover, mutuality does not simply strengthen a relation, since only the creation of a mutual relation is valued by banks and not the maintenance of an existing one. This is in stark contrast to earlier findings in other applications of the SAOM regarding human relations, where enduring mutual relations with 'investments' already undertaken from both sides are extremely unlikely to be terminated. However, we have to bear in mind that a credit contract is in itself already a mutual contract so that the motivations of our actors is certainly different from that of humans forming friendships. The diverging findings on mutuality might, thus, be interpreted as a preference for borrowing to former creditors (to whom some link already exists) but there is no point for banks in keeping up intentionally credit relationships in both directions over time. Furthermore, while we find a high level of the directed clustering coefficient for the complete network we do not observe that banks would put any value on such triadic intermediate relations in our disaggregate analysis. Hence, it appears that the high clustering statistics emerges unintendedly from the interaction of the individual banks. Another interesting aspect is that we find support for the relevance of the core-periphery distinction since the binary categorisation of banks along the results of Fricke and Lux (2014) enters as a significant effect although the model also controls for the size of banks and many other effects capturing hierarchical aspects of the network.

The rest of the paper is organized as follows. Section 5.2 introduces the e-MID data, provides details on how to aggregate trades into an interbank network, and defines basic network terminology and the covariates used. In section 5.3 the general specification of the SAOM, the implemented effects and the estimation approach are explained. In section 5.4 the results for the Italian interbank network are presented and discussed.⁵ Section 5.5 contains final remarks and conclusions.

5.2 Data Set and Network Construction

The Italian e-MID market is the only existing electronic broker market for interbank deposits, while all other segments of interbank lending worldwide are over-the-counter. Detailed information on the latter type is typically unavailable to regulators and researchers. Publications of the European Cen-

⁵ The estimation results for every single period and all specifications mentioned are available upon request from the corresponding author.

tral Bank stated repeatedly in recent years that the e-MID data should be representative for the dynamics of the complete money market, so that one might expect to obtain representative results on interbank trading activity from this data set.⁶ The data set contains tick-by-tick transaction data on Euro denominated interbank deposits from 1999-2010. Only overnight deposits, which account for more than 80% of all trades, are considered in the following.⁷

The 353 different banks that were active during this time span are known to us only by their unique ID number used in the data set. We omit non-Italian banks trading in the e-MID platform because they were mostly only trading within their relatively small group and their number among the e-MID participants showed large variability over time (cf. Finger *et al.*, 2013). Therefore, only the active 255 Italian banks are considered. Only the transactions from 2001-2010 are used in the empirical estimation, since fluctuations of banks entering and leaving the e-MID platform have been very high prior to this period. The market is quote driven and these quotes include bank-ID, the interest rate offered and the identification as sell or buy quote. These characteristics are normally visible to all participating banks. It is possible to engage in bilateral trade, where a quote is only seen by a pre-specified counterparty, but this happens very rarely. The quoting bank in any case gets to know the aggressor before the trade is executed.

Before explaining how to construct our observed networks from this data set we introduce some basic network terminology. A network consists of N nodes (order) and S links (size). It is represented as a $N \times N$ adjacency matrix (AM) x , where the element x_{ij} provides information on whether credit has been extended from bank i to bank j in the current period or not. In our application of the SAOM to interbank data, we only consider a binary adjacency matrix (i.e. $x_{ij} \in \{0, 1\}$) while in other studies we have also used weighted adjacency matrices containing information on the aggregated size of interbank loans. The network is directed and so x_{ij} does not need to be equal to x_{ji} , leading to an asymmetric adjacency matrix. The diagonal elements x_{ii} are zero by assumption, since banks are not trading with themselves. The sum of outgoing links $\sum_{j=1}^N x_{ij}$ is the outdegree of actor i and the sum of incoming links $\sum_{j=1}^N x_{ji}$ defines the indegree of actor i .

The next step is to aggregate single trades into interbank networks. To transform the transactions into a network it is necessary to aggregate them over a specified period. In our main application the data is aggregated on

⁶ See e.g. Beaupain and Durré (2011, 2012).

⁷ Maturities up to 1 year and transactions in Dollars, Sterling and Zlotys are possible.

a quarterly basis, leading to 40 observations.⁸ The most common approach in the literature is to use networks estimated from daily data, but we argue that for the present analysis a longer aggregation period is needed, since daily networks are too noisy and many existing links will simply not be activated on any particular day. Finger *et al.* (2013) scrutinize the question of the ‘appropriate’ aggregation period and show that most network statistics are stable and saturate around the quarterly level. Table 5.1 displays basic statistics of our quarterly networks which will be defined and referred to in the presentation of the empirical model and the estimation results.

Before turning to the SAOM we introduce the covariates implemented as exogenous variables in the model. Note that since the banks’ identity is unknown to us all covariates are derived from the underlying data set of interbank lending and its network representation itself. In general there exist two types of covariates: actor specific covariates and dyadic covariates. All covariates are centered by subtracting the mean prior to the estimation process. Inactive banks in case of actor covariates and inactive dyads in case of dyadic covariates are left out of this procedure and are assigned a zero (the new average). Hence, they will not affect the results and the interpretation of all effects.

The Q actor covariates v_{im}^q are specific characteristics of each actor and are stored in vectors of length N (number of active banks), where q indicates the corresponding covariate, i the actor and m the period. The values are allowed to change over time. In our model we consider four actor covariates:(1) a proxy for the size of each bank v_{im}^1 is calculated as a categorical variable using the total transaction volume each bank was involved in during the current quarter. The banks are classified into quintiles $v_{im}^1 = \{1, 2, 3, 4, 5\}$, where the 20% of banks with the highest volume are attributed size 5 etc.⁹,(2) core is a dummy variable v_{im}^2 and takes the value 1 if the bank had been found to be in the core in the pertinent period according to Fricke and Lux (2014) and zero otherwise,(3) the lending rate v_{im}^3 is calculated as the difference between the volume weighted average interest rate that banks received for their loans and the mean over all active banks which have lent money in this quarter,(4) analogous, the borrowing rate v_{im}^4 is calculated as the volume weighted average interest rate from all transactions in which i had borrowed money minus the mean over all borrowers.

Dyadic covariates w_{ijm}^d measure characteristics specific for each ordered pair (i, j) of banks, where d is the index of the dyadic covariate and m the

⁸ Only 37 of them are used in the application, since we use the trading activity of the last year as a covariate.

⁹ After the centering procedure the values change to $v_{im}^1 = \{-2, -1, 0, 1, 2\}$.

period. They are stored in non symmetric ($N * N$) matrices. The first dyadic covariate, called past trades, is the square root of the number of interbank loans extended from i to j during the last year

$$w_{ijm}^1 = \sqrt{B_{ij}} - \frac{1}{(N-1)^2} \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N \sqrt{B_{ij}}. \quad (5.1)$$

It proxies the intensity of the directed relationship in the past and subtracting the average over all potential links (second term) centers the covariate. Using the square root rather than the raw numbers is motivated by the right skewness of the distribution of the numbers of transactions. The second is the relative deviation of the volume weighted average interest rate between each ordered pair from the volume weighted average among all transactions during the last quarter:

$$w_{ijm}^2 = \frac{\sum_{c=1}^{C_{ij}} a_c i_c / \sum_{c=1}^{C_{ij}} a_c}{\sum_{d=1}^D a_d i_d / \sum_{d=1}^D a_d} - 1/P \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N \frac{\sum_{c=1}^{C_{ij}} a_c i_c / \sum_{c=1}^{C_{ij}} a_c}{\sum_{d=1}^D a_d i_d / \sum_{d=1}^D a_d} \quad (5.2)$$

where a is volume and i the interest rate for each directed transaction, C_{ij} and D are the number of transactions between the ordered pair (i, j) and the total number of all transactions during the last quarter and P is the number of directed pairs of banks which traded during the last quarter. The second term is centering the covariate by subtracting the average of the first term over all active dyads. We call this variable the past rate. It will allow us to measure whether banks prefer to lend to other banks if their previously agreed upon interest rates were relatively high or low. In addition to these actor specific and dyadic covariates, the model will include a number of structural network characteristics in order to assess in how far banks' link formation is tending towards the creation of a particular type of network configuration (see below).

5.3 The Stochastic Actor Oriented Model of Network Formation

The SAOM proposed by Snijders (1996) is a flexible approach to model and analyze longitudinal network data.¹⁰ The actors are supposed to have

¹⁰ It belongs to the agent-based family of models, but since Snijders chose the term actor we stick to this terminology as well. The SAOM approach has some similarity to discrete

full knowledge of the present network including all actors, relations and covariates. This is the case for the e-MID market, since all banks participating have access in real time to at least all the information we use in our analysis. Furthermore, the model assumes that links have the tendency to endure over time which is also crucial for the convergence of the estimation algorithm. Snijders *et al.* (2010) articulate a rule of thumb that a Jaccard index of above .3 indicates a stable link structure, while our observed networks have on average a Jaccard Index of .54 (cf. Table 5.1).¹¹ The evolution is considered to be the result of purposeful actions by independent actors having control of their outgoing links. Certainly, the lending bank has full control of its transactions and should follow a distinct rationale when deciding about its lending relationships. Finally, the actor is supposed to evaluate the network structure and try to obtain an advantageous configuration based on the current state of the network, while modifying its outgoing links. Since banks should try to maximize some kind of objective function by managing their interbank lending operations this seems not restrictive in our case. Hence, our interbank data are in conformity with the basic setting for the application of the SAOM, a network structure undergoing changes in time through deliberate formation of new links, and abortion of existing ones.

5.3.1 Model Specification

A continuous time Markov setting with time parameter $t \in T$ is used and so the actor is solely considering the current state of the network during his decision making process. The M observed networks are embedded in this continuous set of time $T = [t_1, t_M]$, where the unobserved changes take place in between two observations. In our case the network observations are proxied by quarterly aggregated transactions.¹² The N banks which were active during at least two subsequent observations form the set of actors

choice models with social interactions developed in economics (cf. Brock and Durlauf, 2001). One of the main differences is, however, that the latter typically assumes an equilibrium configuration, i.e. every agent takes correctly into account the social influence of himself and all other agents on their peers and the system of interacting choices is observed in a equilibrium of social interactions under rational expectations. With its non-equilibrium setting, the SAOM also bears some resemblance to the type of models of social interactions pioneered by Weidlich and Haag (1983).

¹¹ The Jaccard index $J = \frac{S_{11}}{S_{01} + S_{10} + S_{11}}$ is used to quantify the similarity of two sample sets, where S_{ab} counts the number of relations having status $x_{ij} = a$ at the first instance and $x_{ij} = b$ at the second. It measures the surviving links as a fraction of the links established at any of two adjacent points in time. See Finger *et al.* (2013.) for Jaccard indices for the data under scrutiny measured over different aggregation frequencies.

¹² This is called event data and is e.g. also applied when e-mails are used to proxy human relations.

$\Gamma = N_1, \dots, N_M$. Since we observe entry in and exit from the e-MID over time, the number of actors at time m , N_m , are the subsets of those banks that have been operating in the electronic platform at both t_m and t_{m+1} .¹³ Since the number of banks per quarter is not of much relevance, we will simply drop the time subscript in the following. All possible network configurations constitute the space of potential outcomes.¹⁴ The so-called rate function (RF) and the objective function (OF) together determine the complete model.

The RF $\lambda_i(\rho)$ indicates how often a single bank i will on average consider to change one of its outgoing links (inaction possible) in between two observations $t_m \leq t < t_{m+1}$. Similarly like in other models of social interaction (cf. Lux, 2009), it will be assumed that each actor becomes active with a Poisson rate λ_i that might be varying with time or depending on exogenous factors. The conditional independence of the actors implies that at each point in time t , the time until the next change by any actor follows a negative exponential distribution with parameter $\lambda_+ = \sum_{i=1}^n \lambda_i$. Hence, the expected waiting time is $1/\lambda_+$ and the probability for actor i to be the next actor to change its network position is λ_i/λ_+ . A first step to allow heterogeneity in the rate function is to let the rate of change differ in between the periods ρ_m . The rate could also depend on the actor itself due to actor covariates. Additional parameters would be needed to quantify the strength of the influence of these actor covariates on the activation frequency of agents. Following Snijders and van Duijn (1997) we model the influence of actor covariates by an exponential function

$$\lambda_i(\rho_m, \alpha) = \rho_m \exp \sum_{e=1}^E \alpha_e s_{im}^e, \quad (5.3)$$

to ensure non-negative rates of change for all actors, where s_{im}^e are the E actor covariates as defined for the observation m of the network at time t_m . We will consider only the size of bank i , i.e. v_{im}^1 as defined in sec. 5.2, as covariate in the rate function, with α_e the corresponding coefficient measuring its influence. Note that $E \neq Q$, because not all covariates are implemented in the RF.

The objective function (OF) determining the incremental changes of the actor's network position upon activation is formalized as:

$$o_i(\beta, x, j) = f_i(\beta, x, j) + u_i(x, j). \quad (5.4)$$

¹³ While such movements in and out of the network could be incorporated in the SAOM, we are not covering them in our behavioral modelling. This is because entry and exit are basically consequences of mergers and acquisitions, and therefore are driven by a rationale outside the scope of our model.

¹⁴ There exist $2^{N(N-1)}$ different outcomes defined by whether a link exists or not for each ordered pair of actors.

It indicates the preference of bank i when it is chosen by the RF and determines its action, where x and j denote the dependence on the current state of the network and the counterparty, respectively. It is defined as the sum of the so-called evaluation function (EVF) plus a random element u_i , where β is the vector of parameters of the effects included in the EVF. The EVF expresses the satisfaction or utility of actor i depending on the current network. Possible changes are evaluated in terms of their potential to increase or decrease this level of satisfaction. Some effects focus exclusively on the contribution of already existing links and, hence, measure the change of utility between maintaining an established credit relation or dissolving it. This allows to evaluate whether continuation of a relationship has value by itself on top of contributing to the configuration of connections favored by the agent.¹⁵ The random term in the OF represents the idiosyncratic part of the actor's preference that is not covered by the EVF. It is assumed to follow a Gumbel distribution with mean 0 and scale parameter 1.¹⁶ These random variables are assumed to be independent and identically distributed for all i, j, x and t_m . The new network configuration after the change of an element x_{ij} is denoted by $x(i \rightsquigarrow j)$. The chosen actor faces, therefore, a discrete choice problem with N possibilities (linking to or dissolving the link to the $N - 1$ different j and the option of inactivity denoted by $x(i \rightsquigarrow i)$). The probabilities for the new states, following Maddala (1983), are given by the multinomial logit expression

$$p_{ij}(\beta, x, j) = \frac{\exp(o_i(x(i \rightsquigarrow j)))}{\sum_{n=1}^N \exp(o_i(x(i \rightsquigarrow n)))}. \quad (5.5)$$

This decision making process maximizes the short-term utility of the actor and can be considered as a form of myopic optimization. Hence, the actor is not taking the future ramifications of their own action into account or trying to anticipate actions of others. The EVF is modeled as the sum of a number of effects

$$f_i(\beta, x, j) = \sum_{k=1}^K \beta_k s_{ik}, \quad (5.6)$$

where the weights β_k are statistical parameters indicating the strength of the corresponding effect s_{ik} , and K is the number of effects. These effects are the substantive part of the model and represent possibly meaningful aspects of

¹⁵ In the sociology literature these effects are incorporated in an additional component of the OF called gratification function. However, since only one such effect is implemented in our model we neglect this distinction for the sake of economy in our presentation.

¹⁶ This is the standard assumption for the error term in random utility models, e.g. Maddala (1983).

| Statistic | mean | sd | min | max |
|---------------|---------|--------|------|------|
| Distance | 1520.44 | 434.39 | 818 | 2735 |
| Jaccard Index | .54 | .03 | .47 | .60 |
| Active Banks | 108.97 | 13.62 | 90 | 139 |
| Links | 2535.92 | 725.81 | 1449 | 4062 |
| Density | .21 | .02 | .17 | .25 |
| Mutual Links | 520.50 | 195.49 | 146 | 870 |
| Reciprocity | .20 | .04 | .10 | .29 |
| CC^1 | .58 | .02 | .52 | .61 |
| CC^2 | .15 | .04 | .07 | .24 |

Tab. 5.1: Descriptive statistics calculated for the 36 quarters from 2002-2010.

the network from the viewpoint of the actor. There are two main categories of effects:(1) structural effects depending solely on the network, which can be regarded as endogenous network effects (e.g. reciprocity) and (2) exogenous effects which depend on covariates.

The OF and the RF together define the transition probabilities of the Markov process

$$q_{ij} = \lambda_i(\rho, \alpha)p_{ij}(\beta, x, j). \quad (5.7)$$

To sum up, the RF determines how frequently agents are chosen and the OF defines which action they will undertake with a certain probability to maximize their short-term utility obtained by the network.

5.3.2 Network Effects

The SAOM had been proposed in the first place as a tool to analyze human relationships like friendship networks. In sociology, it is mostly applied to test specific hypothesis (i.e. one or two parameters) and the other effects are implemented as controls to specify the ‘complete’ model. Since this is the first application of a SAOM to interbank data, we will allow for a relatively large number of plausible effects suggested both by the interbank scenario and the former SAOM literature.

The only effect implemented in the flexible part of the rate function is the size v_{im}^1 of the banks. In the form of a categorised variable explained in section 5.2. The effect is included in the RF according to Eq. (5.3):

$$\lambda_i(\rho_m, \alpha) = \rho_m \exp(\alpha v_{im}^1), \quad (5.8)$$

and measures if banks of different size classes are more/less active in creating and ending lending relations.

The objective function is the centrepiece of the model and determines which links are altered by actor i . It consists of the EVF and all implemented effects are summarized in Table 5.2. In the following the terms lending relation or borrowing relation are used as synonyms for outgoing links and incoming links to make the economic implications more transparent. The *outdegree* or *density* effect is defined by the outgoing links of (bank) i ¹⁷

$$s_{i1} = x_{i+} = \sum_{\substack{j=1 \\ j \neq i}}^n x_{ij}, \quad (5.9)$$

i.e. the number of banks i is lending money to. Note that since credit relationships of a bank with itself are never meaningful we will drop $j \neq i$ in the following.¹⁸ Eq. 5.9 represents the basic tendency to establish links. In a very simple model with only the *density* effect a positive parameter would indicate that banks prefer creating new links over ending existing ones and hence would attempt to attract more and more customers in the interbank credit market.¹⁹ We would expect certain limitations of capacity to be present in such a scenario. Indeed, the average density of our network is .20 and since this is well below .5 and quite stable over time we would expect a negative parameter for the *density* effect.

The second effect is the *reciprocity* effect which is defined as:

$$s_{i2} = \sum_{j=1}^n x_{ij}x_{ji}, \quad (5.10)$$

i.e. the number of banks to which i has at the same time a borrowing and a lending relation. A positive parameter of this effect would indicate that banks c.p. are more likely to lend money to another bank if it had already borrowed money from this bank in the same quarter and vice versa. In order to investigate the importance of lasting reciprocal relations, we introduce another related effect called *reciprocity persistence*. This allows us to assess whether banks value already existing mutual credit relations differently from newly created one. This effect is covered by:

¹⁷ We use *italic* letters referring to effects to avoid misunderstandings, since some terms correspond to common network measures.

¹⁸ This does not affect the calculation since all x_{ii} are zero anyway.

¹⁹ In the very long run the density, defined as the fraction of established links out of all possible links of the network would in the absence of countervailing effects tend to 1.

| effect | category |
|----------------------------|------------------|
| density (outdegree) | structural |
| reciprocity | structural |
| reciprocity persistence | structural |
| transitive triplets | structural |
| 3-cycles | structural |
| indegree popularity (sqrt) | structural |
| outdegree activity (sqrt) | structural |
| past trades | dyadic covariate |
| past rates | dyadic covariate |
| covariate popularity | actor covariate |
| covariate activity | actor covariate |
| covariate similarity | actor covariate |

Tab. 5.2: The effects used in the objective function. The columns state the name of the effect and the category. For all four actor covariates (size, core, lending rate, borrowing rate) the last three effects are implemented.

$$s_{i3} = \sum_{j=1} x_{ij}^b x_{ji}, \quad (5.11)$$

where the only difference is the inclusion of x_{ij}^b instead of x_{ij} , which denotes the prevalent state in the previous quarter. Therefore, only if there has existed a link $x_{ij}^b = 1$ at time t_{m-1} could this effect become non zero. A positive (negative) parameter for this effect would indicate that banks are reluctant (positively inclined) to break up a mutual relation. Note that this effect evaluates the persistence of mutual links while the previous *reciprocity* effect indicates whether there is in general preference towards mutual relationships. The intuition is that established mutual relations should be worth more due to investment already undertaken in the past which holds true for human relations.²⁰ However, this reasoning is not as obvious for interbank relations, since credit relations already constitute a mutual contract. The average level of reciprocity, defined as the fraction of reciprocated relations among all established relations, is 20 percent (cf. Table 5.1) indicating that reciprocity in credit relations might be of minor importance.

The *transitive triplets* effect and the *3-cycles* effect investigate different aspects of transitivity, with transitivity referring to the influence of indirect

²⁰ See e.g. Schaefer *et al.* (2010) and Agneessens and Wittek (2012).

relations. The *transitive triplets* effect for i is defined as²¹

$$s_{i4} = \sum_{j,k=1}^n x_{ij}x_{jk}x_{ik}, \quad (5.12)$$

or the number of triadic relations in which i has a lending relation to j and k and additionally bank j has a lending relation to k . The corresponding parameter indicates, whether the indirect relation between j and k influences the probability of a lending relation to be formed from i to k or j . The *3-cycles* effect is defined by the sum of paths of length three completing a cycle between banks (i,j,k) ²²

$$s_{i5} = \sum_{j,k=1}^n x_{ij}x_{jk}x_{ki}, \quad (5.13)$$

i.e. the number of triadic relations for bank i in which the flow of money completes a circle of length 3. It measures whether the two intermediate relations $x_{jk} = 1$ and $x_{ki} = 1$ make the link between i and j more likely. Note that the *transitive triplets* effect is in line with a locally hierarchical structure, while the *3-cycles* effect captures more of an extended reciprocity with a tendency towards emergence of clusters of mutual lenders and borrowers and, hence, is not in line with local hierarchy. The two triadic structures as shown in Figure 5.1 can be used to calculate two types of directed clustering coefficients.²³ The statistics CC^1 in Table 5.1 with an average of 58 percent of all possible triplets corresponds to the *transitive triplets* effect and CC^2 with an average of 15 percent to the *3-cycles* effect. These values in relation to the density indicate that links organized in a triadic structure as in the *transitive triplets* effect appear 2.74 more often than a random link, while links completing a cycle are 27% less likely to be formed. It is this heterogeneity in the occurrence of different types of triplets in the data that motivates us to include the pertinent effects in our model.

The final two structural effects implemented in the model are capturing the influence of the in- and outdegrees of agents. The *indegree-popularity* effect is defined as the sum of the square roots of the indegrees from the banks to which bank i has an outgoing link or the square root of the sum of the borrowing relations of those banks to whom bank i has a lending

²¹ In the following it is understood that $i \neq j \neq k$ holds.

²² A connection along (directed) links between two nodes i and j is called a path and the length of the path is the number of links crossed.

²³ For the calculations see Finger *et al.* (2013).

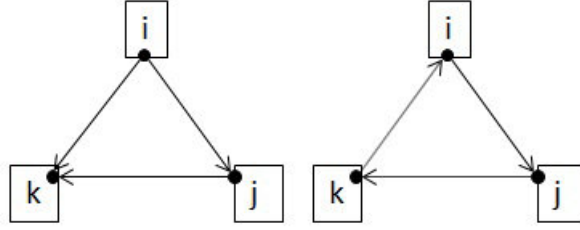


Fig. 5.1: The left triadic structure corresponds to the *transitive triplets* effect CC^1 . The right triadic structure illustrates the *3-cycles* effect CC^2 and is non-hierarchical, since all nodes have the same roles as lenders and borrowers.

relationship:

$$s_{i6} = \sum_{j=1}^N x_{ij} \sqrt{\sum_{h=1}^n x_{hj}}. \quad (5.14)$$

It investigates whether banks with many borrowers are more ‘popular’ and tend to be accepted more easily as borrowers by other banks, where the square root takes capacity constraints into account.²⁴ This will lead to or sustain a certain heterogeneity of the indegrees and would make high degree nodes persistent over time. Therefore, it would also indicate hierarchy at the global level. The *outdegree-activity* effect is defined as the sum of the outdegree of bank i times the square root of this sum or the square root of lending relations of bank i times the lending relations of bank i

$$s_{i7} = \sum_{j=1}^N x_{ij} \sqrt{\sum_{j=1}^n x_{ij}}, \quad (5.15)$$

where the square root again takes capacity constraints into account.²⁵ The effect measures if there is a cumulative effect of ‘active’ banks with already many lending relations tending to be more eager to form additional connections. Analogous to the *indegree-popularity* effect a positive parameter would exert an influence towards a high dispersion among the outdegrees and persistence of their distribution.

²⁴ We could also design this effect as a linear one. However, the linear *indegree-popularity* effect is most of the time insignificant if both effects are implemented.

²⁵ The linear effect without the square root is again left out, because it is most of the time insignificant if both effects are implemented.

The dyadic covariates are the the square root of trades (*past trades*) and the average rate (*past rates*) each for the last year. In both cases only the so-called main effect is implemented, which is defined as the sum of i 's outdegree weighted with the corresponding value for the dyadic covariate. The past trades effect is

$$s_{i8} = \sum_{j=1}^N x_{ij} w_{ij}^1. \quad (5.16)$$

It measures whether lending relations are more likely if banks had already often exchanged liquidity in this direction during the last year. Note that the centering ensures (for all covariates) that a positive parameter indicates at the same time banks favouring regularly (above average) used relations and avoiding sparsely (below average) used ones. The *past rates* effect

$$s_{i9} = \sum_{j=1}^N x_{ij} w_{ij}^2, \quad (5.17)$$

measures the influence of interest rates between each ordered pair of banks of the last quarter, with w_{ij}^2 as defined in eq. (5.2).

The actor covariates are the size of the banks (as categorical variable as introduced above), whether the bank was assigned to the core or not according to Fricke and Lux (2014), the standardized lending rate and the standardized borrowing rate. For each of these actor covariates the same three effects are implemented in the model. The *size popularity* (or alter) effect is defined as the sum of i 's outgoing links weighted with the size v_j^1 of bank j ,

$$s_{i10} = \sum_{j=1}^N x_{ij} v_j^1. \quad (5.18)$$

A positive parameter for this effect indicates that banks prefer lending relations to large banks which themselves would have more borrowing relations or a higher indegree and were therefore more 'popular'. Due to the centering a positive parameter indicates that banks prefer lending to banks which are above the average in terms of size. Hence, at the same time a positive parameter indicates that smaller than average banks are less attractive borrowers. Without the centering such a distinction would not be possible and a positive parameter would imply that all banks would be (to different degrees) more attractive.²⁶ The *size activity* (or ego) effect is the sum of i 's outdegree times

²⁶ The interpretation would not be affected, since larger banks are in both cases more attractive. However, the logic that all banks independently of their actual size are more attractive because of their size is problematic and other effects (mainly *density*) would have to be adjusted for this.

its own size v_i^1

$$s_{i11} = \sum_{j=1}^N v_i^1 x_{ij} \quad (5.19)$$

and measures if the size of banks has an influence on their ‘activity’, i.e. the number of lending relations. The *size similarity* effect is defined as the sum of i 's outdegree each weighted with a similarity score

$$s_{i12} = \sum_{j=1}^n x_{ij} \text{sim}(v_i^1, v_j^1), \quad (5.20)$$

where $\text{sim}(v_i, v_j) = 1 - \frac{|v_i - v_j|}{R_v}$ and $R_v = \max|v_i - v_j|$. Thus, the similarity is calculated as 1 minus the absolute dissimilarity of the pair (i, j) with respect to the range R_v of the normalized covariate and is therefore restricted to the range $[0, 1]$. A positive parameter for this effect indicates that banks with a similar size prefer to have relations with each other. The intuition would suggest that larger banks have more lending as well as borrowing relations and so we expect to find a positive parameter for the first two effects. The *size similarity* effect could be motivated by banks of the same size class being hit by similar or dissimilar liquidity shocks which would lead to preferences for connecting to banks of a similar size or of a different size category. For the core we also implemented the *core popularity*, *core activity* and *core similarity* effects

$$s_{i13} = \sum_{j=1}^N x_{ij} v_j^2 \quad (5.21)$$

$$s_{i14} = \sum_{j=1}^N v_i^2 x_{ij} \quad (5.22)$$

$$s_{i15} = \sum_{j=1}^n x_{ij} \text{sim}(v_i^2, v_j^2). \quad (5.23)$$

As with the size we would expect core banks to have more lending and borrowing relations and therefore positive parameters for s_{i12} and s_{i13} . In the case of the *core similarity* effect two countervailing forces could be at work. On the one hand, core banks are expected mostly to trade with each other, while peripheral banks are supposed not to. The larger number of peripheral banks might lead to a negative result.

Also for the lending rate the *lending rate popularity*, *lending rate activity* and *lending rate similarity* (with lending rates normalised as volume weighted

deviations from the average a defined in sec. 5.2)

$$s_{i16} = \sum_{j=1}^N x_{ij} v_j^3 \quad (5.24)$$

$$s_{i17} = \sum_{j=1}^N v_i^3 x_{ij} \quad (5.25)$$

$$s_{i18} = \sum_{j=1}^n x_{ij} \text{sim}(v_i^3, v_j^3) \quad (5.26)$$

are considered as effects in the EVF. The intuition for these covariates could be that banks acting as market makers in the interbank market are able to charge their customers for this service and since market makers should be large intermediaries we expect again positive coefficients for the *popularity* and *activity* effects. If we adapt the same reasoning for the *lending rate similarity* effect we might expect a negative coefficient, since the market maker able to charge a high lending rate should trade with smaller banks rather than with other market makers.

The last actor covariate is the borrowing rate and again the same three effects

$$s_{i19} = \sum_{j=1}^N x_{ij} v_j^4 \quad (5.27)$$

$$s_{i20} = \sum_{j=1}^N v_i^4 x_{ij} \quad (5.28)$$

$$s_{i21} = \sum_{j=1}^n x_{ij} \text{sim}(v_i^4, v_j^4) \quad (5.29)$$

are included in our analysis. The most interesting effect should be the *borrowing rate popularity*, since a high borrowing rate should indicate a higher likelihood of default and, hence, higher counterparty risk would be associated with granting liquidity to this bank. Therefore, we expect a negative coefficient (β_{18}) for this effect, especially for the time after the start of the GFC. Not so obvious is what implication a high borrowing rate should have on the lending activity of banks. In case of the *borrowing rate similarity* one could argue similarly to the *lending rate similarity* effect that some kind of hierarchical ordering makes it more likely that more dissimilar banks are trading with each other.

5.3.3 Model Estimation

The huge number of possible network states ($2^{N(N-1)}$) makes an analytical solution of probabilities and expected values impossible, but Snijders (1996, 2001) introduced a method of simulated moments (MoM) approach to overcome this problem. Additionally, Snijders *et al.* (2010) implemented a Maximum Likelihood approach (ML) and Koskinen and Snijders (2007) a Bayesian method. The reason for sticking to the MoM is twofold. First, simulation studies in Snijders *et al.* (2010) have shown that the higher efficiency of the ML method is only relevant for small data sets, whereas ours can be considered as large in terms of the numbers of actors and the numbers of the changing relations in between two observations (distance) in comparison to previous applications. Second, the likelihood and Bayesian estimation methods require much more computing time which is particularly relevant for large networks like ours.²⁷ The remaining part of this section will briefly introduce the MoM approach.

The general idea is that sample moments (statistics) are the natural estimators of population moments (expected values of statistics). Since the test of Lospinoso *et al.* (2011) detects time heterogeneity in our application we estimate the model in a first step separately for each period and so we will obtain estimates of the period-specific set of parameters $\theta_m(\rho_m, \alpha_m, \beta_m)$ with a total number of parameters $L = 1 + E + K$. The parameters are estimated by equating the sample statistics z and the expected values of Z .²⁸ Hence, it is necessary to choose L statistics to solve the L -dimensional moment equation

$$\mathbb{E}_\theta Z = z. \quad (5.30)$$

The statistics Z have to be chosen such that the expected value $\mathbb{E}_\theta Z$ is sensitive to the parameter vector θ_m . The first quarterly network is only used to condition on and, hence, no parameter estimates are obtained for this observation. Note that our empirical data will not be assumed to be observed in a stationary state or equilibrium of the actor-based choice process. As a consequence, we condition on the previous observation allowing for transient adjustment prior to convergence to a stationary distribution. Hence, each observation $t_m (m \leq M - 1)$ is used as a conditioning event for the distribution of $x(t_{m+1})$ and so the simulated network is reset after each period to the observed network. This makes the estimation per period and

²⁷ Snijders *et al.* (2010) report a computational time for the ML approach for a smaller data set with much less effects that is about 17 times longer than for the MoM approach.

²⁸ Capital letters refer to the expected statistics Z and networks X constructed for simulations of the model, whereas z and x indicate observed statistics and networks.

the implementation of time dummies for all periods and effects equivalent. Following this, the moment equation for period t_{m+1} can be written as

$$\mathbb{E}_\theta\{Z(X(t_m), X(t_{m+1}))|X(t_m) = x(t_m)\} = z(x(t_m), x(t_{m+1})). \quad (5.31)$$

A sensitive statistic for the rate of change ρ_m which determines the expected ‘amount of change’ is

$$C = \sum_{i,j}^N |X_{ij}(t_{m+1}) - X_{ij}(t_m)|, \quad (5.32)$$

the ‘observed total number of changes’. For the structural characteristics additionally included in the RF an appropriate statistic is

$$C_e = \sum_{i,j}^N s_{im}^e |X_{ij}(t_{m+1}) - X_{ij}(t_m)|, \quad (5.33)$$

where the ‘observed total amount of change’ is weighted with s_{im}^e and, hence, a positive parameter estimate α_e in Eq. (5.3) indicates that banks with a higher value of s_{im}^e are more active.

For the parameter β_k in the EVF a higher value of β_k means that all actors try to target a higher value of s_{ik} . An adequate statistic is

$$S_k = \sum_{i=1}^N s_{ik}(X(t_{m+1})), \quad (5.34)$$

because it is summing up the effect over all actors for the second observation period ($m+1$). The statistic for the *reciprocity* effect for example counts the number of reciprocal ties at time $m+1$. At the beginning of each simulation of the network, the structural statistics and the covariates are assigned their values obtained in t_m , while during the simulation the network and the structural statistics may change while the covariates are fixed. For instance the first altered credit relation affects many structural statistics like out- and indegree and, thus, the next actor drawn by the RF faces a slightly different discrete choice problem. Note that in case of the *reciprocity persistence* effect x_{ij}^b remains fixed and so only the surviving links are counted in this statistic. For instance, a positive value for the *reciprocity persistence* effect would yield more mutual links in general, but would achieve this by a higher number of surviving links and not through the number of new mutual links created during this period.

For the numerical solution of the moment equations a version of the iterative Robbins-Monro algorithm (Robbins and Monro, 1951) is applied. After

each run (r) the new parameter vector $\hat{\theta}_{r+1}$ is obtained from the old values $\hat{\theta}_r$ and the difference between simulated and observed statistics. One such step is defined by

$$\hat{\theta}_{r+1} = \hat{\theta}_r - a_\tau D^{-1}(Z_r - z), \quad (5.35)$$

where Z_r is the simulated value with parameters $\hat{\theta}_r$ for run r and $a_\tau D^{-1}$ determines the speed of adjustment of the parameter vector θ . The coefficient a_τ is decreased towards zero over the estimation process and defines the general speed of the adjustments. D is the estimate of the diagonal matrix $\frac{\partial}{\partial \theta} \mathbb{E}_\theta Z$ measuring how much $\mathbb{E}_\theta Z$ is changing with respect to the specific parameters in θ_r . Hence, after each run the parameters are changed depending on the discrepancy ($Z_r - z$) between the simulated Z_r and observed statistic z .²⁹

The complete estimation procedure consists of three phases. The first phase contains a small number of 50 runs to get a ‘rough estimate’ of $\frac{\partial}{\partial \theta} \mathbb{E}_\theta Z$ to define D . In this phase θ is fixed at pre-specified values. The starting values for the very first simulation are set equal to zero for all parameters except for the basic rate parameter ρ_m .³⁰

In the second phase the Robbins-Monro algorithm is applied to estimate θ using D . It consists of 4 subphases, where $a_\tau = \{.4, .2, .1, .05\}$ remains constant during subphases and decreases between subphases.³¹ At the start of a new subphase the average parameter values of the last subphase are used as the starting point ($\hat{\theta}_r$) in the Robbins-Monro algorithm (Eq. 5.35). Accordingly, the final estimate ($\hat{\theta}$) is the average of the parameter values obtained during the 4th subphase. In the third and final phase θ is kept constant at the estimated values for 5000 runs. The results of this phase are used to calculate standard errors and check the convergence of the algorithm.³²

²⁹ The covariance matrix is approximated as in Schweinberger and Snijders (2007) by applying the delta method and the implicit function theorem.

³⁰ ρ_m is set equal to the total change in between t_m and t_{m+1} , because it is necessary to have a positive rate function to initiate a simulation.

³¹ The end of each subphase is reached when the quasi autocorrelations $(Z_r - z) * (Z_{r-1} - z)$ of two subsequent runs become negative for all statistics. This means that the adjustment direction for every parameter has been changing for the last two updates of θ and so one can assume that the trajectory had reached a limit point.

³² If the convergence turns out to be not sufficient another run of the Robbins-Monro algorithm is initiated. In this new run, phase 1 is skipped and instead the estimated parameter values of the previous run are used to restart the estimation procedure with phase 2.

5.4 Estimation Results

5.4.1 Full Sample Estimation

In our application all effects introduced above are implemented.³³ The period investigated extends from the first quarter of 2002 until the end of 2010.³⁴ Since, the test of Lospinoso *et. al* (2011) points to time heterogeneity for almost all effects and periods we estimate the model per period rather than introducing time dummies, while both procedures should yield the same results. The t-statistics for convergence are always below .1 for all effects and indicate excellent convergence of the algorithm according to Snijders *et al.* (2010).³⁵

To analyze the common significance of each effect for the whole sample, we apply the two-step multilevel analysis proposed by Snijders and Baerveldt (2003). First, a preliminary mean is estimated by taking the average of the 36 estimates. This preliminary estimate is unbiased but not efficient and its statistical error stems from two sources: the variance in every period for which we have an estimate (the variance of each parameter estimate itself) and the variance across the periods. Next, we use the known variance of the single estimates to get an unbiased estimator for the variance across periods. Now we can determine the statistical error for each period separately and we use this in a weighted least squares regression in which the weights depend negatively on the variance of the single estimates. The results for the complete sample are summarized in Table 5.3.³⁶ The first column states the effect and the second to which category it belongs. The third to fifth column display the estimates for the mean parameter, its standard error and the corresponding p value. The last column shows for how many estimation periods the effect was different from zero at a 5% significance level, where the first term in brackets indicates positive significance, the second negative significance and the third lack of significance.³⁷

³³ All calculations in this paper are done using R (cf. R Development Core Team, 2012). For estimating the SAOM the routines of Ripley *et. al* (2013) were applied.

³⁴ The transactions of the year 2001 are used to calculate the first set of dyadic covariates (*past trades*, *past rates*) and the network of the 4th quarter 2001 is used for initialization.

³⁵ They are calculated as average standardised deviations of the simulated and observed statistics of the MoM approach over the 5000 runs in phase 3, i.e. dividing the mean differences by the standard deviation over all simulations.

³⁶ As mentioned above, the model has been estimated in each period by considering only those banks active in t_m and t_{m+1} . Market entry and exit are, therefore, neglected as they will certainly be driven by forces outside of our present framework, i.e. mergers and acquisitions. We estimated the model also in different specifications, where e.g. banks were allowed to enter and leave the market, but got always very similar results.

³⁷ The results are in general qualitatively the same for significance levels of 1% and 10%

Interpreting the results it is necessary to keep in mind the differences between interbank credit relations and human relationships. For instance, all credit relations are mutual contracts, yet we define the flow of money and the corresponding risk taken by the lending bank as the direction of the link.

The rate is 36.055 on average and is necessarily positive, since two subsequent networks are always different. Figure 5.2 shows that the estimated rate of activation and the number of banks both decrease during the sample period and are highly correlated (with a correlation coefficient of .689). However, the rate tends to be much more volatile and we especially observe outbursts of activity before and at the beginning of the GFC, which could be seen as indicators for a ‘nervous’ market environment. The rate parameter drops sharply after the last and most pronounced spike at the beginning of the GFC and assumes its lowest values over the whole time horizon during the continuation of the GFC in 2009/10 which very likely is the consequence of overall reduced activity in the interbank market and a stronger concentration on well-established links. The *size* effect implemented in the RF has an estimated mean of .1566 and is significant at the 1% level. For instance, a bank of size category 5 acts on average about 2 times as often as a bank of size= 1.³⁸ Note that the RF covers activity both in creation and termination of links. The spike at the beginning of the GFC was accompanied by a strong decline in the density (Fricke and Lux (2014)), however, a high rate could also come along with an increasing or stable density. Hence, this effect allows by itself no prediction of the development of such network statistics like the density.

Now we turn to the OF, which as stated above determines why banks establish or dissolve outgoing links. The *density* effect is always negatively significant with a mean of -4.201 indicating capacity constraints in the accumulation of links. Note, however, that many other effects do also include the outdegree of actor i , so that the coefficient alone does not explain completely the evolution of the density.³⁹

Since both the *reciprocity* effect and the *reciprocity persistence* effect are implemented we can distinguish between the tendencies for creation and deletion of mutual links. The coefficients are .680 for the *reciprocity* effect and $-.737$ for the *reciprocity persistence* effect, and both effects are significant at the 1% level.⁴⁰ This shows that banks prefer to create mutual relations,

and all results referred to but not explicitly shown are available from the corresponding author upon request.

³⁸ To calculate this it is necessary to insert the centered covariate values in Eq. (5.8). The effect was four times insignificant.

³⁹ It acts similar to an intercept in a regression framework.

⁴⁰ The single effects per period are 13 and 28 times not significant, but Fisher’s combi-

| effect | category | mean | s.e. | p-value | significance |
|-------------------------|-------------|---------|----------|---------|--------------|
| rate | | 36.055 | | | |
| size | rate | .1566 | (.00084) | .001*** | [32 - 4] |
| density (outdegree) | structural | -4.2010 | (.0647) | .001*** | [- 36 -] |
| reciprocity | structural | .6797 | (.0508) | .001*** | [23 - 13] |
| reciprocity persistence | structural | -.7371 | (.0957) | .001*** | [- 8 28] |
| transitive triplets | structural | -.0029 | (.0010) | .006*** | [- 9 27] |
| 3-cycles | structural | -.0236 | (.0026) | .001*** | [- 34 2] |
| indegree pop. (sqrt) | structural | .3580 | (.0061) | .001*** | [36 - -] |
| outdegree act. (sqrt) | structural | .2157 | (.0056) | .001*** | [36 - -] |
| past trades | dyadic cov. | .1946 | (.0086) | .001*** | [36 - -] |
| past rates | dyadic cov. | .4087 | (.3959) | .309 | [9 7 20] |
| size popularity | actor cov. | .0065 | (.0078) | .410 | [8 6 22] |
| size activity | actor cov. | -.0437 | (.0115) | .001*** | [4 21 11] |
| size similarity | actor cov. | .0536 | (.0170) | .003*** | [6 - 30] |
| core popularity | actor cov. | .2174 | (.0239) | .001*** | [27 - 9] |
| core activity | actor cov. | .2497 | (.0315) | .001*** | [27 1 8] |
| core similarity | actor cov. | -.0062 | (.0129) | .632 | [3 6 27] |
| lending rate pop. | actor cov. | -.0478 | (.0754) | .530 | [2 4 30] |
| lending rate act. | actor cov. | -.2855 | (.1831) | .128 | [5 9 22] |
| lending rate sim. | actor cov. | -.1581 | (.0618) | .015** | [3 11 22] |
| borrowing rate pop. | actor cov. | -.3885 | (.2820) | .177 | [8 13 15] |
| borrowing rate act. | actor cov. | .0991 | (.1179) | .407 | [5 3 28] |
| borrowing rate sim. | actor cov. | -.1146 | (.0547) | .044** | [3 7 26] |

Tab. 5.3: The estimation results for the 36 quarters from 2002 until 2010. The columns state the name of the effect, the category it belongs to, the estimated mean, its standard error and p-value from the two-step weighted least squares approach following Snijders and Baerveldt (2003) and in how many periods the estimated parameters were significant on the 5% level [positive significance|negative significance|lack of significance].

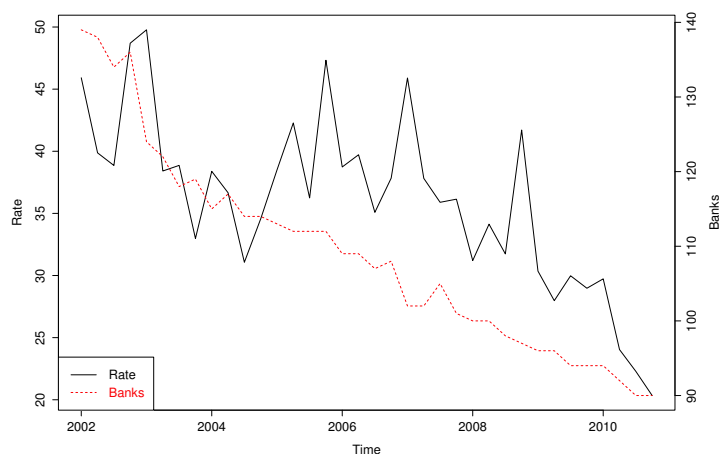


Fig. 5.2: The evolution of the basic rate parameter and the number of banks (dashed) from 2002Q1 to 2010Q4.

i.e. already borrowing money from a bank makes it more likely to also lend money to it. On the other hand banks value maintaining such mutual relations less than creating them and on balance, i.e. summing up both effects ($.6797 + (-.7371) = -.0574$), they appear practically indifferent between dissolving a mutual or a non-mutual relation.⁴¹ The negative result for the *reciprocity persistence* effect stands in stark contrast to all former applications of the SAOM. The reason is that humans highly value mutual relations and even more those in which they already ‘invested’ effort and time.⁴² However, in case of the interbank market relations predominantly go in one direction with the extreme case of (mainly) small banks using the market almost exclusively to lend or borrow. Hence, the intuition for this results should be that often strong one-sided relations exist, which are relatively seldomly reciprocated. To give an example, bank i might be borrowing money from bank j ($x_{ji} = 1$) on a regular basis. Because of this relation it becomes more likely that, if bank i has an excess of liquidity bank j will again be the counterparty ($x_{ij} = 1$). In the next period x_{ij} is likely to be zero again, since bank i might

nation of one-sided p-values indicates that it can be excluded that the parameters have the opposite sign of their estimated means.

⁴¹ If we introduce only the *reciprocity* effect without distinguishing between formation and abortion of a mutual link it is insignificant for almost all specifications tested. This observation supports our interpretation in the main text.

⁴² Schaefer *et al.* (2010) and Agneessens and Wittek (2012) included solely the *reciprocity* effect, whereas Snijders *et al.* (2010) implemented the *reciprocity persistence* effect as well.

not lend money at all. This behaviour leads to the finding that banks have a certain preference for mutual relations, but are not inclined to purposefully maintain them over time.

The *transitive triplets* effect ($-.0029$) and the *3-cycles* effect ($-.0197$) are both negatively significant at the 1% level. Note that interpreting the ‘strength’ of these effects one has to keep in mind that, in contrast to e.g. *reciprocity*, a link can close several triadic structures at once. The negativity of the *3-cycles* effect indicates local hierarchy and is in line with the corresponding low directed clustering coefficient CC^2 . On the other hand the high value for the other directed clustering coefficient CC^1 stands in contrast to the significantly negative coefficient for this configuration in the evaluation function. It, thus, appears that the high empirical value of CC^1 is not generated consciously by the agents, but apparently emerges out of the interplay of other effects.⁴³ Hence, indirect links do not work by themselves as an incentive for banks to form a directed link, which supports the idea that banks are mainly interested in their direct counterparty. This is again in contrast to human relations, where the *transitive triplets* effect is often found to be positive, while the results concerning the *3-cycles* effect are mixed suggesting that some groups are hierarchically ordered but others are not.⁴⁴

The last two entirely structural effects are the *indegree popularity* effect and the *outdegree activity* effect. Both are positively significant at the 1% level with respective coefficients of .358 and .216.⁴⁵ Hence, there appear positive feedback effects to be working in both cases. As a consequence both degree distributions should display persistent dispersion in the sense that banks with many (few) lending or borrowing relations are likely to have many (few) in the future as well. Moreover, this provides evidence for an enduring hierarchical structure with respect to degrees, while it does not indicate that banks with a high indegree would also automatically have a high outdegree.⁴⁶

After considering the structural effects we turn to the covariate related effects. The square root of trades in the last year is a dyadic covariate and thus specific for each ordered pair (i, j) . The corresponding *past trades* effect

⁴³ Estimating a very simple model with just *density*, *reciprocity*, *transitive triplets* and *3-cycles* we get higher values and always significant results in the expected direction to the corresponding clustering coefficients for both triadic effects.

⁴⁴ See for a positive *transitive triplets* effect but negative *3-cycles* effect in human relations Cheadle and Schwadel (2012) and Agneessens and Wittek (2012), whereas Lazega *et al.* (2012) found both effects to be positive.

⁴⁵ This is often found for human relations as well. See e.g. Lazega *et al.* (2012).

⁴⁶ The correlation between both degree distributions is on average only .12 and for some periods even negative.

is positively significant with a mean of .194 and shows that banks prefer enduring lending relations. Hence, trading history is strongly affecting the likelihood of future trades even in an electronic broker market such as the e-MID.⁴⁷ The *past rates* effect on the other hand is insignificant. Hence, a specifically low or high rate realized for the ordered pair (i,j) in the past is not systematically affecting the decisions of bank i with respect to lending money to bank j in the future.⁴⁸

The parameter values of the actor covariate effects for the single periods often change their sign and except for the size *similarity* effect and the core *popularity* effect are even positively and negatively significant at different points in time. Applying a Wald test to the three covariate related effects combined indicates that they are often significant. To be precise, the according test statistics shows 27 times for the size, 36 times for the core, 19 times for the lending rate and 25 times for the borrowing rate significance at the 5% level. There is no clear pattern visible for the significance for any specific time period or interplay among different effects.

The size and core covariates might measure similar attributes of the banks, since it could be argued that mostly large banks might belong to the core. The results, however, disagree to some extent with this intuition. The respective estimated coefficients are .007 for the *size popularity* effect and $-.044$ for the *size activity* effect and only the latter is significant at the 1% level. So larger banks are surprisingly not more popular and even less active. The *similarity* effect (.054) is positively significant and shows that banks of a more similar size are more likely to trade with each other. Core banks on the other hand are more popular (.142) and more active (.293) and both estimates are significant at the 1% level. The significance of these effects, keeping in mind that we additionally control for e.g. *indegree popularity* and *outdegree activity* which also detect hierarchical structures, shows that the core-periphery model is able to retrieve additional information in detecting intermediaries beyond a mere categorisation of the size. The *core similarity* effect ($-.006$) is insignificant, which should be due to its two opposing forces namely that core banks are supposed to trade with each other, while the opposite should hold for the peripheral banks.

All interest rate related effects for the single periods are often insignificant and repeatedly switch their sign. Both *similarity* effects are negatively significant at the 5% level. This indicates that banks prefer to trade with

⁴⁷ To get an understanding of how strong this effect actually is we have to consider that the covariate assumes values up to 17.52. We alternatively considered other covariates to measure the strength of past relations, e.g. following Affinito (2012) the duration, and got very similar results.

⁴⁸ The effect is 7 times negatively and 9 times positively significant per period.

banks with a different profile, which again could be interpreted as some kind of hierarchical ordering in the interbank market. Note that during a transaction the rate is obviously the same for both counterparties, yet for one it is the borrowing rate and for the other the lending rate and the low correlation ($-.075$) between the two average interest rates across banks indicates that a high lending rate reveals almost no information about the borrowing rate and vice versa. The other results are surprising, since no other effect is significant and so both interest rates seem not to systematically affect the popularity or activity of a bank. For instance, a higher borrowing rate that arguably should indicate a higher likelihood of default does not influence other banks directly in their decisions of lending money to this bank.⁴⁹ In total only 2 out of 6 effects related to interest rate actor covariate are significant and some of the results are counter-intuitive. However, this might be expected for the tranquil period until the GFC, where the interest rates were overall very similar.

5.4.2 Before and After the Financial Crisis

In order to see whether the structure of the interbank market experienced significant changes with the onset of the financial crisis we split the sample into two parts. Again, the two-step weighted least squares approach from Snijders and Baerveldt (2003) is used to estimate the model for 26 periods from the first quarter 2002 until the second quarter 2008 and the remaining 10 periods from the 2nd quarter of 2008 until the 4th quarter of 2010. While it might be somewhat hard to say when the GFC exactly started, we believe that letting the second sample begin with the default of Lehman Brothers constitutes the most plausible way of splitting up the data into a pre-crisis and crisis period.

The results in Table 5.4 for the pre-crisis period show that the overall results differ only in very few aspects from the results for the complete sample, which might have been expected since it is the longer and more tranquil period. Thus, we will briefly address the only change before we are going to compare the two subperiods with each other.⁵⁰ The *lending rate activity* effect ($-.462$) is now becoming significant at the 10% level and so in the tranquil period we see the expected economic relationship of a lower lending rate leading to a higher number of borrowers (i.e. higher demand for credit).

The next step is to compare these results to the ones in Table 5.5 for the time after the onset of GFC. Again, many results are very similar to the

⁴⁹ It is 8 times even positive significant indicating high variation.

⁵⁰ Additionally, *transitive triplets* and the *lending rate similarity* effect are now only significant at the 10% level (complete sample: 1% and 5%, respectively).

| effect | category | mean | s.e. | p-value | significance |
|-------------------------|-------------|---------|---------|---------|--------------|
| rate | | 38.879 | | | |
| size | rate | .1613 | (.0078) | .001*** | [25 - 1] |
| density (outdegree) | structural | -4.2733 | (.0610) | .001*** | [- 26 -] |
| reciprocity | structural | .7124 | (.0558) | .001*** | [20 - 6] |
| reciprocity persistence | structural | -.8384 | (.0960) | .001*** | [- 7 19] |
| transitive triplets | structural | -.0018 | (.0010) | .081* | [- 6 20] |
| 3-cycles | structural | -.0175 | (.0011) | .001*** | [- 25 1] |
| indegree pop. (sqrt) | structural | .3535 | (.0068) | .001*** | [26 - -] |
| outdegree act. (sqrt) | structural | .2125 | (.0065) | .001*** | [26 - -] |
| past trades | dyadic cov. | .2025 | (.0083) | .001*** | [26 - -] |
| past rates | dyadic cov. | .6710 | (.5393) | .225 | [9 5 12] |
| size popularity | actor cov. | .0000 | (.0100) | .998 | [6 6 14] |
| size activity | actor cov. | -.0595 | (.0114) | .001*** | [1 17 8] |
| size similarity | actor cov. | .0474 | (.0152) | .004*** | [4 - 22] |
| core popularity | actor cov. | .1890 | (.0289) | .001*** | [18 - 8] |
| core activity | actor cov. | .2175 | (.0359) | .001*** | [18 1 7] |
| core similarity | actor cov. | -.0210 | (.0146) | .164 | [2 6 18] |
| lending rate pop. | actor cov. | -.0210 | (.0776) | .789 | [1 2 23] |
| lending rate act. | actor cov. | -.4620 | (.2389) | .065* | [3 6 17] |
| lending rate sim. | actor cov. | -.1670 | (.0826) | .054* | [3 9 14] |
| borrowing rate pop. | actor cov. | -.5774 | (.3833) | .145 | [7 10 9] |
| borrowing rate act. | actor cov. | -.0387 | (.1196) | .749 | [2 3 21] |
| borrowing rate sim. | actor cov. | -.1476 | (.0604) | .022** | [1 5 20] |

Tab. 5.4: The estimation results for the 26 quarters from the 1st quarter of 2002 until the 2nd quarter of 2008. The columns state the name of the effect, the category it belongs to, the estimated mean, its standard error and p-value from the two-step weighted least squares approach following Snijders and Baerveldt (2003) and in how many periods the estimated parameters were significant on the 5% level [positive significance|negative significance|lack of significance].

| effect | category | mean | s.e. | p-value | significance |
|-------------------------|-------------|---------|---------|---------|--------------|
| rate | | 28.713 | | | |
| size | rate | .1421 | (.0252) | .001*** | [7 -3] |
| density (outdegree) | structural | -4.0008 | (.1197) | .001*** | [- 10 -] |
| reciprocity | structural | .5508 | (.1166) | .001*** | [3 -7] |
| reciprocity persistence | structural | -.3445 | (.2346) | .176 | [- 19] |
| transitive triplets | structural | -.0080 | (.0023) | .007*** | [- 3 7] |
| 3-cycles | structural | -.0485 | (.0062) | .001*** | [- 9 1] |
| indegree pop. (sqrt) | structural | .3722 | (.0130) | .001*** | [10 - -] |
| outdegree act. (sqrt) | structural | .2255 | (.0112) | .001*** | [10 - -] |
| past trades | dyadic cov. | .1742 | (.0219) | .001*** | [10 - -] |
| past rates | dyadic cov. | -.2728 | (.3172) | .412 | [- 2 8] |
| size popularity | actor cov. | .0232 | (.0077) | .015** | [2 -8] |
| size activity | actor cov. | -.0006 | (.0255) | .982 | [3 4 3] |
| size similarity | actor cov. | .0697 | (.0532) | .223 | [2 -8] |
| core popularity | actor cov. | .3005 | (.0291) | .001*** | [9 -1] |
| core activity | actor cov. | .3386 | (.0584) | .001*** | [9 -1] |
| core similarity | actor cov. | .0453 | (.0210) | .059* | [1 -9] |
| lending rate pop. | actor cov. | -.2075 | (.2002) | .327 | [1 2 7] |
| lending rate act. | actor cov. | .1316 | (.2329) | .586 | [2 3 5] |
| lending rate sim. | actor cov. | -.1279 | (.0878) | .179 | [- 2 8] |
| borrowing rate pop. | actor cov. | .0501 | (.2591) | .851 | [1 3 6] |
| borrowing rate act. | actor cov. | .4481 | (.2664) | .127 | [3 -7] |
| borrowing rate sim. | actor cov. | -.0196 | (.1186) | .872 | [2 2 6] |

Tab. 5.5: The estimation results for the 10 quarters from the 2nd quarter of 2008 until the 4th quarter of 2010. The columns state the name of the effect, the category it belongs to, the estimated mean, its standard error and p-value from the two-step weighted least squares approach following Snijders and Baerveldt (2003) and in how many periods the estimated parameters were significant on the 5% level [positive significance|negative significance|lack of significance].

pre-crisis period.⁵¹ One major difference is that post-Lehman, all interest rate related effects have become insignificant. In Table 5.6 we investigate whether the coefficients in Table 5.4 and 5.5 are significantly different by applying a two sample t-test.⁵² The *reciprocity persistence* effect is becoming insignificant and we also find that the difference of the pertinent parameter between the two subperiods (.4939) is significant at the 10% level.⁵³ Thus, taking both *reciprocity* effects together, banks seem not so indifferent anymore to giving up existing mutual relationships. However, interpreting this as a sign for overall favouring mutual relations more might be too far fetched, since the the level of reciprocity fell during the period. In case of the *core similarity* effect becoming positively significant at the 10% level one has to keep in mind that two opposing forces are measured here at the same time, since core banks should be attracted to each other while the opposite by definition holds for those classified as peripheral banks. Thus, this could either indicate that peripheral banks avoid each other less than before or that core banks are relying even more heavily on each other. The fact that some banks dropped out of the core⁵⁴ due to the GFC might indicate that these ‘new’ peripheral banks still relatively frequently trade with each other and drive this finding.⁵⁵

More interesting is that the size seems to play a completely different role, since all three related effects change with the onset of the financial crisis. The *size activity* and *similarity* effect are becoming insignificant. Thus, larger banks are not less active and banks being more equal in terms of size are not more likely to trade with each other anymore. On the other hand the *size popularity* effect is positively significant for the second period, which indicates that banks prefer to transmit their excess liquidity to larger banks. Furthermore, the difference in the two estimates for *activity* and *popularity* is positively significant. This indicates that the attractiveness of larger banks improved after the GFC and this might be due to the belief that the largest banks entail less counterparty risk. However, we are not able to determine to which extent (if any) this is driven by the believe that the government would have to rescue them, which would reduce their counterparty risk to

⁵¹ Note that the *transitive triplets* effect is now significant at the 1% level (pre-crisis 10%).

⁵² The test statistic is $t = \frac{\theta_{PRE} - \theta_{GFC}}{\sqrt{se_{PRE}^2 + se_{GFC}^2}}$, where the subscripts PRE and GFC refer to the period before and after the start of the financial crisis.

⁵³ The level of reciprocity drops in the quarter after the Lehman collapse and Finger *et al.* (2013) detected a structural break in this descriptive variable.

⁵⁴ Before the crisis on average 28% of the banks belonged to the core, afterwards only 23%.

⁵⁵ See Fricke and Lux (2014) for the exact results.

| effect | category | difference | p-value |
|-------------------------|-------------|------------|---------|
| size | rate | -.0191 | .485 |
| density (outdegree) | structural | .2725 | .073* |
| reciprocity | structural | -.1616 | .243 |
| reciprocity persistence | structural | .4939 | .083* |
| transitive triplets | structural | -.0062 | .037** |
| 3-cycles | structural | -.0309 | .001*** |
| indegree pop. (sqrt) | structural | .0187 | .234 |
| outdegree act. (sqrt) | structural | .0130 | .341 |
| past trades | dyadic cov. | -.0283 | .258 |
| past rates | dyadic cov. | -.9438 | .166 |
| size popularity | actor cov. | .0232 | .100* |
| size activity | actor cov. | .0589 | .064* |
| size similarity | actor cov. | .0223 | .696 |
| core popularity | actor cov. | .1115 | .024** |
| core activity | actor cov. | .1211 | .111 |
| core similarity | actor cov. | .0663 | .029** |
| lending rate pop. | actor cov. | -.1865 | .408 |
| lending rate act. | actor cov. | .6036 | .116 |
| lending rate sim. | actor cov. | -.0390 | .753 |
| borrowing rate pop. | actor cov. | .6275 | .208 |
| borrowing rate act. | actor cov. | .4868 | .130 |
| borrowing rate sim. | actor cov. | .1280 | .361 |

Tab. 5.6: The results for the two sample t-test comparing the estimation results for the two subperiods before (in Table 5.4) and after (in Table 5.5) the GFC. The columns state the name of the effect, the category it belongs to, the difference between the two estimates and the pertinent p-value.

literally zero. This is additionally supported by the fact that the difference of the always significant *core popularity* effect is positively significant after the GFC, as well. Hence, larger and core banks have become even more attractive as borrowers which might be related to their seemingly lower default risk. The two triadic effects, *transitive triplets* and *3-cycles* are negatively significant in both subperiods. However, their significant differences indicate that banks avoid both kinds of intermediate relations even more after the crisis.

Finally, all three previously significant interest rate related effects are becoming insignificant. This is surprising since Raddant (2012) found that the spread widened after the GFC and so high borrowing rates should eventually indicate a higher likelihood of default and become an important signal. On the other hand, banks which have to pay high borrowing rates might be (more often) not able to acquire the desired liquidity from a single lender or avoid such deals on an electronic platform such as the e-MID and this could offset the aforementioned reasoning. Overall, interest rates seem to have no systematic impact on banks creating or ending their lending relations after the GFC. In any case, the results indicate that banks did not selectively restrict lending to counterparties with high borrowing rates. With the complete lack of significance of interest rate related effects in the crisis period, the contract rate does not appear to be an *ex ante* decisive causal criterion to enter a credit transaction which supports the relevance of established relations between lenders and borrowers after the onset of the financial crisis.

5.5 Concluding Remarks

We have investigated interbank lending relations documented in the e-MID electronic market via the actor-oriented model of network formation developed in sociology. Such an approach is motivated by previous statistical evidence for relationship lending (e.g. Cocco *et al.*, 2009) and provides the opportunity to study more directly the factors determining the formation (or dissolution) of lasting business relationships.

The most intriguing result is that the behaviour of the banks managing their lending relations was relatively robust throughout the complete sample period, which is supported by the very similar results for the periods before and after the GFC. Throughout every single period the finding that the past trades effect was significant highlights the importance of relationship banking. Hence, even on an electronic platform like e-MID where every bank has the opportunity to observe all the quotes and so the cost to enter into a contract with ‘new’ counterparties is very low banks heavily rely on well established business relations.

Rather surprising is the role of (average) interest rates on the behaviour of banks, because after the GFC there is no significant effect remaining and never did a high borrowing rate, indicating a higher default probability, make banks less attractive as counterparty for lending banks. However, the public character of the e-MID market may bias these results, since a bank would perhaps shy away from indicating publicly urgent liquidity needs by quoting a very high interest rate. So one might argue that in such a scenario the bank would rather try to get the liquidity via OTC markets not sharing this ‘negative’ information with all participating banks. Nevertheless, activating privately (over-the-counter) an established link would support the relevance of relationship banking. Additionally, banks associated with high (counterparty) risk could have more problems to acquire the desired level of liquidity from a single lender. On the other hand, the two most intriguing changes due to the GFC support that the perception of counterparty risk changed. First, larger banks became more popular and stopped being less active and core banks became even more popular and even more active. Here it could be reasoned that larger and core banks have a lower probability of default. However, we are not able to address the reason for this directly and so we can not distinguish whether these banks seem less risky due to their business performance and/or the believe that the government would have to step in in the case of their default to prevent a meltdown of the complete financial system. Second, banks avoid both kinds of intermediate relations more and by this reduce their *indirect* counterparty risk and the overall connectivity of the system (as it is reflected in a decreasing density).

An interesting aspect of the SAOM model is that we can scrutinize the interrelation between aggregate properties and individual behaviour. As concerns intermediate links both triadic effects never work as an incentive for banks to create a lending relation, but often have the contrary effect. This indicates that banks focus on their direct counterparty or try to avoid too much indirect counterparty risk. However, at the same time this means that the high directed clustering coefficient (CC^1) apparently emerges out of the interplay of other effects and seems to be an unintended feature of the overall network.

Another case where we observe unexpected behaviour is the *reciprocity* effect. Here the global statistic shows very similar values to the density (i.e. a random number of reciprocal links) and so one would expect that banks do not care whether a relation is mutual or not. However, banks prefer to form mutual relations, while they seem to be indifferent to maintaining their relationships as a reciprocal one over time. The intuition is that relations are often predominantly pointing in one direction, but the existence of this stable relation makes it more likely that the familiar counterparties sometimes

act in the opposite direction as well, creating a transient mutual relation. The strength of this one sidedness of many relations is remarkable, since we aggregate networks on a quarterly basis.

For the other bank specific features we see that larger banks become more popular borrowers after the beginning of the GFC. For core banks (as identified independently by Fricke and Lux, 2014, for the same sample) we find that they are both more popular and more active in the interbank market. Hence, the core-periphery model does not simply assign the biggest banks to the core, but includes important additional information in detecting intermediaries.

To sum up, we see that banks display persistent behavioral regularities over time, which should ensure (as it is the case) persistent structural features of banking network in terms of e.g. hierarchy and degree distributions, but also strong and lasting relations. These bonds seemingly helped the network to persist during the GFC. Moreover, the findings regarding, e.g. reciprocity and clustering, indicate that the SAOM model is a good start to understand the incentives and resulting behavior of banks managing their relations in the interbank market, and that an actor-based analysis provides important additional insights on top of purely statistical investigations of interbank lending relations. Revealed behavioral regularities could be integrated in models of bank behavior and could be used as behavioral assumptions in simulation models designed for stress-testing of the resilience of the interbank market.

5.6 Appendix

The Appendix investigates the robustness of the empirical results in section 5.4 and presents some complementary results which are referred to in the main text but not explicitly shown. First, we show the results if we allow the banks to enter and leave during the estimation (composition change). Secondly, the estimation results for each quarter are shown. At last we present the quarterly test statistic per actor covariate for the joint significance of the three actor covariate related effects (*popularity, acticity, similarity*).

5.6.1 Composition Change

Tables 5.7 to 5.9 show the results for the complete sample and the two subperiods, if we allow the banks to leave during the estimation (composition change), while in the main application only banks being active in both quarters are taken into account. The results are always very similar. In general, the level of the *reciprocity* effect and *reciprocity persistence* effect increase and the *reciprocity persistence* effect (after the GFC) is becoming significant at the 5% level, while it was insignificant before. The intuition behind this should be that banks right before a merger or acquisition are more likely to have a reciprocal relation, while they are forced (with composition change allowed) to end their relations after one of them ceases to exist. Moreover, for the complete sample the *lending rate activity* effect is becoming significant at the 10% level, while it is slightly insignificant with a p-value of .128 in the main application.

5.6.2 Quarterly Results

For completeness the quarterly estimation results and the Wald test for the joint significance of the actor covariate related effects are presented. Table 5.10 shows the results of the SAOM for each quarter from 2002Q1 until 2010Q4, which are the foundation for the estimations presented in the main application.

| effect | category | mean | s.e. | p-value | significance |
|-------------------------|-------------|---------|----------|---------|--------------|
| rate | | 34.7016 | | | |
| size | rate | .1753 | (.0096) | .001*** | [34 - 2] |
| density (outdegree) | structural | -4.2096 | (.0561) | .001*** | [- 36 -] |
| reciprocity | structural | .9041 | (.0669) | .001*** | [27 - 9] |
| reciprocity persistence | structural | -1.1661 | (.1265) | .001*** | [- 17 19] |
| transitive triplets | structural | -.0029 | (.0001) | .001*** | [- 8 28] |
| 3-cycles | structural | -.0240 | (.0026) | .001*** | [- 34 2] |
| indegree pop. (sqrt.) | structural | .3542 | (.0060) | .001*** | [36 - -] |
| outdegree act. (sqrt.) | structural | .2151 | (.0055) | .001*** | [36 - -] |
| past trades | dyadic cov. | .1905 | .0089 | .001*** | [36 - -] |
| past rates | dyadic cov. | .3967 | (.3991) | .327 | [9 8 19] |
| size popularity | actor cov. | .0090 | (.0081) | .274 | [9 6 21] |
| size activity | actor cov. | -.0455 | (.0121) | .001*** | [5 22 9] |
| size similarity | actor cov. | .0560 | (.0177) | .003*** | [6 - 30] |
| core popularity | actor cov. | .2236 | (.0246) | .001*** | [27 - 9] |
| core activity | actor cov. | .2590 | (.0324) | .001*** | [27 1 8] |
| core similarity | actor cov. | -.0024 | (-.0132) | .857 | [4 5 27] |
| lending rate pop. | actor cov. | -.0441 | (.0794) | .582 | [1 4 31] |
| lending rate act. | actor cov. | -.2940 | (.1892) | .091* | [5 10 21] |
| lending rate sim. | actor cov. | -.1627 | (.0651) | .017** | [2 12 22] |
| borrowing rate pop. | actor cov. | -.3934 | (.2903) | .184 | [7 14 15] |
| borrowing rate act. | actor cov. | .0994 | (.1218) | .420 | [5 3 28] |
| borrowing rate sim. | actor cov. | -.1163 | (.0565) | .047** | [2 8 26] |

Tab. 5.7: The estimation results for the 36 quarters from 2002 until 2010 with composition change. The columns state the name of the effect, the category it belongs to, the estimated mean, its standard error and p-value from the two-step weighted least squares approach following Snijders and Baerveldt (2003) and in how many periods the estimated parameters were significant on the 5% level [positive significance|negative significance|lack of significance].

| effect | category | mean | s.e. | p-value | significance |
|-------------------------|-------------|---------|---------|---------|--------------|
| rate | | 37.5197 | | | |
| size | rate | .1795 | (.0089) | .001*** | [26 - -] |
| density (outdegree) | structural | -4.2776 | (.0639) | .001*** | [- 26 -] |
| reciprocity | structural | .959 | (.0758) | .001*** | [23 - 3] |
| reciprocity persistence | structural | -1.2849 | (.1361) | .001*** | [- 16 10] |
| transitive triplets | structural | -.0017 | (.0010) | .092* | [- 5 21] |
| 3-cycles | structural | -.0181 | (.0010) | .001*** | [- 25 1] |
| indegree pop. (sqrt) | structural | .3493 | (.0070) | .001*** | [26 - -] |
| outdegree act. (sqrt) | structural | .2117 | (.0066) | .001*** | [26 - -] |
| past trades | dyadic cov. | .1976 | (.0091) | .001*** | [26 - -] |
| past rates | dyadic cov. | .6635 | (.5452) | .235 | [9 6 11] |
| size popularity | actor cov. | .0024 | (.0101) | .818 | [7 6 13] |
| size activity | actor cov. | -.0621 | (.0120) | .001*** | [2 18 6] |
| size similarity | actor cov. | .0493 | (.0157) | .004*** | [4 - 22] |
| core popularity | actor cov. | .1947 | (.0297) | .001*** | [18 - 8] |
| core activity | actor cov. | .2286 | (.0372) | .001*** | [18 1 7] |
| core similarity | actor cov. | -.0165 | (.0151) | .285 | [3 5 18] |
| lending rate pop. | actor cov. | -.0178 | (.0801) | .826 | [- 2 24] |
| lending rate act. | actor cov. | -.4762 | (.2479) | .066* | [3 7 16] |
| lending rate sim. | actor cov. | -.1727 | (.0878) | .060* | [2 10 14] |
| borrowing rate pop. | actor cov. | -.5805 | (.3954) | .155 | [6 11 9] |
| borrowing rate act. | actor cov. | -.0453 | (.1235) | .717 | [2 3 21] |
| borrowing rate sim. | actor cov. | -.1490 | (.0611) | .022** | [- 6 20] |

Tab. 5.8: The estimation results for the 26 quarters from the 1st quarter of 2002 until the 2nd quarter of 2008 with composition change. The columns state the name of the effect, the category it belongs to, the estimated mean, its standard error and p-value from the two-step weighted least squares approach following Snijders and Baerveldt (2003) and in how many periods the estimated parameters were significant on the 5% level [positive significance|negative significance|lack of significance].

| effect | category | mean | s.e. | p-value | significance |
|-------------------------|-------------|---------|---------|---------|--------------|
| rate | | 27.3746 | | | |
| size | rate | .1634 | (.0271) | .001*** | [8 -2] |
| density (outdegree) | structural | -4.0204 | (.0966) | .001*** | [- 10 -] |
| reciprocity | structural | .7050 | (.1305) | .001*** | [4 -6] |
| reciprocity persistence | structural | -.676 | (.2489) | .024** | [- 1 9] |
| transitive triplets | structural | -.0085 | (.0022) | .004*** | [- 3 7] |
| 3-cycles | structural | -.0493 | (.0068) | .001*** | [- 9 1] |
| indegree pop. (sqrt) | structural | .3713 | (.0105) | .001*** | [10 - -] |
| outdegree act. (sqrt) | structural | .2264 | (.0092) | .001*** | [10 - -] |
| past trades | dyadic cov. | .1721 | (.0217) | .001*** | [10 - -] |
| past rates | dyadic cov. | -.2721 | (.3234) | .422 | [- 2 8] |
| size popularity | actor cov. | .0256 | (.0082) | .012** | [2 -8] |
| size activity | actor cov. | -.0007 | (.0270) | .980 | [3 4 3] |
| size similarity | actor cov. | .0698 | (.0541) | .229 | [2 -8] |
| core popularity | actor cov. | .3068 | (.0295) | .001*** | [9 -1] |
| core activity | actor cov. | .3441 | (.0602) | .001*** | [9 -1] |
| core similarity | actor cov. | .0471 | (.0215) | .057* | [1 -9] |
| lending rate pop. | actor cov. | -.2210 | (.2171) | .335 | [1 -9] |
| lending rate act. | actor cov. | .1450 | (.2369) | .555 | [2 3 5] |
| lending rate sim. | actor cov. | -.1285 | (.0899) | .187 | [- 2 8] |
| borrowing rate pop. | actor cov. | .0471 | (.2625) | .862 | [1 3 6] |
| borrowing rate act. | actor cov. | .4579 | (.2727) | .127 | [3 -7] |
| borrowing rate sim. | actor cov. | -.0230 | (.1275) | .861 | [2 2 6] |

Tab. 5.9: The estimation results for the 10 quarters from the 2nd quarter of 2008 until the 4th quarter of 2010 with composition change. The columns state the name of the effect, the category it belongs to, the estimated mean, its standard error and p-value from the two-step weighted least squares approach following Snijders and Baerveldt (2003) and in how many periods the estimated parameters were significant on the 5% level [positive significance|negative significance|lack of significance].

Tab. 5.10: The estimation results for each quarter from 2002Q1 until 2002Q4. The columns state the name of the effect, the category it belongs to, the estimated value, the standard error, t-value and the significance (* 10%, ** 5%, *** 1%).

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|-------|----------|---------|
| rate | | 45.9144 | | | |
| size | rate | .1842 | .0288 | 6.4001 | *** |
| density (outdegree) | structural | -4.4139 | .1241 | -35.5768 | *** |
| reciprocity | structural | .6140 | .2670 | 2.2999 | ** |
| reciprocity persistence | structural | -.7568 | .4816 | -1.5715 | |
| transitive triplets | structural | -.0078 | .0021 | -3.6365 | *** |
| 3-cycles | structural | -.0222 | .0033 | -6.7259 | *** |
| indegree pop. (sqrt.) | structural | .3492 | .0131 | 26.7539 | *** |
| outdegree act. (sqrt.) | structural | .2129 | .0108 | 19.7849 | *** |
| past trades | dyadic cov. | .1480 | .0095 | 15.5912 | *** |
| past rates | dyadic cov. | -2.1295 | .2805 | -7.5908 | *** |
| size popularity | actor cov. | .0456 | .0154 | 2.9639 | *** |
| size activity | actor cov. | -.0703 | .0158 | -4.4567 | *** |
| size similarity | actor cov. | .0344 | .0587 | .5849 | |
| core popularity | actor cov. | .3162 | .0539 | 5.8635 | *** |
| core activity | actor cov. | .3198 | .0486 | 6.5814 | *** |
| core similarity | actor cov. | .0843 | .0426 | 1.9794 | ** |
| lending rate pop. | actor cov. | .1694 | .3358 | .5045 | |
| lending rate act. | actor cov. | .2176 | .3552 | .6125 | |
| lending rate sim. | actor cov. | -.0035 | .0014 | -2.4598 | ** |
| borrowing rate pop. | actor cov. | -1.7130 | .4434 | -3.8636 | *** |
| borrowing rate act. | actor cov. | -.2617 | .3721 | -.7032 | |
| borrowing rate sim. | actor cov. | -.0003 | .0015 | -.1853 | |

The estimation results for 2002Q1.

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|-------|----------|---------|
| rate | | 39.8719 | | | |
| size | rate | .1274 | .0262 | 4.8731 | *** |
| density (outdegree) | structural | -4.7258 | .1398 | -33.8021 | *** |
| reciprocity | structural | .6628 | .2516 | 2.6345 | *** |
| reciprocity persistence | structural | -1.0135 | .4597 | -2.2047 | ** |
| transitive triplets | structural | -.0070 | .0024 | -2.9632 | *** |
| 3-cycles | structural | -.0113 | .0038 | -2.9525 | *** |
| indegree pop. (sqrt.) | structural | .3372 | .0146 | 23.0651 | *** |
| outdegree act. (sqrt.) | structural | .2495 | .0125 | 19.9174 | *** |
| past trades | dyadic cov. | .1564 | .0100 | 15.6782 | *** |
| past rates | dyadic cov. | 2.098 | .6013 | 3.4889 | *** |
| size popularity | actor cov. | .0229 | .0182 | 1.2556 | |
| size activity | actor cov. | -.0520 | .0178 | -2.9259 | *** |
| size similarity | actor cov. | .0373 | .0646 | .5773 | |
| core popularity | actor cov. | .2663 | .0613 | 4.3439 | *** |
| core activity | actor cov. | -.0028 | .0531 | -.0536 | |
| core similarity | actor cov. | -.1013 | .0455 | -2.226 | ** |
| lending rate pop. | actor cov. | .2825 | .3486 | .8102 | |
| lending rate act. | actor cov. | -1.5457 | .4932 | -3.1342 | *** |
| lending rate sim. | actor cov. | .0073 | .0037 | 1.9887 | ** |
| borrowing rate pop. | actor cov. | -3.4811 | .5960 | -5.8409 | *** |
| borrowing rate act. | actor cov. | 1.6347 | .5801 | 2.818 | *** |
| borrowing rate sim. | actor cov. | -.0058 | .0017 | -3.3874 | *** |

The estimation results for 2002Q2.

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|--------|----------|---------|
| rate | | 38.8571 | | | |
| size | rate | .1502 | .0333 | 4.5115 | *** |
| density (outdegree) | structural | -3.9556 | .1375 | -28.7743 | *** |
| reciprocity | structural | .8913 | .2675 | 3.3325 | *** |
| reciprocity persistence | structural | -1.5354 | .5466 | -2.8089 | *** |
| transitive triplets | structural | -.0052 | .0027 | -1.9332 | * |
| 3-cycles | structural | -.0199 | .0041 | -4.8101 | *** |
| indegree pop. (sqrt.) | structural | .3033 | .0157 | 19.3542 | *** |
| outdegree act. (sqrt.) | structural | .1917 | .0129 | 14.9145 | *** |
| past trades | dyadic cov. | .2161 | .0108 | 2.0648 | *** |
| past rates | dyadic cov. | -.3551 | .8426 | -.4214 | |
| size popularity | actor cov. | .0523 | .0180 | 2.913 | *** |
| size activity | actor cov. | -.0453 | .0176 | -2.5675 | ** |
| size similarity | actor cov. | .0082 | .0659 | .1249 | |
| core popularity | actor cov. | .2582 | .0686 | 3.7659 | *** |
| core activity | actor cov. | .2442 | .0629 | 3.8825 | *** |
| core similarity | actor cov. | .0352 | .0523 | .6723 | |
| lending rate pop. | actor cov. | .2567 | .9677 | .2653 | |
| lending rate act. | actor cov. | -1.4985 | 1.076 | -1.3927 | |
| lending rate sim. | actor cov. | .0011 | .0018 | .6191 | |
| borrowing rate pop. | actor cov. | -1.0688 | 1.4925 | -.7161 | |
| borrowing rate act. | actor cov. | 1.5672 | 1.3885 | 1.1287 | |
| borrowing rate sim. | actor cov. | -.0023 | .0013 | -1.7812 | * |

The estimation results for 2002Q3.

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|--------|----------|---------|
| rate | | 48.6957 | | | |
| size | rate | .1592 | .0866 | 1.8374 | * |
| density (outdegree) | structural | -4.221 | .1466 | -28.7873 | *** |
| reciprocity | structural | 1.6483 | 1.5741 | 1.0471 | |
| reciprocity persistence | structural | -2.6006 | 3.2428 | -.8020 | |
| transitive triplets | structural | .0034 | .0022 | 1.5440 | |
| 3-cycles | structural | -.0187 | .0033 | -5.6472 | *** |
| indegree pop. (sqrt.) | structural | .3600 | .0168 | 21.4673 | *** |
| outdegree act. (sqrt.) | structural | .1657 | .0127 | 13.0452 | *** |
| past trades | dyadic cov. | .1711 | .0154 | 11.114 | *** |
| past rates | dyadic cov. | -.001 | 1.8527 | .0000 | |
| size popularity | actor cov. | -.0235 | .0166 | -1.4106 | |
| size activity | actor cov. | .0038 | .0177 | .2139 | |
| size similarity | actor cov. | -.0673 | .0640 | -1.0515 | |
| core popularity | actor cov. | .1367 | .0677 | 2.0200 | ** |
| core activity | actor cov. | .3216 | .0642 | 5.0058 | *** |
| core similarity | actor cov. | .0105 | .0686 | .1535 | |
| lending rate pop. | actor cov. | .4064 | .2470 | 1.6453 | * |
| lending rate act. | actor cov. | -.5166 | .7077 | -.7300 | |
| lending rate sim. | actor cov. | .0010 | .002 | .4664 | |
| borrowing rate pop. | actor cov. | -1.4028 | .2999 | -4.6775 | *** |
| borrowing rate act. | actor cov. | -.5472 | .258 | -2.121 | ** |
| borrowing rate sim. | actor cov. | -.0029 | .0019 | -1.5247 | |

The estimation results for 2002Q4.

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|-------|----------|---------|
| rate | | 49.7744 | | | |
| size | rate | .1989 | .0360 | 5.5317 | *** |
| density (outdegree) | structural | -4.259 | .1442 | -29.5401 | *** |
| reciprocity | structural | .4224 | .2827 | 1.4939 | |
| reciprocity persistence | structural | -.3111 | .5397 | -.5765 | |
| transitive triplets | structural | .0025 | .0023 | 1.0991 | |
| 3-cycles | structural | -.0208 | .0040 | -5.2366 | *** |
| indegree pop. (sqrt.) | structural | .3592 | .0147 | 24.4853 | *** |
| outdegree act. (sqrt.) | structural | .1853 | .0129 | 14.3616 | *** |
| past trades | dyadic cov. | .1358 | .0095 | 14.2946 | *** |
| past rates | dyadic cov. | -2.4922 | .5603 | -4.4483 | *** |
| size popularity | actor cov. | .0031 | .0146 | .2124 | |
| size activity | actor cov. | -.0401 | .0151 | -2.6594 | *** |
| size similarity | actor cov. | -.0176 | .0598 | -.2939 | |
| core popularity | actor cov. | .0528 | .0542 | .9740 | |
| core activity | actor cov. | .2014 | .0510 | 3.9458 | *** |
| core similarity | actor cov. | -.0828 | .0408 | -2.0328 | ** |
| lending rate pop. | actor cov. | .0777 | .1640 | .4735 | |
| lending rate act. | actor cov. | -.2197 | .1971 | -1.1149 | |
| lending rate sim. | actor cov. | -.0023 | .0015 | -1.5431 | |
| borrowing rate pop. | actor cov. | .9531 | .2524 | 3.7761 | *** |
| borrowing rate act. | actor cov. | -.2077 | .1731 | -1.1996 | |
| borrowing rate sim. | actor cov. | .0045 | .0019 | 2.3424 | ** |

The estimation results for 2003Q1.

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|-------|----------|---------|
| rate | | 38.4114 | | | |
| size | rate | .1963 | .0317 | 6.1999 | *** |
| density (outdegree) | structural | -4.4908 | .1541 | -29.1486 | *** |
| reciprocity | structural | .7293 | .2160 | 3.3763 | *** |
| reciprocity persistence | structural | -.9118 | .3992 | -2.284 | ** |
| transitive triplets | structural | .0017 | .0025 | .6890 | |
| 3-cycles | structural | -.0129 | .0041 | -3.1164 | *** |
| indegree pop. (sqrt.) | structural | .3851 | .0157 | 24.5157 | *** |
| outdegree act. (sqrt.) | structural | .1967 | .0144 | 13.6691 | *** |
| past trades | dyadic cov. | .1583 | .0092 | 17.2545 | *** |
| past rates | dyadic cov. | -.0088 | .4646 | -.0190 | |
| size popularity | actor cov. | -.0620 | .0163 | -3.81 | *** |
| size activity | actor cov. | -.0332 | .0170 | -1.9587 | * |
| size similarity | actor cov. | -.0407 | .0681 | -.5967 | |
| core popularity | actor cov. | -.0286 | .0622 | -.4605 | |
| core activity | actor cov. | .2506 | .0588 | 4.2596 | *** |
| core similarity | actor cov. | -.1210 | .0465 | -2.6003 | *** |
| lending rate pop. | actor cov. | -.2404 | .1770 | -1.3584 | |
| lending rate act. | actor cov. | -.7395 | .2199 | -3.3629 | *** |
| lending rate sim. | actor cov. | .0047 | .0014 | 3.303 | *** |
| borrowing rate pop. | actor cov. | -1.4179 | .2458 | -5.7684 | *** |
| borrowing rate act. | actor cov. | .3073 | .2331 | 1.3184 | |
| borrowing rate sim. | actor cov. | .0021 | .0013 | 1.5923 | |

The estimation results for 2003Q2.

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|-------|----------|---------|
| rate | | 38.8631 | | | |
| size | rate | .098 | .0349 | 2.8119 | *** |
| density (outdegree) | structural | -5.0278 | .1728 | -29.0935 | *** |
| reciprocity | structural | 1.0428 | .4149 | 2.5136 | ** |
| reciprocity persistence | structural | -1.3096 | .7404 | -1.7687 | * |
| transitive triplets | structural | -.0073 | .0033 | -2.2553 | ** |
| 3-cycles | structural | -.0210 | .0054 | -3.9124 | *** |
| indegree pop. (sqrt.) | structural | .4002 | .0172 | 23.241 | *** |
| outdegree act. (sqrt.) | structural | .2694 | .0167 | 16.1079 | *** |
| past trades | dyadic cov. | .1948 | .0107 | 18.1508 | *** |
| past rates | dyadic cov. | -3.2745 | .4675 | -7.0044 | *** |
| size popularity | actor cov. | .0264 | .0191 | 1.3859 | |
| size activity | actor cov. | -.0774 | .0197 | -3.9385 | *** |
| size similarity | actor cov. | .1025 | .0781 | 1.3122 | |
| core popularity | actor cov. | .2982 | .0728 | 4.0939 | *** |
| core activity | actor cov. | -.1568 | .0652 | -2.4065 | ** |
| core similarity | actor cov. | -.0933 | .0562 | -1.6603 | * |
| lending rate pop. | actor cov. | .1890 | .1254 | 1.5063 | |
| lending rate act. | actor cov. | -.0143 | .1685 | -.0847 | |
| lending rate sim. | actor cov. | -.0009 | .0018 | -.5112 | |
| borrowing rate pop. | actor cov. | 1.7679 | .6549 | 2.6996 | *** |
| borrowing rate act. | actor cov. | .4961 | .6300 | .7875 | |
| borrowing rate sim. | actor cov. | -.0067 | .0069 | -.9613 | |

The estimation results for 2003Q3.

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|-------|----------|---------|
| rate | | 32.9798 | | | |
| size | rate | .1605 | .0340 | 4.7211 | *** |
| density (outdegree) | structural | -3.8676 | .1594 | -24.2576 | *** |
| reciprocity | structural | .5226 | .2056 | 2.5415 | ** |
| reciprocity persistence | structural | -.658 | .3906 | -1.6848 | * |
| transitive triplets | structural | .0043 | .0033 | 1.2941 | |
| 3-cycles | structural | -.0103 | .0046 | -2.2581 | ** |
| indegree pop. (sqrt.) | structural | .3174 | .0158 | 2.0265 | *** |
| outdegree act. (sqrt.) | structural | .1679 | .0173 | 9.6942 | *** |
| past trades | dyadic cov. | .2342 | .0112 | 2.9578 | *** |
| past rates | dyadic cov. | .0330 | .4179 | .0789 | |
| size popularity | actor cov. | .0179 | .0196 | .9128 | |
| size activity | actor cov. | -.0730 | .0205 | -3.5574 | *** |
| size similarity | actor cov. | -.0735 | .0751 | -.9786 | |
| core popularity | actor cov. | .1932 | .0730 | 2.6452 | *** |
| core activity | actor cov. | .2328 | .0670 | 3.4745 | *** |
| core similarity | actor cov. | -.0636 | .0522 | -1.2189 | |
| lending rate pop. | actor cov. | -.3952 | .2095 | -1.8865 | * |
| lending rate act. | actor cov. | .0020 | .2444 | .0083 | |
| lending rate sim. | actor cov. | -.0058 | .0016 | -3.5913 | *** |
| borrowing rate pop. | actor cov. | -.4365 | .2382 | -1.8327 | * |
| borrowing rate act. | actor cov. | .4636 | .2052 | 2.2597 | ** |
| borrowing rate sim. | actor cov. | 7e-04 | .0027 | .2495 | |

The estimation results for 2003Q4.

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|--------|----------|---------|
| rate | | 38.3857 | | | |
| size | rate | .0974 | .0358 | 2.7243 | *** |
| density (outdegree) | structural | -4.3911 | .1702 | -25.7965 | *** |
| reciprocity | structural | .4407 | .3249 | 1.3564 | |
| reciprocity persistence | structural | -.3968 | .5732 | -.6923 | |
| transitive triplets | structural | .0014 | .0036 | .3959 | |
| 3-cycles | structural | -.0177 | .0047 | -3.7323 | *** |
| indegree pop. (sqrt.) | structural | .3417 | .0167 | 2.4343 | *** |
| outdegree act. (sqrt.) | structural | .2265 | .0183 | 12.3669 | *** |
| past trades | dyadic cov. | .2008 | .0110 | 18.2816 | *** |
| past rates | dyadic cov. | -.1876 | .3064 | -.6124 | |
| size popularity | actor cov. | .0205 | .0202 | 1.0150 | |
| size activity | actor cov. | -.0317 | .0192 | -1.6553 | * |
| size similarity | actor cov. | .0599 | .0722 | .8295 | |
| core popularity | actor cov. | .2602 | .0713 | 3.6506 | *** |
| core activity | actor cov. | -.0719 | .0616 | -1.1672 | |
| core similarity | actor cov. | -.0285 | .0534 | -.5344 | |
| lending rate pop. | actor cov. | 2.2661 | 1.1002 | 2.0596 | ** |
| lending rate act. | actor cov. | .9893 | 1.1651 | .8491 | |
| lending rate sim. | actor cov. | -.0016 | .0016 | -1.017 | |
| borrowing rate pop. | actor cov. | 1.9008 | 1.1271 | 1.6865 | * |
| borrowing rate act. | actor cov. | 1.2423 | .9209 | 1.3490 | |
| borrowing rate sim. | actor cov. | .0028 | .0020 | 1.4243 | |

The estimation results for 2004Q1.

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|--------|----------|---------|
| rate | | 36.6593 | | | |
| size | rate | .1165 | .0365 | 3.1904 | *** |
| density (outdegree) | structural | -4.1476 | .1555 | -26.6656 | *** |
| reciprocity | structural | 1.0422 | .3086 | 3.3771 | *** |
| reciprocity persistence | structural | -1.1326 | .6005 | -1.8862 | * |
| transitive triplets | structural | .0042 | .0031 | 1.3595 | |
| 3-cycles | structural | -.0201 | .0051 | -3.9576 | *** |
| indegree pop. (sqrt.) | structural | .3522 | .0177 | 19.9213 | *** |
| outdegree act. (sqrt.) | structural | .1755 | .0159 | 11.0487 | *** |
| past trades | dyadic cov. | .2019 | .0128 | 15.7832 | *** |
| past rates | dyadic cov. | 6.2265 | 1.6368 | 3.804 | *** |
| size popularity | actor cov. | -.0504 | .0206 | -2.4451 | ** |
| size activity | actor cov. | -.0214 | .0183 | -1.1727 | |
| size similarity | actor cov. | .0564 | .0740 | .7623 | |
| core popularity | actor cov. | -.1295 | .0675 | -1.9184 | * |
| core activity | actor cov. | .4510 | .0659 | 6.8416 | *** |
| core similarity | actor cov. | -.0820 | .0509 | -1.6105 | |
| lending rate pop. | actor cov. | .8702 | .6447 | 1.3497 | |
| lending rate act. | actor cov. | -1.4076 | .7259 | -1.9392 | * |
| lending rate sim. | actor cov. | -.0071 | .0016 | -4.3992 | *** |
| borrowing rate pop. | actor cov. | -.4147 | .8941 | -.4638 | |
| borrowing rate act. | actor cov. | -.1187 | .8974 | -.1322 | |
| borrowing rate sim. | actor cov. | -.0015 | .0031 | -.4796 | |

The estimation results for 2004Q2.

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|--------|----------|---------|
| rate | | 31.0661 | | | |
| size | rate | .1899 | .0326 | 5.8206 | *** |
| density (outdegree) | structural | -4.1503 | .1772 | -23.4159 | *** |
| reciprocity | structural | .8432 | .2506 | 3.3645 | *** |
| reciprocity persistence | structural | -1.1356 | .4665 | -2.4342 | ** |
| transitive triplets | structural | -.0019 | .0030 | -.6432 | |
| 3-cycles | structural | -.0181 | .0051 | -3.579 | *** |
| indegree pop. (sqrt.) | structural | .3214 | .0197 | 16.2958 | *** |
| outdegree act. (sqrt.) | structural | .2186 | .0161 | 13.5551 | *** |
| past trades | dyadic cov. | .2355 | .0133 | 17.6827 | *** |
| past rates | dyadic cov. | -.7435 | .5240 | -1.419 | |
| size popularity | actor cov. | -.0222 | .0206 | -1.0737 | |
| size activity | actor cov. | -.0605 | .0204 | -2.9618 | *** |
| size similarity | actor cov. | .0663 | .0765 | .8660 | |
| core popularity | actor cov. | .3622 | .0737 | 4.9133 | *** |
| core activity | actor cov. | .2172 | .0653 | 3.3251 | *** |
| core similarity | actor cov. | -.1135 | .0542 | -2.0951 | ** |
| lending rate pop. | actor cov. | .2460 | .3192 | .7708 | |
| lending rate act. | actor cov. | -.3917 | .3271 | -1.1976 | |
| lending rate sim. | actor cov. | -.0048 | .0025 | -1.8838 | * |
| borrowing rate pop. | actor cov. | -5.8959 | 1.307 | -4.511 | *** |
| borrowing rate act. | actor cov. | 1.4341 | 1.2445 | 1.1524 | |
| borrowing rate sim. | actor cov. | -.0026 | .0028 | -.9224 | |

The estimation results for 2004Q3.

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|-------|---------|---------|
| rate | | 34.5565 | | | |
| size | rate | .1272 | .0332 | 3.8347 | *** |
| density (outdegree) | structural | -4.0081 | .1679 | -23.878 | *** |
| reciprocity | structural | .8714 | .2198 | 3.9644 | *** |
| reciprocity persistence | structural | -.8731 | .4557 | -1.9158 | * |
| transitive triplets | structural | .0036 | .0029 | 1.2440 | |
| 3-cycles | structural | -.0174 | .0045 | -3.893 | *** |
| indegree pop. (sqrt.) | structural | .3187 | .0194 | 16.4393 | *** |
| outdegree act. (sqrt.) | structural | .1868 | .0153 | 12.2213 | *** |
| past trades | dyadic cov. | .2636 | .0131 | 2.1614 | *** |
| past rates | dyadic cov. | -.8447 | .5632 | -1.4997 | |
| size popularity | actor cov. | -.1276 | .0206 | -6.2004 | *** |
| size activity | actor cov. | -.0319 | .0199 | -1.6038 | |
| size similarity | actor cov. | .0787 | .0739 | 1.0645 | |
| core popularity | actor cov. | .3607 | .0687 | 5.2541 | *** |
| core activity | actor cov. | .1489 | .0618 | 2.4098 | ** |
| core similarity | actor cov. | -.1404 | .0524 | -2.6795 | *** |
| lending rate pop. | actor cov. | .1789 | .5779 | .3096 | |
| lending rate act. | actor cov. | -.8814 | .5944 | -1.4828 | |
| lending rate sim. | actor cov. | -.0041 | .0020 | -2.0481 | ** |
| borrowing rate pop. | actor cov. | -.7228 | .6283 | -1.1503 | |
| borrowing rate act. | actor cov. | -.8734 | .6170 | -1.4155 | |
| borrowing rate sim. | actor cov. | -.0006 | .0023 | -.2425 | |

The estimation results for 2004Q4.

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|--------|----------|---------|
| rate | | 38.4954 | | | |
| size | rate | .168 | .0366 | 4.5966 | *** |
| density (outdegree) | structural | -3.9325 | .1684 | -23.3527 | *** |
| reciprocity | structural | .9464 | .3071 | 3.0814 | *** |
| reciprocity persistence | structural | -1.0376 | .548 | -1.8934 | * |
| transitive triplets | structural | .0000 | .0031 | .0068 | |
| 3-cycles | structural | -.0231 | .0050 | -4.6205 | *** |
| indegree pop. (sqrt.) | structural | .3159 | .0184 | 17.1441 | *** |
| outdegree act. (sqrt.) | structural | .1849 | .0165 | 11.1782 | *** |
| past trades | dyadic cov. | .1943 | .0117 | 16.6051 | *** |
| past rates | dyadic cov. | -1.1093 | .4620 | -2.401 | ** |
| size popularity | actor cov. | -.0172 | .0194 | -.8898 | |
| size activity | actor cov. | -.1049 | .0193 | -5.4377 | *** |
| size similarity | actor cov. | .1809 | .0731 | 2.4732 | ** |
| core popularity | actor cov. | .3250 | .0671 | 4.8439 | *** |
| core activity | actor cov. | .3278 | .0605 | 5.4222 | *** |
| core similarity | actor cov. | .0936 | .0502 | 1.8635 | * |
| lending rate pop. | actor cov. | -.5726 | 1.1767 | -.4866 | |
| lending rate act. | actor cov. | -3.3323 | 1.5048 | -2.2144 | ** |
| lending rate sim. | actor cov. | -.0063 | .0091 | -.6946 | |
| borrowing rate pop. | actor cov. | -5.0661 | 1.3742 | -3.6866 | *** |
| borrowing rate act. | actor cov. | -1.1141 | 1.2687 | -.8781 | |
| borrowing rate sim. | actor cov. | -.0060 | .0028 | -2.1647 | ** |

The estimation results for 2005Q1.

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|--------|----------|---------|
| rate | | 42.2798 | | | |
| size | rate | .1492 | .0361 | 4.1326 | *** |
| density (outdegree) | structural | -4.0892 | .1614 | -25.3329 | *** |
| reciprocity | structural | 1.1586 | .5821 | 1.9905 | ** |
| reciprocity persistence | structural | -1.735 | 1.2032 | -1.442 | |
| transitive triplets | structural | .0030 | .0028 | 1.0448 | |
| 3-cycles | structural | -.0163 | .0053 | -3.0719 | *** |
| indegree pop. (sqrt.) | structural | .3601 | .0197 | 18.2764 | *** |
| outdegree act. (sqrt.) | structural | .1725 | .0144 | 12.0111 | *** |
| past trades | dyadic cov. | .2237 | .0128 | 17.5435 | *** |
| past rates | dyadic cov. | -4.3461 | 1.2474 | -3.4843 | *** |
| size popularity | actor cov. | -.0545 | .0184 | -2.9677 | *** |
| size activity | actor cov. | -.0729 | .0193 | -3.7876 | *** |
| size similarity | actor cov. | -.0539 | .0774 | -.6960 | |
| core popularity | actor cov. | .0568 | .0721 | .7880 | |
| core activity | actor cov. | .4818 | .0720 | 6.6935 | *** |
| core similarity | actor cov. | -.0181 | .0548 | -.3313 | |
| lending rate pop. | actor cov. | .8814 | .5361 | 1.6439 | |
| lending rate act. | actor cov. | -5.1153 | 1.0055 | -5.0871 | *** |
| lending rate sim. | actor cov. | .0044 | .0045 | .9689 | |
| borrowing rate pop. | actor cov. | 4.6433 | 1.6571 | 2.8021 | *** |
| borrowing rate act. | actor cov. | -.9112 | 1.412 | -.6453 | |
| borrowing rate sim. | actor cov. | -.0049 | .0025 | -1.9469 | * |

The estimation results for 2005Q2.

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|-------|----------|---------|
| rate | | 36.2412 | | | |
| size | rate | .1399 | .0341 | 4.1007 | *** |
| density (outdegree) | structural | -4.3892 | .1836 | -23.9034 | *** |
| reciprocity | structural | .3259 | .2086 | 1.562 | |
| reciprocity persistence | structural | -.3804 | .3914 | -.9719 | |
| transitive triplets | structural | .0013 | .0034 | .3992 | |
| 3-cycles | structural | -.0148 | .0050 | -2.9608 | *** |
| indegree pop. (sqrt.) | structural | .3791 | .0202 | 18.7719 | *** |
| outdegree act. (sqrt.) | structural | .2192 | .0178 | 12.3298 | *** |
| past trades | dyadic cov. | .2439 | .0128 | 18.9954 | *** |
| past rates | dyadic cov. | 1.0693 | .6915 | 1.5464 | |
| size popularity | actor cov. | -.0262 | .0184 | -1.4234 | |
| size activity | actor cov. | -.1415 | .0198 | -7.1481 | *** |
| size similarity | actor cov. | -.0085 | .0737 | -.1156 | |
| core popularity | actor cov. | .2693 | .0757 | 3.5588 | *** |
| core activity | actor cov. | .2352 | .0700 | 3.3593 | *** |
| core similarity | actor cov. | .0119 | .0557 | .2144 | |
| lending rate pop. | actor cov. | .1662 | .7127 | .2332 | |
| lending rate act. | actor cov. | -2.0043 | .8135 | -2.4637 | ** |
| lending rate sim. | actor cov. | -.0049 | .0023 | -2.1268 | ** |
| borrowing rate pop. | actor cov. | .7931 | .2203 | 3.6009 | *** |
| borrowing rate act. | actor cov. | -.6690 | .2144 | -3.1199 | *** |
| borrowing rate sim. | actor cov. | -.0056 | .0023 | -2.4587 | ** |

The estimation results for 2005Q3.

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|--------|----------|---------|
| rate | | 47.3275 | | | |
| size | rate | .1115 | .0549 | 2.0296 | ** |
| density (outdegree) | structural | -4.2624 | .1897 | -22.4734 | *** |
| reciprocity | structural | 1.366 | .5376 | 2.5412 | ** |
| reciprocity persistence | structural | -2.234 | 1.1257 | -1.9844 | ** |
| transitive triplets | structural | -.0002 | .0031 | -.0650 | |
| 3-cycles | structural | -.0124 | .0046 | -2.7245 | *** |
| indegree pop. (sqrt.) | structural | .3525 | .0196 | 17.9414 | *** |
| outdegree act. (sqrt.) | structural | .2121 | .0178 | 11.9154 | *** |
| past trades | dyadic cov. | .2454 | .0130 | 18.9088 | *** |
| past rates | dyadic cov. | -.0141 | .6646 | -.0212 | |
| size popularity | actor cov. | -.0295 | .0184 | -1.6024 | |
| size activity | actor cov. | -.1063 | .0193 | -5.4986 | *** |
| size similarity | actor cov. | .0345 | .0746 | .4631 | |
| core popularity | actor cov. | .3444 | .0732 | 4.7065 | *** |
| core activity | actor cov. | .1923 | .0626 | 3.0722 | *** |
| core similarity | actor cov. | -.0524 | .0515 | -1.0163 | |
| lending rate pop. | actor cov. | .1996 | .2279 | .8759 | |
| lending rate act. | actor cov. | -.3291 | .2685 | -1.226 | |
| lending rate sim. | actor cov. | -.0043 | .0015 | -2.7702 | *** |
| borrowing rate pop. | actor cov. | .3196 | .2799 | 1.1416 | |
| borrowing rate act. | actor cov. | .4277 | .2272 | 1.8827 | * |
| borrowing rate sim. | actor cov. | -.0029 | .0017 | -1.6588 | * |

The estimation results for 2005Q4.

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|-------|---------|---------|
| rate | | 38.7363 | | | |
| size | rate | .1627 | .0437 | 3.723 | *** |
| density (outdegree) | structural | -4.5509 | .1866 | -24.385 | *** |
| reciprocity | structural | .5866 | .2667 | 2.199 | ** |
| reciprocity persistence | structural | -.2569 | .4838 | -.5311 | |
| transitive triplets | structural | -.0054 | .0040 | -1.3519 | |
| 3-cycles | structural | -.0229 | .0055 | -4.1338 | *** |
| indegree pop. (sqrt.) | structural | .3537 | .0192 | 18.4056 | *** |
| outdegree act. (sqrt.) | structural | .2609 | .0198 | 13.1894 | *** |
| past trades | dyadic cov. | .2519 | .0124 | 2.2480 | *** |
| past rates | dyadic cov. | 1.6438 | .4074 | 4.0346 | *** |
| size popularity | actor cov. | .0596 | .0197 | 3.0328 | *** |
| size activity | actor cov. | -.1792 | .0225 | -7.9829 | *** |
| size similarity | actor cov. | .0477 | .0786 | .6070 | |
| core popularity | actor cov. | .2305 | .0809 | 2.8511 | *** |
| core activity | actor cov. | -.0255 | .0663 | -.3840 | |
| core similarity | actor cov. | .0517 | .0591 | .8749 | |
| lending rate pop. | actor cov. | .2149 | .4774 | .4502 | |
| lending rate act. | actor cov. | -.7528 | .4353 | -1.7292 | * |
| lending rate sim. | actor cov. | -.0046 | .0016 | -2.8435 | *** |
| borrowing rate pop. | actor cov. | .8239 | .4810 | 1.7129 | * |
| borrowing rate act. | actor cov. | -.4959 | .4165 | -1.1907 | |
| borrowing rate sim. | actor cov. | .0014 | .0015 | .9043 | |

The estimation results for 2006Q1.

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|-------|----------|---------|
| rate | | 39.7191 | | | |
| size | rate | .1348 | .0395 | 3.4098 | *** |
| density (outdegree) | structural | -4.1285 | .1638 | -25.2082 | *** |
| reciprocity | structural | .9805 | .4424 | 2.2161 | ** |
| reciprocity persistence | structural | -1.2719 | .8730 | -1.457 | |
| transitive triplets | structural | -.0031 | .0032 | -.9876 | |
| 3-cycles | structural | -.0128 | .0048 | -2.6904 | *** |
| indegree pop. (sqrt.) | structural | .3338 | .0185 | 18.0902 | *** |
| outdegree act. (sqrt.) | structural | .2199 | .0162 | 13.5492 | *** |
| past trades | dyadic cov. | .2289 | .0128 | 17.8564 | *** |
| past rates | dyadic cov. | 4.2744 | .7418 | 5.7622 | *** |
| size popularity | actor cov. | .0815 | .0207 | 3.9461 | *** |
| size activity | actor cov. | -.1026 | .0206 | -4.9826 | *** |
| size similarity | actor cov. | .0937 | .0757 | 1.2383 | |
| core popularity | actor cov. | .0645 | .0718 | .8984 | |
| core activity | actor cov. | .4494 | .0664 | 6.768 | *** |
| core similarity | actor cov. | .0432 | .0532 | .8123 | |
| lending rate pop. | actor cov. | .1740 | .2978 | .5844 | |
| lending rate act. | actor cov. | -.8161 | .3097 | -2.6349 | *** |
| lending rate sim. | actor cov. | -.0025 | .0014 | -1.768 | * |
| borrowing rate pop. | actor cov. | -1.672 | .4679 | -3.5735 | *** |
| borrowing rate act. | actor cov. | .1949 | .4218 | .4619 | |
| borrowing rate sim. | actor cov. | -.004 | .0014 | -2.8138 | *** |

The estimation results for 2006Q2.

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|-------|----------|---------|
| rate | | 35.084 | | | |
| size | rate | .2453 | .0435 | 5.6392 | *** |
| density (outdegree) | structural | -4.2892 | .1729 | -24.8058 | *** |
| reciprocity | structural | .7428 | .3182 | 2.3345 | ** |
| reciprocity persistence | structural | -.9176 | .5985 | -1.5332 | |
| transitive triplets | structural | -.007 | .0034 | -2.0418 | ** |
| 3-cycles | structural | -.0162 | .0049 | -3.2851 | *** |
| indegree pop. (sqrt.) | structural | .3560 | .0191 | 18.6811 | *** |
| outdegree act. (sqrt.) | structural | .2469 | .0172 | 14.34 | *** |
| past trades | dyadic cov. | .2324 | .0122 | 19.0513 | *** |
| past rates | dyadic cov. | 3.419 | .7637 | 4.4772 | *** |
| size popularity | actor cov. | .0707 | .0188 | 3.7508 | *** |
| size activity | actor cov. | -.1463 | .0200 | -7.3267 | *** |
| size similarity | actor cov. | .0982 | .0765 | 1.2849 | |
| core popularity | actor cov. | .2591 | .0722 | 3.5890 | *** |
| core activity | actor cov. | .1178 | .0659 | 1.7885 | * |
| core similarity | actor cov. | .0166 | .0523 | .3176 | |
| lending rate pop. | actor cov. | -.4409 | .3793 | -1.1626 | |
| lending rate act. | actor cov. | -.4192 | .4113 | -1.0191 | |
| lending rate sim. | actor cov. | -.0023 | .0013 | -1.716 | * |
| borrowing rate pop. | actor cov. | -.7660 | .4023 | -1.904 | * |
| borrowing rate act. | actor cov. | .1610 | .3575 | .4502 | |
| borrowing rate sim. | actor cov. | -.0014 | .0014 | -.9752 | |

The estimation results for 2006Q3.

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|-------|---------|---------|
| rate | | 37.8351 | | | |
| size | rate | .1613 | .0376 | 4.2864 | *** |
| density (outdegree) | structural | -3.6951 | .1724 | -21.435 | *** |
| reciprocity | structural | .6397 | .2395 | 2.6704 | *** |
| reciprocity persistence | structural | -.6163 | .4658 | -1.3229 | |
| transitive triplets | structural | -.0044 | .0029 | -1.5026 | |
| 3-cycles | structural | -.0141 | .0045 | -3.1639 | *** |
| indegree pop. (sqrt.) | structural | .3008 | .0193 | 15.5654 | *** |
| outdegree act. (sqrt.) | structural | .1909 | .0160 | 11.9686 | *** |
| past trades | dyadic cov. | .2997 | .0141 | 21.2983 | *** |
| past rates | dyadic cov. | 2.8613 | .7977 | 3.5872 | *** |
| size popularity | actor cov. | -.0031 | .0185 | -.1677 | |
| size activity | actor cov. | -.004 | .0192 | -.2058 | |
| size similarity | actor cov. | .1939 | .075 | 2.5864 | *** |
| core popularity | actor cov. | .1912 | .0663 | 2.8862 | *** |
| core activity | actor cov. | .3804 | .0636 | 5.9782 | *** |
| core similarity | actor cov. | .0095 | .0510 | .1869 | |
| lending rate pop. | actor cov. | .0303 | .2867 | .1056 | |
| lending rate act. | actor cov. | .8982 | .3150 | 2.8517 | *** |
| lending rate sim. | actor cov. | -.0007 | .0016 | -.4585 | |
| borrowing rate pop. | actor cov. | 1.4095 | .3333 | 4.2284 | *** |
| borrowing rate act. | actor cov. | -1.2235 | .3509 | -3.4865 | *** |
| borrowing rate sim. | actor cov. | -.0023 | .0018 | -1.2664 | |

The estimation results for 2006Q4.

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|--------|----------|---------|
| rate | | 45.8994 | | | |
| size | rate | .2249 | .0488 | 4.611 | *** |
| density (outdegree) | structural | -4.6015 | .1845 | -24.9455 | *** |
| reciprocity | structural | 1.3899 | .7346 | 1.8919 | * |
| reciprocity persistence | structural | -2.1342 | 1.4517 | -1.4701 | |
| transitive triplets | structural | -.0110 | .0032 | -3.400 | *** |
| 3-cycles | structural | -.0158 | .0047 | -3.3987 | *** |
| indegree pop. (sqrt.) | structural | .3841 | .0204 | 18.8366 | *** |
| outdegree act. (sqrt.) | structural | .2725 | .0168 | 16.2387 | *** |
| past trades | dyadic cov. | .1843 | .0116 | 15.8488 | *** |
| past rates | dyadic cov. | 6.8250 | .8968 | 7.6101 | *** |
| size popularity | actor cov. | .0302 | .0181 | 1.6692 | * |
| size activity | actor cov. | -.1211 | .0195 | -6.2211 | *** |
| size similarity | actor cov. | .1152 | .0743 | 1.5497 | |
| core popularity | actor cov. | .2048 | .0669 | 3.0597 | *** |
| core activity | actor cov. | .2283 | .0616 | 3.7078 | *** |
| core similarity | actor cov. | .0041 | .0516 | .0801 | |
| lending rate pop. | actor cov. | -.5107 | .5089 | -1.0036 | |
| lending rate act. | actor cov. | -.0825 | .4468 | -.1847 | |
| lending rate sim. | actor cov. | -.0043 | .0017 | -2.5408 | ** |
| borrowing rate pop. | actor cov. | -.3939 | .5202 | -.7573 | |
| borrowing rate act. | actor cov. | -.0636 | .3491 | -.1822 | |
| borrowing rate sim. | actor cov. | -.0033 | .0020 | -1.6975 | * |

The estimation results for 2007Q1.

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|-------|----------|---------|
| rate | | 37.8194 | | | |
| size | rate | .2540 | .0421 | 6.0254 | *** |
| density (outdegree) | structural | -4.5349 | .1823 | -24.8827 | *** |
| reciprocity | structural | 1.2947 | .3704 | 3.4958 | *** |
| reciprocity persistence | structural | -1.8533 | .7863 | -2.3571 | ** |
| transitive triplets | structural | -.0012 | .0032 | -.3915 | |
| 3-cycles | structural | -.0081 | .0046 | -1.7617 | * |
| indegree pop. (sqrt.) | structural | .3987 | .0204 | 19.5858 | *** |
| outdegree act. (sqrt.) | structural | .2367 | .0169 | 13.9724 | *** |
| past trades | dyadic cov. | .1848 | .0128 | 14.3802 | *** |
| past rates | dyadic cov. | -.2112 | .8579 | -.2461 | |
| size popularity | actor cov. | .0372 | .0198 | 1.8738 | * |
| size activity | actor cov. | -.0782 | .0202 | -3.8724 | *** |
| size similarity | actor cov. | .0322 | .0809 | .3983 | |
| core popularity | actor cov. | -.0950 | .0655 | -1.4514 | |
| core activity | actor cov. | .2655 | .0659 | 4.0281 | *** |
| core similarity | actor cov. | -.0441 | .0524 | -.8408 | |
| lending rate pop. | actor cov. | -.6706 | .2662 | -2.5197 | ** |
| lending rate act. | actor cov. | -.0252 | .2727 | -.0925 | |
| lending rate sim. | actor cov. | -.0050 | .0021 | -2.3647 | ** |
| borrowing rate pop. | actor cov. | -1.3777 | .3004 | -4.5868 | *** |
| borrowing rate act. | actor cov. | .1096 | .2778 | .3944 | |
| borrowing rate sim. | actor cov. | .0024 | .0019 | 1.3157 | |

The estimation results for 2007Q2.

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|-------|---------|---------|
| rate | | 35.8918 | | | |
| size | rate | .1743 | .0350 | 4.9808 | *** |
| density (outdegree) | structural | -3.7586 | .1877 | -2.0251 | *** |
| reciprocity | structural | .2903 | .2255 | 1.2871 | |
| reciprocity persistence | structural | -.1465 | .4124 | -.3553 | |
| transitive triplets | structural | -.001 | .0035 | -.3031 | |
| 3-cycles | structural | -.0304 | .0055 | -5.5632 | *** |
| indegree pop. (sqrt.) | structural | .3316 | .0208 | 15.9627 | *** |
| outdegree act. (sqrt.) | structural | .1906 | .0180 | 1.5893 | *** |
| past trades | dyadic cov. | .1652 | .0123 | 13.3973 | *** |
| past rates | dyadic cov. | 2.5878 | .7492 | 3.4541 | *** |
| size popularity | actor cov. | .0614 | .0206 | 2.9821 | *** |
| size activity | actor cov. | -.0945 | .022 | -4.2839 | *** |
| size similarity | actor cov. | .1574 | .0797 | 1.9743 | ** |
| core popularity | actor cov. | .0910 | .0733 | 1.2414 | |
| core activity | actor cov. | .6462 | .0757 | 8.5328 | *** |
| core similarity | actor cov. | .0242 | .0543 | .4456 | |
| lending rate pop. | actor cov. | -.4835 | .1342 | -3.6034 | *** |
| lending rate act. | actor cov. | .7547 | .1986 | 3.7999 | *** |
| lending rate sim. | actor cov. | .0024 | .0021 | 1.1286 | |
| borrowing rate pop. | actor cov. | .5373 | .2730 | 1.9681 | ** |
| borrowing rate act. | actor cov. | .1030 | .2488 | .4139 | |
| borrowing rate sim. | actor cov. | -.0074 | .0029 | -2.5299 | ** |

The estimation results for 2007Q3.

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|--------|----------|---------|
| rate | | 36.1447 | | | |
| size | rate | .1624 | .0369 | 4.4048 | *** |
| density (outdegree) | structural | -4.3576 | .1809 | -24.0921 | *** |
| reciprocity | structural | 1.4395 | .6778 | 2.1237 | ** |
| reciprocity persistence | structural | -2.3901 | 1.2921 | -1.8497 | * |
| transitive triplets | structural | .0047 | .0041 | 1.1538 | |
| 3-cycles | structural | -.0256 | .0057 | -4.4907 | *** |
| indegree pop. (sqrt.) | structural | .4189 | .0208 | 2.1605 | *** |
| outdegree act. (sqrt.) | structural | .1885 | .0195 | 9.682 | *** |
| past trades | dyadic cov. | .1396 | .0144 | 9.6917 | *** |
| past rates | dyadic cov. | -1.7065 | .8778 | -1.944 | * |
| size popularity | actor cov. | -.0580 | .0199 | -2.912 | *** |
| size activity | actor cov. | .0407 | .0239 | 1.7053 | * |
| size similarity | actor cov. | .2011 | .0801 | 2.5123 | ** |
| core popularity | actor cov. | .0379 | .0743 | .5100 | |
| core activity | actor cov. | .1311 | .0741 | 1.7685 | * |
| core similarity | actor cov. | -.1322 | .0538 | -2.4585 | ** |
| lending rate pop. | actor cov. | -.3850 | .4634 | -.8307 | |
| lending rate act. | actor cov. | 2.6351 | 1.0131 | 2.6011 | *** |
| lending rate sim. | actor cov. | .0213 | .0069 | 3.1056 | *** |
| borrowing rate pop. | actor cov. | .5701 | .2889 | 1.9736 | ** |
| borrowing rate act. | actor cov. | .1268 | .2744 | .4619 | |
| borrowing rate sim. | actor cov. | -.0033 | .0024 | -1.3648 | |

The estimation results for 2007Q4.

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|-------|---------|---------|
| rate | | 31.2033 | | | |
| size | rate | .1413 | .0429 | 3.2943 | *** |
| density (outdegree) | structural | -4.2138 | .2061 | -2.4500 | *** |
| reciprocity | structural | .8234 | .3439 | 2.3939 | ** |
| reciprocity persistence | structural | -.9803 | .6306 | -1.5545 | |
| transitive triplets | structural | -.0161 | .0042 | -3.7837 | *** |
| 3-cycles | structural | -.0200 | .0063 | -3.1946 | *** |
| indegree pop. (sqrt.) | structural | .3358 | .023 | 14.6076 | *** |
| outdegree act. (sqrt.) | structural | .2690 | .0205 | 13.1163 | *** |
| past trades | dyadic cov. | .2142 | .0145 | 14.7887 | *** |
| past rates | dyadic cov. | .7994 | .8011 | .9979 | |
| size popularity | actor cov. | .0176 | .0212 | .8329 | |
| size activity | actor cov. | .0392 | .0215 | 1.8242 | * |
| size similarity | actor cov. | .0073 | .0845 | .0870 | |
| core popularity | actor cov. | .4406 | .0747 | 5.8971 | *** |
| core activity | actor cov. | -.0204 | .0632 | -.3233 | |
| core similarity | actor cov. | .1161 | .0530 | 2.1894 | ** |
| lending rate pop. | actor cov. | -.1386 | .3913 | -.3541 | |
| lending rate act. | actor cov. | .1721 | .4704 | .3660 | |
| lending rate sim. | actor cov. | -.0006 | .0022 | -.2544 | |
| borrowing rate pop. | actor cov. | -2.207 | .7121 | -3.0994 | *** |
| borrowing rate act. | actor cov. | -1.0169 | .6940 | -1.4652 | |
| borrowing rate sim. | actor cov. | .0009 | .0019 | .4785 | |

The estimation results for 2008Q1.

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|--------|----------|---------|
| rate | | 34.1326 | | | |
| size | rate | .2190 | .0430 | 5.0932 | *** |
| density (outdegree) | structural | -4.6693 | .1848 | -25.2701 | *** |
| reciprocity | structural | 1.2397 | .3982 | 3.1137 | *** |
| reciprocity persistence | structural | -1.8402 | .8068 | -2.2809 | ** |
| transitive triplets | structural | -.0019 | .0036 | -.5360 | |
| 3-cycles | structural | -.0225 | .0062 | -3.595 | *** |
| indegree pop. (sqrt.) | structural | .4429 | .0220 | 2.1438 | *** |
| outdegree act. (sqrt.) | structural | .2359 | .0172 | 13.6783 | *** |
| past trades | dyadic cov. | .1480 | .0123 | 12.0785 | *** |
| past rates | dyadic cov. | 5.593 | 1.7701 | 3.1597 | *** |
| size popularity | actor cov. | -.0760 | .0215 | -3.5334 | *** |
| size activity | actor cov. | .0596 | .0228 | 2.6117 | *** |
| size similarity | actor cov. | -.0066 | .0898 | -.0733 | |
| core popularity | actor cov. | .1789 | .0711 | 2.5155 | ** |
| core activity | actor cov. | .1194 | .0671 | 1.7784 | * |
| core similarity | actor cov. | .0566 | .0565 | 1.0019 | |
| lending rate pop. | actor cov. | -.0855 | .3706 | -.2306 | |
| lending rate act. | actor cov. | -.7514 | .5369 | -1.3996 | |
| lending rate sim. | actor cov. | -.0003 | .0023 | -.1203 | |
| borrowing rate pop. | actor cov. | -1.4141 | .3702 | -3.8197 | *** |
| borrowing rate act. | actor cov. | -.0986 | .2986 | -.3303 | |
| borrowing rate sim. | actor cov. | .0013 | .0023 | .5894 | |

The estimation results for 2008Q2.

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|-------|----------|---------|
| rate | | 31.7372 | | | |
| size | rate | .1537 | .0397 | 3.8679 | *** |
| density (outdegree) | structural | -3.7097 | .1976 | -18.7715 | *** |
| reciprocity | structural | 1.142 | .3434 | 3.3253 | *** |
| reciprocity persistence | structural | -1.5188 | .7077 | -2.1461 | ** |
| transitive triplets | structural | -7e-04 | .0043 | -.1746 | |
| 3-cycles | structural | -.0236 | .0068 | -3.4841 | *** |
| indegree pop. (sqrt.) | structural | .2945 | .0250 | 11.8015 | *** |
| outdegree act. (sqrt.) | structural | .1947 | .0201 | 9.6997 | *** |
| past trades | dyadic cov. | .2161 | .0131 | 16.4564 | *** |
| past rates | dyadic cov. | 1.739 | 1.117 | 1.5568 | |
| size popularity | actor cov. | .0123 | .0228 | .5398 | |
| size activity | actor cov. | -.0608 | .0220 | -2.7676 | *** |
| size similarity | actor cov. | -.0871 | .0855 | -1.0180 | |
| core popularity | actor cov. | .2036 | .0814 | 2.5008 | ** |
| core activity | actor cov. | .3719 | .0723 | 5.1447 | *** |
| core similarity | actor cov. | .0280 | .0574 | .4881 | |
| lending rate pop. | actor cov. | -.0699 | .3317 | -.2107 | |
| lending rate act. | actor cov. | .5686 | .341 | 1.6678 | * |
| lending rate sim. | actor cov. | -.0009 | .0024 | -.3716 | |
| borrowing rate pop. | actor cov. | 1.0454 | .2459 | 4.252 | *** |
| borrowing rate act. | actor cov. | .0165 | .2451 | .0672 | |
| borrowing rate sim. | actor cov. | -.0033 | .0030 | -1.0896 | |

The estimation results for 2008Q3.

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|-------|----------|---------|
| rate | | 41.7128 | | | |
| size | rate | .1632 | .0466 | 3.5003 | *** |
| density (outdegree) | structural | -4.8514 | .1811 | -26.7826 | *** |
| reciprocity | structural | 1.1375 | .5252 | 2.1657 | ** |
| reciprocity persistence | structural | -1.1919 | .9896 | -1.2045 | |
| transitive triplets | structural | -.0167 | .0041 | -4.0242 | *** |
| 3-cycles | structural | -.0434 | .0095 | -4.5808 | *** |
| indegree pop. (sqrt.) | structural | .4532 | .0219 | 2.6575 | *** |
| outdegree act. (sqrt.) | structural | .2771 | .0170 | 16.3388 | *** |
| past trades | dyadic cov. | .0596 | .0109 | 5.4609 | *** |
| past rates | dyadic cov. | 1.1072 | .7973 | 1.3886 | |
| size popularity | actor cov. | .0008 | .0198 | .0410 | |
| size activity | actor cov. | .0778 | .0201 | 3.8678 | *** |
| size similarity | actor cov. | .3893 | .0867 | 4.4884 | *** |
| core popularity | actor cov. | .3555 | .0734 | 4.8413 | *** |
| core activity | actor cov. | .2656 | .0645 | 4.1149 | *** |
| core similarity | actor cov. | .0652 | .0546 | 1.1954 | |
| lending rate pop. | actor cov. | .0605 | .0521 | 1.1616 | |
| lending rate act. | actor cov. | .2260 | .0659 | 3.4294 | *** |
| lending rate sim. | actor cov. | -.0040 | .0013 | -3.2201 | *** |
| borrowing rate pop. | actor cov. | -.3626 | .0764 | -4.7444 | *** |
| borrowing rate act. | actor cov. | .0930 | .0692 | 1.3439 | |
| borrowing rate sim. | actor cov. | -.0036 | .0013 | -2.7895 | *** |

The estimation results for 2008Q4.

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|-------|----------|---------|
| rate | | 3.3537 | | | |
| size | rate | .0849 | .0442 | 1.9201 | * |
| density (outdegree) | structural | -4.1898 | .1722 | -24.3272 | *** |
| reciprocity | structural | .3771 | .3348 | 1.1264 | |
| reciprocity persistence | structural | .1677 | .6017 | .2787 | |
| transitive triplets | structural | -.0184 | .0065 | -2.8096 | *** |
| 3-cycles | structural | -.0602 | .0147 | -4.1023 | *** |
| indegree pop. (sqrt.) | structural | .3945 | .0231 | 17.0741 | *** |
| outdegree act. (sqrt.) | structural | .2453 | .0204 | 12.016 | *** |
| past trades | dyadic cov. | .0867 | .0102 | 8.4683 | *** |
| past rates | dyadic cov. | -2.3857 | .2724 | -8.758 | *** |
| size popularity | actor cov. | .0542 | .0233 | 2.3278 | ** |
| size activity | actor cov. | .0053 | .0221 | .2395 | |
| size similarity | actor cov. | .0181 | .0914 | .1979 | |
| core popularity | actor cov. | .3560 | .1018 | 3.4975 | *** |
| core activity | actor cov. | .2595 | .0798 | 3.2530 | *** |
| core similarity | actor cov. | .0843 | .0696 | 1.2112 | |
| lending rate pop. | actor cov. | .3449 | .1335 | 2.5845 | *** |
| lending rate act. | actor cov. | -.4828 | .1353 | -3.5672 | *** |
| lending rate sim. | actor cov. | -.0008 | .0017 | -.4579 | |
| borrowing rate pop. | actor cov. | -.3899 | .1444 | -2.7009 | *** |
| borrowing rate act. | actor cov. | .0743 | .1336 | .5561 | |
| borrowing rate sim. | actor cov. | -.0009 | .0021 | -.4438 | |

The estimation results for 2009Q1.

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|-------|----------|---------|
| rate | | 27.9755 | | | |
| size | rate | .0447 | .0420 | 1.0637 | |
| density (outdegree) | structural | -3.8450 | .1752 | -21.9462 | *** |
| reciprocity | structural | .1078 | .3310 | .3256 | |
| reciprocity persistence | structural | -.0715 | .6017 | -.1188 | |
| transitive triplets | structural | -.0093 | .0070 | -1.3347 | |
| 3-cycles | structural | -.0444 | .0173 | -2.561 | ** |
| indegree pop. (sqrt.) | structural | .3604 | .0246 | 14.6209 | *** |
| outdegree act. (sqrt.) | structural | .2111 | .0208 | 1.1288 | *** |
| past trades | dyadic cov. | .1681 | .0109 | 15.3598 | *** |
| past rates | dyadic cov. | -.7618 | .1875 | -4.0618 | *** |
| size popularity | actor cov. | .0266 | .0236 | 1.1246 | |
| size activity | actor cov. | .1229 | .0246 | 4.9873 | *** |
| size similarity | actor cov. | -.0364 | .0939 | -.3872 | |
| core popularity | actor cov. | .4145 | .0995 | 4.1641 | *** |
| core activity | actor cov. | .1818 | .0832 | 2.1858 | ** |
| core similarity | actor cov. | .1094 | .0727 | 1.5038 | |
| lending rate pop. | actor cov. | -.0085 | .2011 | -.0422 | |
| lending rate act. | actor cov. | .9787 | .2264 | 4.3228 | *** |
| lending rate sim. | actor cov. | .0040 | .0021 | 1.8833 | * |
| borrowing rate pop. | actor cov. | .3080 | .2760 | 1.1158 | |
| borrowing rate act. | actor cov. | .7208 | .2540 | 2.8379 | *** |
| borrowing rate sim. | actor cov. | -.0062 | .0019 | -3.2517 | *** |

The estimation results for 2009Q2.

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|-------|----------|---------|
| rate | | 29.9781 | | | |
| size | rate | .0327 | .0479 | .6837 | |
| density (outdegree) | structural | -3.8869 | .1787 | -21.7545 | *** |
| reciprocity | structural | .8945 | .5052 | 1.7708 | * |
| reciprocity persistence | structural | -1.3356 | 1.095 | -1.2196 | |
| transitive triplets | structural | -.0089 | .0061 | -1.4495 | |
| 3-cycles | structural | -.0251 | .0131 | -1.9205 | * |
| indegree pop. (sqrt.) | structural | .3674 | .0252 | 14.5935 | *** |
| outdegree act. (sqrt.) | structural | .2163 | .0202 | 1.7261 | *** |
| past trades | dyadic cov. | .1784 | .0120 | 14.915 | *** |
| past rates | dyadic cov. | -.1515 | .1685 | -.8987 | |
| size popularity | actor cov. | .0214 | .0234 | .9110 | |
| size activity | actor cov. | -.0797 | .0245 | -3.2542 | *** |
| size similarity | actor cov. | .0076 | .0941 | .0806 | |
| core popularity | actor cov. | .1497 | .0884 | 1.6936 | * |
| core activity | actor cov. | .4200 | .0768 | 5.4715 | *** |
| core similarity | actor cov. | -.1022 | .0702 | -1.4554 | |
| lending rate pop. | actor cov. | -.5748 | .4711 | -1.2203 | |
| lending rate act. | actor cov. | .7298 | .4920 | 1.4834 | |
| lending rate sim. | actor cov. | .0001 | .0029 | .0500 | |
| borrowing rate pop. | actor cov. | .1757 | .5920 | .2968 | |
| borrowing rate act. | actor cov. | 2.0046 | .4898 | 4.093 | *** |
| borrowing rate sim. | actor cov. | .0032 | .0023 | 1.3857 | |

The estimation results for 2009Q3.

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|--------|----------|---------|
| rate | | 28.9752 | | | |
| size | rate | .1587 | .0391 | 4.0591 | *** |
| density (outdegree) | structural | -4.1119 | .1768 | -23.2511 | *** |
| reciprocity | structural | .5258 | .4018 | 1.3086 | |
| reciprocity persistence | structural | -.4774 | .7550 | -.6323 | |
| transitive triplets | structural | .0002 | .0062 | .0255 | |
| 3-cycles | structural | -.0723 | .0168 | -4.3082 | *** |
| indegree pop. (sqrt.) | structural | .3742 | .0262 | 14.2986 | *** |
| outdegree act. (sqrt.) | structural | .2289 | .0196 | 11.6667 | *** |
| past trades | dyadic cov. | .1753 | .0124 | 14.0861 | *** |
| past rates | dyadic cov. | .3200 | .2421 | 1.3221 | |
| size popularity | actor cov. | -.0194 | .0245 | -.7930 | |
| size activity | actor cov. | .0987 | .0255 | 3.863 | *** |
| size similarity | actor cov. | -.1816 | .1001 | -1.8143 | * |
| core popularity | actor cov. | .4124 | .1003 | 4.1117 | *** |
| core activity | actor cov. | .0803 | .0832 | .9660 | |
| core similarity | actor cov. | .0592 | .0717 | .8248 | |
| lending rate pop. | actor cov. | -2.4602 | .9975 | -2.4662 | ** |
| lending rate act. | actor cov. | -.2496 | 1.0997 | -.2270 | |
| lending rate sim. | actor cov. | -.0031 | .0023 | -1.3602 | |
| borrowing rate pop. | actor cov. | -2.0534 | .9180 | -2.2367 | ** |
| borrowing rate act. | actor cov. | -.0922 | .8231 | -.1121 | |
| borrowing rate sim. | actor cov. | .0019 | .0032 | .5887 | |

The estimation results for 2009Q4.

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|-------|----------|---------|
| rate | | 29.7315 | | | |
| size | rate | .3123 | .0627 | 4.9808 | *** |
| density (outdegree) | structural | -4.0936 | .1845 | -22.1818 | *** |
| reciprocity | structural | .3898 | .2955 | 1.3191 | |
| reciprocity persistence | structural | .4689 | .5914 | .7928 | |
| transitive triplets | structural | -.0158 | .0058 | -2.7427 | *** |
| 3-cycles | structural | -.0463 | .0125 | -3.7105 | *** |
| indegree pop. (sqrt.) | structural | .3879 | .0247 | 15.7144 | *** |
| outdegree act. (sqrt.) | structural | .2457 | .0203 | 12.1048 | *** |
| past trades | dyadic cov. | .2249 | .0127 | 17.6866 | *** |
| past rates | dyadic cov. | -.5095 | .2753 | -1.8507 | * |
| size popularity | actor cov. | .0166 | .0227 | .7313 | |
| size activity | actor cov. | -.1027 | .0269 | -3.8211 | *** |
| size similarity | actor cov. | .2328 | .0954 | 2.4410 | ** |
| core popularity | actor cov. | .2302 | .0816 | 2.8221 | *** |
| core activity | actor cov. | .5776 | .0726 | 7.9558 | *** |
| core similarity | actor cov. | .1317 | .0633 | 2.0814 | ** |
| lending rate pop. | actor cov. | .44 | .7735 | .5689 | |
| lending rate act. | actor cov. | -1.6967 | .8095 | -2.0958 | ** |
| lending rate sim. | actor cov. | .0006 | .0025 | .2425 | |
| borrowing rate pop. | actor cov. | 1.0379 | .8328 | 1.2463 | |
| borrowing rate act. | actor cov. | 2.2273 | .8048 | 2.7675 | *** |
| borrowing rate sim. | actor cov. | .0045 | .0020 | 2.2342 | ** |

The estimation results for 2010Q1.

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|--------|----------|---------|
| rate | | 24.0392 | | | |
| size | rate | .2405 | .0626 | 3.8442 | *** |
| density (outdegree) | structural | -3.7885 | .1957 | -19.3575 | *** |
| reciprocity | structural | 1.5669 | 1.2133 | 1.2914 | |
| reciprocity persistence | structural | -2.7249 | 2.674 | -1.0190 | |
| transitive triplets | structural | -.0034 | .0059 | -.5739 | |
| 3-cycles | structural | -.0704 | .0117 | -6.0191 | *** |
| indegree pop. (sqrt.) | structural | .3714 | .0315 | 11.7947 | *** |
| outdegree act. (sqrt.) | structural | .2031 | .0215 | 9.4402 | *** |
| past trades | dyadic cov. | .1624 | .0142 | 11.3974 | *** |
| past rates | dyadic cov. | -.4892 | .3523 | -1.3887 | |
| size popularity | actor cov. | .0482 | .0269 | 1.7921 | * |
| size activity | actor cov. | .0184 | .0327 | .5639 | |
| size similarity | actor cov. | .1686 | .1180 | 1.4287 | |
| core popularity | actor cov. | .3799 | .1027 | 3.6975 | *** |
| core activity | actor cov. | .7051 | .0910 | 7.7495 | *** |
| core similarity | actor cov. | .0273 | .0674 | .4043 | |
| lending rate pop. | actor cov. | .0213 | .4306 | .0494 | |
| lending rate act. | actor cov. | .5199 | .4935 | 1.0534 | |
| lending rate sim. | actor cov. | .0009 | .0039 | .2424 | |
| borrowing rate pop. | actor cov. | 1.2568 | .7464 | 1.6839 | * |
| borrowing rate act. | actor cov. | .9758 | .7572 | 1.2887 | |
| borrowing rate sim. | actor cov. | .0047 | .0018 | 2.5317 | ** |

The estimation results for 2010Q2.

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|-------|----------|---------|
| rate | | 22.2861 | | | |
| size | rate | .1461 | .0449 | 3.2556 | *** |
| density (outdegree) | structural | -4.0584 | .2033 | -19.9649 | *** |
| reciprocity | structural | .3177 | .2656 | 1.1963 | |
| reciprocity persistence | structural | .1022 | .4958 | .2061 | |
| transitive triplets | structural | -.0028 | .0072 | -.3850 | |
| 3-cycles | structural | -.0801 | .0148 | -5.4226 | *** |
| indegree pop. (sqrt.) | structural | .3546 | .0265 | 13.3866 | *** |
| outdegree act. (sqrt.) | structural | .2655 | .0254 | 1.4495 | *** |
| past trades | dyadic cov. | .1683 | .0152 | 11.0403 | *** |
| past rates | dyadic cov. | -.0195 | .1026 | -.1903 | |
| size popularity | actor cov. | .0599 | .0284 | 2.1095 | ** |
| size activity | actor cov. | -.0785 | .0293 | -2.6802 | *** |
| size similarity | actor cov. | .1198 | .1080 | 1.1092 | |
| core popularity | actor cov. | .3043 | .1026 | 2.9668 | *** |
| core activity | actor cov. | .2718 | .0872 | 3.117 | *** |
| core similarity | actor cov. | -.0366 | .0747 | -.4899 | |
| lending rate pop. | actor cov. | -1.1609 | .3195 | -3.6339 | *** |
| lending rate act. | actor cov. | .4555 | .3209 | 1.4193 | |
| lending rate sim. | actor cov. | -.0051 | .0020 | -2.5413 | ** |
| borrowing rate pop. | actor cov. | -.5385 | .3882 | -1.3873 | |
| borrowing rate act. | actor cov. | -.5754 | .3346 | -1.7196 | * |
| borrowing rate sim. | actor cov. | .0002 | .0017 | .1324 | |

The estimation results for 2010Q3.

| effect | category | value | s.e. | t-value | signif. |
|-------------------------|-------------|---------|-------|----------|---------|
| rate | | 2.3403 | | | |
| size | rate | .1463 | .0453 | 3.2323 | *** |
| density (outdegree) | structural | -3.4112 | .2095 | -16.2826 | *** |
| reciprocity | structural | .7579 | .3548 | 2.1364 | ** |
| reciprocity persistence | structural | -.9580 | .6838 | -1.4011 | |
| transitive triplets | structural | -.0015 | .0069 | -.2208 | |
| 3-cycles | structural | -.0392 | .0118 | -3.3302 | *** |
| indegree pop. (sqrt.) | structural | .3519 | .029 | 12.1495 | *** |
| outdegree act. (sqrt.) | structural | .1555 | .0255 | 6.0964 | *** |
| past trades | dyadic cov. | .3097 | .0194 | 15.9602 | *** |
| past rates | dyadic cov. | .0253 | .1797 | .1406 | |
| size popularity | actor cov. | .0392 | .0280 | 1.399 | |
| size activity | actor cov. | -.0117 | .0277 | -.4224 | |
| size similarity | actor cov. | .0715 | .106 | .6743 | |
| core popularity | actor cov. | .2838 | .0968 | 2.9307 | *** |
| core activity | actor cov. | .2548 | .0869 | 2.9323 | *** |
| core similarity | actor cov. | .0591 | .0723 | .8166 | |
| lending rate pop. | actor cov. | -.2328 | .2668 | -.8727 | |
| lending rate act. | actor cov. | -.7142 | .3013 | -2.3708 | ** |
| lending rate sim. | actor cov. | -.0006 | .0020 | -.3062 | |
| borrowing rate pop. | actor cov. | -.0698 | .3036 | -.2298 | |
| borrowing rate act. | actor cov. | .2332 | .2660 | .8765 | |
| borrowing rate sim. | actor cov. | -.0018 | .0022 | -.8347 | |

The estimation results for 2010Q4.

Finally, the χ^2 distributed test statistics for the joint significance of the three actor covariate related effects (popularity, activity, similarity) for each actor covariate (size, core, lending rate, borrowing rate) are presented in Table 5.11. There is no pattern visible in the significance among covariates or with respect to a specific period whatsoever.

| period | size | core | lending rate | borrowing rate |
|--------|-----------|-----------|--------------|----------------|
| 2002Q1 | 31.163*** | 67.257*** | 6.215 | 17.824*** |
| 2002Q2 | 10.703** | 38.708*** | 17.439*** | 56.136*** |
| 2003Q3 | 16.819*** | 28.733*** | 2.591 | 6.273* |
| 2002Q4 | 3.225 | 35.441*** | 3.908 | 41.501*** |
| 2003Q1 | 7.470* | 27.641*** | 4.595 | 19.613*** |
| 2003Q2 | 19.200*** | 30.844*** | 19.31*** | 35.715*** |
| 2003Q3 | 18.991*** | 39.743*** | 3.050 | 11.142** |
| 2003Q4 | 15.452*** | 24.520*** | 15.367*** | 15.429*** |
| 2004Q1 | 4.547 | 19.562*** | 4.781 | 11.956*** |
| 2004Q2 | 6.977* | 61.475*** | 29.602*** | 0.385 |
| 2004Q3 | 9.574** | 49.486*** | 8.421** | 44.629*** |
| 2004Q4 | 40.462*** | 55.951*** | 9.714** | 3.153 |
| 2005Q1 | 34.123*** | 45.922*** | 5.755 | 15.172*** |
| 2005Q2 | 25.965*** | 47.352*** | 28.872*** | 15.081*** |
| 2005Q3 | 53.190*** | 24.672*** | 31.568*** | 29.524*** |
| 2005Q4 | 32.312*** | 36.321*** | 14.001*** | 4.091 |
| 2006Q1 | 70.000*** | 9.140** | 10.397** | 5.870 |
| 2006Q2 | 40.719*** | 47.384*** | 10.337** | 20.288*** |
| 2006Q3 | 66.313*** | 15.722*** | 5.300 | 5.396 |
| 2006Q4 | 6.765* | 42.965*** | 9.065** | 41.412*** |
| 2007Q1 | 43.629*** | 20.917*** | 6.724* | 4.225 |
| 2007Q2 | 20.576*** | 21.585*** | 10.410** | 22.201*** |
| 2007Q3 | 30.026*** | 73.481*** | 26.824*** | 9.992** |
| 2007Q4 | 14.758*** | 12.138*** | 29.112*** | 4.601 |
| 2008Q1 | 3.967 | 36.556*** | 0.379 | 11.513*** |
| 2008Q2 | 21.705*** | 8.305** | 1.983 | 15.372*** |
| 2008Q3 | 10.033** | 32.374*** | 2.944 | 28.723*** |
| 2008Q4 | 34.809*** | 32.813*** | 23.854*** | 31.731*** |
| 2009Q1 | 5.753 | 19.886*** | 26.188*** | 8.297** |
| 2009Q2 | 25.650*** | 19.805*** | 18.818*** | 23.237*** |
| 2009Q3 | 11.673*** | 48.634*** | 6.344* | 17.845*** |
| 2009Q4 | 19.670*** | 18.230*** | 6.629* | 8.559** |
| 2010Q1 | 22.213*** | 67.186*** | 6.990* | 8.744** |
| 2010Q2 | 6.914* | 65.111*** | 1.984 | 7.468* |
| 2010Q3 | 13.178*** | 23.813*** | 20.234*** | 4.692 |
| 2010Q4 | 2.642 | 15.154*** | 6.304* | 1.544 |

Tab. 5.11: The χ^2 distributed test statistics for the joint significance of the actor covariate related effects (popularity, activity, similarity) for each for each covariate. The columns state the name of the covariate and the significance is indicated by the * (* 10%, ** 5%, *** 1%).

6. WHO ARE THE SPECULATORS ON COMMODITY
FUTURES MARKETS?

6.1 Introduction

The price boom on almost all commodity markets between 2001 and the middle of 2008 has often been associated with the financialization of these markets and the corresponding increase of speculative activity. When it comes to speculation the public perception is biased towards potential adverse distributional effects. For instance, one focus of the German media is the dramatic effects higher prices of agricultural commodities have for the population in developing countries and the important role of two German financial institutions, namely the Deutsche Bank and Allianz, in the corresponding future markets. Foodwatch (2013) recently published several internal documents from both institutions, in which the authors admit that speculation increases the commodity price volatility in specific periods and acknowledge the potentially negative effects for farmers and consumers. These effects are not only limited to developing countries as distortions in the price formation process might provide adverse signals for the real sector and reduce the efficiency of resource allocation.

Contrary to the public perception, in the scientific community the ramifications of speculation in this respect are highly controversial, because fundamental explanations - most prominently the soaring demand due to the rapid economic growth in developing countries or the expansive monetary policy in the US - are also associated with the price boom. In case of agricultural commodities additional reasons might be the use of grains as biofuel¹ and bad harvest due to foul weather conditions in important agricultural areas.

The financialization of commodity markets is an on-going process with the future markets at the center, since they enable traders to make speculative profits without actually participating in the physical trade of the underlying commodity. A cornerstone in the process was the discovery that investors are able to reduce their portfolio risk by including commodities.² This is due to the weak correlation between the returns of commodities and typical asset classes like stock and bond markets. However, a necessity to realize the possible gains was the reduction of the trading costs due to the commodity future markets becoming deeper and more liquid. This finding leads to two results: First, a rise in investment vehicles covering a broad spectrum of commodity futures, typically long only, called commodity index

¹ Hertel and Beckman (2011) investigate the influence of biofuel on the volatility of food commodity prices by strengthening the link to energy commodity and Trostle (2010) identifies several demand and supply factors including biofuel determining food commodity prices.

² See e.g. Gorton and Rouwenhorst (2006), Erb and Harvey (2006) and Conover *et al.* (2010).

funds (CIF). Second, an increasing number of individual investors participating in commodity markets since financial institutions started to offer CIFs to their customers. Today the Deutsche Bank advises their customers to hold 5-10% of commodities in their portfolios. Note that Commodity index trading is not limited to publicly traded CIF, yet also used by many institutional investors to reduce their portfolio risk.

The dramatic increase of funds invested this way in recent years has also made the research community focus on CIF when investigating the effects of speculation. Gilbert (2010b, 2010a) uses the activity of CIF as a proxy for speculation and finds a significant impact on returns for commodity markets. In contrast, Stoll and Whaley (2010), Sanders and Irwin (2010) and Irwin *et al.* (2009) do not detect evidence that CIF influence the prices of commodity future markets. Sanders and Irwin (2011) conclude in their survey that no final judgment on the effects of CIF on commodity future markets has yet been possible. However, Stoll and Whaley (2010) question whether CIF should at all be seen as speculators, because the funds are typically passively managed and buy only long contracts. Brunetti *et al.* (2011) and Büyüksahin and Robe (2014) use open positions of swap dealers, which are similar to those of CIF as we will see later, and find no significant impact on the returns of commodity markets. Bohl and Stephan (2012) compare the conditional volatility for the periods of 1992-2002 and 2002-2012 and conclude that the spot markets are not destabilized by increasing speculative activity.

This paper takes an approach similar to Sanders *et al.* (2004) to measure speculation using data from the Commodity Futures Trading Commission (CFTC) published in their weekly Commitments of Traders (COT) reports displaying the aggregate positions for specified trader classes. Our sample consists of 10 commodities and covers the period June 2006 to June 2013. We approximate the expected price change of the respective class by taking the difference of the long position and the short position (net-long). In the following the net-long positions are used to investigate if there is a Granger causal link between the speculative activities on future markets of the different trader classes and the nearest future price. The Toda and Yamamoto (1995) procedure is applied to test for Granger non causality.

The most robust finding of the paper is that the class called ‘Money Manager’, which are not involved with the physical commodity and trade funds on behalf of their clients, rarely influence the price, but often react to (are Granger caused by) it. The same is true for the ‘Producer/Merchant/Processor/User’ of the commodities, yet their adjustments to a price change point typically in the opposite direction. Given that and the very strong negative correlation between the net-long positions of both classes we conclude that Money Manager provide the hedging possibilities needed by the producers

in the first place, while (probably) creating profits doing so. For the commodity index traders we overall find little evidence of causal relations with the price in any direction and additionally relatively low correlation with the other classes. Hence, they seemingly follow their investment strategy reducing their overall portfolio risk relatively independent of the price or of other trader classes. All in all, this article does not detect any robust adverse effect of speculative activity on commodity markets. However, these findings make a further investigation of the different classes and their interrelations a promising field for future research.

The remaining part of this paper is structured as follows. In the next section the data set used, the future markets and commodities investigated are briefly introduced. Afterwards, the trader classes filed in the different reports and their connections are explained and possible candidates for speculative activity are identified. In the following empirical study the Granger-causal links between the trader classes and the prices are investigated. Finally, we draw a brief conclusion of the most interesting results and give an outlook for future research.

6.2 Data

In their weekly COT report the CFTC publishes Tuesday's aggregated positioning for specific futures and options on these futures. Reporting firms have to hand in daily reports about traders holding positions above specific levels set by the CFTC for the different expiration dates and commodities. If this is the case and one trader exceeds one of the limits, the reporting firm has to report all positions of the trader regarding the commodity to the CFTC. The biggest drawback of the data set is that traders are always assigned to a single class reflecting the purpose of the *majority* of their positions. However, it is often the case that traders hold positions for different reasons. Nevertheless, the COT reports should be able to give a conclusive picture of the future and underlying option markets, since the fraction of reported positions of the total open interest is very high, with on average .90 for the long positions and .85 for the short positions for the 10 commodities.

Overall, the CFTC releases five such reports every Friday categorizing traders in different classes and distinguishing between positions on future markets alone and combined reports also including trader's option positions. In the process of assigning the traders into categories, which they normally do by themselves, the CFTC reserves the right to change this self-classification. The combined reports recalculate option positions to replicate equivalent positions in the underlying future market using delta-factors supplied by the

exchanges.³

The legacy report solely distinguishing between commercial (COM) and non-commercial (NC) trader is available since 1986 for some commodities and the respective report for futures-and-options combined is available since 1995. More recently, the CFTC started to file new disaggregated reports for futures-only and futures-and-options combined, respectively, on September the 4th 2009. In these reports the broad classes of COM and NC are split up into two categories to increase the level of transparency. Furthermore, the CFTC used their historical data (but recent trader classifications) to back-cast the reports until the 13th of June 2006. As a fifth report the CFTC publishes the commodity index trader supplement since January 2007 and used again historical data to back-cast the supplement until the beginning of 2006. It files the new trader group 'commodity index trader' (CIT), while it additionally states the positions for 'COM non CIT' and 'NC non CIT'. The supplement is only available as a futures-and-options combined report and since this format also contains more information we restrict our analysis to the combined reports in the following. Moreover, to make the results comparable among different reports we use for all categories 365 weekly observations, from the 13th of June 2006 until the 4th of June 2013.

The 10 commodities we are investigating are listed in all three combined reports and consist of 8 agricultural commodities, namely corn, soybean, soybean oil, wheat traded at the Chicago Board of Trade (CBT), wheat traded at the Kansas City Board of Trade (KCBT), cocoa, coffee and sugar plus two live-stock commodities with feeder cattle and live cattle. All are listed in the two most important commodity indices Standard and Poor's Goldman Sachs Commodity Index (S&P-GSCI) and Dow Jones-UBS Commodity Index (DJ-UBSCI), which nowadays serve as benchmarks. Therefore, we use the S&P-GSCI spot price index for the single commodities, which tracks the price of nearby future contracts as price. In 6.1 a brief summary about these commodities is provided, where the future exchange, the sector and the share of CIT are listed. The average share of CIT is with .272 quite high and varies with a standard deviation of .070 substantially between the commodities.

On future markets a buyer of a future contract is called being long, while a seller of a future contract is called being short. Thus, the holder of the long position promises to buy from the holder of the short position the respective commodity at the pre-agreed price. Therefore, the value of a long contract increases with the price of the commodity, while the value of a short contract decreases if the price of the commodity increases. On commodity future ex-

³ The delta factor measures the sensitivity of an option in respect to a price change of the underlying in our case the commodity future.

| Commodity | Future Exchange | Contract Size | Sector | CIT share of long positions | Share of CIT listed as NC (long pos.) |
|-----------------|-----------------|----------------|--------|-----------------------------|---------------------------------------|
| Cocoa | CSC | 10 METRIC TONS | Agric. | .158 (.043) | .321 (.107) |
| Coffee 'C' | CSC | 37,500 POUNDS | Agric. | .259 (.047) | .161 (.060) |
| Corn | CBT | 1000 BUSHELS | Agric. | .248 (.039) | .174 (.056) |
| Soybeans | CBT | 5,000 BUSHELS | Agric. | .251 (.039) | .169 (.062) |
| Soybean oil | CBT | 60,000 POUNDS | Agric. | .260 (.042) | .120 (.058) |
| Sugar | CSC | 112,000 POUNDS | Agric. | .292 (.047) | .200 (.064) |
| Wheat (Chicago) | CBT | 1000 BUSHELS | Agric. | .415 (.048) | .158 (.047) |
| Wheat (Kansas) | KCBT | 1000 BUSHELS | Agric. | .251 (.060) | .268 (.110) |
| Feeder cattle | CME | 50,000 POUNDS | Lives. | .233 (.053) | .434 (.079) |
| Live cattle | CME | 40,000 POUNDS | Lives. | .355 (.055) | .194 (.035) |

Tab. 6.1: In the 6th columns are the name of the commodities, the exchange the investigated futures are listed, the contract size of the future, to which sector the commodity belongs, the share of long positions CIT are holding on average of all positions (including non reported trader) and the share of CIT long positions listed as NC in the Legacy Report. The brackets in the last two columns indicate standard deviations.

changes there is no need to actually purchase the physical underlying at the expiration of a long position, since all the futures are cash-settled with the exchange. However, the exchange is not actually engaged in the trading process but only ensures the final settlement. Therefore, the number of long and short contracts is always the same and theoretically unlimited. In the following analysis the short positions are subtracted from the long positions to receive the net-long positions for each class. They should reflect the expected price change, because a positive net-long position indicates that the respective class would profit from a price increase. Prior to the empirical analysis we have a closer look at the individual classes of trading participants to investigate their basic interdependencies and identify the candidates associated with speculative activity.

6.3 Trader Groups

The nine trader categories filed in the three combined reports and the fraction of the total average long and short position relations are shown in Table 6.2 with their average long and short position as fraction of the total interest.⁴ The commercial trader (COM) in the legacy report consist of all traders using the future market primarily for hedging purposes, while the

⁴ Note that only .90 of the long positions and .85 of the short positions regarding the total open interest have to be reported and are therefore included in each report (row).

| | | | | |
|-------------------|---|--|---------------------------------------|---|
| Legacy Report | Commercial (COM) (long=.440 short=.500) | Commercial (COM) Commercial (COM) | Non-Commercial (NC) (.454 .340) | |
| Aggregated Report | Producer, Merchant, Processor, User (Prod) (.204 .441) | Swap-Dealer (Swap) (.235 .059) | Money Manager (MM) (.267 .161) | Other Reportables (Others) (.190 .181) |
| Supplement | Commercial (SCOM) (.226 .476) | Commodity Index Trader (CIT) (.274 .026) | Non-Commercial (SNC) (.394 .338) | |

Tab. 6.2: Trader classifications of the CFTC. The classes are distinguished for the legacy report, the aggregated report and the supplement. In brackets are the average fraction of long and short positions of the total open interest and since non reported are not included the shares do not add up to unity.

non-commercial trader consists of all remaining traders which have to be reported.

In the disaggregated report the ‘Producer/Merchant/Processor/User’ (Prod) class incorporates all traders whose main business involves the physical commodity and uses the future market to hedge the risk associated with this activity. This includes farmers producing the commodity but also firms processing, packing or handling it in different ways. The ‘Swap Dealers’ (Swaps) are classified in the Legacy report as COM entity, but are not engaged in any activity directly related to the physical commodity. They predominantly manage their risk position resulting from swap agreements with commodity futures. In the first place the goal of the disaggregated report was to disentangle the Swaps from the Prods in order to increase the transparency. The ‘Money Managers’ (MM) are specialized entities like commodity pool operators, but also unregistered funds which are engaged in the future markets on behalf of their clients. The fourth class in the disaggregated report, ‘Other Reportables’ (Others) works, similarly to the NC class in the legacy report, as residual and consists of all traders not belonging to the MM class while being listed as NC in the legacy report. The classes COM and NC are thereby exactly split up into their two subordinate categories. Another approach was undertaken in the Supplement, where the new created class of ‘Commodity Index Trader’ (CIT) then consists of traders formerly listed as COM as well as NC. With an average share of 78% the COM constitute the majority. The two remaining classes in the Supplement are the basic categories of COM and NC minus the CIT, which will be referred to as SCOM and SNC, respectively.

To better understand the interdependencies of the net-long positions of the different trader groups and prices we look at the correlation among them displayed in Table 6.3. The most striking result is the high number of values

| | Prices | CIT | SNC | SCOM | MM | Others | Prod | Swap | NC | COM |
|--------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-----|
| Prices | 1 | | | | | | | | | |
| CIT | .1280 (.2188) | 1 | | | | | | | | |
| SNC | .2215 (.1966) | .1188 (.2932) | 1 | | | | | | | |
| SCOM | -.2055 (.2184) | -.4901 (.1406) | -.8753 (.1157) | 1 | | | | | | |
| MM | .3240 (.1922) | .1945 (.2501) | .9631 (.0172) | -.8770 (.0991) | 1 | | | | | |
| Others | -.0119 (.3423) | .0896 (.1841) | .1497 (.2735) | -.1650 (.2918) | -.0068 (.2708) | 1 | | | | |
| Prod | -.1656 (.2405) | -.4993 (.1399) | -.8510 (.1751) | .9818 (.0225) | -.8467 (.1522) | .1495 (.2949) | 1 | | | |
| Swap | -.0697 (.3055) | .8058 (.0481) | .1172 (.2949) | -.4124 (.1502) | .1186 (.2717) | .01998 (.2639) | -.5023 (.1326) | 1 | | |
| NC | .3053 (.2071) | .1975 (.2776) | .9786 (.0128) | -.8918 (.1092) | .9791 (.0114) | .1818 (.2629) | -.8568 (.1688) | .1096 (.3037) | 1 | |
| COM | -.2822 (.2328) | -.1363 (.2081) | -.9500 (.0429) | 0.8980 (.0992) | -.9450 (.0631) | -.1891 (.2665) | .8597 (.1423) | -.0478 (.2369) | -.9690 (.0481) | 1 |

Tab. 6.3: Average correlations and standard deviations (in brackets) between the net-long positions of the trader classes with each other and the S&P GSCI Spot Price Indexes for the 10 commodities.

above $|.8|$, though the reasons for this are very disperse. The high negative correlation between COM and NC from the legacy report might have been expected, since these are the only two classes considered and we know that on average 90% (85%) of the long (short) positions have to be reported. Therefore, it is very likely that not only these two groups are often engaged in mutual contracts, but that a large fraction of their net-long positions is the result of such trades.⁵

Interestingly, the subordinate classes of the COM in the disaggregated report are very diverse correlated with their superior class and with each other, showing that the split increased the transparency by revealing two distinct trading strategies. As expected, the Prods are highly positively correlated (.860), while the Swaps are almost uncorrelated ($-.048$). The reason for this is that the changes in the net-long positions from one week to the other are much higher in absolute terms in case of the Prods, and so they dominate the changes of the Swaps when they are jointly observed in the COM class. Furthermore, we observe that the Swaps are highly correlated with the CIT and looking at the average positions in Table 6.2 are almost long only as well. Hence, it is likely that most of the COM traders included in the CIT are listed as Swap in the disaggregated report. The SCOM are very highly correlated with the Prods which again indicates that the commercial trader categorized as Swaps in the disaggregated report are likely to be listed

⁵ Note that it is possible that two traders out of the same class trade with each other, yet these trades have no influence on the net-long position.

as CIT in the Supplement.

On the ‘non commercial’ side of the reports we see again that the two subordinate classes in the disaggregated report follow different trading strategies, since the MM are almost perfectly correlated (.979) with the NC, while the Others are only weakly correlated (.182). The dominance of the MM in respect to the superior class NC is again due to the larger absolute changes of the net-long position. The SNC are almost perfectly correlated with both the MM (.982) and the NC (.979). This is not surprising regarding the smaller fraction (22%) of CIT formerly listed as NC. This is also supported by the low correlations of the CIT with the MM (.195) and the Others (.090).

In total we observe three distinct groups of classes being strongly correlated. The Prod, the SCOM and the COM being short on average and using the future market primarily for hedging purposes and not for speculation. The CIT and Swap follow a predominantly ‘long only’ long-term investment strategy and are represented by the CIT in the remaining part. The MM, NC and SNC trade in the market to generate speculative profits (being long on average) and are represented by the MM in the empirical analysis, because of the most precise definition in this respect. Since the Others are the residual class and they also have an almost balanced position, they are left out of the empirical analysis. Therefore, the investigation is focused on the CIT and the MM as possible candidates for speculative activity.

Looking at the correlation of the net-long positions with the associated prices, it is on average much higher for the MM (.321) compared to the CIT (.128). However, the high standard deviation for both values suggests that the correlations are very different for the single commodities. To illustrate this, Figure 6.1 shows in the first row the S&P GSCI Price index for soybeans in the left panel and the S&P GSCI Price index for cocoa in the right panel, with the corresponding net-long position for MM. In the second row the same prices are plotted now including the CIT net-long position. The net-long position for MM is highly correlated with the price for soybeans (.701), while even slightly negatively correlated with the price of cocoa (−.105).

The net-long positions for CIT on the other hand are more highly correlated with the price of cocoa (.466) than with soybeans (.234). Additionally, we observe that the net-long position of MM sometimes becomes negative, while the strategy of CIT is always predominantly long and, hence, would imply an expected price increase. Regarding the CIT, the simultaneous drop in both positions in 2008 at the beginning of the global financial crises might be the most prominent feature. The overall correlation between the CIT net-long positions among all commodities is with .394 relatively low, since CIT are generally considered to buy/sell future contracts in a fixed ratio for all commodities in their portfolio, while these ratios are rarely adjusted.

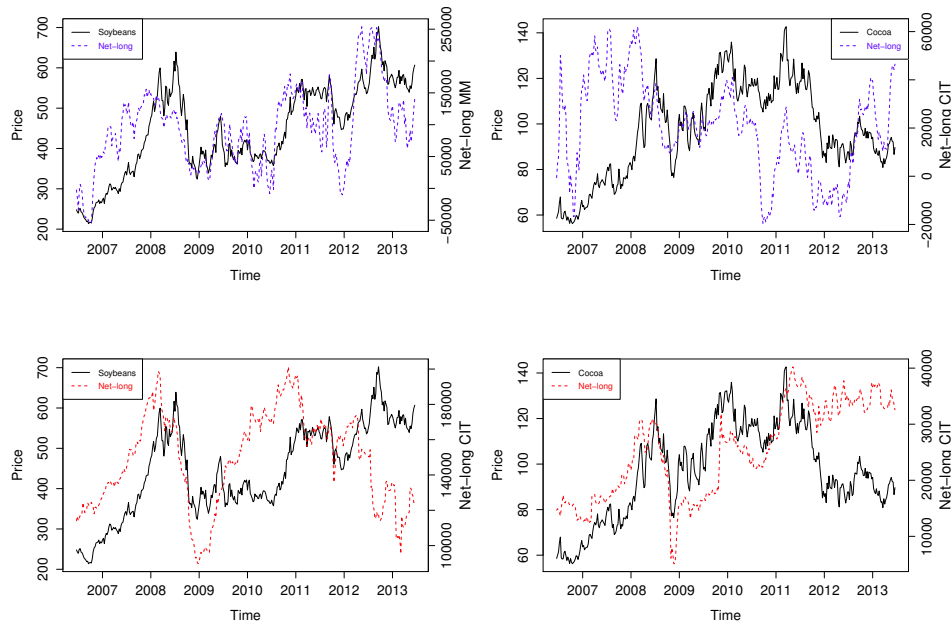


Fig. 6.1: Plotted Time series of the S&P GSCI spot price index (solid line) for soybeans (left) and cocoa (right) and the corresponding net-long positions (dashed line) of MM (top) and CIT (bottom).

After considering the basic setup of future markets and the general investment strategies of the different classes, we analyse the market behaviour of CIT and MM by investigating whether their trading activities do Granger cause prices or vice versa.

6.4 Empirical Analysis

In testing for Granger non-causality between the prices and the net-long positions we follow the ‘surplus lag’ procedure introduced by Toda and Yamamoto (1995) using (bivariate) Vector autoregressions (VARs) in levels and adding additional lags according to the maximum order of integration of the two time series. The application of this procedure instead of an analysis using a Vector Error Correction Model (VECM) is necessary, since the net-long positions and the prices are not always integrated of the same order. The implementation of the additional lag(s) ensures that the Wald test statistics are asymptotically chi-square distributed under the null. As a first

| Variable | Test | Cocoa | Coffee | Sugar | Corn | Wheat C. |
|--------------|-------------------|------------|------------|------------|-------------|------------|
| Price | ADF in levels | -2.099 | -1.546 | -1.482 | -1.777 | -2.477 |
| | ADF in 1st diff. | -10.541*** | -3.954*** | -21.407*** | -19.026*** | -19.631*** |
| | KPSS in levels | .902*** | .1147*** | 1.475*** | 1.482*** | .321 |
| | KPSS in 1st diff. | .237 | .1917 | .096 | .061 | .080 |
| CIT | ADF in levels | -1.863 | -2.49 | -1.828 | -3.488*** | -1.581 |
| | ADF in 1st diff. | -15.479*** | -14.440*** | -5.625*** | -4.150*** | -16.439*** |
| | KPSS in levels | 1.698*** | .123 | .220 | .155 | .242 |
| | KPSS in 1st diff. | .038 | .099 | .158 | .092 | .113 |
| MM | ADF in levels | -2.806** | -2.436 | -1.858 | -3.449*** | -3.213** |
| | ADF in 1st diff. | -12.075*** | -10.821** | -7.296*** | -5.919*** | -13.854*** |
| | KPSS in levels | .789*** | .400* | .439* | .195 | .372* |
| | KPSS in 1st diff. | .069 | .162 | .190 | .041 | .022 |
| Variable | Test | Wheat K. | Soybeans | Soyb. oil | Feeder cat. | Live cat. |
| Price | ADF in levels | -2.192 | -1.717 | -2.039 | -.923 | -1.262 |
| | ADF in 1st diff. | -12.444*** | -18.958*** | -18.388*** | -9.279*** | -9.899*** |
| | KPSS in levels | .375* | 1.419*** | .918*** | 1.650*** | 1.676*** |
| | KPSS in 1st diff. | .072 | .053 | .115 | .109 | .062 |
| CIT | ADF in levels | -1.215 | -2.523 | -1.830 | -2.6701* | -2.560 |
| | ADF in 1st diff. | -16.757*** | -4.962*** | -15.604*** | -17.002*** | -4.152*** |
| | KPSS in levels | 1.169*** | .225 | 1.061*** | .171 | .371* |
| | KPSS in 1st diff. | .085 | .176 | .063 | .095 | .238 |
| MM | ADF in levels | -2.7351* | -3.190** | -2.403 | -3.020** | -2.300 |
| | ADF in 1st diff. | -13.303*** | -15.120*** | -8.011*** | -12.198*** | -14.377*** |
| | KPSS in levels | .155 | .5673** | .623** | .275 | .598** |
| | KPSS in 1st diff. | .047 | .036 | .033 | .069 | .0778 |

Tab. 6.4: The test statistics for the Augmented Dickey Fuller (ADF) and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test in levels and first differences. The time series investigated are the S&P GSCI spot price index and the net-long positions of CIT and MM for the 10 commodities. ***, ** and * indicate statistical significance at .01%, .05% and .10% level respectively.

step the Augmented Dickey-Fuller Test (ADF) and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test are applied to the time series in levels and first differences to determine the maximum order of integration. The ADF test has non-stationarity as the null hypothesis, while the KPSS test has stationarity as the null hypothesis. The lag-lengths for the ADF regressions were chosen using the Akaike information criterion. The results for all time series under investigation are shown in Table 6.4 and can be summarized as follows.

In case of the price series the ADF test always signals that the process is integrated of order one or $I(1)$, because it never rejects the null having a unit root in levels and always rejects the null at the 1% significance level in first differences. The KPSS test rejects for 8 out of 10 commodities the null of stationarity at the 1% level and for wheat KBT at the 10% level, yet for wheat CBT it does not reject the null. Therefore, the wheat CBT price would

be considered as I(0) according to the KPSS test and I(1) according to the ADF test. In the following we treat it as an I(1) process, yet the main result is robust to the choice of the order of integration. In first differences the test never rejects the null, so that it is fair to conclude in combination with the results from the ADF test that none of the price series should be treated as an I(2) process. For the net-long positions of both classes the results are only robust in the sense that none of the time series are integrated of order two or higher and hence the number of additional lags m which have to be included is always 1. However, the results of the unit root test regarding whether the time series are I(0) or I(1) are not clear cut. The next step is to set up a bivariate VAR,

$$P_t = c_1 + a_1 P_{t-1} + \dots + a_p P_{t-p} + b_1 N_{t-1} + \dots + b_p N_{t-p} + u_t \quad (6.1)$$

$$N_t = c_2 + d_1 N_{t-1} + \dots + d_p N_{t-p} + e_1 P_{t-1} + \dots + e_p P_{t-p} + v_t \quad (6.2)$$

where t indexes time, c_1 and c_2 are constants, u_t and v_t are white noise, P_t denotes the price series, N_t is the net-long position of the respective trader class and p is the number of lags. To determine the optimal lag length p we rely on the minimum of four standard criteria, namely the Akaike information criterion, the Schwarz information criterion, the Hannan-Quinn information criterion and the final prediction error. In case of conflicting results we estimate several VAR(p) with all proposed lag length and chose the number of lags by looking at the results of the Breusch-Godfrey LM test for remaining autocorrelation.⁶ The next step is to include the corresponding m additional lags derived from the maximum order of integration analysis in our preferred VAR(p) models and estimate the new VAR($p + m$) system

$$P_t = c_1 + \sum_{i=1}^p a_i P_{t-i} + \sum_{j=1}^m a_{p+j} P_{t-p-j} + \sum_{i=1}^p b_i N_{t-i} + \sum_{j=1}^m b_{p+j} N_{t-p-j} + u_t \quad (6.3)$$

$$N_t = c_2 + \sum_{i=1}^p d_i N_{t-i} + \sum_{j=1}^m d_{p+j} N_{t-p-j} + \sum_{i=1}^p e_i P_{t-i} + \sum_{j=1}^m e_{p+j} P_{t-p-j} + v_t \quad (6.4)$$

using OLS.⁷ In this model the null hypothesis $b_1(e_1) = \dots = b_p(e_1)$ are jointly zero implies that N(P) does not Granger cause P(N). The inclusion of the additional m lags (omitted in the test) makes the Wald test statistic asymptotically χ^2 distributed with p degrees of freedom, but not efficient. However,

⁶ The results for the lag length criteria and the Breusch-Godfrey LM test for the estimated VAR(p) are available from the corresponding author upon request.

⁷ The complete estimation results are available from the corresponding author upon request.

since our sample size is reasonably large and only one lag is added in all scenarios the loss is relatively small. The results for the Granger non-causality tests together with the selected lag lengths are shown in Table 6.5.

In case of the net-long positions of CIT the null of no Granger causality cannot be rejected in both directions. Meaning that neither the lagged values of N_t help us to predict the value of P_t nor the other way around. Only for sugar, feeder cattle and live cattle the null that the price does not Granger cause the net-long position of CIT is rejected at the 10% significance level. On the other hand solely the net-long position in the corn future market rejects the null on the 5% level indicating that CIT could cause price movements.

The most robust result is that the net-long positions of MM seem to be 7 out of 10 times Granger caused by the prices, because the null of no Granger causality is rejected for wheat CBT at the 10% level, for corn and soybean oil at the 5% level and for cocoa at the 1% level. The commodities cocoa and live cattle are special in the sense that the test for Granger non causality is also rejected at the 5% respectively 1% level in the other direction, meaning that the net-long position additionally seem to Granger cause the price process. Note that the net-long position of the Prods (not stated here) feature qualitatively similar results regarding the test for Granger non-causality, except for corn, soybean oil and wheat KBT where the null is not rejected in both directions. However, the coefficients of the lagged terms mostly have the opposing sign.

| Commodity | Cocoa | Coffee | Sugar | Corn | Wheat C. |
|---|---------------------|-----------------|-------------------|---------------------|---------------------|
| Dependent variable: CIT, Price | | | | | |
| lags | 5 | 2 | 3 | 2 | 2 |
| H0: Does CIT not Granger cause the price | 2.633 (.756) | .884 (.643) | 3.619 (.306) | 6.161** (.046) | 0.973 (.615) |
| H0: Does the price not Granger cause CIT | 6.028 (.304) | .756 (.685) | 6.847* (.077) | 4.316 (.116) | 3.838 (.147) |
| Dependent variable: MM, Price | | | | | |
| lags | 2 | 6 | 2 | 2 | 2 |
| H0: Does MM not Granger cause the price | 6.456** (.040) | 3.655 (.723) | .014 (.993) | .432 (.806) | .980 (.613) |
| H0: Does the price not Granger cause MM | 16.635*** (.000) | 7.931 (.243) | 2.639 (.267) | 6.798** (.033) | 5.863* (.053) |
| Commodity | Wheat K. | Soybean | Soyb. oil | Feeder cat. | Live cat. |
| Dependent variable: CIT, Price | | | | | |
| lags | 1 | 2 | 2 | 4 | 4 |
| H0: Does CIT not Granger cause the price | 0.612 (.434) | 0.079 (.961) | 3.978 (.137) | .725 (.948) | 3.905 (.419) |
| H0: Does the price not Granger cause CIT | .251 (.617) | 4.486 (.106) | .749 (.688) | 8.390* (.078) | 8.501* (.075) |
| Dependent variable: MM, Price | | | | | |
| lags | 2 | 2 | 2 | 4 | 4 |
| H0: Does MM not Granger cause the price | .296 (.863) | .270 (.874) | .034 (.983) | 2.947 (.567) | 16.903*** (.002) |
| H0: Does the price not Granger cause MM | 10.116*** (.006) | .340 (.844) | 6.753** (.034) | 41.663*** (.001) | 16.636*** (.002) |

Tab. 6.5: The optimal lag length indicated by the Akaike information criterion, the Schwarz information criterion, the Hannan-Quinn information criterion and the final prediction error for the bivariate VAR's without including the extra lag. The number of lags is also the number of d.o.f. in the Wald test-statistic, when testing for Granger non causality between the net-long positions of the CIT or MM and the S&P GSCI spot price index. ***, ** and * indicate statistical significance at .01, .05 and .10 level respectively. P-values are presented in parentheses.

6.5 *Conclusion*

The question whether speculation has a negative impact on commodity markets by e.g. increasing the volatility or causing price bubbles has many different aspects which all need to be answered to give a conclusive statement. First of all it is necessary to define what speculation is or who the speculators are. A seemingly adequate answer regarding the investment horizon is that short-term strategies are more speculative than long-term strategies and in this sense commodity index trader should not be considered as speculators at all. Then again, the effects of trading and not their characteristics should be of special interest, since every trader engaged on future markets has speculative incentives to do so.

Our analysis of Granger causal effects of trading activity of specific classes on the price indicates that no such relationship exists. However, money manager and producer seem to adjust their positions in respect to the prices. Additionally their net-long positions are highly negatively correlated, supporting the fact that money manager provide the hedging possibilities for the producer, while looking for speculative profits at the same time. The commodity index traders seem to trade largely unaffected by prices or other traders following the strategy of reducing their portfolio risk. This might even reduce the price for hedgers by increasing the demand for short contracts in normal times, but add to the volatility in times of distress when they withdraw the majority of their funds (as observed in 2008). In doing so they might even create the ‘missing link’ (low correlation of returns) to other financial markets they try to exploit. Hence, there are still many open topics for future research, for example the investigation of intertemporal causality. Using data with an even higher frequency than the weekly data this study is based upon might represent another promising avenue for future research when investigating the effects of trader positions.

7. SPECULATIVE ACTIVITY ON FUTURES MARKETS
FOR ENERGY COMMODITIES AND PRECIOUS METALS

7.1 Introduction

Two of the most important commodities for the world economy - gold and crude oil - have both undergone tremendous price spikes during the last years, but not all of them at the same time. The oil price peaked in July 2008 with more than 140\$ per barrel, while the gold price reached its all time high above 1900\$ per troy ounce in September 2011. However, both developments are closely related to the global financial crisis (GFC). The oil price rose until the verge of the GFC and plummeted afterwards, as the trust in financial markets and economic prospects started to deteriorate. On the other side the gold price (after a swift decline) profited from this uncertainty and gave a comeback as an asset of last resort during the most turbulent times. Although the changes are at least partially backed by fundamental factors the ongoing financilization of commodity markets during this period raises the question to which extent excessive speculation may have contributed to these price trends.

The academic world agrees on the perception that significant changes happened due to the financial integration, but not on its qualitative ramifications.¹ Among others, Alquist and Kilian (2010) and Büyükşahin and Harris (2011) find that since 2003 ever more financial investors have entered commodity markets. According to Gorton and Rouwenhorst (2006), Erb and Harvey (2006) and Conover *et al.* (2010) one of their incentives might have been to reduce their portfolio risk by including commodities (as a hedge). Tang and Xiong (2010) show that as a result the co-movement of the oil price and prices for other commodities increased, while Büyükşahin and Robe (2014) find the same effect for the returns of equity indices and commodities. The further impact of speculation on commodity markets is highly debated. On the one hand, Büyükşahin and Harris (2011), Fattouh *et al.* (2013) and Knittel and Pindyck (2013) conclude that speculation played (if any) only a minor role in driving up the price for crude oil. On the other hand Kaufmann (2011) and Singleton (2013) find that speculation had a significant impact. Baur and Glover (2014) state that the surge in the gold price was at least partially caused by trend following speculative activity. Moreover, Zhang, Zhao, Cai, Liu, and Zhou (2010) find that there is a cointegration relation between the price for crude oil and gold and that the volatility of the price for gold is linearly Granger caused by changes of the crude oil price.

This study uses the approach of Finger *et al.* (2013), which is similar to Sanders *et al.* (2004), to investigate the impact of speculation. The expected

¹ See Fattouh, Kilian, and Mahadeva (2013) for an extensive survey of this strand of literature.

price change is measured as the long position minus the short position (or net-long) of the respective trader group. The data about the trader position is from the weekly Commitments of Traders (COT) report of the Commodity Futures Trading Commission (CFTC). The investigated period spans from June 2006 until December 2013. The ‘surplus lag’ procedure of Toda and Yamamoto (1995) is applied to determine whether the net-long position (speculative activity) of trader groups Granger cause the price of the commodity or vice versa. The study incorporates two other energy commodities, namely heating oil and natural gas and silver and platinum as two additional precious metals to investigate whether speculators might have a stronger impact on the ‘little brothers’ of crude oil and gold. An OLS CUSUM test (Ploberger and Krämer, 1992) is applied to detect common structural breaks in the price series of the six commodities under investigation. Accordingly, the sample period is split into three subperiods.

The results support the CFTC in publishing an additional ‘disaggregated’ report, since the four trader groups (instead of two) clearly follow distinct trading strategies and this should increase the transparency of the data. Out of these groups Money Managers and Swap Dealers are identified as the two candidates for engaging in speculative activity. The interdependencies between prices and net-long positions are weak for the complete sample. If there is any relation whatsoever, the results indicate that the traders adjust their positions with respect to the price. However, Money Managers seem to have at least partially caused the price surge of crude oil before the global financial crisis (GFC) and the price increase for silver after the onset of the crisis. Hence, in line with the literature, I find some evidence for speculation having an impact on the prices of energy commodities and precious metals.

The rest of the paper is organized as follows. In section 2 the data set used and the commodities under investigation are introduced. Thereafter the relations among the different trader groups are analyzed. Section 4 consists of the tests and the empirical analysis for Granger non-causality, while section 5 concludes.

7.2 *Data*

The CFTC publishes positions for various futures and options on these futures In the COT reports every Friday. These reports rely on the Tuesday’s position from reporting firms about traders whose holdings exceed thresholds set by the CFTC. A trader passing one of these limits is obligated to hand in all his positions regarding the respective commodity. For the six commodities we are investigating the reportable positions account for 88.5 percent of all

long positions and 93.9 percent of the short positions. Therefore, the reports should give a representative picture of the overall market activity.

In total, the CFTC issues five reports which differ with respect to the trader classes and to the considerations of positions on futures-only or futures-and-options combined. In the latter reports the corresponding delta-factors (supplied by the exchanges) are used to transform options into equivalent futures positions.² The reportable traders usually assign themselves into a trader class for each report, while the CFTC is enabled to change this classification. However, the classification of all positions of one trader into a single group might often be an oversimplification, e.g. a bank trades on its own account and in the name of its customers at the same time, reducing the informational value of the data set.

The so called ‘legacy’ report distinguishes between commercial (COM) and non-commercial (NC) traders and has been made available since 1986 for futures-only and since 1995 for futures-and-options combined. In order to increase the transparency the CFTC started to file new ‘disaggregated’ reports for futures-only and futures-and-options combined, respectively, on September the 4th 2009. To achieve this, each superordinate group of COM and NC is split up into two categories. Furthermore, the CFTC used historical data regarding the positions and recent trader classifications to extend the time span covered by the data to the 13th of June 2006. The CFTC additionally publishes a ‘supplemental’ report since January 2007 focusing on the new group of ‘commodity index trader’(CIT), which is especially prominent in the debate regarding price spikes of agricultural commodities. However, it does not cover any of the commodities investigated in this study and therefore the influence of CIT cannot be investigated. In the following only the futures-and-options combined format of the ‘legacy’ report and the ‘disaggregated’ report are considered, since they should incorporate more information than the futures-only reports. The sample period in this study consists of the 365 weekly observations gathered from the 13th of June 2006 until the 31st of December 2013.

Each of the six commodities is listed in the ‘legacy’ report and the ‘disaggregated’ report and consists of 3 energy commodities crude oil (WTI), heating oil (No. 2, N.Y. harbor) and natural gas plus three precious metals including gold, platinum and silver. All but platinum are part of the two most important commodity indices, namely the Standard and Poor’s Goldman Sachs Commodity Index (*S&P – GSCI*) and the Dow Jones-UBS Commodity Index (DJ-UBSCI), which can be considered as benchmarks.

² The delta-factor measures the responsiveness of the price of an option with respect to a price change of the underlying asset.

| Commodity | Futures Exchange | Contract Size | Sector | Share of reported long positions | Share of reported short positions |
|-------------|------------------|---------------------|-------------|----------------------------------|-----------------------------------|
| crude oil | NYME | 1,000 barrels | energy | .964 (.006) | .966 (.006) |
| heating oil | NYME | 42,000 U.S. gallons | energy | .869 (.019) | .902 (.023) |
| natural gas | NYME | 10,000 mmBtu's | energy | .930 (.016) | .961 (.009) |
| gold | CMX | 100 troy ounces | prec. metal | .894 (.021) | .951 (.010) |
| platinum | NYME | 50 troy ounces | prec. metal | .841 (.056) | .938 (.032) |
| silver | CMX | 5,000 troy ounces | prec. metal | .814 (.036) | .913 (.014) |

Tab. 7.1: In the columns presented are the name of the commodities, the exchange on which the investigated futures are listed, the contract size, to which sector the commodity belongs, the average share of reported long positions and the average share of reported short positions. The brackets in the last two columns indicate standard deviations.

Moreover, gold (with a target weight of 11.53%), natural gas (9.45%) and WTI crude oil (8.49%) are the biggest contributors to the DJ-UBSCI. In the following the *S&P – GSCI* spot price index for single commodities is used to investigate the influence of different trader groups on the price level. Table 7.1 summarizes basic information regarding the commodities, like the respective futures exchange, the contract size, the sector, and the average share of reportable long and short positions.

Before analyzing the interdependencies between the trader groups in more detail, a brief introduction of the functioning of commodity futures markets in general is given. The buyer of a futures contract is called being long and promises to buy the fixed contract size (or multiple of it) of the respective commodity from the seller of the futures contract (being short) at the expiration date for the price agreed upon at the beginning of the contract. Hence, if the price of the underlying commodity increases, so does the value of the long contract, while the opposite is true for the short contract. All the futures investigated are cash-settled with the respective exchange and hence, there is no need to actually purchase or deliver the physical commodity at the expiration of the contract. This facilitates the trading process especially for traders looking for speculative profits, which shall ensure all the demand for hedging to be met. In this process the exchange only ensures the final settlement, but does not hold contracts by itself. Hence, a new futures contract is only created if two traders agree upon a price and expiration date which is why the number of long and short contracts has to be the same. Hence, in theory, the futures market constitutes an unlimited (in size) zero sum game for the participants. However, single traders or trader groups might realize profits, while others suffer a loss. In the following the net-long positions, meaning long positions minus short positions are calculated for each group

| Legacy Report | Commercial (COM) (long=.378 short=.589) | | Non-Commercial (NC) (.508 .350) | |
|-------------------|---|-------------------------------------|--------------------------------------|--|
| Aggregated Report | Producer, Merchant, Processor, User (Prod) (.142 .337) | Swap-Dealer (Swap) (.235 .251) | Money Manager (MM) (.314 .181) | Other Reportables (Other) (.193 .169) |

Tab. 7.2: Trader classifications of the CFTC. The classes are distinguished for the ‘legacy’ report and the ‘disaggregated’ report. In brackets are the average fraction of long and short positions of the total open interest. Since non reported positions are not included the shares do not add up to unity.

and commodity. Since only these contracts are creating profits or losses (the others cancel out), they should represent the expected price change of the respective group. For instance, a negative net-long position shows that the entire trader group will profit if the price is decreasing.

7.3 Trader Groups

The seven trader groups filed in the futures-and-options combined of the ‘legacy’ report and the ‘disaggregated’ report, as well as the fraction of long positions and short positions that they on average hold are shown in Table 7.2. The commercial traders (COM) in the ‘legacy’ report are said to use the futures market mainly for hedging purposes, while the non-commercial traders (NC) are the residual class of all remaining reportable traders. This very simplistic classification of whether a trader uses the market primarily for hedging or is likely to put heterogeneous trader into one group.

Therefore, the aim of the ‘disaggregated’ report is to increase the transparency by defining more precise trading purposes for various trader groups. The main recommendation by the CFTC staff report (2008) is to at least include an additional group of swap dealers, who may use the market for hedging purposes, but are not associated with the physical commodity. To achieve this, the group of the COM is split into ‘Swap Dealers’ (Swap) and ‘Producer/Merchant/Processor/User’ (Prod). The latter group consists of all reportable traders actually concerned with the underlying physical commodity. Therefore these trader should use futures markets to hedge the risk of their primary business activity. On the other side the Swap hedge the risk associated with swap agreements on commodity futures. Mining companies (in case of precious metals) but also hedge funds might serve as counterparties to these agreements. Hence, both groups are more homogeneous than their superordinate group, but whether they might be considered as speculators

| | Prices | NC | COM | Prod | Swap | MM | Other |
|--------|-------------------|-------------------|-------------------|-------------------|-------------------|------------------|-------|
| Prices | 1 | | | | | | |
| NC | .2152 (.2548) | 1 | | | | | |
| COM | -.2277 (.2471) | -.9830 (.0148) | 1 | | | | |
| Prod | -.0484 (.2839) | -.5021 (.6109) | .5017 (.5845) | 1 | | | |
| Swap | -.2288 (.3000) | -.6336 (.4213) | .6609 (.4032) | -.1156 (.6174) | 1 | | |
| MM | .2621 (.3463) | .9050 (.0982) | -.9085 (.0799) | -.5705 (.4536) | -.5731 (.3325) | 1 | |
| Other | .0275 (.3747) | .4819 (.2404) | -.4544 (.2658) | -.0647 (.6411) | -.3779 (.3505) | .1226 (.4404) | 1 |

Tab. 7.3: Average correlations and standard deviations (in brackets) between the net-long positions of the individual trader classes and the S&P GSCI Spot Price Indexes for the 6 commodities.

or not depends on their unobserved counterparties in case of the Swap.

Also, the NC of the ‘legacy’ report are split into two subordinate groups. The ‘Money Manager’ (MM) for example are registered commodity trading advisors or unregistered funds trading with futures in the name of their clients. The last group ‘Other Reportables’ (Other) functions as a residual and includes all reportable traders which have been listed as NC in the ‘legacy’ report, but do not fit into the MM.

To get a better understanding of the relations between the different trader groups, the correlations among their net-long positions plus the S&P GSCI Spot Price Index are investigated. Table 7.3 shows the average correlation coefficients for the 6 commodities.

In terms of the ‘legacy’ report the results confirm the expectations. The COM are on average short, i.e. have a negative net-long position ($.378 - .589 = -.211$), and the NC are on average long (.158), while they are highly negatively correlated (correlation coefficient = $-.983$). Hence, the excess need of the COM to hedge their price risk is met by the NC providing this service. However, since the two groups account together for around 90 percent of the long and short positions, it is ‘quasi’ inevitable that trades between the COM and the NC are causing their opposing net-long positions.

The next step is to look at the interdependencies between the two groups

of the ‘legacy’ report and the subordinate groups in the ‘disaggregated’ report. In case of the NC, which act as a residual class in the ‘legacy’ report, the MM are similarly long on average (.133) as their superior class and also the evolution of the net-long positions is almost perfectly correlated (.905). In contrast to the residual class of NC the MM are a precise classification of financial funds acting on the commodity futures market. On top of this, the net-long position of the MM also have the strongest correlation with the price series of the respective commodity, making them the first candidate for destabilizing speculative activity. On the other side the ‘new’ residual class Other is also long on average (.025), but to a lesser degree and the correlation (.482) with the NC is much lower as well. Moreover, the Other do not have a cc above |.5| with any of these groups and are almost uncorrelated with the price series (.028) and the MM (.123). Hence, they only serve as a ‘new’ residual class for making the MM distinguishable, but will not be explicitly considered in the empirical analysis.

The Prod have a strongly negative net-long position (−.195) on average which accounts for the main part of the net-long position of the NC, but the correlation (.502) is relatively low with respect to this observation. In contrast the average net-long position of the Swap is only marginally negative (−.016). Yet the correlation (.661) with the superior class is stronger. Together with the slightly negative correlation (−.116) between the two trader groups, it is safe to conclude that the split created two distinct trading strategies, where the Prod seem to use futures markets to guard themselves against increasing prices. However, the standard deviations in Table 7.3 are very high for both groups and so the correlations seem to be very diverse for the single commodities. Therefore, as a next step the correlation coefficients and net-long positions for the subsamples of energy commodities and precious metals are considered.³ In case of the energy commodities the net-long position of the Swap is positive (.072) and the correlation with the NC is reduced (.549), yet on a comparable level. The Prod are on average short (−.102) but the correlation with the NC is very low (.115). The correlation between the Swap and Prod net-long position is even strongly negative (−.650). For precious metals the picture is very different. Both groups are on average short (Prod= −.288|Swap= −.104) and the correlation between all three groups is strongly positive. Hence, especially the Swap still seem to consist of very heterogeneous traders with respect to the specific commodity and are therefore the second candidate for speculative activity on futures markets. This is supported by Finger *et al.* (2013) analyzing ten agricultural and live stock commodities. For this set of commodities the Swap follow an almost

³ The complete Tables can be found in the Appendix.

long-only strategy with an average net-long position of $.235 - 059 = .176$ and seem to consist primarily of commodity index trader, which is supported by the high correlation (.806) between the two trader groups.

In summary, the four trader groups listed in the ‘disaggregated’ report follow distinct trading strategies, the differences between the sectors and single commodities still being significant. The MM are the prime candidate for engaging in potentially destabilizing speculative activity, while the Swap show a very inconsistent trading behavior with respect to the individual commodities and deserve a closer look as well. The average correlation with the corresponding price is much higher in the case of MM (.262) than in the case of the Swap ($-.229$) but the standard deviations show that they are very dispersed. Figure 7.1 illustrates this for crude oil and silver. The correlation among the price of crude oil and the respective net-long position of the MM (top left panel) is strongly positive (.612), but the price of silver (top right) evolves almost independently (.127). The respective net-long position of the Swap is also highly, yet negatively correlated ($-.516$) with the price of crude oil (bottom left). On the other hand the Swap are positively correlated (.234) with the price of silver (bottom right).

Moreover, at first glance the two price series seem to evolve very differently during the sample period. However, a closer look reveals that the general pattern is to a certain extent similar, which is supported by the relatively high correlation coefficient of .526. At the beginning, both price series increase until the onset of the GFC and even the (relative) price drop in respect to the crisis is of a comparable scale. Yet the magnitude of the following ‘recovery’ is substantially different. The price for silver (similar to the gold price) seems to be strongly driven upwards due to its safe haven property and accordingly, while the price for crude oil also quickly increases but never reaches its maximum values of the period before the crisis.

7.4 *Empirical Analysis*

In this study the ‘surplus lag’ procedure following Toda and Yamamoto (1995) is applied to test for Granger non-causality. According to Clarke and Mirza (2006) this procedure is preferable over the more common Vector Error Correction Model (VECM). The basic model setup consists of bivariate Vector autoregressions (VARs), where the two time series in levels are the price of the commodity and the net-long position of the trader group under investigation. The only extension to this framework is that additional (surplus) lags equal to the maximum order of integration have to be included, but are not taken into account during the Wald test for Granger non-causality. The

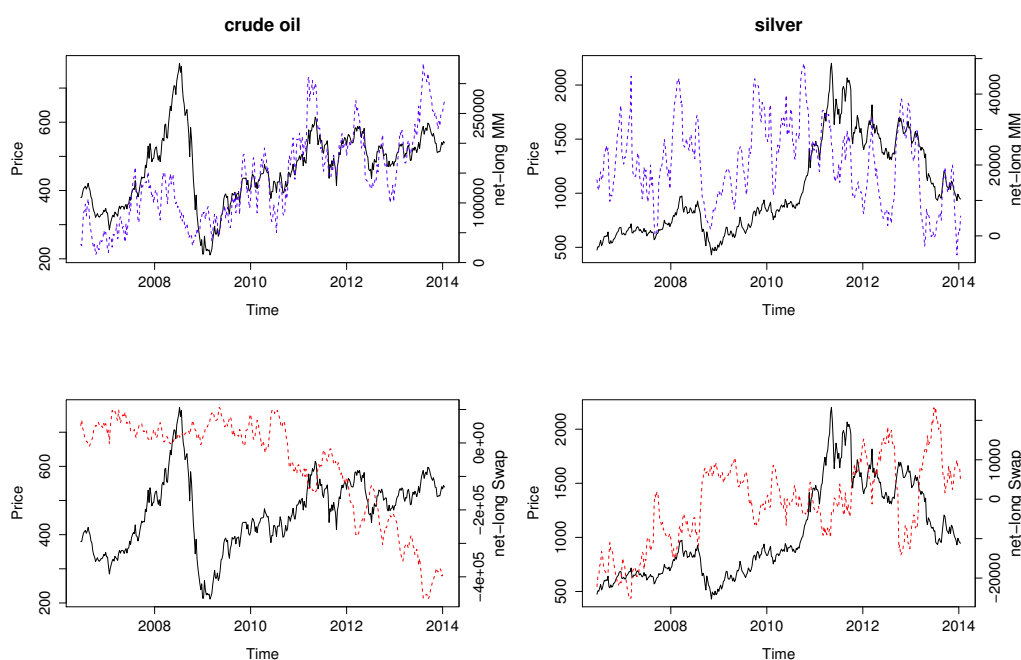


Fig. 7.1: Plotted Time series of the S&P GSCI spot price index (solid line) for crude oil (left) and silver (right) and the corresponding net-long positions (dashed line) of MM (top) and CIT (bottom).

implementation of the additional lag(s) results in asymptotically chi-square distributed Wald test statistics. This outcome is robust for different orders of integration among the time series, which is the case here, but would be problematic in a VECM setup.

Before applying the VAR model it is necessary to determine the order of the integration of all time series. For this task two standard tests for unit roots, namely the Augmented Dickey-Fuller test (ADF) and the Kwiatkowski-Phillips-Schmidt-Shin test (KPSS) are used for each time series in levels and first differences. Note that the null hypothesis for the ADF is non-stationarity and the Schwarz information criterion was used to chose the appropriate lag length. On the contrary, the null hypothesis for the KPSS test is stationarity.⁴

In Table 7.4 all results for the prices and the net-long positions of the MM and the Swap are summarized.⁵ For all price series, except for gold,

⁴ Note that tests for unit roots have in general low power.

⁵ In the Appendix the results for the net-long positions of Prod and Other are presented.

| Variable | Test | Crude Oil | Heating Oil | Natural Gas |
|--------------|-------------------|------------|-------------|-------------|
| Price | ADF in levels | -1.968 | -1.527 | -1.574 |
| | ADF in 1st diff. | -20.113*** | -18.480*** | -18.665*** |
| | KPSS in levels | .585** | .942*** | 1.556*** |
| | KPSS in 1st diff. | .042 | .057 | .060 |
| Swap | ADF in levels | -.452 | -1.695 | -2.219 |
| | ADF in 1st diff. | -14.860*** | -16.751*** | -5.625*** |
| | KPSS in levels | 1.896*** | .297 | .854*** |
| | KPSS in 1st diff. | .147 | .109 | .092 |
| MM | ADF in levels | -2.343 | -4.823*** | -2.586* |
| | ADF in 1st diff. | -18.360*** | -14.504*** | -16.294*** |
| | KPSS in levels | 1.784*** | .762*** | .323 |
| | KPSS in 1st diff. | .003 | .032 | .166 |
| Variable | Test | Gold | Platinum | Silver |
| Price | ADF in levels | -1.5892 | -2.189 | -1.673 |
| | ADF in 1st diff. | -20.166*** | -19.674*** | -18.891*** |
| | KPSS in levels | 2.104*** | .435* | 1.563*** |
| | KPSS in 1st diff. | .370* | .080 | .178 |
| Swap | ADF in levels | -3.199** | -3.883*** | -2.793* |
| | ADF in 1st diff. | -19.629*** | -16.024*** | -17.482*** |
| | KPSS in levels | .297 | 1.301*** | 1.297*** |
| | KPSS in 1st diff. | .0327 | .058 | .0462 |
| MM | ADF in levels | -2.312 | -2.938** | -4.395*** |
| | ADF in 1st diff. | -18.342*** | -13.338*** | -15.619*** |
| | KPSS in levels | .466** | 1.761*** | .291 |
| | KPSS in 1st diff. | .103 | .045 | .033 |

Tab. 7.4: The test statistics for the Augmented Dickey Fuller (ADF) and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test in levels and first differences. The time series investigated are the S&P GSCI spot price index and the net-long positions of the Swap and the MM for the 6 commodities. ***, ** and * indicate statistical significance at .01%, .05% and .10% level, respectively.

both tests agree that the series in levels is not stationary, but the series of first differences is. Therefore, they are integrated of order one or I(1). In case of the gold price the KPSS test in first differences still rejects the null hypothesis of stationarity at the 10% significance level, while the ADF test for first differences rejects the null hypothesis of being non-stationary at the 1% significance level. The KPSS test statistic of second differences is insignificant. Hence, the gold price is I(1) according to the ADF test and I(2) according to the KPSS test. In the following analysis it is considered to be I(1) (like all other prices). Yet the qualitative results of the later analysis are not affected by this choice.⁶

The results for the net-long positions of both trader groups are not as clear cut. For the Swap only the net-long positions of crude oil and natural gas are I(1) according to both test statistics. For the remaining commodities either the ADF or the KPSS test indicates that the time series is I(0) or that it is I(1), except for gold for which the tests agree on an I(0) process. In the case of the MM only the net-long positions of crude oil and gold are conclusively I(1) and natural gas and silver are I(0). For heating oil and platinum the two test give conflicting evidence for either being a I(0) or I(1) process. However, all tests support that none of these time series are I(2) or integrated of an even higher order. Hence, the maximum order of integration is always 1 for all commodities. After determining the order of integration the next step is to set up the bivariate VAR for the two pairs of price and net-long position of either the MM or the Swap for each commodity.

The basic structure of the VAR is,

$$P_t = c_1 + a_1 P_{t-1} + \dots + a_p P_{t-p} + b_1 N_{t-1} + \dots + b_p N_{t-p} + u_t \quad (7.1)$$

$$N_t = c_2 + d_1 N_{t-1} + \dots + d_p N_{t-p} + e_1 P_{t-1} + \dots + e_p P_{t-p} + v_t \quad (7.2)$$

where c_1 and c_2 are constants, t is the time index and p the number of lags included, u_t and v_t are white noise, P_t denotes the price and N_t is the net-long position of either the MM or the Swap. The optimal lag length p is determined by the minimum of four standard criteria (final prediction error, Akaike information criterion, Schwarz information criterion, and the Hannan-Quinn information criterion). If the results for the criteria are inconclusive, all candidate VARs(p) are estimated using OLS and the optimal lag order is chosen according to the Breusch-Godfrey LM test for remaining autocorrelation in the residuals.⁷ The final step in the model setup is to

⁶ The results for considering the gold price to be I(2), which necessitates to add a second 'surplus' lag, are shown in the Appendix.

⁷ The VAR(p) is also estimated if the criteria give the same result to check whether

incorporate the m additional lags with respect to the maximum order of integration (for each pair of equations) in the VAR(p) models and estimate the corresponding VAR($p + m$) systems

$$P_t = c_1 + \sum_{i=1}^p a_i P_{t-i} + \sum_{j=1}^m a_{p+j} P_{t-p-j} + \sum_{i=1}^p b_i N_{t-i} + \sum_{j=1}^m b_{p+j} N_{t-p-j} + u_t \quad (7.3)$$

$$N_t = c_1 + \sum_{i=1}^p d_i N_{t-i} + \sum_{j=1}^m d_{p+j} N_{t-p-j} + \sum_{i=1}^p e_i P_{t-i} + \sum_{j=1}^m e_{p+j} P_{t-p-j} + v_t \quad (7.4)$$

using OLS, where the second and fourth sum in each equation represent the m additional lags.⁸ In order to test for Granger non-causality the null hypothesis is defined as $b_1 = \dots = b_p$ ($e_1 = \dots = e_p$) being jointly equal to zero. This implies that the net long position N does not Granger cause the price P (or vice versa). Note that a time series (X) is defined to Granger cause another time series (Y), if past realizations of X improve the forecast of Y above and beyond considering only past realizations of Y . The additional m lags (excluded from the test) lead to an asymptotically χ^2 distributed Wald test statistic with p degrees of freedom, but it is not efficient. The corresponding loss of efficiency in our case, however, is relatively small due to the fairly large sample size and the necessity to always include only one extra lag.⁹ The Wald test statistics for Granger non-causality plus the chosen lag length are presented in Table 7.5.

The findings for the net-long position with respect to the prices of energy commodities are rather mixed for the two trader groups. In the case of the Swap for all three commodities (crude oil, heating oil, and natural gas) the null hypothesis that the price does not Granger cause the net-long position is rejected at the 5% significance level and at the 1% level for heating oil. The null is only rejected at the 5% significance level for natural gas in the opposite direction. The price for heating oil also seems to Granger cause the net-long positions of the MM (null rejected at 1% level), but this is the only rejection of the null in both directions. Hence, the Swap seem to react to the price of energy commodities, while the evidence that their trading behavior might affect the price is restricted to one out of three cases. The MM as well do not seem to influence the price.

the Breusch-Godfrey LM test shows low autocorrelation in the residuals. The complete results for all criteria and the Breusch-Godfrey LM test are available from the author upon request.

⁸ The estimation results for all VARs are available from the author upon request.

⁹ Zapata and Rambaldi (1997) show that the ‘surplus lag’ is accurate for sample sizes of 100 or larger and emphasize the few restrictive assumptions and robustness of this procedure.

| Commodity | Crude Oil | Heating Oil | Natural Gas |
|--|---------------------|---------------------|-------------------|
| Dependent variable: Swap, Price | | | |
| lags | 2 | 2 | 3 |
| H0: Does Swap not Granger cause the price | 2.182 (.336) | 1.789 (.409) | 9.798** (.020) |
| H0: Does the price not Granger cause Swap | 21.578*** (.000) | 6.506** (.039) | 8.357** (.039) |
| Dependent variable: MM, Price | | | |
| lags | 7 | 6 | 3 |
| H0: Does MM not Granger cause the price | 7.462 (.382) | 4.562 (.601) | 3.223 (.359) |
| H0: Does the price not Granger cause MM | 8.500 (.291) | 12.096*** (.060) | 5.783 (.123) |
| Commodity | Gold | Platinum | Silver |
| Dependent variable: Swap, Price | | | |
| lags | 1 | 3 | 4 |
| H0: Does Swap not Granger cause the price | .008 (.929) | 6.663* (.084) | 5.285 (.259) |
| H0: Does the price not Granger cause Swap | 2.531 (.112) | 2.049 (.562) | 5.504 (.239) |
| Dependent variable: MM, Price | | | |
| lags | 2 | 3 | 2 |
| H0: Does MM not Granger cause the price | 1.325 (.516) | 4.611 (.203) | 1.988 (.370) |
| H0: Does the price not Granger cause MM | 3.267 (.195) | 3.353 (.340) | 1.796 (.407) |

Tab. 7.5: The optimal lag length indicated by the Akaike information criterion, the Schwarz information criterion, the Hannan-Quinn information criterion and the final prediction error for the bivariate VAR's without including the extra lag. The number of lags is also the number of d.o.f. in the Wald test-statistic, when testing for Granger non-causality between the net-long positions of the CIT or MM and the S&P GSCI spot price index. ***, ** and * indicate statistical significance at .01, .05 and .10 level, respectively. P-values are presented in parentheses.

For the three precious metals the results are more consistent. Only the null hypothesis that the net-long position of the Swap does not Granger cause the price of platinum is rejected at the 10% level. Therefore, the interdependencies between price and trading behavior are overall weak, especially for the MM, which are the main suspects to engage in destabilizing speculative activity. The Swap on the other side seem to react to prices in case of energy commodities, while their net-long positions in two out of six cases (natural gas and platinum) seem to influence the price of the commodity.

However, if we consider that the sample period includes the GFC one might argue that the interdependencies change over time. A look at the time series in Figure 7.1 supports this idea. To determine structural breaks in the prices of the commodities the OLS-CUSUM test following Ploberger and Krämer (1992) is applied.¹⁰ In general, the corresponding test statistics detect at least one structural break for each price series. For heating oil, crude oil, natural gas, and platinum the price drop due to the onset of the GFC is identified as a structural break. A second common break point this time for heating oil, crude oil, gold, and silver is around the beginning of the year 2011. Here, the intuition might be that the most turbulent times with respect to the GFC ended somewhere around this time, which is reflected in the less trending price movements. In case of heating oil and crude oil the price recovered from the aforementioned drop. For gold and silver the price increase most likely due to their safe haven character stopped around this time. Therefore, the sample is split into three subsamples according to the break points set by the Lehman collapse and the beginning of 2011.

The first subsample (pre-GFC) from the 13th of June 2006 until the 9th of September 2008 consists of 118 weekly observations and the steep price increase for heating oil, crude oil, natural gas, and platinum might indicate destabilizing speculation during this period. Also, the gold and silver price increased, but the increase was moderate compared to the one after the GFC. The next period (GFC) from the 16th of September 2008 until the 28th of December 2010 consists of 120 weekly observations and incorporates price drops for all commodities in the beginning. While the energy commodities do not reach the price level before the GFC until the end of the sample. The gold and silver prices have surged upwards and already surpassed the pre-GFC maximum in late 2009. The last subsample from the 4th of January 2010 until the 31st of December 2013 consists of 157 weekly observations and appears to display random price movements for all commodities, yet on

¹⁰ The graphs of the cumulated residuals of the OLS regressions are shown in the Appendix. I used Zeileis, Leisch, Hornik, and Kleiber (2002) to estimate the breakpoints and plot the cumulated residuals.

| period | pre GFC | crisis | after GFC |
|---|----------------------|---------------------------|-----------------------------|
| observations | 118 | 120 | 157 |
| H0 that Swap not Granger cause the price is rejected for | natural gas | natural gas | |
| H0 that the price not Granger cause Swap is rejected for | heating oil, gold | crude oil, heating oil | crude oil, natural gas |
| H0 that MM not Granger cause the price is rejected for | crude oil | silver | silver |
| H0 that the price not Granger cause MM is rejected for | | heating oil | heating oil, natural gas |

Tab. 7.6: The commodities for which the null hypothesis was rejected for the three subsamples.

different levels compared to the complete sample. For instance, the S&P spot price index for natural gas is 198.05 on the 31st of December 2013. This value is below the sample average of 245.23 and the starting value of 301.43. At the same time the S&P spot price index for gold of 699.99 is above the sample average (676.41) and more than double of its starting value (330). Hence, the price series evolve similarly until the onset of the GFC and the following price drop, but not afterwards. The next step is to estimate the VAR model for each individual subsample.

The estimation results of the tests for Granger non-causality are summarized in Table 7.6. The Table only lists those commodities for which the null hypothesis got at least rejected at the 10% significance level.¹¹ The results for the Swap differ only in minor details from the complete sample results. They seem to Granger cause the price for natural gas for the pre-GFC and GFC period, while their trading behavior seems to react to the price after the GFC. For other commodities the trading behaviour of the Swap never affects the price, but in case of gold (pre-GFC), heating oil (pre-GFC and GFC) and crude oil (GFC and after GFC) they seem to adjust their positions to the price.

In case of the MM the interdependencies are stronger compared to the complete sample results. They still seem to adjust their trading behaviour to the prices of heating oil (GFC and after GFC) and natural gas (after GFC). However, their net-long positions seem to affect the prices of crude oil (pre-GFC) and silver (GFC and after GFC). Hence, they potentially cause two out of the six commodities during the periods in which the respective prices

¹¹ The complete results for all subsamples and trader groups are included in the Appendix.

showed the largest price increases.

7.5 *Conclusion*

A look at the time series of almost all commodities in the last ten years immediately raises the question whether the dramatic price changes and periods of strong price trends occurred solely due to changing fundamentals or whether the financial integration of these markets played an important role. To answer this question, it is necessary to identify potential speculators and therefore I rely on the COT report from the CFTC. The ‘disaggregated’ report separates four trader groups and I identify Money Manager and Swap Dealer as the two prime candidates for speculative activity. This study investigates the effect of speculation on prices of energy commodities and precious metals during the period from June 2006 until December 2013. The former experienced a dramatic price increase until the onset of the GFC and a huge price drop directly after the onset of the crisis. The same is true for the latter, but precious metals, i.e. gold and silver, experienced a surge in prices after their crisis-related drop and reached all time highs. These observations and corresponding tests for (common) structural breaks in the price series led to splitting the sample into three subperiods. The first sample includes observations until the beginning of the GFC in September 2008 and the second until the end of December 2011, which marked the end of the most turbulent times for the financial system and commodity price movements.

The analysis for Granger causal effects among the speculative activity (measured as net-long position) of the two classes and the prices revealed mixed evidence. Investigating the complete sample, the MM did not influence any price, while the Swap seem to have caused the price of natural gas and platinum, but also adjusted their trading with respect to all three energy commodities. However, the MM seem to have Granger caused the price of crude oil during the price surge before the GFC and the price of silver during the two subperiods after the onset of the GFC, in which the price had its most dramatic increase. They also seem to have reacted to price changes of heating oil (GFC, after GFC) and natural gas (after GFC). The Swap seem to only cause the price of natural gas before and during the GFC, while they always adjusted their trading to two alternating commodities. Hence, the prime candidate of MM might have had an influence on 2 out of 6 commodities during periods of upward trending prices. The evidence might not be robust enough to justify harsher regulations of futures markets, the most transparent segment of non physical trading activity of commodities. Yet regulators should focus on over-the-counter trades, which are also used

for speculation and less regulated.¹² However, it supports close surveillance of all of these markets and future research in this direction.

¹² See Parsons (2010) for a more detailed description of over-the-counter trading activity in commodity markets.

| | Prices | NC | COM | Prod | Swap | MM | Other |
|--------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------|
| Prices | 1 | | | | | | |
| NC | .1643 (.3420) | 1 | | | | | |
| COM | -.2121 (.2997) | -.9806 (.0173) | 1 | | | | |
| Prod | .1983 (.0866) | -.0946 (.6563) | .1148 (.6274) | 1 | | | |
| Swap | -.3165 (.2571) | -.5393 (.6360) | .5485 (.6032) | -.6497 (.1354) | 1 | | |
| MM | .2251 (.5171) | .8357 (.0980) | -.8629 (.0935) | -.2691 (.4881) | -.4199 (.4274) | 1 | |
| Other | .0093 (.5215) | .4321 (.2643) | -.3563 (.3129) | .4085 (.5219) | -.3819 (.5391) | -.0094 (.5026) | 1 |

Tab. 7.7: Average correlations and standard deviations (in brackets) between the net-long positions of the trader classes with each other and the S&P GSCI Spot Price Indexes for the 3 energy commodities.

7.6 Appendix

7.6.1 Correlation

The average correlation coefficients among the trader groups and the prices for the three energy commodities are shown in Table 7.7 and in Table 7.8 for the three precious metals.

7.6.2 Complementary Results

The complete results for the tests of Granger non-causality for the three subperiods are shown in Table 7.9 for the pre-GFC period, in Table 7.10 for the crisis period and in Table 7.11 for the period after the GFC.

7.6.3 Structural Breaks

To determine (common) structural breaks in the price series of the commodities during the sample period, the OLS-CUSUM test introduced by Ploberger and Krämer (1992) was applied. The plotted results for the statistics of cumulative OLS residuals of the 6 commodities are illustrated in Figure 7.2.

| | Prices | NC | COM | Prod | Swap | MM | Other |
|--------|-------------------|-------------------|-------------------|-------------------|-------------------|------------------|-------|
| Prices | 1 | | | | | | |
| NC | .2660 (.1937) | 1 | | | | | |
| COM | -.2433 (.2491) | -.9854 (.0152) | 1 | | | | |
| Prod | -.2951 (.1069) | -.9096 (.0641) | .8886 (.1063) | 1 | | | |
| Swap | -.1412 (.3685) | -.7279 (.1122) | .7733 (.0691) | .4185 (.2807) | 1 | | |
| MM | .2991 (.1685) | .9744 (.0077) | -.9541 (.0314) | -.8720 (.0595) | -.7264 (.1525) | 1 | |
| Other | .0458 (.2795) | .5318 (.2592) | -.5524 (.2234) | -.5378 (.2891) | -.3738 (.1280) | .3392 (.3026) | 1 |

Tab. 7.8: Average correlations and standard deviations (in brackets) between the net-long positions of the trader classes with each other and the S&P GSCI Spot Price Indexes for the 3 precious metals commodities.

7.6.4 Robustness Check

As mentioned in the main text the ADF test and KPSS test disagree on the order of integration of the S&P GSCI spot price index of gold. To show that the qualitative result is not affected by either choosing the process to be I(1) or I(2) the results for the latter of all 4 trader groups are presented in Table 7.12. None of the null hypotheses can be rejected neither for the MM nor the Swap. However, the result that the futures net-long position of the Swap is affected by the price attains a p-value of .101 and is almost significant. Yet already the result for the gold price being an I(1) process is a borderline case (p-value=.112) in terms of significance.

7.6.5 Other Trader Groups

For completeness, this section contains all the findings for the two other trader groups of the ‘disaggregated’ report, namely the Prod and the Other. The results for the unit-root tests (Table 7.13) and the results of the Granger non-causality test of the complete sample (Table 7.14) and the subsamples (Tables 7.15 to 7.17).

| Commodity | Crude Oil | Heating Oil | Natural Gas |
|--|--------------------|------------------|---------------------|
| Dependent variable: Swap, Price | | | |
| lags | 1 | 1 | 3 |
| H0: Does Swap not Granger cause the price | .050 (.824) | .185 (.668) | 24.656*** (.001) |
| H0: Does the price not Granger cause Swap | 2.612 (.106) | 2.911* (.088) | 3.865 (.795) |
| Dependent variable: MM, Price | | | |
| lags | 6 | 1 | 3 |
| H0: Does MM not Granger cause the price | 14.356** (.026) | 1.428 (.232) | 3.718 (.294) |
| H0: Does the price not Granger cause MM | 10.192 (.117) | .097 (.755) | 2.259 (.521) |
| Commodity | Gold | Platinum | Silver |
| Dependent variable: Swap, Price | | | |
| lags | 1 | 3 | 2 |
| H0: Does Swap not Granger cause the price | .028 (.867) | 2.707 (.439) | .139 (.933) |
| H0: Does the price not Granger cause Swap | 3.749** (.053) | 2.601 (.4573) | .538 (.764) |
| Dependent variable: MM, Price | | | |
| lags | 1 | 3 | 2 |
| H0: Does MM not Granger cause the price | .000 (.984) | .349 (.950) | .067 (.967) |
| H0: Does the price not Granger cause MM | 1.418 (.234) | 2.257 (.521) | 1.994 (.369) |

Tab. 7.9: The results for the 118 weekly observations from 13.06.2006 until 09.09.2008. The optimal lag length indicated by the Akaike information criterion, the Schwarz information criterion, the Hannan-Quinn information criterion and the final prediction error for the bivariate VAR's without including the extra lag. The number of lags is also the number of d.o.f. in the Wald test-statistic, when testing for Granger non-causality between the net-long positions of the CIT or MM and the S&P GSCI spot price index. ***, ** and * indicate statistical significance at .01, .05 and .10 level respectively. P-values are presented in parentheses.

| Commodity | Crude Oil | Heating Oil | Natural Gas |
|--|---------------------|--------------------|-------------------|
| Dependent variable: Swap, Price | | | |
| lags | 2 | 1 | 2 |
| H0: Does Swap not Granger cause the price | 1.787 (.824) | .552 (.458) | 5.475* (.065) |
| H0: Does the price not Granger cause Swap | 10.295*** (.006) | 4.465** (.035) | 3.26 (.196) |
| Dependent variable: MM, Price | | | |
| lags | 6 | 4 | 1 |
| H0: Does MM not Granger cause the price | 10.253 (.114) | 2.210 (.697) | 1.210 (.271) |
| H0: Does the price not Granger cause MM | 3.641 (.725) | 10.323** (.035) | .040 (.842) |
| Commodity | Gold | Platinum | Silver |
| Dependent variable: Swap, Price | | | |
| lags | 1 | 1 | 4 |
| H0: Does Swap not Granger cause the price | .039 (.844) | .096 (.757) | 6.184 (.186) |
| H0: Does the price not Granger cause Swap | .709 (.400) | .303 (.582) | 2.971 (.563) |
| Dependent variable: MM, Price | | | |
| lags | 2 | 2 | 2 |
| H0: Does MM not Granger cause the price | .470 (.791) | 1.134 (.567) | 4.8071* (.090) |
| H0: Does the price not Granger cause MM | .453 (.797) | .639 (.727) | .232 (.890) |

Tab. 7.10: The results for the 120 weekly observations from 16.09.2008 until 28.12.2010. The optimal lag length indicated by the Akaike information criterion, the Schwarz information criterion, the Hannan-Quinn information criterion and the final prediction error for the bivariate VAR's without including the extra lag. The number of lags is also the number of d.o.f. in the Wald test-statistic, when testing for Granger non-causality between the net-long positions of the CIT or MM and the S&P GSCI spot price index. ***, ** and * indicate statistical significance at .01, .05 and .10 level respectively. P-values are presented in parentheses.

| Commodity | Crude Oil | Heating Oil | Natural Gas |
|---|-------------------|-------------------|---------------------|
| Dependent variable: | | | |
| Swap, Price | | | |
| lags | 2 | 2 | 2 |
| H0: Does Swap not Granger cause the price | 1.447 (.485) | 1.005 (.605) | 1.002 (.606) |
| H0: Does the price not Granger cause Swap | 7.193** (.027) | 1.305 (.521) | 16.025*** (.000) |
| Dependent variable: | | | |
| MM, Price | | | |
| lags | 3 | 9 | 2 |
| H0: Does MM not Granger cause the price | 5.868 (.118) | 12.442 (.190) | .504 (.777) |
| H0: Does the price not Granger cause MM | 5.783 (.123) | 14.872* (.095) | 12.860*** (.002) |
| Commodity | Gold | Platinum | Silver |
| Dependent variable: | | | |
| Swap, Price | | | |
| lags | 1 | 2 | 2 |
| H0: Does Swap not Granger cause the price | .318 (.573) | .094 (.954) | .794 (.672) |
| H0: Does the price not Granger cause Swap | .833 (.361) | 3.951 (.139) | 1.024 (.599) |
| Dependent variable: | | | |
| MM, Price | | | |
| lags | 1 | 2 | 2 |
| H0: Does MM not Granger cause the price | .011 (.915) | 2.836 (.242) | 4.807* (.099) |
| H0: Does the price not Granger cause MM | .886 (.347) | 2.487 (.288) | 1.633 (.442) |

Tab. 7.11: The results for the 157 weekly observations from 04.01.2010 until 31.12.2013. The optimal lag length indicated by the Akaike information criterion, the Schwarz information criterion, the Hannan-Quinn information criterion and the final prediction error for the bivariate VAR's without including the extra lag. The number of lags is also the number of d.o.f. in the Wald test-statistic, when testing for Granger non-causality between the net-long positions of the CIT or MM and the S&P GSCI spot price index. ***, ** and * indicate statistical significance at .01, .05 and .10 level respectively. P-values are presented in parentheses.

| Commodity | Gold |
|--|-------------------|
| Dependent variable: | |
| Swap, Price | |
| lags | 1(2) |
| H0: Does Swap not Granger cause the price | .015 (.903) |
| H0: Does the price not Granger cause Swap | 2.694 (.101) |
| Dependent variable: | |
| MM, Price | |
| lags | 2(2) |
| H0: Does MM not Granger cause the price | 1.967 (.3739) |
| H0: Does the price not Granger cause MM | 3.410 (.1818) |
| Commodity | Gold |
| Dependent variable: | |
| Prod, Price | |
| lags | 2(2) |
| H0: Does Prod not Granger cause the price | 1.877 (.391) |
| H0: Does the price not Granger cause Prod | 3.025 (.220) |
| Dependent variable: | |
| Other, Price | |
| lags | 1(2) |
| H0: Does Other not Granger cause the price | .341 (.559) |
| H0: Does the price not Granger cause Other | 2.934* (.0867) |

Tab. 7.12: The results for gold if a second ‘surplus’ lag is added into the VAR equations. The optimal lag length indicated by the Akaike information criterion, the Schwarz information criterion, the Hannan-Quinn information criterion and the final prediction error for the bivariate VAR’s without including the extra lag. The number of lags is also the number of d.o.f. in the Wald test-statistic, when testing for Granger non-causality between the net-long positions of the Swap or MM and the S&P GSCI spot price index. ***, ** and * indicate statistical significance at .01, .05 and .10 level respectively. P-values are presented in parentheses.

| Variable | Test | Crude Oil | Heating Oil | Natural Gas |
|--------------|-------------------|------------|-------------|-------------|
| Prod | ADF in levels | -.812 | -2.499 | -1.603 |
| | ADF in 1st diff. | -20.963*** | -14.969*** | -16.701*** |
| | KPSS in levels | .878*** | .446* | .966*** |
| | KPSS in 1st diff. | .257 | .154 | .098 |
| Other | ADF in levels | -1.938 | -1.249 | -1.057 |
| | ADF in 1st diff. | -22.119*** | -18.251*** | -18.579*** |
| | KPSS in levels | 1.813*** | 1.410*** | 1.9180*** |
| | KPSS in 1st diff. | .108 | .236 | .237 |
| Variable | Test | Gold | Platinum | Silver |
| Prod | ADF in levels | -1.809 | -1.673 | -3.155** |
| | ADF in 1st diff. | -17.324*** | -16.088*** | -16.797*** |
| | KPSS in levels | .392* | 2.055*** | .411* |
| | KPSS in 1st diff. | .186 | .034 | .112 |
| Other | ADF in levels | -3.902*** | -1.229 | -3.764*** |
| | ADF in 1st diff. | -20.942*** | -17.710*** | -17.423*** |
| | KPSS in levels | .133 | 1.955*** | .526** |
| | KPSS in 1st diff. | .033 | .030 | .029 |

Tab. 7.13: The test statistics for the Augmented Dickey Fuller (ADF) and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test in levels and first differences. The time series investigated are the S&P GSCI spot price index and the net-long positions of Swap and MM for the 6 commodities. ***, ** and * indicate statistical significance at .01%, .05% and .10% level respectively.

| Commodity | Crude Oil | Heating Oil | Natural Gas |
|---|--------------------|---------------------|-------------------|
| Dependent variable: Prod, Price | | | |
| lags | 7 | 5 | 2 |
| H0: Does Prod not Granger cause the price | 5.317 (.621) | 7.869 (.164) | 1.789 (.409) |
| H0: Does the price not Granger cause Prod | 17.203** (.016) | 14.211** (.014) | 4.582 (.101) |
| Dependent variable: Other, Price | | | |
| lags | 1 | 2 | 5 |
| H0: Does Other not Granger cause the price | .255 (.613) | .972 (.615) | 4.843 (.435) |
| H0: Does the price not Granger cause Other | .113 (.737) | 21.259*** (.000) | 4.929 (.425) |
| Commodity | Gold | Platinum | Silver |
| Dependent variable: Prod, Price | | | |
| lags | 2 | 3 | 2 |
| H0: Does Prod not Granger cause the price | 1.125 (.570) | 5.366 (.147) | 2.177 (.337) |
| H0: Does the price not Granger cause Prod | 3.149 (.207) | 7.335* (.062) | 5.014* (.082) |
| Dependent variable: Other, Price | | | |
| lags | 1 | 4 | 11 |
| H0: Does Other not Granger cause the price | .320 (.572) | 2.832 (.586) | 10.073 (.524) |
| H0: Does the price not Granger cause Other | 2.945* (.086) | 8.526* (.074) | 17.686* (.089) |

Tab. 7.14: The optimal lag length indicated by the Akaike information criterion, the Schwarz information criterion, the Hannan-Quinn information criterion and the final prediction error for the bivariate VAR's without including the extra lag. The number of lags is also the number of d.o.f. in the Wald test-statistic, when testing for Granger non-causality between the net-long positions of the CIT or MM and the S&P GSCI spot price index. ***, ** and * indicate statistical significance at .01, .05 and .10 level respectively. P-values are presented in parentheses.

| Commodity | Crude Oil | Heating Oil | Natural Gas |
|--|-----------------|------------------|-------------------|
| Dependent variable: | | | |
| Prod, Price | | | |
| lags | 4 | 1 | 1 |
| H0: Does Prod not Granger cause the price | 6.922 (.140) | 1.064 (.302) | 1.223 (.269) |
| H0: Does the price not Granger cause Prod | 5.789 (.215) | .008 (.931) | .249 (.618) |
| Dependent variable: | | | |
| Other, Price | | | |
| lags | 1 | 1 | 2 |
| H0: Does Other not Granger cause the price | .018 (.892) | 1.340 (.247) | 2.019 (.364) |
| H0: Does the price not Granger cause Other | .464 (.496) | .007 (.935) | 1.948 (.378) |
| Commodity | Gold | Platinum | Silver |
| Dependent variable: | | | |
| Prod, Price | | | |
| lags | 2 | 3 | 2 |
| H0: Does Prod not Granger cause the price | .554 (.758) | 4.723 (.193) | .047 (.829) |
| H0: Does the price not Granger cause Prod | .015 (.993) | 5.694 (.127) | 5.664** (.017) |
| Dependent variable: | | | |
| Other, Price | | | |
| lags | 1 | 5 | 1 |
| H0: Does Other not Granger cause the price | .000 (1.000) | 4.049 (.542) | .390 (.532) |
| H0: Does the price not Granger cause Other | .563 (.453) | 9.841* (.080) | .913 (.339) |

Tab. 7.15: The results for the 118 weekly observations from 13.06.2006 until 09.09.2008. The optimal lag length indicated by the Akaike information criterion, the Schwarz information criterion, the Hannan-Quinn information criterion and the final prediction error for the bivariate VAR's without including the extra lag. The number of lags is also the number of d.o.f. in the Wald test-statistic, when testing for Granger non-causality between the net-long positions of the CIT or MM and the S&P GSCI spot price index. ***, ** and * indicate statistical significance at .01, .05 and .10 level respectively. P-values are presented in parentheses.

| Commodity | Crude Oil | Heating Oil | Natural Gas |
|--|--------------------|------------------|------------------|
| Dependent variable: | | | |
| Prod, Price | | | |
| lags | 1 | 2 | 1 |
| H0: Does Prod not Granger cause the price | .029 (.865) | 1.528 (.466) | 1.081 (.299) |
| H0: Does the price not Granger cause Prod | .003 (.958) | 6.032 (.049) | .005 (.941) |
| Dependent variable: | | | |
| Other, Price | | | |
| lags | 1 | 2 | 1 |
| H0: Does Other not Granger cause the price | 3.176* (.075) | .095 (.954) | .578 (.447) |
| H0: Does the price not Granger cause Other | 2.653 (.103) | 6.980* (.031) | .376 (.540) |
| Commodity | Gold | Platinum | Silver |
| Dependent variable: | | | |
| Prod, Price | | | |
| lags | 2 | 3 | 3 |
| H0: Does Prod not Granger cause the price | .469 (.791) | 3.543 (.315) | 5.549* (.062) |
| H0: Does the price not Granger cause Prod | 4.336 (.114) | 3.625 (.305) | 4.127 (.127) |
| Dependent variable: | | | |
| Other, Price | | | |
| lags | 1 | 1 | 1 |
| H0: Does Other not Granger cause the price | .021 (.884) | 1.109 (.292) | .063 (.802) |
| H0: Does the price not Granger cause Other | 8.046*** (.005) | .028 (.867) | .000 (.989) |

Tab. 7.16: The results for the 120 weekly observations from 16.09.2008 until 28.12.2010. The optimal lag length indicated by the Akaike information criterion, the Schwarz information criterion, the Hannan-Quinn information criterion and the final prediction error for the bivariate VAR's without including the extra lag. The number of lags is also the number of d.o.f. in the Wald test-statistic, when testing for Granger non-causality between the net-long positions of the CIT or MM and the S&P GSCI spot price index. ***, ** and * indicate statistical significance at .01, .05 and .10 level respectively. P-values are presented in parentheses.

| Commodity | Crude Oil | Heating Oil | Natural Gas |
|--|-------------------|---------------------|---------------------|
| Dependent variable: | | | |
| Prod, Price | | | |
| lags | 1 | 2 | 2 |
| H0: Does Prod not Granger cause the price | 1.049 (.306) | .920 (.631) | .151 (.927) |
| H0: Does the price not Granger cause Prod | .000 (.997) | 5.849* (.054) | 25.632*** (.000) |
| Dependent variable: | | | |
| Other, Price | | | |
| lags | 1 | 3 | 2 |
| H0: Does Other not Granger cause the price | 1.886 (.170) | 1.679 (.642) | .2.896 (.235) |
| H0: Does the price not Granger cause Other | 4.450** (.035) | 16.588*** (.001) | 8.050** (.018) |
| Commodity | Gold | Platinum | Silver |
| Dependent variable: | | | |
| Prod, Price | | | |
| lags | 1 | 2 | 2 |
| H0: Does Prod not Granger cause the price | .009 (.925) | 5.426* (.066) | 6.056** (.048) |
| H0: Does the price not Granger cause Prod | .663 (.416) | 8.600** (.014) | 1.102 (.576) |
| Dependent variable: | | | |
| Other, Price | | | |
| lags | 1 | 7 | 2 |
| H0: Does Other not Granger cause the price | .079 (.779) | 8.140 (.320) | 1.479 (.477) |
| H0: Does the price not Granger cause Other | .131 (.717) | 10.278 (.173) | 2.419 (.298) |

Tab. 7.17: The results for the 157 weekly observations from 04.01.2010 until 31.12.2013. The optimal lag length indicated by the Akaike information criterion, the Schwarz information criterion, the Hannan-Quinn information criterion and the final prediction error for the bivariate VAR's without including the extra lag. The number of lags is also the number of d.o.f. in the Wald test-statistic, when testing for Granger non-causality between the net-long positions of the CIT or MM and the S&P GSCI spot price index. ***, ** and * indicate statistical significance at .01, .05 and .10 level respectively. P-values are presented in parentheses.

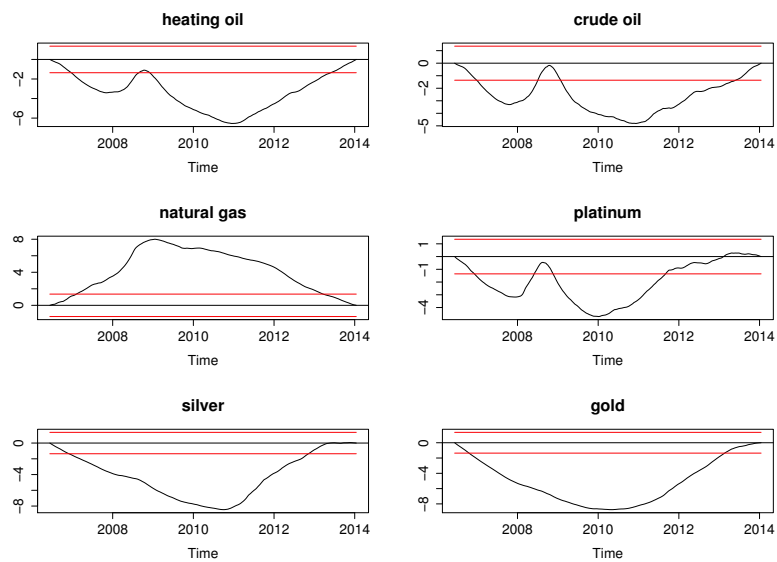


Fig. 7.2: Plotted statistics for the cumulative OLS residuals of the 6 commodities. The horizontal lines (red) define the statistical significance at the 5% level.

8. CONCLUDING REMARKS

All the studies show how complex the interactions on the micro-level (agent, bank, trader group) are and shed some light on how they determine the observed outcomes (price, network) on the macro-level. The two main findings in case of the artificial financial market are: (1) the introduction of a second asset triggers alternating herding behavior of the informed traders. This results from simple profit considerations by the individual agents rather than any form of coordination between them. (2) The FTT ‘bans’ the informed traders already for very small tax rates and in contrast the FAT has almost no effect on them with respect to which market they submit their orders to. On account of this, the FTT is able to reduce the distortions from the fundamental value in the taxed market for very small tax rates, while the FAT is affecting the overall trading process less and therefore an introduction might bear less risk. Interesting future projects for an agent-based financial market could incorporate findings of behavioral finance, like overconfidence, anchoring and adjustment or home bias to observe how such fallacies work together in a market setup.

For the banking network, the most important finding is that banks show persistent behavior over the complete sample (including the GFC) and strong preference for continued lending relations. Another interesting finding is that in terms of intermediate relations the high directed clustering coefficient (macro-level) seems not to be preferred by the banks, but emerges unintendedly out of their behavior which is steered by other preferences. However, these are only first steps and it is necessary to look for robust universal results for different banking networks and at the same time overcome the temptation of ‘nice’ findings such as a power-law in the degree distribution which are observed in other disciplines. Moreover, a more profound knowledge of the behavior of banks is needed to finally tackle the task of modelling the financial system and provide reliable policy recommendations.

In the investigation of the agricultural commodity markets it is interesting to see that the commodity index traders follow a clear ‘long only’ strategy. They most likely seek to meet the demand of investors who want to include these assets for diversification purposes in their portfolio; a strategy that should make them relatively independent from the developments on these markets. Therefore, it makes sense that they neither systematically affect the price nor react to it. The ‘typical’ trader groups of commercials and non-commercials on the other side also do not influence the price but react to it (in opposing directions). The results for energy commodities and precious metals show that the Money Managers (non-commercials) seem to Granger cause the most extreme price increase in the case of two out of six commodities, namely crude oil and silver. Hence, there is some evidence that excessive speculation might have influenced the prices of commodities. However, the

regulators, in my view, should be primarily concerned with the less regulated off exchange trading than with futures markets. In order to add new insights to the existing literature in the future, a more detailed data set would be required.

The lack of appropriate data is overall a very important limitation for many research projects and is often seen as irrevocable. However, it is necessary to distinguish between different economic disciplines and research questions whether there is an actual lack of data or the data is not made available. For many macroeconomic settings the difficulty to measure fundamentals like inflation or GDP growth and the overall poor collection of such data in the past generates a real lack of data. However, in finance there is no general lack of data, but researchers are often prohibited to use the data or important aspects of it.

In case of my own thesis these limitations potentially hinder much deeper insights in all but the agent based model and this is simply because it generates its own data. The interbank loan data set is anonymized. Therefore, it is impossible to associate the banks with balance sheet measures or other specific features, which would most likely help us to better understand their behavior. In this case e-MID as a privately owned company sells the data to produce some auxiliary income, but naturally secures the concerns of their primary customers (participating financial institutions). Moreover, interbank loans are not the only market in which banks are exposed to each other (e.g. derivatives) and the study of banking networks would benefit enormously if it would be possible to incorporate these segments as well. On the other hand for the investigation of the effects of speculative activity on commodity prices the data is made publicly available by the CFTC, but is again restrictive in many ways. The traders are classified into up to four groups and their positions are aggregated, hence we do not know the single traders nor their asset holdings. Moreover, the weekly aggregation, despite the CFTC receiving the data on a daily basis, is obscuring potentially important information. This is most likely concealing important insights, since on financial markets every trade or even order reveals information. Another general complaint is that the CFTC only issues data regarding products traded on public exchanges, while it is a well known fact that in many markets bilateral over-the-counter trades are of great significance.

It is unlikely that the availability of any given data alone would satisfyingly solve all corresponding research questions, but simply looking at the enormous (not only monetary) costs of the GFC makes it difficult to understand why the public should value the secrecy of this data higher than the potential benefit.

Bibliography

- AFFINITO, M. (2012): “Do interbank customer relationships exist? And how did they function in the crisis? Learning from Italy,” *Journal of Banking and Finance*, 36(12), 3163 – 3184.
- AGNEESSENS, F., AND R. WITTEK (2012): “Where do intra-organizational advice relations come from? The role of informal status and social capital in social exchange,” *Social Networks*, 34(3), 333 – 345.
- ALLEN, F., AND D. GALE (2000): “Financial Contagion,” *Journal of Political Economy*, 108(1), 1–33.
- ALQUIST, R., AND L. KILIAN (2010): “What do we learn from the price of crude oil futures?,” *Journal of Applied Econometrics*, 25(4), 539–573.
- BARABASI, A.-L., AND R. ALBERT (1999): “Emergence of Scaling in Random Networks,” *Science*, 286(5439), 509–512.
- BASLE COMMITTEE ON BANKING SUPERVISION (2008): “Principles for Sound Liquidity Risk Management and Supervision,” Framework publication, Bank of International Settlements.
- (2011): “Global Systemically Important Banks: Assessment Methodology and the Additional Loss Absorbency Requirement,” Final Report, Bank for International Settlements.
- BATTISTON, S., D. D. GATTI, M. GALLEGATI, B. GREENWALD, AND J. E. STIGLITZ (2012): “Default Cascades: When Does Risk Diversification Increase Stability?,” *Journal of Financial Stability*, 8(3), 138–149.
- BAUR, D. G., AND K. J. GLOVER (2014): “Heterogeneous expectations in the gold market: Specification and estimation,” *Journal of Economic Dynamics and Control*, 40(0), 116 – 133.
- BEAUPAIN, R., AND A. DURRÉ (2011): “Inferring trading dynamics for an OTC market: The case of the euro area overnight money market,” *Quantitative Finance*, 11(9), 1285–1295.

-
- BEAUPAIN, R., AND A. DURRÉ (2012): “Nonlinear Liquidity Adjustments in the Euro Area Overnight Money Market,” Working Paper Series 1500, European Central Bank.
- BECH, M., AND E. ATALAY (2010): “The Topology of the Federal Funds Market,” *Physica A*, 389(22), 5223–5246.
- BOHL, M. T., AND P. M. STEPHAN (2012): “Does Futures Speculation Destabilize Spot Prices? New Evidence for Commodity Markets,” *SSRN Electronic Journal*.
- BORGATTI, S. P., AND M. G. EVERETT (2000): “Models of Core/Periphery Structures,” *Social Networks*, 21(4), 375 – 395.
- BOSS, M., H. ELSINGER, M. SUMMER, AND S. THURNER (2004): “Network Topology of the Interbank Market,” *Quantitative Finance*, 4(6), 677–684.
- BRANDES, U., AND D. WAGNER (2004): “visone-Analysis and Visualization of Social Networks,” in *Graph Drawing Software*, ed. by M. Jünger, and P. Mutzel, pp. 321–340. Springer-Verlag.
- BROCK, W. A., AND S. N. DURLAUF (2001): “Discrete choice with social interactions,” *The Review of Economic Studies*, 68(2), 235–260.
- BRODT, A. I. (1978): “A Dynamic Balance Sheet Management Model for a Canadian Chartered Bank,” *Journal of Banking & Finance*, 2(3), 221–241.
- BRUNETTI, C., B. BUYUKSAHIN, AND J. HARRIS (2011): “Speculators, prices and market volatility,” *Available at SSRN 1736737*.
- BRUNNERMEIER, M. K. (2009): “Deciphering the Liquidity and Credit Crunch 2007-2008,” *Journal of Economic Perspectives*, 23(1), 77–100.
- BRUNNERMEIER, M. K., AND L. H. PEDERSEN (2009): “Market Liquidity and Funding Liquidity,” *The Review of Financial Studies*, 22(6), 2201–2238.
- BUTTS, C., AND K. CARLEY (2001): “Multivariate Methods for Interstructural Analysis,” CASOS Working Paper, Carnegie Mellon University.
- BÜYÜKŞAHIN, B., AND J. H. HARRIS (2011): “Do speculators drive crude oil futures prices?,” *Energy Journal*, 32(2).

- BÜYÜKŞAHİN, B., AND M. A. ROBE (2014): “Speculators, commodities and cross-market linkages,” *Journal of International Money and Finance*, 42(0), 38 – 70, Understanding International Commodity Price Fluctuations.
- CARLIN, B. I., M. S. LOBO, AND S. VISWANATHAN (2007): “Episodic Liquidity Crises: Cooperative and Predatory Trading,” *The Journal of Finance*, 62(5), 2235–2274.
- CFTC (2008): “Commodity Swap Dealers & Index Traders with Commission Recommendations,” Staff report, Commodity Futures Trading Commission.
- CHEADLE, J. E., AND P. SCHWADEL (2012): “The ‘friendship dynamics of religion,’ or the ‘religious dynamics of friendship’? A social network analysis of adolescents who attend small schools,” *Social Science Research*, 41(5), 1198 – 1212.
- CHECKLEY, M., AND C. STEGLICH (2007): “Partners in power: job mobility and dynamic deal-making,” *European Management Review*, 4(3), 161–171.
- CLARKE, J. A., AND S. MIRZA (2006): “A comparison of some common methods for detecting Granger noncausality,” *Journal of Statistical Computation and Simulation*, 76(3), 207–231.
- 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.
- COHEN-COLE, E., E. PATACCINI, AND Y. ZENOU (2011): “Systemic risk and network formation in the interbank market,” .
- COLANDER, D., M. GOLDBERG, A. HAAS, K. JUSELIOUS, A. KIRMAN, T. LUX, AND B. SLOTH (2009): “The Financial Crisis and the Systemic Failure of Academic Economics,” *Critical Review*, 21(2-3), 249–267.
- COLLIARD, J.-E., AND P. HOFFMANN (2013): “Sand in the chips: Evidence on taxing transactions in an electronic market,” Discussion paper, Working paper European Central Bank.
- CONOVER, C. M., G. R. JENSEN, R. R. JOHNSON, AND J. M. MERCER (2010): “Is now the time to add commodities to your portfolio?,” *The Journal of Investing*, 19(3), 10–19.
- COVAL, J., J. JUREK, AND E. STAFFORD (2009): “The Economics of Structured Finance,” *The Journal of Economic Perspectives*, 23(1), 3–26.

-
- CRAIG, B., AND G. VON PETER (2010): “Interbank Tiering and Money Center Banks,” Discussion Paper, Series 2: Banking and Financial Studies 12/2010, Deutsche Bundesbank.
- DE MASI, G., G. IORI, AND G. CALDARELLI (2006): “Fitness Model for the Italian Interbank Money Market,” *Phys. Rev. E*, 74(6), 66112.
- ELSINGER, H., A. LEHAR, AND M. SUMMER (2006): “Using Market Information for Banking System Risk Assessment,” *International Journal of Central Banking*, 2(1).
- ERB, C. B., AND C. R. HARVEY (2006): “The strategic and tactical value of commodity futures,” *Financial Analysts Journal*, pp. 69–97.
- ERDÖS, P., AND A. RENYI (1959): “On Random Graphs,” *Publicationes Mathematicae*, 6, 290–297.
- FARMER, J. D., L. GILLEMOT, F. LILLO, S. MIKE, AND A. SEN (2004): “What Really Causes Large Price Changes?,” *Quantitative Finance*, 4(4), 383–397.
- FATTOUH, B., L. KILIAN, AND L. MAHADEVA (2013): “The role of speculation in oil markets: What have we learned so far?,” *Energy Journal*, 34(3).
- FINGER, K., D. FRICKE, AND T. LUX (2013): “Network analysis of the e-MID overnight money market: the informational value of different aggregation levels for intrinsic dynamic processes,” *Computational Management Science*, 10(2-3), 187–211.
- FINGER, K., M. HAAS, A. KLOS, AND S. REITZ (2013): “Who are the Speculators on Commodity Markets?,” QBER Discussion Paper 9-2013, Institute for Quantitative Business and Economics Research.
- FINGER, K., AND T. LUX (2013): “Friendship between Banks: An Application of the Actor-Oriented Model of Network Formation to Interbank Credit Relations,” unpublished manuscript.
- FRICKE, D. (2012): “Trading Strategies in the Overnight Money Market: Correlations and Clustering on the e-MID Trading Platform,” *Physica A*, (391), 6528–6542.
- FRICKE, D., AND T. LUX (2013a): “The effects of a financial transaction tax in an artificial financial market,” *Journal of Economic Interaction and Coordination*, pp. 1–32.

-
- (2013b): “On the Distribution of Links in the Interbank Network: Evidence from the e-MID Overnight Money Market,” Kiel working paper 1819, Kiel Institute for the World Economy.
- (2014): “Core-Periphery Structure in the Overnight Money Market: Evidence from the e-MID Trading Platform,” *Computational Economics*, in press.
- GALLEGATI, M., G. GIULIONI, A. KIRMAN, AND A. PALESTRINI (2011): “What’s that Got to Do With the Price of Fish? Buyers Behavior on the Ancona Fish Market,” *Journal of Economic Behavior & Organization*, 80(1), 20–33.
- GILBERT, C. L. (2010a): “How to Understand High Food Prices,” *Journal of Agricultural Economics*, 61(2), 398–425.
- GILBERT, C. L. (2010b): “Speculative influences on commodity futures prices 2006-2008,” United Nations Conference on Trade and Development.
- GOH, K.-I., B. KAHNG, AND D. KIM (2001): “Universal Behavior of Load Distribution in Scale-Free Networks,” *Phys. Rev. Lett.*, 87, 278701.
- GORTON, G., AND K. G. ROUWENHORST (2006): “Facts and Fantasies about Commodity Futures,” *Financial Analysts Journal*, 62, pp. 47–68.
- HALDANE, A. G., AND R. M. MAY (2011): “Systemic risk in banking ecosystems,” *Nature*, 469(7330), 351–355.
- HANKE, M., J. HUBER, M. KIRCHLER, AND M. SUTTER (2010): “The Economic Consequences of a Tobin Tax - An Experimental Analysis,” *Journal of Economic Behavior and Organization*, 74, 58–71.
- HAU, H. (2006): “The Role of Transaction Costs for Financial Volatility: Evidence from the Paris Bourse,” *Journal of the European Economic Association*, 4(4), 862–890.
- HERTEL, T. W., AND J. BECKMAN (2011): “Commodity Price Volatility in the Biofuel Era: An Examination of the Linkage Between Energy and Agricultural Markets,” Working Paper 16824, National Bureau of Economic Research.
- HO, T. S. Y., AND A. SAUNDERS (1985): “A Micro Model of the Federal Funds Market,” *Journal of Finance*, 40(3), 977–88.

- HOLME, P., C. R. EDLING, AND F. LILJEROS (2004): "Structure and Time Evolution of an Internet Dating Community," *Social Networks*, 26(2), 155 – 174.
- IIF SPECIAL COMMITTEE ON LIQUIDITY RISK (2007): "Principles of Liquidity Risk Management," Framework publication, Institute of International Finance.
- IMF (2010): "A Fair and Substantial Contribution by the Financial Sector," Final Report for the G-20, International Monetary Fund.
- INAOKA, H., T. NINOMYIA, K. TANIGUCHI, T. SHIMIZU, AND H. TAKAYASU (2004): "Fractal Network Derived from Banking Transaction – An Analysis of Network Structures Formed by Financial Institutions," *Bank of Japan Working Papers*, 04-E-04, 1–22.
- IORI, G., G. DE MASI, O. V. PRECUP, G. GABBI, AND G. CALDARELLI (2008): "A Network Analysis of the Italian Overnight Money Market," *Journal of Economic Dynamics and Control*, 32(1), 259–278.
- IORI, G., S. JAFAREY, AND F. G. PADILLA (2006): "Systemic Risk in the Interbank Market," *Journal of Economic Behavior & Organization*, 61(4), 525–542.
- IORI, G., R. N. MANTEGNA, L. MAROTTA, S. MICCICHE, J. PORTER, AND M. TUMMINELLO (2014): "Networked relationships in the e-MID Interbank market: A trading model with memory," *arXiv preprint arXiv:1403.3638*.
- IRWIN, S. H., D. R. SANDERS, R. P. MERRIN, AND OTHERS (2009): "Devil or angel? The role of speculation in the recent commodity price boom (and bust)," *Journal of Agricultural and Applied Economics*, 41(2), 377–391.
- JACKSON, M. O. (2008): *Social and Economic Networks*. Princeton University Press.
- JOHNSON, J. C., J. J. LUCZKOVICH, S. P. BORGATTI, AND T. A. SNIJDERS (2009): "Using social network analysis tools in ecology: Markov process transition models applied to the seasonal trophic network dynamics of the Chesapeake Bay," *Ecological Modelling*, 220(22), 3133 – 3140.
- JONES, C. M., AND P. J. SEGUIN (1997): "Transaction Costs and Price Volatility: Evidence from Commission Deregulation," *American Economic Review*, 87(4), 728–37.

- KAUFMANN, R. K. (2011): “The role of market fundamentals and speculation in recent price changes for crude oil,” *Energy Policy*, 39(1), 105 – 115.
- KIRMAN, A. (1993): “Ants, Rationality, and Recruitment,” *The Quarterly Journal of Economics*, 108(1), 137–56.
- KIRMAN, A. P., AND N. J. VRIEND (2001): “Evolving Market Structure: An ACE Model of Price Dispersion and Loyalty,” *Journal of Economic Dynamics and Control*, 25(3-4), 459–502.
- KNITTEL, C. R., AND R. S. PINDYCK (2013): “The simple economics of commodity price speculation,” Discussion paper, National Bureau of Economic Research.
- KÖNIG, M. D., C. J. TESSONE, AND Y. ZENOU (2009): “A Dynamic Model of Network Formation with Strategic Interactions,” CEPR Discussion Papers 7521, C.E.P.R. Discussion Papers.
- (2010): “From Assortative to Dissortative Networks: The Role of Capacity Constraints,” *Advances in Complex Systems*, 13(04), 483–499.
- KOSKINEN, J. H., AND T. A. SNIJDERS (2007): “Bayesian inference for dynamic social network data,” *Journal of Statistical Planning and Inference*, 137(12), 3930 – 3938.
- KYRIAKOPOULOS, F., S. THURNER, C. PUHR, AND S. W. SCHMITZ (2009): “Network and Eigenvalue Analysis of Financial Transaction Networks,” *The European Physical Journal B - Condensed Matter and Complex Systems*, 71(4), 523–531.
- LAZEGA, E., L. MOUNIER, T. SNIJDERS, AND P. TUBARO (2012): “Norms, status and the dynamics of advice networks: A case study,” *Social Networks*, 34(3), 323 – 332.
- LEE, J.-S., K.-I. GOH, B. KAHNG, AND D. KIM (2006): “Intrinsic degree-correlations in the static model of scale-free networks,” *The European Physical Journal B - Condensed Matter and Complex Systems*, 49(2), 231–238.
- LENGNICK, M., AND H.-W. WOHLTMANN (2012): “Agent-based financial markets and New Keynesian macroeconomics: a synthesis,” *Journal of Economic Interaction and Coordination*, pp. 1–32.

-
- LOSPINOSO, J., M. SCHWEINBERGER, T. SNIJDERS, AND R. RIPLEY (2011): “Assessing and accounting for time heterogeneity in stochastic actor oriented models,” *Advances in Data Analysis and Classification*, 5, 147–176.
- LUX, T. (1998): “The socio-economic dynamics of speculative markets: interacting agents, chaos, and the fat tails of return distributions,” *Journal of Economic Behavior & Organization*, 33(2), 143–165.
- (2009): “Rational forecasts or social opinion dynamics? Identification of interaction effects in a business climate survey,” *Journal of Economic Behavior & Organization*, 72(2), 638–655.
- (2011): “Network Theory is Sorely Required,” *Nature*, (469), 303.
- MADDALA, G. S. (1983): *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge University Press.
- MIKE, S., AND J. D. FARMER (2008): “An Empirical Behavioral Model of Liquidity and Volatility,” *Journal of Economic Dynamics and Control*, 32(1), 200–234.
- MILGRAM, S. (1967): “The Small-World Problem,” *Psychol. Today*, 2, 60–67.
- MONTAGNA, M., AND T. LUX (2013): “Hubs and Resilience: Towards More Realistic Models of the Interbank Markets,” Kiel working paper 1826, Kiel Institute for the World Economy.
- NEWMAN, M. (2010): *Networks - An Introduction*. Oxford University Press.
- NEWMAN, M. E. J. (2002): “Assortative Mixing in Networks,” *Phys. Rev. Lett.*, 89, 208701.
- (2003): “Mixing Patterns in Networks,” *Phys. Rev. E*, 67, 026126.
- 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.
- PARSONS, J. E. (2010): “Black gold and fool’s gold: speculation in the oil futures market,” *Economia*, 10(2), 81–116.
- PASTOR-SATORRAS, R., A. VÁZQUEZ, AND A. VESPIGNANI (2001): “Dynamical and Correlation Properties of the Internet,” *Phys. Rev. Lett.*, 87, 258701.

- PIRAVEENAN, M., M. PROKOPENKO, AND A. ZOMAYA (2012): "Assortative mixing in directed biological networks," *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)*, 9(1), 66–78.
- PLOBERGER, W., AND W. KRÄMER (1992): "The CUSUM test with OLS residuals," *Econometrica: Journal of the Econometric Society*, pp. 271–285.
- PRÉFONTAINE, J., J. DESROCHERS, AND L. GODBOUT (2010): "The Analysis Of Comments Received By The BIS On "Principles For Sound Liquidity Risk Management And Supervision",," *International Business & Economics Research Journal (IBER)*, 9(7).
- R DEVELOPMENT CORE TEAM (2012): "R: A Language and Environment for Statistical Computing," *R Foundation for Statistical Computing*.
- RADDANT, M. (2012): "Structure in the Italian overnight loan market," Discussion Paper 1772, Kiel Working Papers.
- REKA, A., AND BARABÁSI (2002): "Statistical Mechanics of Complex Networks," *Rev. Mod. Phys.*, 74, 47–97.
- RIPLEY, R., K. BOITMANIS, AND T. A. SNIJDERS (2013): "RSiena: Siena - Simulation Investigation for Empirical Network Analysis," *R package version 1.1-232*.
- ROBBINS, H., AND S. MONRO (1951): "A stochastic approximation method," *The Annals of Mathematical Statistics*, 22(3), 400–407.
- ROLL, R. (1989): "Price Volatility, International Market Links, and Their Implications for Regulatory Policies," *Journal of Financial Service Research*, 3, 211–246.
- SACHS, A. (2013): "Completeness, interconnectedness and distribution of interbank exposures – a parameterized analysis of the stability of financial networks," *Quantitative Finance*, forthcoming.
- SANDERS, D. R., K. BORIS, AND M. MANFREDO (2004): "Hedgers, funds, and small speculators in the energy futures markets: an analysis of the CFTC's Commitments of Traders reports," *Energy Economics*, 26(3), 425–445.
- SANDERS, D. R., AND S. H. IRWIN (2010): "A speculative bubble in commodity futures prices? Cross-sectional evidence," *Agricultural Economics*, 41(1), 25–32.

-
- (2011): “New Evidence on the Impact of Index Funds in U.S. Grain Futures Markets,” *Canadian Journal of Agricultural Economics/Revue canadienne d’agroeconomie*, 59(4), 519–532.
- SCHAEFER, D. R., J. M. LIGHT, R. A. FABES, L. D. HANISH, AND C. L. MARTIN (2010): “Fundamental principles of network formation among preschool children,” *Social Networks*, 32(1), 61 – 71.
- SCHWEINBERGER, M., AND T. A. SNIJDERS (2007): “Markov models for digraph panel data: Monte Carlo-based derivative estimation,” *Computational Statistics & Data Analysis*, 51(9), 4465–4483.
- SHAVIRO, D. (2012): “The Financial Transactions Tax versus (?) the Financial Activities Tax,” *NYU Law and Economics Research Paper No. 12-04*.
- SINGLETON, K. J. (2013): “Investor flows and the 2008 boom/bust in oil prices,” *Management Science*.
- SNIJDERS, T., AND M. VAN DUIJN (1997): “Simulation for statistical inference in dynamic network models,” in *Simulating social phenomena*, pp. 493–512. Springer Berlin Heidelberg.
- SNIJDERS, T. A., AND C. BAERVELDT (2003): “A multilevel network study of the effects of delinquent behavior on friendship evolution,” *Journal of mathematical sociology*, 27(2-3), 123–151.
- SNIJDERS, T. A., J. KOSKINEN, AND M. SCHWEINBERGER (2010): “Maximum likelihood estimation for social network dynamics.,” *The Annals of Applied Statistics*, 4(2), 567–588.
- SNIJDERS, T. A., G. G. VAN DE BUNT, AND C. E. STEGLICH (2010): “Introduction to stochastic actor-based models for network dynamics,” *Social Networks*, 32(1), 44 – 60.
- SNIJDERS, T. A. B. (1996): “Stochastic Actor-Oriented Dynamic Network Analysis,” *Journal of Mathematical Sociology*, 21, 149–172.
- (2001): “The Statistical Evaluation of Social Network Dynamics,” *Sociological Methodology*, 31(1), 361–395.
- SORAMAKI, K., M. L. BECH, J. ARNOLD, R. J. GLASS, AND W. BEYELER (2007): “The Topology of Interbank Payment Flows,” *Physica A*, (379), 317–333.

- STOLL, H. R., AND R. E. WHALEY (2010): "Commodity index investing and commodity futures prices," *Journal of Applied Finance*, 20(1), 7–46.
- STUMPF, M. P. H., C. WIUF, AND R. M. MAY (2005): "Subnets of Scale-free Networks are not Scale-free: Sampling Properties of Networks," *Proceedings of the National Academy of Sciences of the United States of America*, 102(12), 4221–4224.
- TANG, K., AND W. XIONG (2010): "Index investment and financialization of commodities," Discussion paper, National Bureau of Economic Research.
- TODA, H. Y., AND T. YAMAMOTO (1995): "Statistical inference in vector autoregressions with possibly integrated processes," *Journal of Econometrics*, 66(1-2), 225–250.
- TROSTLE, R. (2010): *Global Agricultural Supply and Demand: Factors Contributing to the Recent Increase in Food Commodity Prices*. DIANE Publishing.
- UPPER, C. (2011): "Simulation Methods to Assess the Danger of Contagion in Interbank Markets," *Journal of Financial Stability*, 7(3), 111–125.
- WATTS, D. J., AND S. H. STROGATZ (1998): "Collective Dynamics of Small-World Networks," *Nature*, 393(6684), 440–442.
- WEIDLICH, W., AND G. HAAG (1983): *Concepts and models of a quantitative sociology: The dynamics of interacting populations*. Springer, Berlin.
- WESTERHOFF, F. H., AND R. DIECI (2006): "The effectiveness of Keynes-Tobin transaction taxes when heterogeneous agents can trade in different markets: A behavioral finance approach," *Journal of Economic Dynamics and Control*, 30(2), 293–322.
- WILKINSON, I., AND L. YOUNG (2002): "On Cooperating: Firms, Relations and Networks," *Journal of Business Research*, 55(2), 123–132.
- XIE, Y., T. ZHOU, AND B. WANG (2008): "Scale-Free Networks Without Growth," *Physica A: Statistical Mechanics and Its Applications*, 387(7), 1683–1688.
- XULVI-BRUNET, R., AND I. M. SOKOLOV (2004): "Reshuffling Scale-Free Networks: From Random to Assortative," *Phys. Rev. E*, 70, 066102.
- YOUSSEFMIR, M., B. A. HUBERMAN, AND T. HOGG (1998): "Bubbles and Market Crashes," *Computational Economics*, 12(2), 97–114.

-
- ZAPATA, H. O., AND A. N. RAMBALDI (1997): “Monte Carlo Evidence on Cointegration and Causation,” *Oxford Bulletin of Economics and Statistics*, 59(2), 285–298.
- ZEILEIS, A., F. LEISCH, K. HORNIK, AND C. KLEIBER (2002): “strucchange: An R Package for Testing for Structural Change in Linear Regression Models,” *Journal of Statistical Software*, 7(2), 1–38.
- ZHANG, H., H. ZHAO, W. CAI, J. LIU, AND W. ZHOU (2010): “Using the k-core decomposition to analyze the static structure of large-scale software systems,” *The Journal of Supercomputing*, pp. 352–369.
- ZHOU, H. (2002): “Scaling Exponents and Clustering Coefficients of a Growing Random Network,” *Phys. Rev. E*, 66, 016125.

