

Essays in Nonlinear Dynamics
in Economics and Econometrics
with Applications to Monetary Policy and Banking

This thesis has been written within the European Doctorate in Economics - Erasmus Mundus (EDE-EM), with the purpose to obtain a joint doctorate degree in economics at the Faculty of Economics and Business at the University of Amsterdam and at the Department of Business Administration and Economics at the Bielefeld University.

Layout: Marcin Wolski

Cover design: Coördesign, Leiden

ISBN 978 90 5629 753 4

NUR 780

© Marcin Wolski, 2014

All rights reserved. Without limiting the rights under copyright reserved above, no part of this book may be reproduced, stored in or introduced into a retrieval system, or transmitted, in any form or by any means (electronic, mechanical, photocopying, recording or otherwise) without the written permission of both the copyright owner and the author of the book.

Essays in Nonlinear Dynamics
in Economics and Econometrics
with Applications to Monetary Policy and Banking

Marcin Wolski

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. D.C. van den Boom
ten overstaan van een door het college voor promoties
ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op vrijdag 27 juni 2014, te 16:00 uur

door

Marcin Wolski

geboren te Lublin, Polen

Promotiecommissie

Promotor

Prof. dr. Herbert Dawid (Universität Bielefeld)

Prof. dr. Cees G.H. Diks (Universiteit van Amsterdam)

Co-promotor

Prof. dr. Cars H. Hommes (Universiteit van Amsterdam)

Overige leden

Prof. dr. Ron J. Berndsen (Universiteit van Tilburg)

Dr. Kees Jan van Garderen (Universiteit van Amsterdam)

Dr. Sander van der Hoog (Universität Bielefeld)

Prof. dr. Franc Klaassen (Universiteit van Amsterdam)

Dr. Valentyn Panchenko (University of New South Wales)

Prof. dr. Jan Tuinstra (Universiteit van Amsterdam)

Dla Karoliny

Acknowledgements

This thesis would have not been written without the support from countless people whom I used to meet during this exciting intellectual journey.

I would like to thank my advisors; Cees for the thought provoking and inspiring conversations; that you have always found time for me (even when coming unexpectedly); and that you polished the ideas I used to bring to you in a messy pile of papers; Herbert for the motivating discussions and professional approach; that you have always supported my career plans and that you made Bielefeld my research home; Cars for opening to me the world of nonlinear economic dynamics and for asking inspiring questions when I was drifting away with my research concepts.

Huge credit goes to my colleagues at the International Monetary Fund, Franziska, Sophia and Karim, for broadening my horizons, stimulating research discussions and motivating debates on the nature of the global economy. The time spent at the Fund was a great research adventure and I feel that I learned tremendously from your experience and discussions even though I have not even touched the top of an iceberg of the topics we were covering.

Thank you to the members of my PhD committee, Franc, Jan, Kees Jan, Ron, Sander and Valentyn for the very detailed and useful feedback on this manuscript.

I would like to highlight the role of the European Doctorate in Economics - Erasmus Mundus (or EDEEM as it is often called), and the Erasmus Mundus Association, in completing this thesis. My most honest thank you for the support you gave to my ideas and for the coordination of my research path.

Peter Hall, from the University of Melbourne, must be singled out for his comments on the technical part of Chapter 3. At the same time all the numerical simulations, referred to over the following chapters, would have not been ready without the support from LISA Compute Cluster, maintained by the SURFsara team in Amsterdam.

Through all the universities I have visited while completing this thesis, the biggest credit shall arrive to the University of Amsterdam (CeNDEF in particular), Bielefeld University, Tilburg University and Warsaw School of Economics. Thank you for creating a stimulating and motivating research environment. My sincere gratitude goes Dan, Marco, Tomasz, David, Te, Domenico, Michiel, Juanxi, Peter, Alek, Florian, Thom, Manu, Marcin and all of the seminar participants and research colleagues who contributed to the discussion explored in this thesis.

Because of many hard working and sleepless nights, Kosma, Kaman, Frank, Andreea and Iulia, thank you for keeping me down to the ground when I was flying away with my PhD. The Amsterdam Football Group and my UvA climbing team deserve my most honest thanks for keeping my mind and body fresh while writing this thesis.

Thank you to my Polish friends Skot and Olga, Bubu and Kasia, Mateo, Madej, Obiad, Pogoda and Kaśka, Cudny, Sosna and Ania, Wiurek, Kuba, Szafa, Marzi, Sawio and Łysy. Cheers for the entertainment and hilarious working environment. You all are parts of this thesis.

I shall thank my family; my parents, Babcia, Gośka and Marcin for the ongoing faith in what I have been doing and for setting a benchmark for top quality research.

Last but not least, many thanks to Karolina who has never given up on me; for her love and patience; for her suggestions and spelling control and that she was always my first reviewer. I am a happy man beside her.

Marcin Wolski

March 2014

Contents

1	General Introduction and Thesis Outline	1
2	Monetary Policy, Banking and Heterogeneous Agents	11
2.1	Introduction	11
2.2	The model	13
2.2.1	The IS curve	15
2.2.2	The Phillips curve	18
2.2.3	The banking sector curve	19
2.3	The influence of heterogeneity	21
2.4	Numerical analysis	23
2.4.1	Formation of expectations	23
2.4.2	Calibration and numerical results	24
2.5	Conclusions and discussion	27
Appendix 2.A	Baseline derivation	30
Appendix 2.B	The influence of heterogeneous agents	33
Appendix 2.B.1	The heterogeneous IS curve	33
Appendix 2.B.2	The heterogeneous Phillips curve	35
Appendix 2.B.3	The heterogeneous banking sector curve	36
Appendix 2.C	Model dynamics	37
3	Nonlinear Granger Causality - Guidelines for Multivariate Analysis	39

CONTENTS

3.1	Introduction	39
3.2	Asymptotic properties of the DP test	41
3.2.1	The dimensionality problem	44
3.3	Data sharpening as a bias reduction method	45
3.3.1	Bandwidth selection	48
3.3.2	Performance of the DS in Granger causality setting	49
3.4	Nonlinear Granger causality in the US grain market	51
3.5	Conclusions and discussion	56
Appendix 3.A	Asymptotics of the sharpened test (Corollary 3.3.1)	58
Appendix 3.B	Application of bias reduction	60
Appendix 3.C	Illustration of the empirical results	63
4	Exploring Nonlinearities in Financial Systemic Risk	69
4.1	Introduction	69
4.2	Methodology of NCoVaR	73
4.2.1	Optimal bandwidth	79
4.3	Numerical simulations	80
4.3.1	Performance of the NCoVaR test	82
4.4	Assessing financial systemic risk	86
4.5	Conclusions and discussion	90
Appendix 4.A	Asymptotic properties of test statistic (Theorem 4.2.1)	92
Appendix 4.A.1	Dependence	93
Appendix 4.B	Optimal bandwidth sequence (Corollary 4.2.1)	94
Appendix 4.C	Data description and results	95
5	Do Safe Havens Make Asset Markets Safer?	101
5.1	Introduction	101
5.2	Data	104

5.3	Defining safe havens	108
5.4	Mapping the network of sovereign bond yields and bank equity	111
5.5	Modeling shock propagation	116
5.6	Feedback loops in shock propagation	118
5.6.1	Sovereign bond yield shock	118
5.6.2	Bank equity price shock	122
5.7	The role of safe havens in shock propagation	124
5.7.1	Sovereign bond yield shock	125
5.7.2	Bank equity price shock	129
5.8	Conclusions and issues for further research	131
	Appendix 5.A Network graphs	132
	Appendix 5.B Shock propagation mechanism	137
6	Summary	139
	Bibliography	156
	Samenvatting (Summary in Dutch)	157

Chapter 1

General Introduction and Thesis Outline

“As a policy maker during the crisis, I found the available models of limited help. In fact, I would go further: in the face of the crisis, we felt abandoned by conventional tools.”

— Jean-Claude Trichet, President of the ECB, 18 November 2010

As pointed out by Jean-Claude Trichet, the financial crisis from the years 2007-2009 revealed imperfections in existing economic modeling techniques. The standard Dynamic Stochastic General Equilibrium (DSGE) models, widely used by central bankers and policy makers around the world, proved not to capture the intriguing complexity of the global financial sector nor could they have reproduced the boom and bust scenarios which are observed in the real world (Buiter, 2009). In fact, paraphrasing Charles Goodhart from the Bank of England, the standard central banking “excludes everything that [we shall be] interested in”.

The failure of these models might be largely attributed to several simplifying assumptions which they are built upon. To the most widely criticized belong the Rational Expectations Hypothesis (REH) and representative agent structure (Frydman and Goldberg, 2007), linear dependencies (Hommes, 2013) and the absence of the well-characterized financial sector (Bernanke et al., 1999; Tovar, 2008). Those shortcomings used to be neglected for many years as the global economy was growing steadily with little fluctuations, making the DSGE models powerful tools which provide a coherent framework for policy discussion and analysis. The beauty

of their simplicity turned, however, into their biggest nightmare as the recent financial crisis erupted. Their forecasting accuracy, highlighted on pre-crisis samples (see e.g. Christoffel et al. (2010)), in terms of Root Mean Square Error (RMSE), proved to be no better than naive forecasts (Edge and Gürkaynak, 2010).

As pointed out by Tovar (2008) “[d]espite the rapid progress made in recent years, at their current stage of development, these [DSGE] models are not fully ready to accomplish all what is being asked from them”. The goal of this thesis is therefore threefold. Firstly, it contributes to the ongoing debate on economic modeling by investigating economic dynamics under heterogeneous market structures. Secondly, it proposes econometric concepts of assessing the influence of nonlinear profiles in economic relationships. Thirdly, it studies the role of the network structures in the shock propagation mechanisms of the global economy. Because of their extraordinary relevance in the real world, a lot of attention is being paid to banking and financial markets.

The role of expectations

The general equilibrium models, like the Real Business Cycle (RBC), developed by Kydland and Prescott (1982), or the new Keynesian framework, pioneered by Clarida et al. (1999) and Woodford (2003), assume at the micro level the utility-maximizing consumers, profit-maximizing companies and market clearing for all goods at all dates in all markets (Hommes, 2013). A subtle assumption of rational expectations helps to solve the models analytically and derive the macro behavior directly from the micro founded principles.

REH has a long history in economics, ranging back to the seminal papers of Muth (1961) and Lucas (1972). It states that on average economic agents act as if they could predict future outcomes perfectly. That means that the incorrect expectations cancel out with each other or are being eliminated by natural selection (Friedman, 1953) and at the aggregate level one observes perfectly accurate foresight. In mathematical terms this is parallel to equalizing a variable today

to its expected value in the market clearing equilibrium tomorrow (Garcia, 2011). Exploiting the mathematical courtesy of REH, studying the macro behavior is as easy as looking at the one representative (or average) agent and associating it with the aggregate decision making process, as in equilibrium everybody shall have the same model consistent expectations without any systematic errors.

Nevertheless, REH oversees the possibility that the incorrect expectations might be self-enforcing instead of being self-mitigating. Indeed, if bad decisions today lead to even worse decisions tomorrow this feedback mechanism might be of great importance for the aggregate economic dynamics (Frydman and Goldberg, 2007), driving the system further away from the fundamentals and creating possible bubbles. This type of feedback structure has been already recognized in the literature; for instance Soros (2003) refers to it as *vicious cycles* and Brunnermeier (2009) calls it by simply *spirals*. Frydman and Goldberg (2007) highlight that REH is very susceptible to this type of expectational dynamics.

Although, in the literature there is no consensus on how to represent economic expectations, their role and especially the influence of their interactions are an extremely important aspect of modern economic modeling (Stanislawska and Tomczyk, 2010; Evans and Honkapohja, 2001; Hommes, 2013). Recently, however, one has observed a paradigm shift from REH to the ideas of bounded rationality and heterogeneous expectations (see e.g. Conlisk (1996); Brock and Hommes (1997); Branch (2004); Branch and McGough (2009)). The reasoning behind boundedly rational agents is attributed to Simon (1955, 1957). Simon points out that because of the lack of information or limited cognitive and computing capacities, individuals might not be perfect forecasters nor optimizers but rather they tend to use simple heuristics in their decision making process when acting under uncertainty. This view has been widely confirmed in laboratory experiments (Tversky and Kahneman, 1974), proving that in reality these simple rules of thumb might lead to significant biases so that the incorrect expectations do not necessarily cancel out as suggested by REH.

This in fact puts in conflict the idea of a representative agent structure, widely present in

DSGE models. In a situation where the agents are boundedly rational they do not have to share the same information set nor use the same heuristics in forming their expectations. The *ex ante* individual prediction might thereof not coincide with the *ex post* aggregate realizations but certainly they affect them. As a consequence, the beliefs of some agents might indirectly influence the beliefs of others so that the economy becomes an expectational feedback system (Hommes, 2013). Heterogeneous expectations have been confirmed both in laboratory experiments (Hommes, 2011, 2013) and in the survey studies (Carroll, 2003; Mankiw et al., 2003) and tend to be an intriguing and thought-provoking phenomenon for economic modeling.

Heterogeneous expectations, together with boundedly rational agents, proved to generate complex structures and interesting nonlinear economic dynamics in the DSGE framework (see e.g. Branch and McGough (2010) or Massaro (2013)). Therefore, they might be an alternative to the standard model assumptions, pointing out a direction for future developments. In this thesis, Chapter 2 is fully devoted to these intriguing phenomena in the DSGE new Keynesian framework with an active banking sector.

Nonlinear dynamics

The standard linear framework fits nicely in globally stable systems which are close to equilibrium. It performed tremendously well from the mid-1980s till 2006, a period often referred to as the Great Moderation, when the global economy was at a stable growth path. In the absence of large shocks, the system was settling down to its local equilibrium and the concerns arising from possible threats and risks were underestimated by both financial markets and macroprudential authorities (Blinder, 2013).

As it is known in the mathematical sciences, the dynamics around a steady state might be approximated by log-linearization. However, moving further away from that point, log-linearization produces less accurate approximations. Consequently, the linear economic models could misperceive the risks which are further away from a given equilibrium point. In fact this

was clearly visible when the US housing bubble collapsed in years 2006/2007 materializing all the risk which the world economy had been accumulating during the Great Moderation (Blinder, 2013). Nobody had expected such a big shock nor the continuing recession in the majority of advanced economies.

Linear models offer attractive mathematical properties, making them relatively easy to solve analytically. These simplifications, however, might have not kept up with the changes in the globalized and heavily digitalized economy. As pointed out by Alan Blinder, the former Vice Chairman of the Board of Governors of the Federal Reserve System, in the years before the crisis “the complexity went amok” (Blinder, 2013). Because of their design, purely linear models cannot capture the sophisticated and complex nature of the modern financial system. The need for new (nonlinear) analytical methods has been therefore widely signalized by professionals (Buiters, 2009).

The role of the financial sector and monetary policy

The importance of the financial sector (often referred to as simply banking) in economic modeling has already been recognized and included in more sophisticated models. Nevertheless, the standard RBC and the new Keynesian models are built around the Efficient Market Hypothesis (EMH), in which no financial disequilibrium is possible (Krugman, 2009). The commonly used view among practitioners highlights the inevitable link between the real economy and its financial side, especially when the presence of the latter provokes frictions and market imperfections (Bernanke et al., 1999), or may even cause significant real disturbances (Blinder, 2013).

The topic of financial frictions has attracted a lot of attention recently (Brunnermeier, 2009). Nevertheless, the recent developments in financial engineering and accounting, like emergence of Structured Investment Vehicles (Taboada, 2010), heavy leverage (Blinder, 2013), novel financial products (Datz, 2013) and global exposures and imbalances (International Monetary Fund, 2013), made it more complex in nature not only for regulators and financial authorities but also

for financial markets themselves (Datz, 2013).

As a consequence, in order to stabilize the markets and to bridle the financial complexity in the aftermath of the crisis 2007-2009, a huge mandate was given to central banks in advanced economies, like Federal Reserve, European Central Bank or Bank of Japan (International Monetary Fund, 2013). The role of standard monetary policy, i.e. stabilization of inflation dynamics (or in the US also the production level) by controlling the nominal short-term interest rates (Woodford, 2003), has evolved into something often referred to as *modern monetary policy*. Under the latter, central banks are allowed to manipulate long-term interest rates and bail-out troubled markets, or more generally as Mario Draghi, the President of the ECB, famously pledged “[to do] whatever it takes”. The implications put central bankers and the modern monetary policy into an urgent need for better tools, designed to capture the complex dynamics of the global economy. This is why the ideas presented in this thesis are assessed through a prism of monetary policy and banking.

Thesis outline

The methods developed and applied in this thesis aim to contribute to the ongoing discussion on the fascinating, rapidly changing and *primo loco* highly nonlinear profile of the financial world, being a potentially attractive standpoint for policy makers and practitioners. Chapter 2 studies the implications of a presence of boundedly rational agents in a monetary policy framework with an active banking sector. Chapters 3 and 4 develop econometric tests of studying nonlinear Granger (1969) causal relations in two different settings. Chapter 5 is a result of my stay at the International Monetary Fund (IMF) in the Summer of 2013 and presents an application of the network modeling to the global banking sector and sovereign bond market and explores the role of safe havens in shock propagation mechanism. Chapters 2-4 are published as working papers at the National Bank of Poland and Center for Nonlinear Dynamics in Economics and Econometrics (CeNDEF) at the University of Amsterdam; Wolski (2013b) is based on Chapter 2, Diks

and Wolski (2013) is based on Chapter 3 and Wolski (2013a) is based on Chapter 4. Chapter 5, co-authored by Franziska Ohnsorge and Y. Sophia Zhang, is forthcoming as an IMF working paper. The ideas contained in this thesis aim at encouraging a thought-provoking discussion on the nature of nonlinear structures in economic dynamics and econometrics and shall not be associated with views of any of the aforementioned institutions nor their policies.

Chapter 2 investigates the phenomenon of heterogeneous expectations, analyzing their role in monetary policy conduct with an active banking sector. In addition to fundamentalists, we assume a constant fraction of boundedly rational agents who use simple heuristics to form their expectations. We focus on two types of heuristics which are most commonly referred to throughout the literature (Hommes, 2013), i.e. adaptive and extrapolative expectations. Both assume that future realizations depend on the past performance of particular variables, however, the former assumes that the influence of past realizations decreases over time whereas the latter manifests the opposite. The impact of those biased beliefs is studied in the aggregate economy framework with an active banking sector, originally developed by Goodfriend and McCallum (2007). We first show that the presence of the banking sector changes the determinacy structure of the system and, depending on the heuristics used, the presence of boundedly rational agents might have either stabilizing or destabilizing effect. In particular, when boundedly rational agents have extrapolative expectations, the range of the stable (determinate) monetary policy instruments is narrowed.

In Chapter 3 we propose an extension of the nonlinear Granger causality test, originally introduced by Diks and Panchenko (2006). We show that the basic test statistic lacks consistency in the multivariate setting. The problem is the result of the kernel density estimator bias, which does not converge to zero at a sufficiently fast rate when the number of conditioning variables is larger than one. In order to overcome this difficulty we apply the data-sharpening method for bias reduction (Hall and Minnotte, 2002). We then derive the asymptotic properties of the sharpened test statistic and we investigate its performance numerically. We conclude with an empirical application to the US grain market, as it creates an ideal environment to test our

methodology. Chapter 3 does not exploit the financial markets explicitly and might be treated as a general introduction to the topics covered in the Chapter 4. Nevertheless, nonparametric Granger causality tests have been widely applied to financial time series (for instance to exchange rates in Bekiros and Diks (2008a) and to crude oil prices in Bekiros and Diks (2008b)) so that one may easily extend our reasoning to a different financial setting. In fact, Chapter 4 is closely related to Chapter 3 and raises the discussion on nonparametric Granger causality testing to the financial environment.

More specifically, Chapter 4 proposes a new methodology of assessing the effects of individual institution's risk on the others and on the system as a whole. We build upon the Conditional Value-at-Risk approach. However, we introduce explicit Granger causal linkages and we account for possible nonlinearities in financial time series. Conditional Value-at-Risk-Nonlinear Granger Causality, or NCoVaR as we call it for simplicity, has regular asymptotic properties which makes it particularly appealing for practical applications. We test our approach empirically and assess the contribution of the euro area financial companies to the overall systemic risk. We find that only a few financial institutions pose a serious *ex ante* threat to systemic stability risk, whereas, given that the system is already in trouble, there are more institutions which hamper its recovery. Moreover, we discover non-negligible nonlinear structures in the systemic risk profile of the euro zone.

In Chapter 5 we create a network of bilateral correlations of changes in sovereign bond yields and individual bank equity price changes. We study the nature and the evolution of this network in the years 2000-2013. We show that, in this context, safe havens have an intuitive representation as countries in which changes in sovereign bond yields and bank equity prices are positively correlated. Safe havens, however, have one additional feature, i.e. their asset prices are highly correlated with those of other countries making them hubs for capital flows. We investigate how these two properties of safe havens have affected the propagation of bank and sovereign shocks in our asset price network since 2000, in a simple shock propagation framework. On balance, we find that the presence of safe havens has amplified shock propagation.

Chapter 6 concludes and offers some ideas for future research on nonlinear dynamics in economics and econometrics.

Each chapter is a self-contained manuscript, with separate introduction, summary and appendices, and might be read independently from other chapters. For the reader's convenience, the common bibliography is collected at the end of the thesis. A digital copy of these pages can be found in the online libraries of the Universiteit van Amsterdam (www.uba.uva.nl) and Universität Bielefeld (www.ub.uni-bielefeld.de).

Chapter 2

Monetary Policy, Banking and Heterogeneous Agents

2.1 Introduction

The need for a framework which would incorporate financial frictions in DSGE models was stressed long before the 2007-2009 financial crisis (Bernanke and Gertler, 1989; Kiyotaki and Moore, 1997). The body of literature in this topic has grown substantially thereafter, bringing significant changes to monetary policy conduct (Rotemberg and Woodford, 1997; Woodford, 2003). It is surprising, as argued by Goodfriend and McCallum (2007) and Casares, Miguel and Poutineau (2010), that the role of the banking sector was left unexplored in the monetary policy analysis until recently.

The framework used in this study clarifies this oversight. Firstly, by introducing profit-maximizing bankers at the micro level, one may explicitly study the impact of their individual behavior on the macro aggregates. Secondly, the differentiation of the capital market allows to investigate the relationship between various types of interest rates (Goodfriend, 2005). Thirdly, by having government bonds which serve for collateral purposes, one observes the direct influence of public policy on the monetary aggregates.

Most noticeably, however, a banking sector *per se* is an important, if not the most important part of each economy (Levine, 1997). Since it is a general source of liquidity, its problems may easily spread over the other sectors, bringing them down eventually. Especially, the recent history proves that banking sector disturbances might result in sovereign crises, as recently took place in the euro zone (Grammatikos and Vermeulen, 2012). Therefore, a detailed study of the banking sector's role in the monetary framework is required in order to (i) understand its transmission mechanism and (ii) endow the monetary authorities with the sufficient preventive tools.

The goal of this chapter is twofold. Firstly, we assess the determinacy properties of different monetary policies in the DSGE model with a banking sector of Goodfriend and McCallum (2007). The model is built within the standard new Keynesian framework where the aggregate dynamics is a direct consequence of individual utility maximizing behavior of forward-looking agents. Secondly, we relax the assumption of agents' homogeneity and investigate how the presence of the backward-looking (or boundedly rational after Hommes (2013)) agents influences the determinacy of the equilibrium. We introduce agents' heterogeneity at the micro level, which means that each agent is solving the individual optimization problem simultaneously. It is an important distinction from a variety of models which neglect this aspect and allow for agents' heterogeneity at the macro level only. Clearly, such a concept violates the Subjective Expected Utility (SEU) theory and in our view is inappropriate. Instead, we follow the classical approach where the macro behavior is a direct consequence of agents' micro optimal plans.

The latter part of this study is motivated by a growing body of research which shows explicitly that agents differ in forming expectations. This phenomenon was confirmed by both survey data analysis (Carroll, 2003; Mankiw et al., 2003; Branch, 2004) as well as laboratory experiments with human subjects (Hommes et al., 2005; Assenza et al., 2011; Hommes, 2011; Pfajfar and Zakelj, 2011). The heterogeneity among agents was proved to have important implications on the determinacy properties in the new Keynesian models (Branch and McGough, 2009; Massaro, 2013). We follow this approach and assess its implication within the framework

with a banking sector.

This chapter is organized as follows. Section 2.2 describes the workhorse model and discusses the implications of the banking sector on monetary policy conduct. In Section 2.3 we relax the assumption of a representative agent structure and introduce boundedly rational backward-looking agents. Section 2.4 presents the numerical results and Section 2.5 concludes.

2.2 The model

In this section we develop the workhorse version of the model. Since the complete derivation, with the first order conditions and aggregation, is described in detail in the original paper of Goodfriend and McCallum (2007), we skip it in the main part of this text. However, for the reader's convenience, the complete derivation is given in Appendix 2.A.

The model space consists of a continuum of farmers who provide labor supply to the production and banking sectors at the same time t (n_t and m_t , respectively). Additionally, each farmer manufactures a differentiated product and sells it in the monopolistically competitive environment. As in the standard new Keynesian framework, it is assumed that only a fraction $(1 - \omega)$ of all farmers can adjust their prices fully flexibly. The remaining part takes the prices from the previous period (Calvo, 1983). Given these conditions, the goal of each farmer is to maximize her expected utility, which is a linear combination of consumption and leisure, over the infinite horizon.

In the utility maximization problem, each farmer has to take into account three constraints: (i) the budget constraint, (ii) the production constraint and (iii) the banking constraint. The first of these is the standard intertemporal budget constraint which ensures that the net income and bond/money holdings in one period are being transmitted to the next period. The second constraint is a direct consequence of the production technology, which in this case is of the Cobb-Douglas type. Assuming market clearing, the production (Y_t) in each period is the consequence of the amount of capital (K_t) and labor (n_t^d) involved, corrected for their output

elasticities: η and $(1 - \eta)$, respectively. The banking constraint assumes that the level of consumption (C_t) has to be rigidly related to the level of deposits held at a bank. One may view this as if all the transactions were being facilitated through the banking sector and each agent may consume a part V of her wealth only. A bank is then allowed to use $(1 - rr)$ fraction of the deposits to produce loans using the Cobb-Douglas production function with collateral (col_t) and labor (m_t^d) as production factors and α and $(1 - \alpha)$ being the output elasticities. The collateral consists of two parts, i.e. the discounted level of real bond holdings $B_{t+1}/(P_t^A(1 + r_t^B))$, with P_t^A being the aggregate price level and r_t^B the interest rate on bonds, and real level of capital $q_t K_{t+1}$, corrected for the inferiority of capital to bonds for collateral purposes, v . The last term results from the fact that bonds, contrary to capital goods, do not require substantial monitoring effort in order to verify their market value (Goodfriend and McCallum, 2007).

Such a banking sector setting captures several important aspects of financial intermediation. Firstly, it enters the consumer utility maximization problem at the micro level. Secondly, it builds a clear link between households and a production sector. Thirdly, because of its dependence on governmental securities, it comprises the monetary policy transmission mechanism (through the repo market).

There are two main simplifications of the original model. Firstly, we abstract from the capital shocks in the loan production function. We assume that the capital level is at its steady state level and the productivity shocks are transmitted through the labor channels only. This simplification does not affect the final results as in the determinacy analysis the stochastic terms do not play a role (Blanchard and Kahn, 1980). Secondly, we assume a zero tax rate. Eventually, the role of government is narrowed to issuing bonds in each period at some exogenously given level, and paying the interest.

Given the specification above, we may now turn to derivation of three model equations: the Investment-Savings (IS) curve, the Phillips curve and the banking curve. The first two of these build the standard new Keynesian model. The last one is the direct consequence of the presence of the banking sector and describes its role in the aggregate dynamics explicitly.

2.2.1 The IS curve

The model implies the presence of two Lagrange multipliers: λ_t for the budget constraint and ξ_t for the production constraint. They represent the shadow values, or the utility gains, of unit values of consumption and production respectively (Casares, Miguel and Poutineau, 2010). In particular, from the banking labor demand optimality condition we know that

$$\lambda_t^i = \frac{\xi_t^i}{\varphi_t^i} = \frac{\frac{\phi}{C_t^i}}{1 + \left(\frac{1-rr}{V}\right) \chi_t^i}, \quad (2.1)$$

where φ_t^i is the individual marginal production cost, ϕ is the utility weight on consumption and we explored the fact that the χ_t^i might be viewed as the individual marginal loan management cost, or simply the marginal banking cost (Goodfriend and McCallum, 2007; Casares, Miguel and Poutineau, 2010)¹. To put it more formally, imagine the cost minimization problem of a representative bank in a situation without collateral cost. The total cost function may be rewritten as $TC_t = m_t^d w_t$, where w_t is the real wage. The minimization problem includes the loan production constraint with a Lagrangian multiplier (here perceived as a marginal cost (Walsh, 2010)), denoted by χ_t . The first order condition implies that $\chi_t = V w_t m_t^d / ((1 - rr)(1 - \alpha)C_t)$. In fact, χ_t^i is parallel to the individual marginal production cost that is being often referred to in the standard new Keynesian framework (Walsh, 2010). One may view that as a general variable describing the situation in the banking sector, i.e. the higher it is the less effective the loan management is. As it is shown later, this variable is of crucial importance as it becomes a link between a standard new Keynesian model and the banking system.

Eq. (2.1) gives the first overview of the model behavior. Firstly, the shadow value of production equals the shadow value of consumption corrected for the marginal production cost. In other words, additional consumption has to turn up in either increased production or decreased production costs. Secondly, λ_t is the marginal utility of consumption corrected for the marginal banking cost. Put differently, each additional unit of consumption requires more deposits, which

¹We include superscript i to underline the individual level of the relationship which is explored in detail later. In the representative agent structure it may be omitted as every agent behaves the same.

may be raised at the cost χ_t . It is straightforward to notice that the lower the marginal banking cost, the relatively cheaper the additional consumption. On the other hand, a highly inefficient banking sector limits the incentives to increase consumption.

Substituting Eq. (2.1) into the bond optimality condition, we finally arrive at the familiar Euler equation

$$\beta E_t^i \left(\frac{\frac{\phi}{C_{t+1}^i}}{1 + \left(\frac{1-rr}{V}\right) \chi_{t+1}^i} \right) = \frac{\frac{\phi}{C_t^i}}{1 + \left(\frac{1-rr}{V}\right) \chi_t^i} (1 + E_t^i \pi_{t+1}) \left(\frac{1 - \frac{1-rr}{V} \chi_t^i \Omega_t^i}{1 + r_t^B} \right), \quad (2.2)$$

where $(1 + E_t^i \pi_{t+1}) = P_{t+1}^A / P_t^A$ is the inflation rate and $\Omega_t^i = \alpha C_t^i / col_t^i$.

Following Goodfriend (2005), let us introduce a one-period default-free security with the nominal rate denoted by r_t^T . Since we additionally assume that it cannot serve for collateral purposes, r_t^T represents a pure intertemporal rate of interest and serves as a benchmark for other interest rates. From the agent optimization problem, we know that $1 + r_t^{i,T} = E_t^i \lambda_t^i P_{t+1}^i / (\beta \lambda_{t+1}^i P_t^i)$ so that it includes the discounted difference between expected changes in shadow prices and actual prices. An important distinction is that the pricing of this fictitious security is done at the individual level which is not strange given its completely artificial and agent-dependent nature. Eventually, the last term of Eq. (2.2) might be rewritten as the reciprocal of $(1 + r_t^{i,T})$.

At the same time, let us assume that each bank can obtain funds from the interbank market at the common rate r_t^{IB} . It can then loan them to agents at the rate $r_t^{i,T}$. The profit maximization of a bank implies that the marginal costs of obtaining funds has to be equal their marginal profit so that

$$(1 + r_t^{IB})(1 + \chi_t^i) = (1 + r_t^{i,T}). \quad (2.3)$$

Inserting Eq. (2.3) into Eq. (2.2) and taking the log approximation around the steady state we have

$$\hat{Y}_t^i = E_t^i \hat{Y}_{t+1}^i + \left(\frac{1-rr}{V} \right) E_t^i \tilde{\chi}_{t+1}^i - \left(\frac{1-rr}{V} + 1 \right) \tilde{\chi}_t^i - (\hat{r}_t^{IB} - E_t^i \pi_{t+1}), \quad (2.4)$$

where tildes and hats denote deviations and percentage deviations from the steady state, respectively, and we explored the market clearing condition².

As in the standard new Keynesian framework, we define the potential output as the output under completely flexible prices and wages (Walsh, 2010). We additionally assume that in such a situation there is a fixed proportion between employment in the production and banking sector, $n_t^d \propto m_t^d$. Following Walsh (2010), price flexibility implies that all agents can adjust their prices immediately, which gives that the marginal cost of production φ_t is equal $(\theta - 1)/\theta$ across all individuals, where θ is the elasticity of substitution between consumption goods. The labor optimality condition implies that the real wage has to be equal the marginal rate of substitution between leisure and consumption, corrected for the presence of the banking sector. Combining the above-mentioned points with Eq. (2.1) and the production constraint, we finally get that under flexible prices and wages, the supply of labor of each individual is fixed so that if the capital stock is in the steady state (as we assume throughout the model) the log deviations of the potential product depend only on exogenous disturbances, $\hat{Y}_t^f = (1 - \eta)(A1_t - \bar{A}1)$. Subtracting them from both sides of Eq. (2.3) and omitting the i superscript, we finally arrive at the aggregate IS curve corrected for the presence of a banking sector

$$x_t = E_t x_{t+1} + \left(\frac{1 - rr}{V} \right) E_t \tilde{\chi}_{t+1} - \left(\frac{1 - rr}{V} + 1 \right) \tilde{\chi}_t - [\hat{r}_t^{IB} - E_t \pi_{t+1}] + u_t, \quad (2.5)$$

where $x_t = \hat{Y}_t - \hat{Y}_t^f$ is the output gap measure and u_t is the disturbance term that depends only on exogenous productivity shocks.

It is straightforward to notice that when skipping the banking sector variables from Eq. (2.5) we obtain the standard new Keynesian IS curve. What is important, is that the aggregate dynamics is affected not only by the current, but also expected future values of the banking variables. In other words, the way the agents form their expectations about future banking sector conditions seems to play a role in determining current production. The impact of the banking sector is limited by (i) the reserve requirement, rr , and (ii) the proportion of consumption that has to

²Following literature, we take the zero inflation steady state.

be covered by deposits, V . Clearly, the lower the minimum reserve requirement, the larger the loan production so that the importance of the banking sector increases, *ceteris paribus*. At the same time, if the consumption-to-deposits coverage ratio is large, relative size of the banking sector is smaller so that its impact decreases.

2.2.2 The Phillips curve

The model allows us also to derive the explicit formula for the Phillips (or Aggregate Supply) curve. We know that all the farmers share the same production technology and face the same constant demand elasticities. We know from the Calvo lottery that a fraction ω of agents cannot adjust their prices in a given period t . Profits of some future date $t+k$ are affected only if an agent did not receive a chance to adjust prices between t and $t+k$. Therefore, the probability of having lower expected profits in period k is ω^k . Having pointed that out, the price optimality condition has to be corrected for the nominal price rigidities in the long run and by iterating forward it might be viewed as

$$E_t^i \sum_{k=0}^{\infty} \beta^k \omega^k \left[(1-\theta) \left(\frac{P_t^i}{P_{t+k}^A} \right) + \theta \left(\frac{\xi_{t+k}^i}{\lambda_{t+k}^i} \right) \right] \left(\frac{1}{P_t^i} \right) \left(\frac{P_t^i}{P_{t+k}^A} \right) C_{t+k}^A = 0. \quad (2.6)$$

Solving for optimal price setting, we arrive at

$$\frac{P_t^i}{P_t^A} = \frac{E_t^i \sum_{k=0}^{\infty} \beta^k \omega^k C_{t+k}^A \varphi_{t+k}^i \left(\frac{P_{t+k}^A}{P_t^A} \right)^\theta}{E_t^i \sum_{k=0}^{\infty} \beta^k \omega^k C_{t+k}^A \left(\frac{P_{t+k}^A}{P_t^A} \right)^{\theta-1}}, \quad (2.7)$$

where $\varphi_t^i = \xi_t^i / \lambda_t^i$ is the individual marginal production cost (Goodfriend and McCallum, 2007).

Skipping the i superscript and taking a log approximation, after some algebra we obtain³

$$\pi_t = \beta E_t \pi_{t+1} + \kappa \hat{\varphi}_t, \quad (2.8)$$

³For a detailed derivation see the appendix of Chapter 8 from Walsh (2010).

where $\kappa = \frac{(1-\omega)(1-\beta\omega)}{\omega}$. We further explore the fact that given the Cobb-Douglas production function, the steady state log deviations of the marginal production cost might be viewed as an output gap measure (Goodfriend and McCallum, 2007). Finally, we arrive at the standard new Keynesian Phillips curve

$$\pi_t = \beta E_t \pi_{t+1} + \kappa x_t. \quad (2.9)$$

What is important is that the situation in the banking sector does not affect the inflation level directly but only through the consumption channel. The absence of the banking variables in Eq. (2.9) is a consequence of the banking sector specification. The level of consumption is rigidly related to the amount of deposits in the banking sector. Therefore, changes in the banking sector would result in a different deposit level, which would shake the consumption eventually. However, there is no direct link to the inflation in the meantime.

2.2.3 The banking sector curve

Since the presence of the banking sector affects the aggregate evolution of the IS and (indirectly) Phillips curves, it is also necessary to describe its dynamics. Observing that $\varphi_t = q_t K_t / (\eta C_t)$, the capital optimality condition implies

$$1 - \frac{v(1-rr)}{V} \Omega_t^i \chi_t^i = \beta(2-\delta) E_t \left[\frac{\left(1 + \left(\frac{1-rr}{V}\right) \chi_t^i\right) \varphi_{t+1}^i}{\left(1 + \left(\frac{1-rr}{V}\right) \chi_{t+1}^i\right) \varphi_t^i} \right]. \quad (2.10)$$

Observe that the LHS of Eq. (2.10) is almost identical with the numerator of the last term in Eq. (2.2). The only difference comes from the inferiority of capital to bonds for collateral purposes, v . Applying the same interest rate reasoning to the log approximation of the LHS of Eq. (2.10), we see that $-v(1-rr)\Omega_t^i \chi_t^i / V = -v(r_t^{IB} - r_t^B + \chi_t^i)$. Since the interbank rate r_t^{IB} and the government bond rate r_t^B are both short-term rates, they should be close to each other around the equilibrium (Goodfriend and McCallum, 2007). Additionally, given the fact that v is relatively small, we neglect the influence of $v(r_t^{IB} - r_t^B)$. Eventually, after taking the deviations

from the steady state of Eq. (2.10), iterating forward and skipping the i superscript, we get

$$\left(v + \frac{1 - rr}{V}\right) \tilde{\chi}_t = \frac{1 - rr}{V} E_t \tilde{\chi}_{t+1} - (E_t x_{t+1} - x_t). \quad (2.11)$$

Given Eq. (2.11) it is clear that the marginal cost of banking depends on (i) expectations about the banking situation in the future and (ii) the current and expected future production. In particular, the expectations about higher next period marginal banking costs work as a self-fulfilling prophecy, increasing also today's cost. This positive feedback structure reflects, to at least some degree, financial market sentiment and herding behavior. When investors see that the banking sector is going to face difficulties the next day, they will adjust their today's positions accordingly. On the other hand, given the link between the banking sector and consumption, high expectations about next period output gap decrease today's marginal banking cost (negative feedback). Imagine that people expect that there will be a decrease in production in the next period. Since the banking sector is a source of funding, there will be gradually less effort involved in the loan production, bringing today's marginal cost down.

The effects on the current banking situation are proportional to the size of the banking sector, expressed by $(1 - rr)/V$, being more prominent for smaller banking sectors. Smaller banking sectors are more vulnerable to changes in the production sector as the relatively higher part of the banking capital is involved. On the other side, a bigger banking sector might be viewed as being more stable in the sense that the production sector affects it to the lower extent. It should be kept in mind, however, that the model does not say that big banks are ultimately stable as a high drop in today's production can cause the marginal banking cost to skyrocket. Eq. (2.11) predicts only that this effect will be more prominent in the environment with a smaller banking sector.

At the same time, the inferiority of capital to bonds for collateral purposes, v , also plays a role in determining the current marginal banking cost. In particular, let us consider the extreme case when capital cannot serve as a collateral, i.e. $v = 0$. Banks do not have access to capital then so that the only link between them and the production sector is through loans. If there

is a production shock, it affects the bond holdings and labor in the banking sector, making it more severe. In this sense, using capital as collateral serves as a hedge against production sector disturbances. When banks can access capital, in the presence of a production shock, its magnitude is being partially absorbed by the capital part.

2.3 The influence of heterogeneity

So far, we assumed that all the agents are the same and each of them faces the same optimization problem. Before turning to the numerical results, let us first consider what happens in the environment with heterogeneous agents. Contrary to the standard representative agent framework, we allow a part $(1-\gamma)$ of agents to be boundedly rational in forming their expectations⁴. In other words, we assume that a constant proportion of agents is uniformed or unable to form rational expectations. This implies that we may divide our continuum of farmers into two groups: those with rational expectations (E^{RE}) producing good $j \in [0, \gamma]$ and those with boundedly rational expectations (E^{BRE}) producing good $j \in [\gamma, 1]$. By rational agents we mean forward-looking fundamentalists who try to analyze the economy and form their expectations accordingly. Both groups of agents behave as if everybody in the economy was of their type.

To be able to aggregate the results over both groups, we follow the methodology proposed by Branch and McGough (2009) and we impose similar seven axioms on expectation operators:

1. expectations operators fix observables,
2. if z is a forecasted variable and has a steady state, then $E^{RE} \bar{z} = E^{BRE} \bar{z} = \bar{z}$,
3. expectations operators are linear,
4. if for all $k \geq 0$, z_{t+k} and $\sum_{k=0}^{\infty} \beta^{t+k} z_{t+k}$ are forecasted variables then

$$E_t^\tau \left(\sum_{k=0}^{\infty} \beta^{t+k} z_{t+k} \right) = \sum_{k=0}^{\infty} \beta^{t+k} E_t^\tau z_{t+k} \text{ for } \tau \in \{RE, BRE\},$$

⁴Throughout this chapter we use the term ‘rational’ to refer to forward-looking whereas ‘boundedly rational’ to express backward-looking expectations.

5. expectation operators satisfy the law of iterative expectations,
6. if z is a forecasted variable at time t and time $t+k$ then $E_t^\tau E_{t+k}^{\tau'} z_{t+k} = E_t^\tau z_{t+k}$ for $\tau \neq \tau'$,
7. all agents have common expectations on expected differences in limiting wealth and marginal banking cost.

Our contribution to the original methodology comprises axiom 7, which describes the limiting behavior of the expectation operators. Since we add the banking sector to the model, we have to include it also in the expectation formation. Branch and McGough (2009) assume that both types of agents have common expectation on their limiting wealth. It allows to represent the aggregate expectations operator as a weighted average of group expectations. Otherwise, there is an extra term on the limiting behavior of expectations that complicates the dynamics (see Eq. (2.41) from Appendix 2.B). A similar pattern might be observed when aggregating the banking sector (Eq. (2.49) from Appendix 2.B). The aggregate dynamics of the system is therefore influenced by how agents predict the banking sector behaves over the infinite horizon.

Axiom 7 might be viewed as an agreement among all agents that in the far future their banking sectors will be equivalent or will at least generate the same marginal costs. From the macroeconomic perspective, one may think of it as if both groups of agents were trying to reach the banking sector technological frontier. Since there is a common technology, both types of agents should be heading towards the same frontier eventually, satisfying axiom 7.

Proposition 2.3.1. *In the presence of fraction $(1 - \gamma)$ of boundedly rational agents, if agents' expectations satisfy axioms 1-7 then the model from Eq. (2.5), (2.9) and (2.11) can be rewritten as*

$$x_t = \bar{E}_t x_{t+1} + \left(\frac{1 - rr}{V} \right) \bar{E}_t \tilde{\chi}_{t+1} - \left(\frac{1 - rr}{V} + 1 \right) \tilde{\chi}_t - [\hat{r}_t^{IB} - \bar{E}_t \pi_{t+1}] + u_t, \quad (2.12)$$

$$\pi_t = \beta \bar{E}_t \pi_{t+1} + \kappa x_t, \quad (2.13)$$

$$\left(v + \frac{1 - rr}{V} \right) \tilde{\chi}_t = \frac{1 - rr}{V} \bar{E}_t \tilde{\chi}_{t+1} - (\bar{E}_t x_{t+1} - x_t), \quad (2.14)$$

where $\bar{E}_t = \gamma E_t^{RE} + (1 - \gamma) E_t^{BRE}$.

The proof of Proposition 2.3.1 can be found in Appendix 2.B.

2.4 Numerical analysis

As opposed to the standard framework, the central bank policy instrument is the interbank interest rate, \hat{r}_t^{IB} (not the bond rate). In fact, this is the monetary policy tool used in practice (Goodfriend and McCallum, 2007). As argued by Bernanke and Woodford (1997), to close the model we use the forward-looking Taylor rule of the form

$$\hat{r}_t^{IB} = \rho_x E_t^{RE} x_{t+1} + \rho_\pi E_t^{RE} \pi_{t+1}, \quad (2.15)$$

where ρ_x and ρ_π are constant weights on output and inflation variability, respectively. We follow a common approach and assume that the central bank does not target the situation in the banking sector directly. Including a banking sector variable in the monetary rule would extend the monetary policy analysis to a three-dimensional problem so that the interpretation of the results would not be straightforward anymore. Instead, the purpose of this study is to observe how the standard monetary policy rule behaves in the environment with a present banking sector.

2.4.1 Formation of expectations

Throughout the model, we assume that the economy consists of two types of agents that are homogeneous within each group. The first type of agents, $i = RE$, are those who form rational expectations. We abstract here from the standard understanding of rationality, where agents have full knowledge and capacities to perfectly predict the future. Instead, we rather view them as being forward-looking fundamentalists, who collect information and form their expectations accordingly. They are not aware of the presence of the other type of agents so that they form their expectations as if everybody in the economy was rational in forming the expectations

(Branch and McGough, 2009).

The second type of agents is not able to form rational expectations and use simple backward-looking heuristics instead to predict the future. Following Evans and Honkapohja (2001) we assume them to have adaptive expectations of the form

$$E_t^{BRE} z_{t+1} = \mu^2 z_{t-1}, \quad (2.16)$$

where z is either x , π or $\tilde{\chi}$. Parameter $\mu > 0$ describes the magnitude and the direction of the expectations. If $\mu > 1$, the influence of the past is being extrapolated to the future so that we would call those expectations extrapolative. On the other hand, when $\mu < 1$, this influence disappears over time and we would call those expectations adaptive⁵. When $\mu = 1$, the boundedly rational agents form naive expectations (Evans and Honkapohja, 2001).

Given the expectation operators for both groups of agents, we may rewrite the aggregate expectations as

$$\bar{E}_t z_{t+1} = \gamma E_t^{RE} z_{t+1} + (1 - \gamma) \mu^2 z_{t-1}, \quad (2.17)$$

with z being either x , π or $\tilde{\chi}$.

2.4.2 Calibration and numerical results

DSGE models often exhibit indeterminacy, i.e. there is no unique path guiding the equilibrium. In such a situation, the quantities and prices might not be even locally determinate, making the monetary policy conduct more unstable (Woodford, 1994). Therefore, it is important to make sure that the monetary tools provide a determinate structure of the economy.

⁵In the literature, adaptive expectations are being recognized as the whole group of operators of the form similar to Eq. (2.16). However, for clarity purposes, we distinguish here between extrapolative and adaptive expectations when $\mu > 1$ and $\mu < 1$, respectively.

Table 2.1: Calibration values for the model parameters.

Parameter	V	rr	ν	κ	β
Value	0.31	0.005	0.2	0.05	0.99

Let us write the complete model in the matrix form

$$\begin{pmatrix} B & 0 \\ 0 & I_3 \end{pmatrix} \begin{pmatrix} y_{t+1} \\ y_t \end{pmatrix} = \begin{pmatrix} F & -C \\ I_3 & 0 \end{pmatrix} \begin{pmatrix} y_t \\ y_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_t \\ 0 \end{pmatrix}, \quad (2.18)$$

where $y = (x, \pi, \tilde{\chi})'$, $\varepsilon = (u, 0, 0)$ is a vector of exogenous shocks and B , F and C are the coefficient matrices described in detail in Appendix 2.C.

To study the determinacy properties, we apply the methodology developed by Blanchard and Kahn (1980). Since it does not depend on the exogenous disturbances, we omit ε in our further analysis. The determinacy is a result of the properties of the solution matrix M , where

$$M = \begin{pmatrix} B^{-1}F & -B^{-1}C \\ I_3 & 0 \end{pmatrix}. \quad (2.19)$$

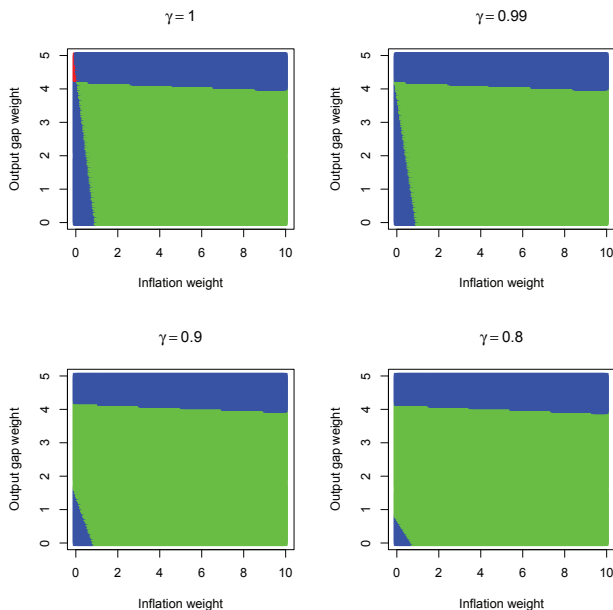
The equilibrium of the system is determinate only if the number of eigenvalues that are outside the unit circle is equal to the number of non-predetermined variables (or the forward-looking variables (Walsh, 2010)), which is 3 in this case. Having more eigenvalues outside the unit circle implies explosiveness and fewer of them implies indeterminacy. The degree of indeterminacy is equal to the number of non-predetermined variables less the number of eigenvalues outside the unit circle (Evans and McGough, 2005).

We calibrate our model accordingly to Goodfriend and McCallum (2007). The detailed values are presented in Table 2.1.

The determinacy properties are studied for extrapolative and adaptive expectations separately. For the former, the μ parameter is set to 1.1 and for the latter to 0.9 (Branch and McGough, 2009). The ranges for policy parameters ρ_x and ρ_π are set from 0 to 5 and 10,

respectively, in order to show the complete behavior of the system. The results are presented in Figs 2.1 and 2.2.

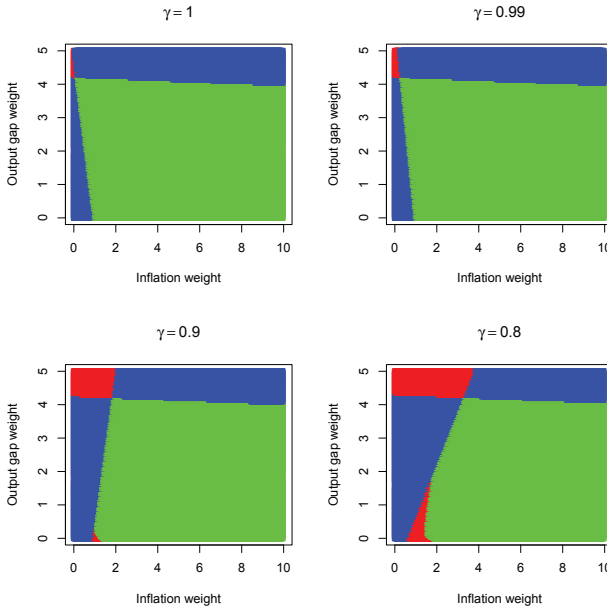
Figure 2.1: Determinacy properties ($\mu = 0.9$). Green color describes determinacy, blue order 1 indeterminacy and red order 2 indeterminacy.



Firstly, the results confirm the 'rotating' behavior of the system from Branch and McGough (2009). With adaptive expectations the system rotates counterclockwise so that the determinacy area increases. With extrapolative expectations the system rotates clockwise decreasing the determinacy area.

Secondly, the location of the indeterminacy of order one and two is in line with the figures presented in Branch and McGough (2009). In fact, the only difference lies in the size of those areas, comparing with the original paper. This, however, is the consequence of the banking calibration parameters and the different specification of the utility function. In fact, if we allow for extra parameter describing the intertemporal substitution elasticity of consumption in the

Figure 2.2: Determinacy properties ($\mu = 1.1$). Green color describes determinacy, blue order 1 indeterminacy and red order 2 indeterminacy.



utility function, σ , the determinacy area is narrowed from the top, being more similar to the results from Branch and McGough (2009) and Bullard and Mitra (2002).

Thirdly, the presence of the banking sector has one important impact on determinacy properties. When agents form extrapolative expectations ($\mu = 1.1$), a new region of indeterminacy of order 2 arises for too lenient inflation targeting. In the case with adaptive expectations ($\mu = 0.9$) there is no similar effect.

2.5 Conclusions and discussion

The goal of this chapter was twofold. Firstly, we derived a workhorse model for monetary policy analysis with the present banking sector. Secondly, we relaxed the assumption of the

representative agent structure and investigated the effects of the presence of boundedly rational agents.

The results suggest that the presence of a banking sector changes the determinacy structure of the equilibrium. Given that agents form adaptive expectations, the determinacy structure rotates counterclockwise, so that more lenient output gap and inflation targeting still guarantees determinacy.

The problem arises when backward-looking agents extrapolate the past performance over their future forecasts. The presence of the banking sector brings additional indeterminacy area for lower inflation targeting parameter. In other words, in the environment with a fraction of extrapolative agents, if the monetary policy does not fight inflation sufficiently well, it may not reach the equilibrium in the long run.

In fact this pattern might have significant consequences for the actual monetary policy conduct. Pfajfar and Zakelj (2011) suggest that the fraction of extrapolative agents might be as high as 30%, even larger than in our analysis. Given the fact that the estimated Taylor rule parameters vary usually in the region of (0,1) for the output gap weight and of (1,2) for the inflation weight (Taylor, 1999; Woodford, 2003), this may suggest that the system is very close to indeterminacy, if not indeterminate already, which arises as a consequence of the banking sector. Therefore, it seems vital for the monetary policy to address the issue of agents' heterogeneity and investigate in detail how they form their forecasts. There could be many solutions to the problem raised above, however, it is beyond the scope of this chapter to discuss them in detail. Assuming that the inflation and output weights are set to satisfy the goals of the monetary policy, there seem to be still ways out of the problem. For instance, one may think of increasing the clarity and flexibility of capital, somehow reducing its inferiority for collateral purposes. This would make current marginal banking cost more robust with respect to the future disturbances and thereof could decrease the influence of destabilizing extrapolative expectations. Another solution would be smaller minimum capital requirement, however, this could translate into higher banking sector leverage and eventually may cause more problems than it originally aimed to

solve.

It is clear that households' expectations play an important role in determining the monetary policy, especially when a banking sector is present. However, this research shows just the top of an iceberg and more study is required in order to fully understand the phenomenon of banking in the modern economy. In particular, a straightforward extension of this study is to endogenize the fraction of rational agents, making it dependent on other systemic variables.

Appendix 2.A Baseline derivation

The utility of a farmer is defined as a weighted average of her consumption and leisure and takes the form

$$U^i(C_t^i, n_t^i, m_t^i) = \phi \log(C_t^i) + (1 - \phi) \log(1 - n_t^i - m_t^i), \quad (2.20)$$

where ϕ is the relative preference weight on consumption and t is the time subscript. C_t^i represents a composite consumption good and is of the standard Constant Elasticity of Substitution (CES) form, as in Dixit and Stiglitz (1977)

$$C_t^i = \left(\int_0^1 c_{jt}^{\frac{\theta-1}{\theta}} dj \right)^{\frac{\theta}{\theta-1}}, \quad (2.21)$$

with θ being the elasticity of substitution.

The farmer's decision problem is to maximize her discounted expected utility subject to the budget and technology constraints. Assuming a cashless limit (Woodford, 2003; Branch and McGough, 2009), we may define the former in real terms as

$$w_t(n_t^i + m_t^i) + q_t(1 - \delta)K_t^i + \frac{Y_t^i P_t^i}{P_t^A} + \frac{B_t^i}{P_t^A} = w_t(n_t^{i,d} + m_t^{i,d}) + C_t^i + q_t K_{t+1}^i + \frac{B_{t+1}}{P_t^A(1 + r_t^B)}, \quad (2.22)$$

where K_t^i is capital level with q_t being its real price and δ the depreciation rate, w_t is the real wage and B_t^i are the nominal bond holdings with the nominal interest equal r_t^B . Y_t^i is the production level, P^i is the price of the individual good and P_t^A is the aggregate price level, as in the Dixit-Stiglitz setup. Superscript d denotes the amount of labor demanded by a given farmer. Superscript i and subscript t relate to the agent and time dimensions, respectively.

Contrary to the standard new Keynesian framework, there is a capital market in the model. Its role is twofold. Firstly, capital serves as a production factor in the farmers' technology. Secondly, it is used as a collateral in the banking sector to produce loans. For simplicity, it is assumed that the aggregate capital stock is on a steady state growth path (Goodfriend and McCallum, 2007). What is important is that farmers are allowed to trade it so that its market

price q_t may fluctuate.

The production constraint requires that

$$Y_t^i = K_t^{i\eta} \left(e^{A_{1t}} n_t^{i,d} \right)^{1-\eta}, \quad (2.23)$$

where A_{1t} is an aggregate productivity disturbance and η is the capital elasticity measure.

A novelty in the model is the presence of the banking sector. Its main role is to facilitate transactions between production and consumption sides of the economy. Since the medium of exchange is the crucial role of the monetary policy analysis, the model does not distinguish between transaction balances and time deposits at the banks. In this simple form, it implies that the farmer's consumption in each period has to be rigidly related to the deposits held at a bank (Goodfriend and McCallum, 2007). In other words, in each period, the level of consumption (C_t^i) has to be covered by some constant fraction of the real deposits (VD_t^i/P_t^A). Since each bank has to hold a given level of reserves at the central bank (rr), the nominal amount of loans it may produce from deposits held by farmer i is constrained by $L_t^i = (1 - rr)D_t^i$. At the same time, the real loan production depends on the collateral and loan monitoring, and is assumed to be of a Cobb-Douglas form

$$\frac{L_t^i}{P_t^A} = F \left(\frac{B_{t+1}^i}{P_t^A(1 + r_t^B)} + vq_t K_{t+1}^i \right)^\alpha \left(e^{A_{2t}} m_t^{i,d} \right)^{1-\alpha}. \quad (2.24)$$

The loan monitoring is assumed to be proportional to the labor supplied to the banking sector by farmer i and A_{2t} is the productivity disturbance similar to the one in the production sector. Since capital stock require a substantial monitoring effort to confirm its physical condition, its inferiority to bonds for collateral purposes is expressed by v (Goodfriend and McCallum, 2007).

The complete intertemporal farmers's maximization problem (with a presence of the banking sector) may be written as

$$\max_{n_t^i, m_t^i, n_t^{i,d}, m_t^{i,d}, P_t^i, K_{t+1}^i, B_{t+1}^i} E_t^i \sum_{k=0}^{\infty} \beta^k \left[\phi \log(C_{t+k}^i) + (1 - \phi) \log(1 - n_{t+k}^i - m_{t+k}^i) \right]. \quad (2.25)$$

subject to the budget constraint (Eq. 2.22) and production constraint (Eq. 2.23).

Before solving the optimization problem, from Eq. (2.24) we know that

$$C_t^i = \frac{VF}{1-rr} (b_{t+1}^i + vq_t K_{t+1}^i)^\alpha \left(e^{A2t} m_t^{i,d} \right)^{1-\alpha}, \quad (2.26)$$

where $b_{t+1}^i = B_{t+1}^i / (P_t^A (1 + r^B))$. Additionally, by imposing market clearing we know that the good produced by farmer i is equal to its demand

$$Y_t^i = \left(\frac{P_t^i}{P_t^A} \right)^{-\theta} C_t^A, \quad (2.27)$$

where C_t^A is the aggregate consumption level that each individual takes as given.

Let the Lagrange multipliers be λ_t and ξ_t for the budget and production constraints respectively. By including Eq. (2.26) and Eq. (2.27) into the maximization problem and assuming market symmetry (Goodfriend and McCallum, 2007), the first order conditions provide

$$\frac{-(1-\phi)}{1-n_t^i - m_t^i} + \lambda_t^i w_t = 0, \quad (2.28)$$

$$-\lambda_t^i w_t + \xi_t^i e^{A1t} (1-\eta) \left(\frac{K_t^i}{e^{A1t} n_t^i} \right)^\eta = 0, \quad (2.29)$$

$$\left(\frac{\phi}{C_t^i} - \lambda_t^i \right) \frac{C_t^i (1-\alpha)}{m_t^i} - \lambda_t^i w_t = 0, \quad (2.30)$$

$$C_t^A \left(\frac{P_t^i}{P_t^A} \right)^{-\theta} \left(\frac{(1-\theta)\lambda_t^i}{P_t^A} + \frac{\theta\xi_t^i}{P_t^i} \right) = 0, \quad (2.31)$$

$$\left(\frac{\phi}{C_t^i \lambda_t^i} - 1 \right) \Omega_t^i v q_t - q_t + \beta(1-\delta) E_t^i \left(\frac{\lambda_{t+1}^i}{\lambda_t^i} q_{t+1} \right) + \beta \eta E_t^i \left(\frac{\xi_{t+1}^i}{\lambda_t^i} \left(\frac{e^{A1t+1} n_{t+1}^i}{K_{t+1}^i} \right)^{1-\eta} \right) = 0, \quad (2.32)$$

$$\left(\frac{\phi}{C_t^i \lambda_t^i} - 1 \right) \Omega_t^i - 1 + \beta E_t^i \left(\frac{\lambda_{t+1}^i}{\lambda_t^i} \frac{P_t^A}{P_{t+1}^A} (1+r^B) \right) = 0, \quad (2.33)$$

where Ω_t^i is the partial derivative of the deposit constraint $C_t^i = \frac{VL_t^i}{(1-rr)P_t^A}$ with respect to

collateral

$$\Omega_t^i = \frac{\alpha C_t^i}{b_{t+1}^i + v q_t K_{t+1}^i}. \quad (2.34)$$

Appendix 2.B The influence of heterogeneous agents

Throughout the following derivation, we assume that each agent belongs to one of the two groups, i.e. $i = \tau \in \{RE, BRE\}$. By superscript A we will refer to the aggregate values.

Appendix 2.B.1 The heterogeneous IS curve

Let us first introduce a benevolent financial institution that helps farmers in hedging the risk associated with the Calvo lottery (Shi, 1999; Mankiw and Reis, 2007). In each period it collects all the income from the market and then redistribute it evenly across farmers. Given this property and assuming cashless limit, the agents' budget constraint becomes

$$w_t(n_t^i + m_t^i) + q_t(1 - \delta)K_t^i + \frac{Y_t^i P_t^i}{P_t^A} + \frac{B_t^i}{P_t^A} + I_{r,t}^i = w_t(n_t^{i,d} + m_t^{i,d}) + C_t^i + q_t K_{t+1}^i + \frac{B_{t+1}^i}{P_t^A(1 + r_t^B)} + I_{p,t}^i, \quad (2.35)$$

where $I_{r,t}^i$ and $I_{p,t}^i$ are the real receipts from and payments to the insurance agency. Each agent maximizes her expected utility over an infinite horizon, subject to Eq. (2.35) instead of (Eq. 2.22).

We know that the average real income (denoted by Ψ_t^τ) and the average marginal banking cost χ_t^τ obtained by rational and boundedly rational agents are

$$\Psi_t^{RE} = \frac{1}{\gamma P_t^A} \int_0^\gamma P_t^i Y_t^i di \quad \text{and} \quad \Psi_t^{BRE} = \frac{1}{(1 - \gamma) P_t^A} \int_\gamma^1 P_t^i Y_t^i di, \quad (2.36)$$

$$\chi_t^{RE} = \frac{1}{\gamma} \int_0^\gamma \chi_t^i di \quad \text{and} \quad \chi_t^{BRE} = \frac{1}{1 - \gamma} \int_\gamma^1 \chi_t^i di. \quad (2.37)$$

From the above equations it is clear that we may view the aggregate production and aggregate real marginal banking cost as a weighted average of their components, i.e. $Y_t^A = \gamma Y_t^{RE} + (1 -$

$\gamma)Y_t^{BRE}$ and $\chi_t^A = \gamma\chi_t^{RE} + (1 - \gamma)\chi_t^{BRE}$.

Following Branch and McGough (2009), if an agent is of type τ , then her real receipts from and payments to the insurance agency are $I_{r,t}^i = \Psi_t^\tau$ and $I_{p,t}^i = Y_t^i P_t^i / P_t^A$. By market clearing and axiom A2 the steady states of consumption and production are equal at individual and group levels. By imposing market symmetry, the budget constraint (Eq. 2.35) yields

$$\hat{C}_t^\tau = \hat{\Psi}_t^\tau + \frac{B_t^\tau / P_t^A}{\bar{Y}_t^A} - \frac{B_{t+1}^\tau / (P_t^A(1 + \bar{r}_t^B))}{\bar{Y}_t^A} + \frac{q_t(1 - \delta)K^\tau}{\bar{Y}_t^A} - \frac{q_t K_{t+1}^\tau}{\bar{Y}_t^A}, \quad (2.38)$$

where the bars indicate the steady state levels. Bond and capital market clearing require that $\alpha B_t^{BRE} = -(1 - \alpha)B_t^{BRE}$ and $\alpha K_t^{RE} = -(1 - \alpha)K_t^{BRE}$. After multiplying Eq. (2.38) by γ for rational and by $(1 - \gamma)$ for boundedly rational agents and summing up, we arrive at

$$\hat{Y}_t^A = \gamma\hat{\Psi}_t^{RE} + (1 - \gamma)\hat{\Psi}_t^{BRE}. \quad (2.39)$$

From Eq. (2.4), (2.37) and (2.38) we have

$$\hat{\Psi}_t^\tau = E_t^\tau \hat{\Psi}_{t+1}^\tau + \left(\frac{1 - rr}{V}\right) E_t^\tau \tilde{\chi}_{t+1}^\tau - \left(\frac{1 - rr}{V} + 1\right) \tilde{\chi}_t^\tau - (\hat{r}_t^{IB} - E_t^\tau \pi_{t+1}). \quad (2.40)$$

Iterating this equation forward and substituting into Eq. (2.39) we finally get

$$\begin{aligned} \hat{Y}_t^A &= \bar{E}_t \hat{Y}_{t+1}^A + \left(\frac{1 - rr}{V}\right) \bar{E}_t \tilde{\chi}_{t+1}^A - \left(\frac{1 - rr}{V} + 1\right) \tilde{\chi}_t^A - (\hat{r}_t^{IB} - \bar{E}_t \pi_{t+1}) + \\ &+ \left(\gamma\hat{\Psi}_\infty^{RE} + (1 - \gamma)\hat{\Psi}_\infty^{BRE}\right) - \bar{E}_t \left(\gamma\hat{\Psi}_\infty^{RE} + (1 - \gamma)\hat{\Psi}_\infty^{BRE}\right), \end{aligned} \quad (2.41)$$

with $\bar{E}_t = \gamma E_t^{RE} + (1 - \gamma)E_t^{BRE}$ and $\hat{\Psi}_\infty^\tau = \lim_{k \rightarrow \infty} E_t^\tau \hat{\Psi}_{t+k}^\tau$. In fact, Eq. (2.41) is of exactly the same form as in Branch and McGough (2009) but with a banking sector present. Axiom 7 indicates that agents predict their limiting wealth identically, which makes

$$\left(\gamma\hat{\Psi}_\infty^{RE} + (1 - \gamma)\hat{\Psi}_\infty^{BRE}\right) = \bar{E}_t \left(\gamma\hat{\Psi}_\infty^{RE} + (1 - \gamma)\hat{\Psi}_\infty^{BRE}\right). \quad (2.42)$$

Subtracting the log deviations of the potential product from both sides, we finally arrive at the heterogeneous IS curve with a present banking sector

$$x_t = \bar{E}_t x_{t+1} + \left(\frac{1-rr}{V} \right) \bar{E}_t \tilde{\chi}_{t+1}^A - \left(\frac{1-rr}{V} + 1 \right) \tilde{\chi}_t^A - [\hat{r}_t^{LB} - \bar{E}_t \pi_{t+1}] + u_t, \quad (2.43)$$

where $x_t = \hat{Y}_t^A - \hat{Y}_t^{f,A}$ is the output gap measure, the expectation operator is the weighted average of the group expectations $\bar{E}_t = \gamma E_t^{RE} + (1-\gamma) E_t^{BRE}$ and u_t is a disturbance term that depends only on exogenous productivity shocks.

Appendix 2.B.2 The heterogeneous Phillips curve

It is important to note that when farmers may hedge against the Calvo risk their production level would be 0 in equilibrium as a result of the free-riding problem. Therefore, following Branch and McGough (2009), we assume that farmers make their pricing decisions as if there was no insuring agency.

Let us take the log approximation of Eq. (2.7)

$$\log P_t^\tau - \log P_t^A = (1-\omega\beta)\hat{\varphi}_t^\tau + \omega\beta E_t^\tau \pi_{t+1} + \omega\beta E_t^\tau \log P_{t+1}^\tau / P_{t+1}^A. \quad (2.44)$$

Branch and McGough (2009) show that the Calvo lottery implies aggregate inflation to follow

$$\pi_t = \frac{1-\omega}{\omega} \left(\gamma \log P_t^{RE} / P_t^A + (1-\gamma) \log P_t^{BRE} / P_t^A \right). \quad (2.45)$$

As long as the pricing decisions are homogeneous within each group τ , by multiplying Eq. (2.44) by γ for rational and by $(1-\gamma)$ for boundedly rational agents and adding up, after some algebra we arrive at the final aggregate heterogeneous Phillips curve

$$\pi_t = \beta \bar{E}_t \pi_{t+1} + \kappa \varphi_t^A, \quad (2.46)$$

where $\kappa = \frac{(1-\omega)(1-\beta\omega)}{\omega}$.

Finally, noting that the aggregate marginal production cost is the aggregate output gap measure, the heterogeneous new Keynesian Phillips curve amended for the banking sector may be viewed as

$$\pi_t = \beta \bar{E}_t \pi_{t+1} + \kappa x_t, \quad (2.47)$$

where $\bar{E}_t = \gamma E_t^{RE} + (1 - \gamma) E_t^{BRE}$.

Appendix 2.B.3 The heterogeneous banking sector curve

Taking the steady state log deviations of Eq. (2.10) and iterating forward we get for each group of agents

$$\tilde{\chi}_t^\tau = \left(\frac{1 - rr}{V} \right)^{-1} \left[-v E_t^\tau \sum_{j=0}^{\infty} \tilde{\chi}_{t+j}^\tau + \left(\frac{1 - rr}{V} \right) \tilde{\chi}_\infty^\tau - (x_\infty^\tau - x_t^\tau) \right], \quad (2.48)$$

where $\hat{\chi}_\infty^\tau = \lim_{k \rightarrow \infty} E_t^\tau \hat{\chi}_{t+k}^\tau$ and $x_\infty^\tau = \lim_{k \rightarrow \infty} E_t^\tau x_{t+k}^\tau$.

Given Eq. (2.37) and (2.48), we get

$$\begin{aligned} \tilde{\chi}_t^A &= \gamma \tilde{\chi}_t^{RE} + (1 - \gamma) \tilde{\chi}_t^{BRE} \\ &= \left(\frac{1 - rr}{V} \right)^{-1} \left[-v \left(\gamma E_t^{RE} \sum_{k=0}^{\infty} \tilde{\chi}_{t+k}^{RE} + (1 - \gamma) E_t^{BRE} \sum_{k=0}^{\infty} \tilde{\chi}_{t+k}^{BRE} \right) + x_t \right. \\ &\quad \left. + \left(\frac{1 - rr}{V} \right) (\gamma \tilde{\chi}_\infty^{RE} + (1 - \gamma) \tilde{\chi}_\infty^{BRE}) - (\gamma x_\infty^{RE} + (1 - \gamma) x_\infty^{BRE}) \right] \\ &= \bar{E}_t \tilde{\chi}_{t+1}^A - \frac{vV}{1 - rr} \tilde{\chi}_t^A - \left(\frac{1 - rr}{V} \right)^{-1} (\bar{E}_t x_{t+1} - x_t) \\ &\quad + [(\gamma \tilde{\chi}_\infty^{RE} + (1 - \gamma) \tilde{\chi}_\infty^{BRE}) - \bar{E}_t (\gamma \tilde{\chi}_\infty^{RE} + (1 - \gamma) \tilde{\chi}_\infty^{BRE})] \\ &\quad - \left(\frac{1 - rr}{V} \right)^{-1} [(\gamma x_\infty^{RE} + (1 - \gamma) x_\infty^{BRE}) - \bar{E}_t (\gamma x_\infty^{RE} + (1 - \gamma) x_\infty^{BRE})]. \end{aligned} \quad (2.49)$$

The last two lines disappear due to Axiom 7, which gives

$$(\gamma \tilde{\chi}_\infty^{RE} + (1 - \gamma) \tilde{\chi}_\infty^{BRE}) = \bar{E}_t (\gamma \tilde{\chi}_\infty^{RE} + (1 - \gamma) \tilde{\chi}_\infty^{BRE}) \quad (2.50)$$

$$(\gamma x_\infty^{RE} + (1 - \gamma)x_\infty^{BRE}) = \bar{E}_t (\gamma x_\infty^{RE} + (1 - \gamma)x_\infty^{BRE}) \quad (2.51)$$

so that the final banking curve equation may be written as Eq. (2.14).

Appendix 2.C Model dynamics

The condensed model can be viewed as

$$\begin{pmatrix} B & 0 \\ 0 & I_3 \end{pmatrix} \begin{pmatrix} y_{t+1} \\ y_t \end{pmatrix} = \begin{pmatrix} F & -C \\ I_3 & 0 \end{pmatrix} \begin{pmatrix} y_t \\ y_{t-1} \end{pmatrix}, \quad (2.52)$$

where $y = (x, \pi, \tilde{\chi})'$ and

$$B = \begin{pmatrix} \gamma - \rho_x & \gamma - \rho_\pi & \frac{\gamma(1-rr)}{V} \\ 0 & \beta\gamma & 0 \\ -\gamma & 0 & \frac{\gamma(1-rr)}{V} \end{pmatrix}, \quad (2.53)$$

$$F = \begin{pmatrix} 1 & 0 & \frac{(1-rr)}{V} + 1 \\ -\kappa & 1 & 0 \\ -1 & 0 & v + \frac{(1-rr)}{V} \end{pmatrix}, \quad (2.54)$$

$$C = \begin{pmatrix} (1-\gamma)\mu^2 & (1-\gamma)\mu^2 & \frac{(1-\gamma)\mu^2(1-rr)}{V} \\ 0 & \beta(1-\gamma)\mu^2 & 0 \\ -(1-\gamma)\mu^2 & 0 & \frac{(1-\gamma)\mu^2(1-rr)}{V} \end{pmatrix}. \quad (2.55)$$

Chapter 3

Nonlinear Granger Causality - Guidelines for Multivariate Analysis

3.1 Introduction

Since the introduction of Granger causality over four decades ago (Granger, 1969), the body of literature on this topic has grown substantially, becoming standard methodology not only among economists and econometricians, but also finding followers in physics or even biology (Guo et al., 2010). Not surprisingly, it alleviated an ongoing discussion on the nature and validity of the concept, pointing out its methodological limitations. Although, the spectrum of arguments against the idea of Granger causality is very broad, the main line of criticism follows from the very simple nature of the dependence relations in the economy, which Granger causality originally assumes (Cartwright, 2007). The scope of this chapter is to contribute to the discussion allowing for a more complex structural setting in the nonparametric Granger causality testing.

Imagine a strictly stationary bivariate process $\{(X_t, Y_t)\}$. We say that $\{X_t\}$ is a Granger cause of $\{Y_t\}$ if past and current values of X contain additional information on future values of Y that is not contained in past and current Y -values alone. If we denote the information

contained in past observations X_s and Y_s , $s \leq t$, by $\mathcal{F}_{X,t}$ and $\mathcal{F}_{Y,t}$, respectively, and let ‘ \sim ’ denote equivalence in distribution, the formal definition is:

Definition 3.1.1. *For a strictly stationary bivariate time series process $\{(X_t, Y_t)\}$, $t \in \mathbb{Z}$, $\{X_t\}$ is a Granger cause of $\{Y_t\}$ if, for some $k \geq 1$,*

$$(Y_{t+1}, \dots, Y_{t+k}) | (\mathcal{F}_{X,t}, \mathcal{F}_{Y,t}) \not\sim (Y_{t+1}, \dots, Y_{t+k}) | \mathcal{F}_{Y,t}.$$

Clearly, such a definition is very simplistic and seems to be inappropriate to apply in complex environments. An obvious shortcoming is the fact that the vectors of interests are assumed to be univariate, making the whole problem detached from reality. In other words, this methodology does not allow to control for *every possible source of variation of every kind [...]* as argued by (Cartwright, 2007). An advantage of such a general definition is, however, that it does not assume any parametric relations between the time series and instead focuses on the conditional distributions only¹.

The most commonly used nonparametric test for the above hypothesis testing (Def. 3.1.1) is the one proposed by Hiemstra and Jones (1994). Its main advantage lies in a very clear and intuitive reasoning together with a strong asymptotic theory, derived even for a multivariate setting² (Bai et al., 2010). Nevertheless, the test can severely over-reject if the null is satisfied (Diks and Panchenko, 2005). Therefore, Diks and Panchenko (2006) (hereafter DP) proposed a new test statistic which corrects for this shortcoming but, as it turns out, because of the large kernel estimator bias the DP test lacks consistency in the multivariate setting.

The goal of this chapter is therefore twofold. Firstly, in order to reduce the kernel estimator bias we apply the data sharpening method (Hall and Minnotte, 2002) and we derive the asymptotic properties for the sharpened DP test in a multivariate setting. Secondly, we investigate its

¹This brings additional modeling flexibility and does not bind us to the linear autoregressive model as originally proposed by Granger (1969).

²Throughout the chapter, we will refer to a multivariate setting by a situation where vector $\{X_t\}$ is allowed to be multidimensional and $\{Y_t\}$ is univariate, i.e. $\{X_t\} = \{X_{1,t}, X_{2,t}, \dots, X_{m,t}\}$, $m \geq 1$. In principle, the dimensions of $\{X_t\}$ vector might describe its corresponding lags, i.e. $X_{1,t} = X_{t-1}$, $X_{2,t} = X_{t-2}$ etc, so that by the bivariate case we refer to the situation where $\{X_t\}$ is univariate.

performance both numerically and empirically on the US grain market. We chose this specific market due to its straightforward causal relationship, where the price of each grain is influenced not only by the other grains, but to a large extent the whole market is driven by weather forecasts. Therefore, it serves as an almost ideal environment to test our new methodology in practice.

This chapter is organized as follows. Section 3.2 discusses the asymptotic properties of the original DP test and shows why it lacks consistency in the multivariate setting. In Section 3.3 we replace the standard kernel density estimator by its sharpened form and we show that the new test statistic is asymptotically normally distributed. We confirm the theoretical results by computer simulations. In Section 3.4 we apply the new test to the US grain market. Section 3.5 summarizes and concludes.

3.2 Asymptotic properties of the DP test

In testing for Granger non-causality, the aim is to detect evidence against the null hypothesis

$$H_0 : \quad \{X_t\} \text{ is not Granger causing } \{Y_t\},$$

with Granger causality defined according to Def. 3.1.1. We limit ourselves to tests for detecting Granger causality for $k = 1$, which is the case considered most often in practice. Under the null hypothesis Y_{t+1} is conditionally independent of X_t, X_{t-1}, \dots , given Y_t, Y_{t-1}, \dots . In a non-parametric setting, conditioning on the infinite past is impossible without a model restriction, such as an assumption that the order of the process is finite. Therefore, in practice conditional independence is tested using finite lags l_X and l_Y , i.e.

$$Y_{t+1} | (X_t^{l_X}; Y_t^{l_Y}) \sim Y_{t+1} | Y_t^{l_Y},$$

where $X_t^{l_X} = (X_{t-l_X+1}, \dots, X_t)$ and $Y_t^{l_Y} = (Y_{t-l_Y+1}, \dots, Y_t)$. For a strictly stationary bivariate time series $\{(X_t, Y_t)\}$ this is a statement about the distribution of the $l_X + l_Y + 1$ -dimensional vector $W_t = (X_t^{l_X}, Y_t^{l_Y}, Z_t)$, where $Z_t = Y_{t+1}$. To keep the notation simple, and to bring about the fact that the null hypothesis is a statement about the invariant distribution of W_t , we often drop the time index and just write $W = (X, Y, Z)$, where the latter is a random vector with the distribution of $(X_t^{l_X}, Y_t^{l_Y}, Y_{t+1})$. Nevertheless, Denker and Keller (1983) and Diks and Panchenko (2006) show that the reasoning holds for weakly-dependent W_t provided that the covariance between the local density estimators is taken into account in the asymptotic variance of the test statistic.

For now, let us consider the simplest setting, where $l_X = l_Y = 1$ so that $W = (X, Y, Z)$ denotes a three-variate random variable, distributed as $W_t = (X_t, Y_t, Y_{t+1})$. (Throughout the chapter we assume that W is a continuous random variable.) The DP test restates the null hypothesis in terms of the joint probability distribution $f_{X,Y,Z}(X, Y, Z)$ and its marginals, i.e.

$$q \equiv E[f_{X,Y,Z}(X, Y, Z)f_Y(Y) - f_{X,Y}(X, Y)f_{Y,Z}(Y, Z)] = 0. \quad (3.1)$$

Given a simple square kernel density estimator

$$\hat{f}_W(W_i) = \frac{(2\varepsilon)^{-d_W}}{n-1} \sum_{j, j \neq i} I\left(\frac{\|W_i - W_j\|}{\varepsilon}\right), \quad (3.2)$$

where ε is a bandwidth, $\|\cdot\|$ is the maximum norm and $I(\cdot)$ is the indicator function taking values 1 for any argument within the unit circle, one readily finds a natural estimator of q being given as

$$T_n(\varepsilon) = \frac{(n-1)}{n(n-2)} \sum_i (\hat{f}_{X,Y,Z}(X_i, Y_i, Z_i) \hat{f}_Y(Y_i) - \hat{f}_{X,Y}(X_i, Y_i) \hat{f}_{Y,Z}(Y_i, Z_i)). \quad (3.3)$$

The asymptotic behavior of $T_n(\varepsilon)$ follows directly from the reasoning originally designed for the MSE (Mean Squared Error) optimal bandwidth selection under the shrinking condi-

tions, developed by Powell and Stoker (1996). The test statistic has a corresponding third order U-statistic representation with a kernel given by $\tilde{K}(W_i, W_j, W_k)$. Let us denote $\tilde{K}_1(w_1) = E[\tilde{K}(w_1, W_2, W_3)]$ and $\tilde{K}_2(w_1, w_2) = E[\tilde{K}(w_1, w_2, W_3)]$, and assume that the rates of convergence of the pointwise bias as well as the second moment kernel expansions depend on the bandwidth size in the following way (in fact these are the conditions imposed by Powell and Stoker (1996))

$$\tilde{K}_1(w_i, \varepsilon) - \lim_{\varepsilon \rightarrow 0} \tilde{K}_1(w_i, \varepsilon) = s(w_i)\varepsilon^\alpha + s^*(w_i, \varepsilon), \quad \alpha > 0, \quad (3.4)$$

$$E \left[(\tilde{K}_2(W_1, W_2))^2 \right] = q_2\varepsilon^{-\gamma} + q_2^*(\varepsilon), \quad \gamma > 0, \quad (3.5)$$

$$E \left[(\tilde{K}(W_1, W_2, W_3))^2 \right] = q_3\varepsilon^{-\delta} + q_3^*(\varepsilon), \quad \delta > 0, \quad (3.6)$$

where the remainder terms are negligible, i.e. $E\|s^*(W_i, \varepsilon)\|^2 = o(\varepsilon^{2\alpha})$, $(q_2^*(\varepsilon))^2 = o(\varepsilon^{-\gamma})$ and $(q_3^*(\varepsilon))^2 = o(\varepsilon^{-\delta})$. Parameters α , γ and δ follow directly from the specification of the kernel function \tilde{K} and might be derived analytically. In fact, it might be verified that α is of the same magnitude as the local kernel estimator bias and Diks and Panchenko (2006) show that two remaining parameters depend on the dimensionality of the system as $\gamma = d_x + d_y + d_z$ and $\delta = d_x + 2d_y + d_z$.

Having pointed that out, the MSE of the test statistic might be expressed as

$$\text{MSE}[T_n(\varepsilon)] = (E[s(W_i)])^2 \varepsilon^{2\alpha} + \frac{9}{n} C_0 \varepsilon^\alpha + \frac{9}{n} \text{Var} \left[\lim_{\varepsilon \rightarrow 0} \tilde{K}_1(W_i, \varepsilon) \right] + \frac{18}{n^2} q_2 \varepsilon^{-\gamma} + \frac{6}{n^3} q_3 \varepsilon^{-\delta}, \quad (3.7)$$

where $C_0 = 2\text{Cov} \left[\lim_{\varepsilon \rightarrow 0} \tilde{K}_1(W_i, \varepsilon), s(W_i) \right]$. In order to guarantee asymptotic normality of $T_n(\varepsilon)$ all the ε -dependent terms in Eq. (3.7) have to be $o(n^{-1})$. Given the bandwidth shrinking condition, i.e. $\varepsilon_n \equiv Cn^{-\beta}$, one may find that this implies

$$\sqrt{n} \frac{T_n(\varepsilon_n) - q}{\sigma} \xrightarrow{d} \mathcal{N}(0, 1) \quad \text{iff} \quad \frac{1}{2\alpha} < \beta < \frac{1}{d_x + d_y + d_z}, \quad (3.8)$$

with σ^2 being the asymptotic variance of $\sqrt{n}(T_n(\varepsilon_n) - q)$.

Clearly, given the standard kernel density estimator with bias of order 2 and the basic model specification with $d_x = d_y = d_z = 1$, the test statistic is asymptotically normally distributed for any positive constant C and $\beta \in (1/4, 1/3)$. Given suitable mixing conditions (see Denker and Keller (1983)) and provided that covariances between local density estimators are taken into account, the result holds also for the weakly dependent time series (Diks and Panchenko, 2006).

3.2.1 The dimensionality problem

Let us now consider what happens if we increase the dimensionality by one. For clarity purposes, imagine that we would like to condition the causal relationship on one additional variable, denoted by Q , so that the null hypothesis of conditional independence becomes³

$$Y_{t+1}|(X_t^{l_X}; Y_t^{l_Y}, Q_t^{l_Q}) \sim Y_{t+1}|(Y_t^{l_Y}, Q_t^{l_Q}).$$

Let us keep $l_X = l_Y = l_Q = 1$. Following the reasoning from the previous section, one may find that the asymptotic normality condition requires β to be in range between $1/(2\alpha)$ and $1/(d_x + d_y + d_z + d_q)$. Given the same standard kernel density estimator with the local bias of order $\alpha = 2$ and $d_x = d_y = d_z = d_q = 1$, one observes that if we increase the dimensionality of the original problem by any number $v \geq 1$, there is no feasible β -region which would endow $T_n(\varepsilon_n)$ with asymptotic normality.

The associated problem results from a too large expected pointwise kernel estimator bias, i.e. $E[s(W_i)]$. By increasing the vector space, we decrease the estimator precision, which seems to play a crucial role in the MSE of the test statistic.

One may relate this problem to the so-called *curse of dimensionality*. As suggested by Scott (1992), in statistics the problem is a consequence of sparsity of data in larger dimensions.

³In practice, it is difficult to find an explicit representation of Q variable. However, one may think of the increased dimensionality problem as conditioning on more than one lag, for instance $Q_t = X_{t-1}$.

Imagine, for instance, a uniform sample over the $[-1, 1]^d$ hypercube, where d is the total number of dimensions. Given arbitrary small region of radius $\mu < 1$, it might be shown that as $d \rightarrow \infty$ the number of points within $[-\mu, \mu]^d$ tends to 0. Straightforward implication suggests that in finite higher dimensional spaces, the smoothing parameter should be larger in order to capture similar number of points. Nevertheless, by increasing the bandwidth window we decrease the precision of the estimator, violating the consistency of the test statistic in this case.

There are several methods which could decrease the dimensionality problem. Scott (1992) suggests principal components method, projection pursuit or informative components analysis. These solutions, however, put additional boundaries on the underlying structure of the data. For instance, they might be of a great advantage when dealing with 100-dimensional spaces where one could assume that the data structure falls into a 20-dimensional manifold. In our example it is very likely, however, that the minimum number of independent manifolds is larger than 3 so that the dimension reduction does not necessary have to improve the test performance. Moreover, as argued by Cartwright (2007), we do not want to decrease the complexity of the environment.

Another solution is a precision improvement, or in other words, reduction of the estimator bias. Since it does not assume any particular underlying data structure, it is of greater advantage in our setting.

3.3 Data sharpening as a bias reduction method

The intuition behind Data Sharpening (DS) is to slightly perturb the original data set by a sharpening function $\psi_p(\cdot)$ in order to obtain the desirable properties of the density estimator \hat{f} (here p is the order of bias reduction). Hall and Minnotte (2002) show that the ‘sharpened’ \hat{f} has smaller bias with variance being of the same order as original \hat{f} . The idea of the perturbation is to tighten the data set, i.e. concentrate points where they were already dense and thin them where they were originally sparse. The explicit form of the sharpening function depends then

on the order of the bias reduction we would like to get but the technique might be in principle applied to obtain arbitrary low levels of bias reduction (Hall and Minnotte, 2002).

There are several advantages of DS among the bias reduction techniques. Firstly, as mentioned before, it allows for very high levels of bias reduction. Since testing for Granger causality is widely recognized for its practical purposes, the universality of a method is of a great importance. Secondly, as we confirm in our study, it does not affect the kernel function directly, which leaves other asymptotic properties of the MSE of the test statistic untouched (see other ε -dependent terms in Eq. (3.7)). Thirdly, it is easy and straightforward to implement, even in a multivariate setting.

With respect to Eq. (3.2), let us consider a sharpened form of the estimator

$$\hat{f}_W^s(W_i) = \frac{\varepsilon^{-d_W}}{n} \sum_j K_{multi} \left(\frac{W_i - \psi_p(W_j)}{\varepsilon} \right), \quad (3.9)$$

where $K_{multi}(W) = (2\pi)^{-d_W/2} \exp(-1/2W^T W)$ is the standard multivariate Gaussian kernel, as described in Wand and Jones (1995) and Silverman (1998).⁴

We obtain the sharpened form of the test statistic, $T_n^s(\varepsilon_n)$, by substituting the sharpened estimators into Eq. (3.3). As we show in Appendix 3.A, the pointwise bias is of order $o(\varepsilon^p)$ with other properties of the kernel \tilde{K} being the same. This in fact makes the bias of $T_n^s(\varepsilon_n)$ (from Eq. (3.4)) being $\alpha = p$ with parameters γ and δ from Eq. (3.5) and Eq. (3.6) unchanged. This reasoning might be summarized in the following corollary, which is a generalization of the theorem in Diks and Panchenko (2006) and proposition in Hall and Minnotte (2002)

Corollary 3.3.1. *For any sufficiently smooth, continuous and infinitely differentiable density, there exist a sharpening function $\psi_p(\cdot)$, where p is the order of bias reduction, for which one may find a sequence of bandwidths $\varepsilon_n = Cn^{-\beta}$ with $C > 0$ and $\beta \in (1/(2p), 1/D)$, where $D < \infty$ is the total dimensionality of the problem, which guarantees that for a weakly-dependent process*

⁴In principle, our reasoning holds for any sufficiently smooth, symmetric and multiplicative probability density as a kernel function. Square kernel, as originally applied by Diks and Panchenko (2006), proves not to be smooth enough which led us to the standard Gaussian kernel.

the sharpened test statistic T_n^s satisfies:

$$\sqrt{n} \frac{(T_n^s(\varepsilon_n) - q)}{S_n} \xrightarrow{d} N(0, 1),$$

where S_n^2 is an autocorrelation consistent estimator of the asymptotic variance of $\sqrt{n}(T_n^s(\varepsilon_n) - q)$.

The proof of Corollary 3.3.1 can be found in Appendix 3.A.

In order to illustrate its practical application, let us consider the same dimensionality problem as described in Section 3.2.1. The original kernel estimator bias of order $o(\varepsilon^2)$, which was effectively blocking the consistency of the test, might be reduced to $o(\varepsilon^4)$ by applying the sharpening function of the form

$$\psi_4(W) = I + h^2 \frac{\kappa_2 \hat{f}'(W)}{2 \hat{f}(W)}, \quad (3.10)$$

where I is the identity function, h is the sharpening bandwidth, κ_2 is the second moment of the kernel and \hat{f}' denotes the first derivative of the density estimator.⁵ For the sake of clarity, the detailed derivations and expressions might be found in the Appendix 3.B. Clearly, it is possible now to find a range for β -values which would guarantee asymptotic normality; in this case it is $\beta \in (1/8, 1/4)$.

There are several other methods of kernel bias reduction. The literature distinguishes *inter alia* among higher order kernels (Granovsky and Mller, 1991), variable bandwidth estimators (Abramson, 1982), variable location estimators (Samiuddin and El-Sayyad, 1990) or parametric transformation methods (Abramson, 1984). Under sufficient smoothness of the underlying density, they all reduce the bias from $o(\varepsilon^2)$ to $o(\varepsilon^4)$ as the sample size increases. Although it is likely that they might be also successfully applied in our setting, their properties do not guarantee a clear-cut asymptotic theory for the test statistic. Therefore, we leave this exercise

⁵We employ the Nadaraya-Watson estimator as a plug-in estimator for sharpening function as suggested by Choi and Hall (1999). This, in fact, makes the optimal sharpening bandwidth h dependent on ε_n .

for future consideration.

3.3.1 Bandwidth selection

The optimal bandwidth, denoted by ε^* , corresponds to the smallest MSE of the test statistic, $T_n^s(\varepsilon_n)$. Following the Diks and Panchenko (2006) methodology, this implies that the sum of dominating squared terms in Eq. (3.7) is minimized, so that under the bandwidth shrinking condition

$$\varepsilon^* = C^* n^{\frac{-2}{2\alpha+7}}, \quad (3.11)$$

with

$$C^* = \left(\frac{18\gamma q_2}{2\alpha E[s(W)]^2} \right)^{\frac{1}{2\alpha+7}}. \quad (3.12)$$

One may readily observe that the general formula for the optimal bandwidth is similar to the one derived in the Diks and Panchenko (2006). DS changes the pointwise bias of the estimator density estimator, intuitively affecting both the rate of convergence, i.e. parameter α , and the leading bias term, i.e. $s(w_i)$.

In order to get more insight into the effects of DS on the optimal bandwidth selection in the DP setting, it is worthwhile to test it in a similar environment as Diks and Panchenko (2006) proposed. Therefore, we consider here an interdependent multivariate ARCH process, however for the sake of presentational purposes, extended to the 3-variate setting and representing the dimensionality problem discussed in the previous section. Consider the ARCH process without instantaneous dependence

$$\begin{aligned} Q_t &\sim N(0, c + aQ_{t-1}^2) \\ X_t &\sim N(0, c + aY_{t-1}^2) \\ Y_t &\sim N(0, c + aQ_{t-1}^2). \end{aligned} \quad (3.13)$$

It is clear that the process satisfies the null that $\{X_t\}$ is not Granger causing $\{Y_t\}$, corrected for the presence of $\{Q_t\}$. Parameters c and a are chosen in order to guarantee stationarity and ergodicity, i.e. $c > 0$ and $0 < a < 1$.

Because of the complexity of the problem, in order to get more insight into the magnitude of the optimal constant C^* , and optimal bandwidth value ε^* , we rely on Monte Carlo simulations. We perform 1000 simulations of process from Eq. (3.13) with $a = 0.4$ and $c = 1$ for different sample sizes. We extract values for \hat{q}_2 and $E[s(W)]$ using standard kernel methods for density and derivative estimation, described in Wand and Jones (1995) and Silverman (1998). The results are presented in Table 3.1.

Table 3.1: Optimal constants and bandwidth values for the $T_n^s(\varepsilon_n)$ test of the 3-variate process from Eq. (3.13) for different sample sizes under the bandwidth shrinking condition. The values represent the mean over 1000 simulations.

sample size (n)	50	100	200	500	1000
C^*	0.83	0.89	0.94	0.97	0.98
ε^*	0.43	0.41	0.39	0.34	0.31

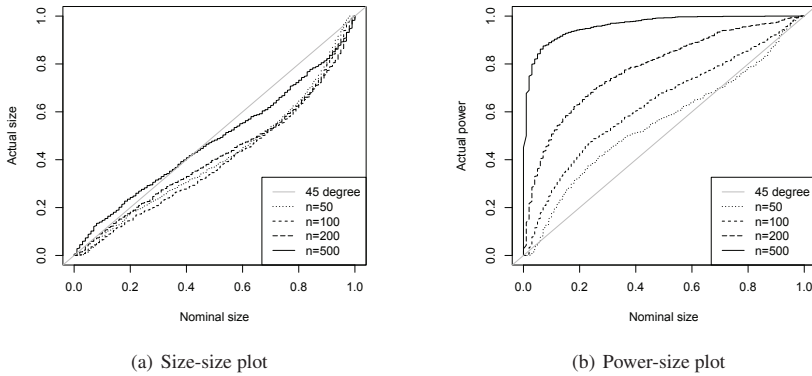
The reported optimal bandwidths are smaller than those from Diks and Panchenko (2006). This is a straightforward result of the DS method. Given that the sharpened estimate has lower bias, the test does not have to include such a wide range of points in order to yield similar properties. This in fact guarantees asymptotic normality of the sharpened test statistic under smaller bandwidth values.

3.3.2 Performance of the DS in Granger causality setting

Given the optimal bandwidth values, we may turn to the assessment of the performance of the DS-augmented DP test. Again we rely here on Monte Carlo simulations. Since process from Eq. (3.13) matches the basic properties of the observed financial time series (like conditional heteroskedasticity), we use it as an underlying behavior for the simulations for our test size assessment. For the test power assessment we use the same process, however, we switch the causality between $\{X_t\}$ and $\{Y_t\}$ so that, even conditioning on $\{Q_t\}$, the null hypothesis of no Granger causality is violated.

The results from 1000 simulations for various time series lengths are summarized by the size-size and power-size plots shown in Fig. 3.1.

Figure 3.1: Size-size and power-size plots of the $T_n^s(\varepsilon_n)$ test of 3-variate process from Eq. (3.13) for different sample sizes under the bandwidth shrinking condition aggregated over 1000 simulations.



One may readily observe that the test demonstrates larger power on larger samples. For 5% significance level, it ranges from 0.05 for $n = 50$ (no power) to 0.82 for $n = 500$ (high power). A simple rule of thumb may suggest that the test yields satisfactory results for samples of length 500 and larger. Interestingly, for the same significance levels and sample sizes, the sharpened DP test offers better power than its original counterpart. In fact, the standard DP test yields power of 0.8 for samples of 1000-2000 length.

At the same time the test tends to be rather conservative for larger nominal p-values, i.e. it under-rejects when the null is satisfied. However, for relatively small significance levels the size-size plot suggests that the larger the sample size, the closer the size is to the ideal rejection probability.

One may view DS as an almost ideal tool for bias reduction. We observe, however, a price for the increased precision of the pointwise estimators. For each point in the distribution the algorithm calculates its sharpened form. This in fact shows up as an additional loop in the procedure, increasing the computational time from $\mathcal{O}(n^2)$ to $\mathcal{O}(n^3)$. For relatively short time series it may not seem as a problem but for n larger than a couple of thousand, computational

time might be a bottle neck of the analysis. Therefore, for larger data sets, we recommend using DS together with multicore or cloud computing.

3.4 Nonlinear Granger causality in the US grain market

In order to show a practical application of the sharpened DP test, we choose the US grain market as it offers an intuitive and straightforward environment for our hypothesis testing. There is a common agreement among professionals that any causal relation between prices of different crops has to be corrected for the weather forecasts at that particular moment (see for instance Popp et al. (2003) and Carreck and Christian (1997)). This conditioning variable suits as a perfect example of Q variable, from the 3-variate example in previous sections.

We consider corn, beans and wheat as being most representative of the US grain market. We consider prices of the 1-month ahead rolling future contracts, traded in USD at the Chicago Board of Trade (CBoT). The weather variable is approximated by the rolling monthly futures on Heating Degree Days (HDD), averaged over Philadelphia, New York, Portland, Chicago and Cincinnati. Daily time series comprise the period from 09/01/2010 till 03/06/2013 making together 633 observations. The data have been obtained from Bloomberg.

We take all variables in logs and evaluate their statistical properties to check whether the time series are stationary. The results are presented in Table 3.2.

Looking at the raw data, only prices of corn prove to be stationary at the 5% significance level. Therefore, in order to assess Granger causality in the market we focus on first differences of all the variables, i.e. log returns.

In the analysis we consider pairwise relations and complete system separately. In the former we take into account the direct relations between two grains only and in the latter we look at the model with all grains included. Since in the system setting, Q variable is two dimensional, by Q_1 we refer to the conditioning on grain and Q_2 to conditioning on weather.

To underpin the results, we relate them with the standard linear Granger causality setting, as

Table 3.2: Unit root tests of the log prices on US grain market in period 09/01/2010 till 03/06/2013 for raw data and for first differences. Test types comprise the Augmented Dickey-Fuller test (ADF) and Phillips-Perron test (PP) as described in Fuller (1995) and Phillips and Perron (1988), respectively. In both tests the null assumes non-stationarity. CV denotes the Critical Value for a given test specification.

Variable	Test type	Trend	5% CV	Raw data		First diff.	
				Test stat.	Unit root	Test stat.	Unit root
Corn	ADF	no	-2.86	-3.585	no	-24.574	no
	PP	no	-2.86	-3.591	no	-24.568	no
	ADF	yes	-3.41	-3.469	no	-24.611	no
	PP	yes	-3.41	-3.493	no	-24.605	no
Bean	ADF	no	-2.86	-2.666	yes	-24.504	no
	PP	no	-2.86	-2.668	yes	-24.496	no
	ADF	yes	-3.41	-2.564	yes	-24.523	no
	PP	yes	-3.41	-2.575	yes	-24.516	no
Wheat	ADF	no	-2.86	-2.299	yes	-24.905	no
	PP	no	-2.86	-2.288	yes	-24.913	no
	ADF	yes	-3.41	-2.272	yes	-24.890	no
	PP	yes	-3.41	-2.261	yes	-24.898	no
HDD	ADF	no	-2.86	-1.247	yes	-21.707	no
	PP	no	-2.86	-1.547	yes	-22.048	no
	ADF	yes	-3.41	-1.409	yes	-21.690	no
	PP	yes	-3.41	-1.739	yes	-22.032	no

proposed by Granger (1969). We also investigate the causality among the VAR-filtered residuals, making sure that discovered causality effects are the results of nonlinearities. We study the explicit role of the weather variable by comparing our results with the original DP test, i.e. without a conditioning variable. In the analysis we assume the lag of each conditioning variable to be 1, as suggested by the Bayesian Information Criterion from the VAR specification. The optimal bandwidth value for the original DP test is equal 1.27 and for the sharpened test it is 0.33. Before running the tests, we standardize the data by either standard normal or uniform marginal transformations. The results for the pairwise relations are presented in Tables 3.3 and 3.4 and for the complete system in Tables 3.5 and 3.6. Graphical illustration of the results can be found in Appendix 3.C in Tables 3.C.1 through 3.C.4.

Table 3.3: Causality results for the pairwise relations of the log returns on the US grain market, without conditioning on Weather (HDD). (*), (**), (***) denotes p-value statistical significance at 10%, 5% and 1%. Period: 09/01/2010-03/06/2013. Nonlinear tests are performed on standardized data, assuming (N)ormal or (U)niform transformation. Number of lags is $l_X = l_Y = 1$ from the Bayesian Information Criterion. Graphical illustration can be found in Table 3.C.1.

Variables	Linear Granger Causality			Nonlinear Granger Causality (N)			Nonlinear Granger Causality (U)		
	Raw data	VAR residuals	VAR residuals	Raw data	VAR residuals	VAR residuals	Raw data	VAR residuals	VAR residuals
X	Y → X	X → Y	Y → X	X → Y	Y → X	Y → X	X → Y	Y → X	X → Y
Corn				***	**	**	***	**	***
Corn	*								*
Beans									
Wheat									

Table 3.4: Causality results for the pairwise relations of the log returns on the US grain market, with conditioning on Weather (HDD). (*), (**), (***) denotes p-value statistical significance at 10%, 5% and 1%. Period: 09/01/2010-03/06/2013. Nonlinear tests are performed on standardized data, assuming (N)ormal or (U)niform transformation. Number of lags is $l_X = l_Y = l_Q = 1$ from the Bayesian Information Criterion. Graphical illustration can be found in Table 3.C.2.

Variables	Linear Granger Causality			Nonlinear Granger Causality (N)			Nonlinear Granger Causality (U)		
	Raw data	VAR residuals	VAR residuals	Raw data	VAR residuals	VAR residuals	Raw data	VAR residuals	VAR residuals
X	Y → X	X → Y	Y → X	X → Y	Y → X	Y → X	X → Y	Y → X	X → Y
Corn				**	**	**	**	**	**
Corn	*								
Beans									
Wheat									

Table 3.5: Causality results for the system setting of the log returns on the US grain market, without conditioning on Weather (HDD). $(*)$, $(**)$, $(***)$ denotes p-value statistical significance at 10%, 5% and 1%. Period: 09/01/2010-03/06/2013. Nonlinear tests are performed on standardized data, assuming (N)ormal or (U)niform transformation. Number of lags is $l_X = l_Y = l_{Q_1} = 1$ from the Bayesian Information Criterion. Graphical illustration can be found in Table 3.C.3.

Variables	Linear Granger Causality			Nonlinear Granger Causality (N)			Nonlinear Granger Causality (U)		
	Raw data	VAR residuals	Raw data	VAR residuals	Raw data	VAR residuals	Raw data	VAR residuals	
X	Y → X	X → Y	Y → X	X → Y	Y → X	X → Y	Y → X	X → Y	
Corn			***	***		**		***	
Corn	*					**		**	
Beans									
Wheat									

Table 3.6: Causality results for the system setting of the log returns on the US grain market, with conditioning on Weather (HDD). $(*)$, $(**)$, $(***)$ denotes p-value statistical significance at 10%, 5% and 1%. Period: 09/01/2010-03/06/2013. Nonlinear tests are performed on standardized data, assuming (N)ormal or (U)niform transformation. Number of lags is $l_X = l_Y = l_{Q_1} = l_{Q_2} = 1$ from the Bayesian Information Criterion. Graphical illustration can be found in Table 3.C.4.

Variables	Linear Granger Causality			Nonlinear Granger Causality (N)			Nonlinear Granger Causality (U)		
	Raw data	VAR residuals	Raw data	VAR residuals	Raw data	VAR residuals	Raw data	VAR residuals	
X	Y → X	X → Y	Y → X	X → Y	Y → X	X → Y	Y → X	X → Y	
Corn			***	***	*	***	*	**	
Corn	*		***	***		***		**	
Beans			***	***		***		**	
Wheat									

One may readily observe that the US grain market does not show much linear Granger causality. The only exception is the possible impact of beans on corn prices in all settings. After VAR filtering this relation disappears, however, as expected.

Interestingly, our results suggest that the relations between US grain prices exhibit a lot of nonlinearities. Looking at the basic pairwise setting, there are strong causal linkages between corn and wheat. If we, however, condition on weather forecasts, some of the relations vanish, in particular in the uniform transformation setting. Moreover, we observe that after conditioning, some new causal relations emerge between corn and beans, which additionally are purely nonlinear in nature. It suggests that weather forecasts have a dual role in the grain market. They do not only drive many of the causal relations themselves but they also mask some of the others in the bivariate setting. From our pairwise results it is clear that weather is masking the corn-beans whereas is driving wheat-corn Granger causality.

In the basic system setting the corn-wheat causal relation is preserved, being significant also after linear filtering. The corn-beans influence is also visible in the uniform transformation setting, confirming the pairwise results. After conditioning on weather forecasts, however, we observe the emergence of the Granger causal relation in the beans-wheat market. Interestingly, the multivariate setting exhibits many regularities from the pairwise study, nevertheless, because of the complexity of the environment, we see new nonlinear relations between all the grains. Prices of corn are Granger causing those of wheat and beans, whereas prices of beans are also influencing those of wheat, conditioning on the weather forecasts.⁶

A straightforward explanation of our results could be that the nonlinear causal relation emerge from bigger to smaller markets. Corn is the most heavily traded grain on the CBoT, followed by beans and wheat. Intuitively, bigger markets should affect those of smaller size as they are deeper and more liquid (Sari et al., 2012). This reasoning is fact in line with our previous finding on the dual role of the weather forecasts in grain market. Since the majority of

⁶Interestingly, conditioning on weather forecast does not fully remove the bias in the corn-wheat relation, as in the pairwise nonlinear test (raw data normal adjustment) and in the system nonlinear tests (raw data normal and uniform adjustments). This, in fact suggests that there can exist additional factors which could influence the corn-wheat price relation, being a potential topic for further investigation.

shocks in the grain market are weather-related, they serve as a common factor and are displayed in all the markets, mitigating the effects of the grain-specific shocks. Correcting for the weather stance allows, therefore, to reveal causal relations between grain-specific shocks, which spread from larger to smaller markets.

3.5 Conclusions and discussion

This study contributes to the ongoing discussion on the validity of the Granger causality concept, allowing it to be applied in more complex environments. We show that the Granger causality test proposed by Diks and Panchenko (2006) lacks consistency in a multivariate setting. The problems arise as a consequence of a too large pointwise estimator bias, which decreases the precision of the tests statistic and affects its asymptotic properties. In order to bring back its desirable properties we propose a sharpened form of the test statistic, which under mild regularity conditions is again asymptotically normal. In fact, we confirm that the sharpening function reduces the original bias of the estimator without any consequences for its further properties, as originally suggested by Hall and Minnotte (2002). We assess properties of the sharpened test numerically, demonstrating that its power is larger than that of the basic DP test.

In order to show the practical side of our study, we apply the test to the US grain market as, because of its weather-dependent structure, it serves as an ideal environment to assess our methodology. We consider Granger causality between corn, beans and wheat, conditioning on the weather forecasts, approximated by the future contracts on Heating Degree Days. Our results suggest that the US grain market exhibits many nonlinear relations. We discover a dual role of the weather forecasts. Firstly, they seem to drive the causal relation from wheat to corn, in the pairwise setting as they serve as a common factor. Secondly, they are masking the causal relations from corn to beans and from beans to wheat in the system setting. Correcting for the common factor, we reveal the true nonlinear Granger casual relations in the US grain market, suggesting that the causality spreads from bigger, i.e. deeper and more liquid, to smaller

markets.

Our results might have important further implications for the food market analysis. As suggested by Gilbert (2010), future contracts are the major transition channel through which macro variables affect food prices. Understanding possible nonlinear economic dynamics in these markets is therefore of a great significance, as it may prevent possible bubbles and instant food price rises, as the ones observed between 2007 and 2008.

Appendix 3.A Asymptotic properties of the sharpened test statistic (Corollary 3.3.1)

We closely follow here the reasoning developed in Diks and Panchenko (2006), however, for the Gaussian kernel and sharpened estimator, as proposed in Hall and Minnotte (2002). We analyze here the case of a random sample as the dependency results follow from the reasoning in Denker and Keller (1983) and Diks and Panchenko (2006). In particular, redefining conditions in Eq. (3.4)-(3.6) for sharpened Gaussian kernel estimators and provided that covariances between local density estimators are taken into account, the asymptotic results hold for weakly dependent time series (see Diks and Panchenko (2006)).

By symmetrization with respect to three different indices i, j, k , for a given ε the sharpened test statistic (Eq. 3.3 with sharpened estimators) might be rewritten in the form of the third order U-statistic as

$$T_n^s(\varepsilon) = \frac{1}{n^3} \sum_{i,j,k} \tilde{K}^s(W_i, W_j, W_k) \quad (3.14)$$

with $W_i = (X_i^{lx}, Y_i^{ly}, Z_i)$, $i = 1, \dots, n$ and sharpened form of the kernel being

$$\begin{aligned} \tilde{K}^s(W_i, W_j, W_k) = & \frac{\varepsilon^{-dx-2dy-dz}}{6} [(G_{ik}^{XYZ} G_{ij}^Y - G_{ik}^{XY} G_{ij}^{YZ}) \\ & + (G_{ij}^{XYZ} G_{ik}^Y - G_{ij}^{XY} G_{ik}^{YZ}) + (G_{jk}^{XYZ} G_{ji}^Y - G_{jk}^{XY} G_{ji}^{YZ}) \\ & + (G_{ji}^{XYZ} G_{jk}^Y - G_{ji}^{XY} G_{jk}^{YZ}) + (G_{ki}^{XYZ} G_{kj}^Y - G_{ki}^{XY} G_{kj}^{YZ}) \\ & + (G_{kj}^{XYZ} G_{ki}^Y - G_{kj}^{XY} G_{ki}^{YZ})]. \end{aligned} \quad (3.15)$$

where $G_{i,j}^W$ is the sharpened form of the multivariate kernel density, i.e.

$$G_{i,j}^W = K_{multi} \left(\frac{W_i - \psi_p(W_j)}{\varepsilon} \right). \quad (3.16)$$

We assume that the density is smooth enough and infinitely differentiable so that it is possi-

ble to find any sharpening function which would guarantee bias reduction of order p , i.e.

$$E_{W_i}[G_i^W] - G_i^W = \varepsilon^p R_p(W_i) + o(\varepsilon^p), \quad (3.17)$$

where $R_p(W_i)$ is the leading bias term associated with ε^p evaluated at point W_i .

Let us define \tilde{K}_1^s and \tilde{K}_2^s as in Conditions in Eq. (3.4)-(3.6). The bias of products of estimated densities, i.e. $s(w_i)$, follows from the properties of the local estimator bias (see the previous section) and identities such as $E[\hat{f}_V \hat{f}_W] = E[(f_V + (\hat{f}_V - f_V))(f_W + (\hat{f}_W - f_W))] = f_V f_W + f_V E[\hat{f}_W - f_W] + f_W E[\hat{f}_V - f_V] + o(\varepsilon^p)$. Therefore, the local bias of the $T_n^s(\varepsilon)$ might be rewritten as proportional to

$$\begin{aligned} \tilde{K}_1^s(w_i, \varepsilon) - \lim_{\varepsilon \rightarrow 0} \tilde{K}_1^s(w_i, \varepsilon) &= \varepsilon^p (f_Y(y_i) R_p(x_i, y_i, z_i) - f_{X,Y}(x_i, y_i) R_p(y_i, z_i) \\ &\quad + f_{X,Y,Z}(x_i, y_i, z_i) R_p(y_i) - f_{Y,Z}(y_i, z_i) R_p(x_i, y_i)) + o(\varepsilon^p). \end{aligned} \quad (3.18)$$

Taking into account Condition in Eq. (3.4), one may find that it holds with $\alpha = p$ and $s(w_i)$ being equal the term in the brackets.

Looking at the Condition in Eq. (3.5), taking the expectations over W_k for each of the contributions to the kernel function \tilde{K}^s , one finds that the dominant terms are proportional to $\varepsilon^{-d_X-2d_Y-d_Z} G_{ij}^{XYZ} G_{ik}^Y$ and $\varepsilon^{-d_X-2d_Y-d_Z} G_{ji}^{XYZ} G_{jk}^Y$, for which we have

$$E_{W_k} [\varepsilon^{-d_X-2d_Y-d_Z} G_{ij}^{XYZ} G_{ik}^Y] = \varepsilon^{-d_X-d_Y-d_Z} G_{ij}^{XYZ} f_Y(Y_i) + o(\varepsilon^{-d_X-d_Y-d_Z}), \quad (3.19)$$

and

$$E_{W_k} [\varepsilon^{-d_X-2d_Y-d_Z} G_{ji}^{XYZ} G_{jk}^Y] = \varepsilon^{-d_X-d_Y-d_Z} G_{ji}^{XYZ} f_Y(Y_j) + o(\varepsilon^{-d_X-d_Y-d_Z}). \quad (3.20)$$

Since all the terms are vanishing ε slower, we can rewrite that $E \left[\left(\tilde{K}_2^s(w_i, w_j) \right)^2 \right]$ is equal

$$\begin{aligned} & \frac{1}{36} E \left[\left(\varepsilon^{-d_X - d_Y - d_Z} E_{W_k} \left[G_{ij}^{XYZ} G_{ik}^Y \right] \right)^2 \right] + o(\varepsilon^{-d_X - d_Y - d_Z}) \\ &= \frac{\varepsilon^{-2d_X - 2d_Y - 2d_Z}}{36} E \left[\left(G_{ij}^{XYZ} f_Y(Y_i) \right)^2 \right] + o(\varepsilon^{-d_X - d_Y - d_Z}) \\ &= \frac{\varepsilon^{-d_X - d_Y - d_Z}}{36} E \left[\left(f_{X,Y,Z}(X_i, Y_i, Z_i) f_Y(Y_i) \right)^2 \right] + o(\varepsilon^{-d_X - d_Y - d_Z}), \end{aligned} \quad (3.21)$$

where we exploited the fact that $G_{ij}^{XYZ} G_{ik}^Y$, $G_{ji}^{XYZ} G_{jk}^Y$ are asymptotically perfectly correlated as ε tends to 0 sufficiently slowly as $n \rightarrow \infty$. This confirms that the γ parameter from the original Diks and Panchenko (2006) methodology is unaffected by the DS, being equal $d_X + d_Y + d_Z$ with

$$q_2 = \frac{1}{36} E \left[\left(f_{X,Y,Z}(X_i, Y_i, Z_i) f_Y(Y_i) \right)^2 \right]. \quad (3.22)$$

Since the variance of \tilde{K}^s is limited by $\varepsilon^{d_X + 2d_Y + d_Z}$ as the sample size increases, condition from Eq. (3.6) holds for $\delta = d_X + 2d_Y + d_Z$, again being the same as in Diks and Panchenko (2006). This brings us to the conclusion that DS decreased the local bias of $T_n^s(\varepsilon)$ only, leaving the further MSE asymptotic properties of the test statistic unchanged (see Eq. (3.7)).

Appendix 3.B Application of bias reduction

For practical purposes, let us assume that the $H = \text{diag}(\varepsilon, \varepsilon, \dots, \varepsilon)$ is a $d_W \times d_W$ bandwidth matrix so that the local density estimator of d_W -variate random vector from Eq. (3.2) becomes

$$\hat{f}_W(W_i) = \frac{\varepsilon^{-d_W}}{n} \sum_j K_H(W_i - W_j), \quad (3.23)$$

where $K_H(W_i - W_j) = K(H^{-1}(W_i - W_j)) = K_{\text{multi}}((W_i - W_j)/\varepsilon)$.

Assume also that the density function is infinitely differentiable and let f' be the vector of first-order partial derivatives of f , f'' be the matrix of second-order partial derivatives of f , $f^{(3)}$ be the cube of third-order partial derivatives of f , $f^{(4)}$ be the 4-dimensional matrix of fourth-

order partial derivatives of f etc, with all the entries being piecewise continuous and square integrable. For presentational purposes, let us also use \int as a shorthand for $\int \cdots \int_{\mathbb{R}^{d_W}}$ and dW as a shorthand for $dW_1 \cdots dW_{d_W}$. By I_{d_W} we denote also the $d_W \times d_W$ identity matrix.

Let us consider the case study example from the chapter, where we extend the basic analysis to the 3-variate causality testing, i.e. $Y_{t+1}|(X_t^{LX}, Y_t^{LY}, Q_t^{LQ}) \sim Y_{t+1}|(Y_t^{LY}, Q_t^{LQ})$. As it is shown in the text, the standard DP test lacks consistency because of the too large pointwise estimator bias. The original bias of the standard kernel density estimator at point W_i might be computed from the second order Taylor expansion around the estimation point (Wand and Jones, 1995)

$$\begin{aligned}
 E \left[\hat{f}_W(W_i) \right] &= \frac{\varepsilon^{-d_W}}{n} \sum_j E [K_H(W_i - W_j)] = \frac{\varepsilon^{-d_W}}{n} \sum_j \int_{-\infty}^{\infty} K_H(W_i - W) f(W) dW \\
 &= \varepsilon^{-d_W} \int_{-\infty}^{\infty} K_H(W_i - W) f(W) dW = \int_{-\infty}^{\infty} K_{multi}(s) f(W - \varepsilon s) ds \\
 &= \int_{-\infty}^{\infty} K_{multi}(s) \left\{ f(W_i) - \varepsilon s^T f'(W_i) + \frac{\varepsilon^2}{2} s^T f''(W_i) s + o(\varepsilon^4) \right\} ds \\
 &= f(W_i) + \frac{\varepsilon^2}{2} \kappa_2 tr \{ f''(W_i) \} + o(\varepsilon^4) = f(W_i) + o(\varepsilon^2),
 \end{aligned} \tag{3.24}$$

where we exploited the fact that $\int s K_{multi}(s) ds = 0_{d_W}$ and $\int s s^T K_{multi}(s) ds = \kappa_2 I_{d_W}$.

The dominant term in the local estimator bias (R_2) is driven by $1/2\kappa_2 tr \{ f''(W_i) \}$, which is of order $o(\varepsilon^2)$. The idea of DS is to eliminate this term by applying appropriate sharpening function. It can be best illustrated by calculating the expected value of the sharpened estimator where the DS function is given by Eq. (3.10) with the sharpening bandwidth h being

ε -dependent

$$\begin{aligned}
 E \left[\hat{f}_W^s(W_i) \right] &= \frac{\varepsilon^{-d_W}}{n} \sum_j E [K_H(W_i - \psi_4(W_j))] = \frac{\varepsilon^{-d_W}}{n} \sum_j \int_{-\infty}^{\infty} K_H(W_i - \psi_4(W_j)) dF(W) \\
 &= \varepsilon^{-d_W} \int_{-\infty}^{\infty} K_H(W_i - \psi_4(W)) dF(W) = \varepsilon^{-d_W} \int_{-\infty}^{\infty} K_H(W_i - V) dF(\psi_4^{-1}(V)) \\
 &= \varepsilon^{-d_W} \int_{-\infty}^{\infty} K_H(W_i - V) f(\psi_4^{-1}(V)) \left| \frac{\partial \psi_4^{-1}(V)}{\partial V} \right| dV \\
 &= \varepsilon^{-d_W} \int_{-\infty}^{\infty} K_H(W_i - V) \left\{ f(V) - \frac{\varepsilon^2}{2} \kappa_2 \text{tr} \{ f''(V) \} + \frac{\varepsilon^4}{4} k_2^2 U(V) + o(\varepsilon^6) \right\} dV \\
 &= \int_{-\infty}^{\infty} K_{multi}(s) \left\{ f(W_i) - \varepsilon s^T f'(W_i) + \frac{\varepsilon^2}{2} s^T f''(W_i) s - \frac{\varepsilon^2}{2} \kappa_2 \text{tr} \{ f''(W_i) \} \right. \\
 &\quad \left. + \frac{\varepsilon^3}{2} \kappa_2 \text{tr} \{ s^T f^{(3)}(W_i) \} - \frac{\varepsilon^4}{4} \kappa_2 \text{tr} \{ s^T f^{(4)}(W_i) s \} + \frac{\varepsilon^4}{4} k_2^2 U(W_i) \right\} ds + o(\varepsilon^6) \\
 &= f(W_i) + \varepsilon^4 R_4(W_i) + o(\varepsilon^6),
 \end{aligned} \tag{3.25}$$

where

$$\begin{aligned}
 U(V) &= \frac{f'(V)^T f'(V) f'(V)^T f'(V)}{f(V)^3} - \frac{5f'(V)^T f''(V) f'(V)}{2f(V)^2} \\
 &\quad - \frac{2f'(V)^T (B_1(f''(V)) - f''(V)) f'(V)}{2f(V)^2} + \frac{\text{tr} \{ f''(V)^T f''(V) \} - \sum |B_2(f''(V))|}{f(V)} \\
 &\quad + \frac{\text{tr} \{ f'(V)^T f^{(3)}(V) \}}{f(V)},
 \end{aligned} \tag{3.26}$$

and

$$R_4(W_i) = \frac{1}{4} (\kappa_2^2 U(W_i) - \kappa_4 \text{tr} \{ f^{(4)}(W_i) \}). \tag{3.27}$$

Matrix transformation $B_1(\cdot)$ puts the trace of the argument on each of the diagonal entries and $B_2(\cdot)$ takes 2x2 submatrix around the diagonal of the argument.

Clearly, the original bias of order $o(\varepsilon^2)$ has decreased to the order $o(\varepsilon^4)$ without any effect on the kernel function K_{multi} , leaving the further properties of Eq. (3.3) the same as in the original reasoning from Diks and Panchenko (2006). Therefore one may calculate optimal bandwidth

values, which endow the test statistic with asymptotic normality, for the 3-variate setting from Eq. (3.11) by plugging in the estimates for $s(w_i)$ and q_2 as

$$\begin{aligned} s(w_i) &= f_{Y,Q}(y_i, q_i)R_4(x_i, y_i, z_i, q_i) - f_{X,Y,Q}(x_i, y_i, q_i)R_4(y_i, z_i, q_i) \\ &\quad + f_{X,Y,Z,Q}(x_i, y_i, z_i, q_i)R_4(y_i, q_i) - f_{Y,Z,Q}(y_i, z_i, q_i)R_4(x_i, y_i, q_i), \end{aligned} \quad (3.28)$$

and

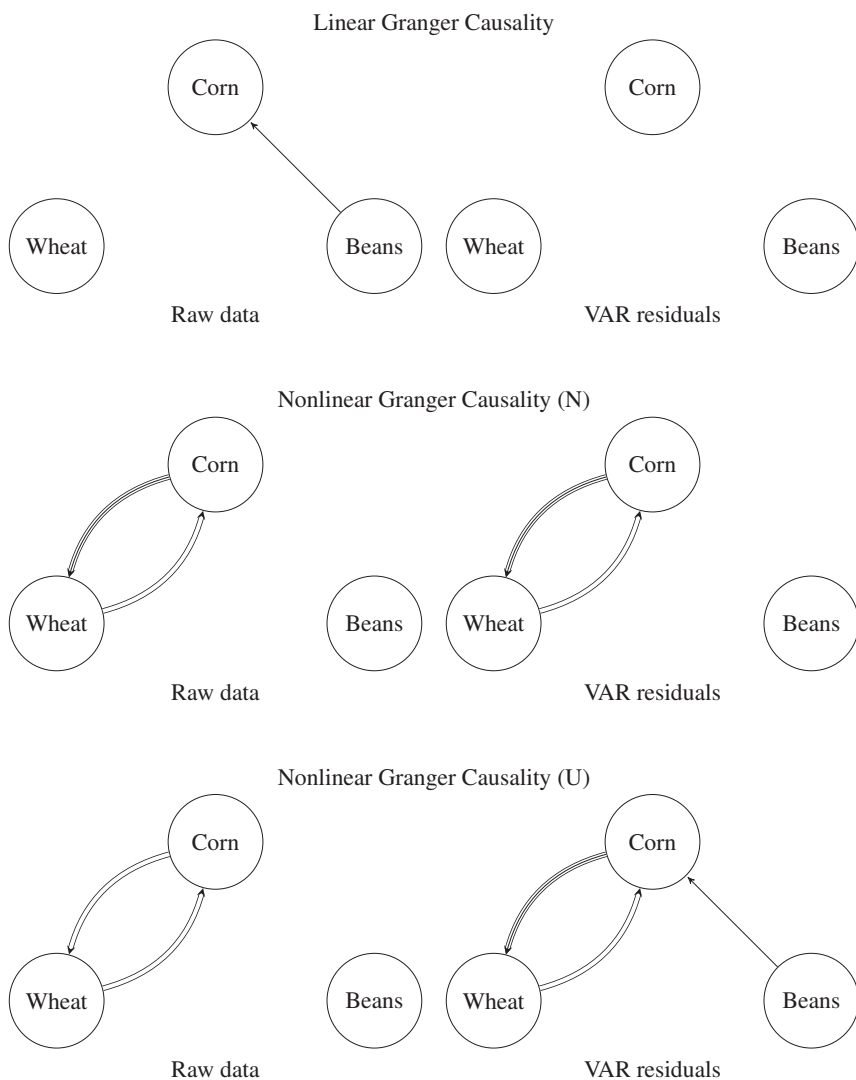
$$q_2 = \frac{1}{36} E [(f_{X,Y,Z,Q}(X_i, Y_i, Z_i, Q_i)) f_{Y,Q}(Y_i, Q_i)^2]. \quad (3.29)$$

In fact, bias reduction from $o(\varepsilon^2)$ to $o(\varepsilon^4)$ allows to include up to 4 additional variables. Any additional conditioning variable would again violate the consistency of the test, requiring more appropriate sharpening function.

Appendix 3.C Illustration of the empirical results

(See Tables 3.C.1 through 3.C.4 on the next pages.)

Table 3.C.1: Causality results for the pairwise relations of the log returns on the US grain market, without conditioning on Weather (HDD). Single, double and triple arrows denote p-value statistical significance at 10%, 5% and 1%. Period: 09/01/2010-03/06/2013. Nonlinear tests are performed on standardized data, assuming (N)ormal or (U)niform transformation. Number of lags is $l_X = l_Y = 1$ from the Bayesian Information Criterion.



APPENDIX 3.C. ILLUSTRATION OF THE EMPIRICAL RESULTS

Table 3.C.2: Causality results for the pairwise relations of the log returns on the US grain market, with conditioning on Weather (HDD). Single, double and triple arrows denote p-value statistical significance at 10%, 5% and 1%. Period: 09/01/2010-03/06/2013. Nonlinear tests are performed on standardized data, assuming (N)ormal or (U)niform transformation. Number of lags is $l_X = l_Y = l_Q = 1$ from the Bayesian Information Criterion.

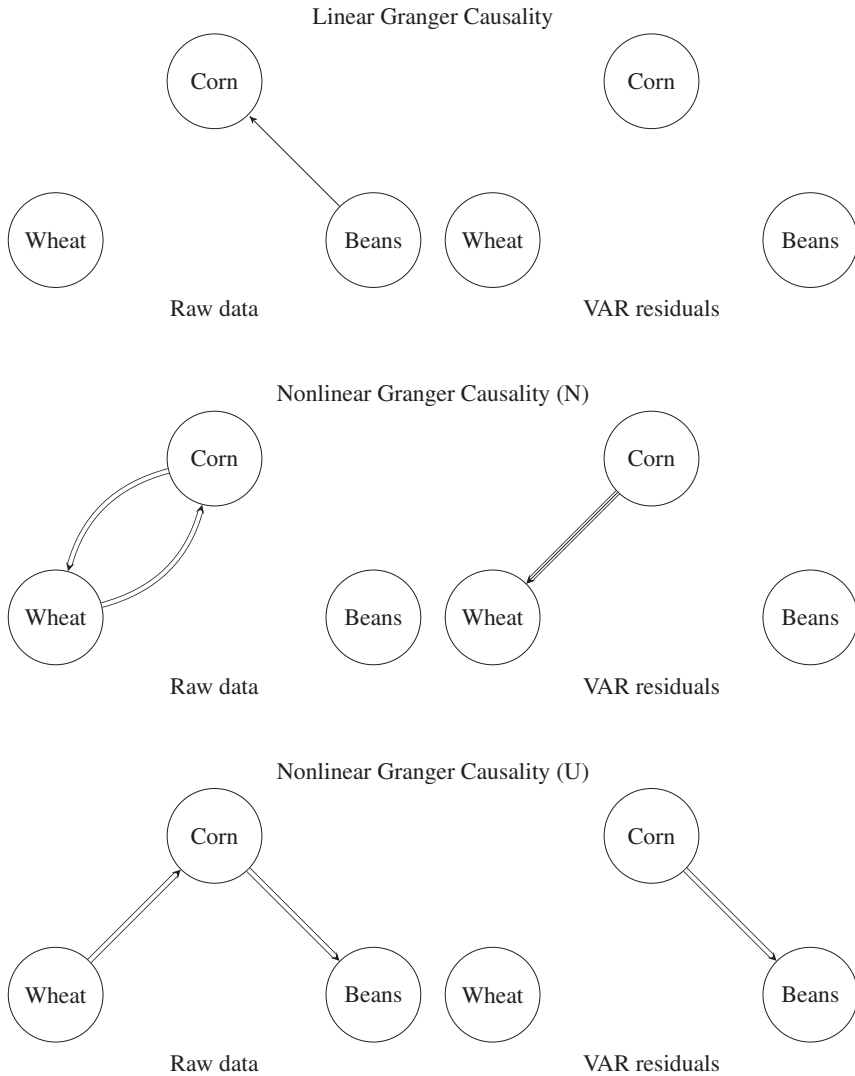
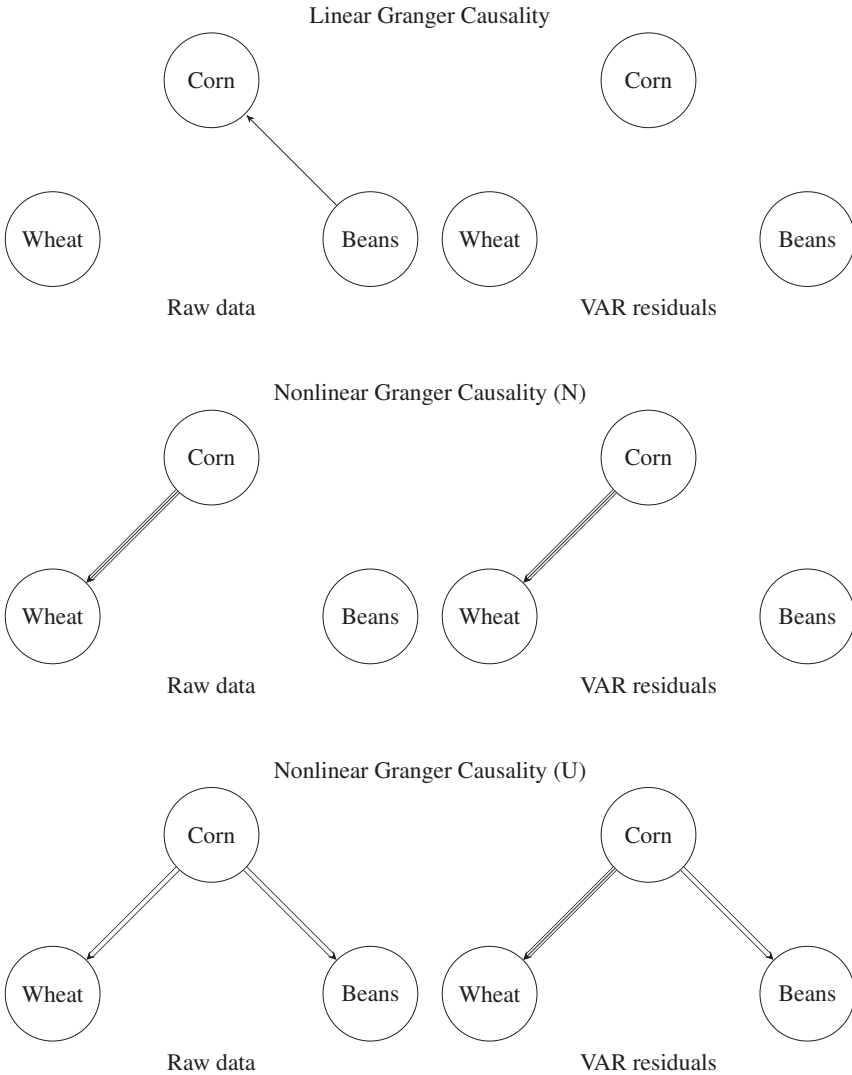
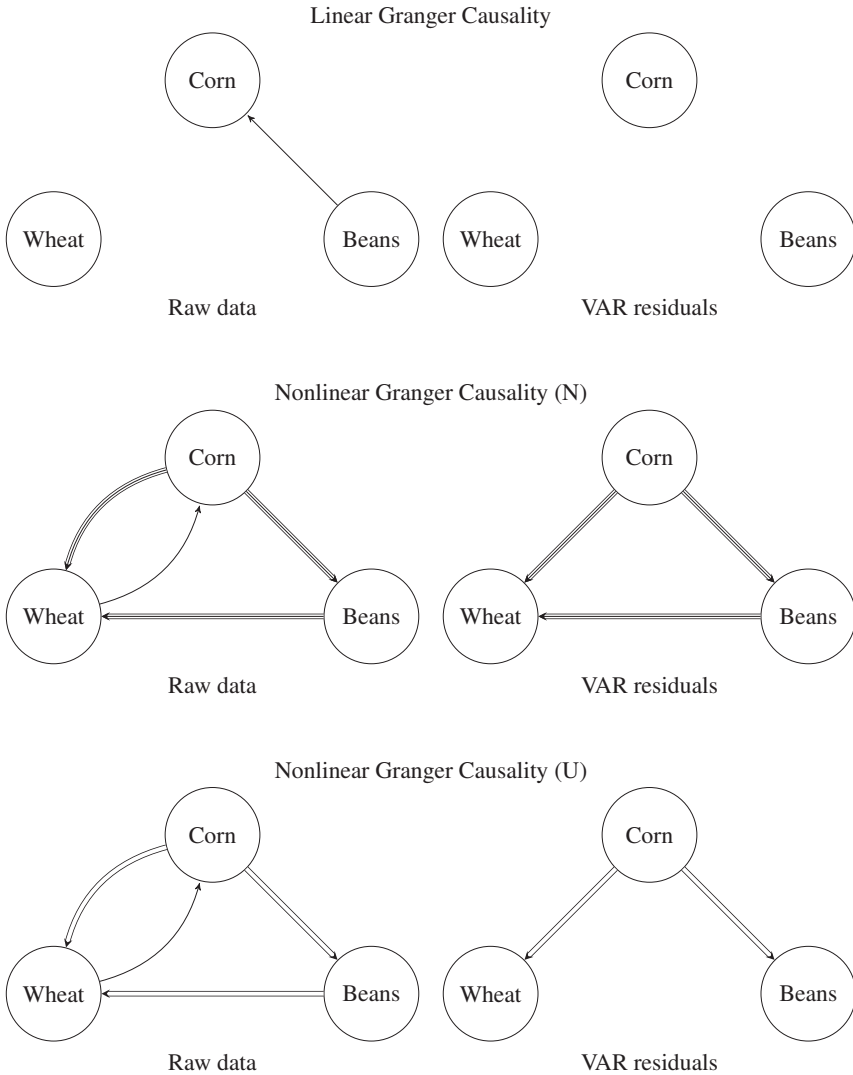


Table 3.C.3: Causality results for the system setting of the log returns on the US grain market, without conditioning on Weather (HDD). Single, double and triple arrows denote p-value statistical significance at 10%, 5% and 1%. Period: 09/01/2010-03/06/2013. Nonlinear tests are performed on standardized data, assuming (N)ormal or (U)niform transformation. Number of lags is $l_X = l_Y = l_{Q_1} = 1$ from the Bayesian Information Criterion.



APPENDIX 3.C. ILLUSTRATION OF THE EMPIRICAL RESULTS

Table 3.C.4: Causality results for the system setting of the log returns on the US grain market, with conditioning on Weather (HDD). Single, double and triple arrows denote p-value statistical significance at 10%, 5% and 1%. Period: 09/01/2010-03/06/2013. Nonlinear tests are performed on standardized data, assuming (N)ormal or (U)niform transformation. Number of lags is $l_X = l_Y = l_{Q_1} = l_{Q_2} = 1$ from the Bayesian Information Criterion.



Chapter 4

Exploring Nonlinearities in Financial Systemic Risk

4.1 Introduction

The 2007-2009 crisis shed new light on the complexity within the financial sector. The linkages and risk exposures between various institutions proved to be of great significance in transmitting distress across the whole financial system. Additionally, during systemic events the malaise spreads across the financial world rapidly through indirect channels, like price effects or liquidity spirals (Brunnermeier, 2009). In effect, market values of various financial assets tend to move closer together, drifting away from their fundamentals. In particular, one observes high regularities in their tail co-movements (Adrian and Brunnermeier, 2011).

Because of its strong adverse effects on the real economy, great attention has been paid to measuring and monitoring systemic risk, i.e. risk of disruption in the entire financial system, and individual risk exposures. The majority of econometric approaches in these fields focus on co-risk measures, where the risk of the financial system is assessed in relation to the risk of individual institutions. The intuition behind these models lies in negative externalities which one institution imposes on the others and on the system as a whole. As argued by Adrian

and Brunnermeier (2011), these externalities are a consequence of excessive risk taking and leverage. Given, for instance, that one institution is facing a liquidity shock, it liquidates its assets at fire-sale prices as given, affecting borrowing constraints of others and actually causing the fire-sale prices. A wonderful summary of research in this field can be found in Acharya (2009), Acharya et al. (2010) or Adrian and Brunnermeier (2011).

A commonly used econometric approach, in the growing body of literature on this topic, is Conditional Value-at-Risk (CoVaR), attributed to Adrian and Brunnermeier (2011). It is built around the concept of Value-at-Risk (VaR), which determines the maximum loss on returns within the γ -percentile confidence interval (Kupiec, 2002). CoVaR assesses VaR_γ of one institution conditional on distress in the other. In particular, if the former represents the system, one may associate CoVaR with a systemic risk measure.

A clear shortcoming of such an approach lies in its susceptibility to model misspecification. Imagine that returns come from an unknown probability distribution F , with density f . Assume now that f is steeper or nonlinear around its VaR_γ . Clearly, standard parametric approaches oversee this irregularity so that even a small variation in VaR_γ might affect co-risk results. In this chapter we develop a methodology which corrects for this shortcoming, contributing to the discussion on nonlinear economic dynamics in systemic risk.

The existence of nonlinearities in the field has been already recognized. Huang et al. (2010) suggest that “a bank’s contribution to the systemic risk is roughly linear in its default probability and highly nonlinear with respect to institution size and asset correlation”. This is supported by empirical observations of the financial markets described by He and Krishnamurthy (2012). In fact, He and Krishnamurthy (2012) built a theoretical model which matches nonlinear dynamics across different economic variables, including systemic risk. XiaoHua and Shiyang (2012) investigated the topic from the neural network perspective and designed an early warning mechanism accordingly. This chapter aims to propose a formal approach to assess the relevance of nonlinearities in driving systemic events.

We build our approach around the intuition of CoVaR. In particular, we focus on the Granger

causal effect that distress in one institution may lead to distress in the other or in the whole system, where distress is defined by being near VaR_γ .

There are two main novelties in our methodology. The first one is the notion of causality. The basic CoVaR notion does not distinguish between direct causal and common factor effects. Adrian and Brunnermeier (2011) treat this as a virtue rather than a problem, arguing that common factor effects are of more importance when dealing with systemic risk, which can be expected to be particularly true for the herding behavior (Brunnermeier et al., 2009). One may, however, want to study the causal relations explicitly. Imagine for instance a group of the biggest financial institutions. Since they do not only trade with each other but also serve as clearing houses or liquidity backstops for smaller parties, they are central to the financial system. Now, imagine that one of them is in trouble. It affects all the banks that are exposed to its risk, but since it is relatively large, its distress might alone translate into problems in the entire financial system. The causal kind of reasoning seems therefore particularly appealing for policy makers and central bankers, who in fact might want to focus on preventing this individual causal relation.

Another justification for considering causality in individual and systemic risk lies in its possible applications to networks and contagion analysis (see for instance Chinazzi and Fagiolo (2013)). Looking at any pair of institutions, the possible risk effects of one on another do not have to be bilaterally equal (as they are assumed to be in a non-causal setting). For instance, a lender has a different kind of risk exposure to a creditor than the other way around. Causality captures that phenomenon explicitly, allowing for a more detailed analysis on network spillovers, cascades and shock propagation.

In our study we employ the general causality of Granger type, i.e. a nonparametric version of the concept originally proposed by Granger (1969), as it is intuitive and does not bring many model restrictions. It has been also successfully applied as a network mapping tool in financial analysis (Gao and Ren, 2013).

The second novelty lies in the definition of financial distress. In our study we assume that an

institution is in trouble when it is around its VaR_γ . Practically speaking, our definition captures the majority of events which fall below VaR_γ together with some of the events above it. The reason why we allow for some variation around VaR_γ lies in its possible nonlinear structure, whose role we want to study explicitly. We recognize that our definition might not capture some of the extreme values from the left tail of the distribution, being potentially susceptible to *black swans* (Taleb, 2010). Our analysis shows, however, that the optimal region around VaR_γ is very slowly decreasing with the sample size, somehow hampering the risk of neglecting the extreme events. Additionally, our setup might be naturally extended to a more general setting, including all the events below VaR_γ . This, however, is behind the scope of this chapter and we leave it for further investigation.

In our analysis we consider two scenarios of potential Granger causality. In the first setting we investigate the role of individual institutions in blocking the recovery of the system which is already under distress. In the second scenario we measure the contribution of individual institutions to the systemic problems. The second setting is more similar to the standard understanding of systemic risk (Acharya, 2009) and might be useful in *ex ante* applications. The first scenario might be perceived either as a kind of a robustness check or a policy relevant tool for *ex post* actions. Indeed, if the system is already in trouble one may want to determine which of its parts are hampering its recovery. In fact, we could think of these two scenarios from a perspective of a doctor who either prescribes precautionary drugs or is trying to heal an already sick patient.

This chapter is organized as follows. In Section 4.2 we explain the methodology of Conditional Value-at-Risk-Nonlinear Granger Causality (or NCoVaR for simplicity). We evaluate the asymptotic properties of the test statistic and we confirm them numerically in Section 4.3. In Section 4.4 we apply our approach to the euro zone financial sector and evaluate which institutions got the most significant impact on the systemic risk in years 2000-2012. Section 4.5 concludes.

4.2 Methodology of NCoVaR

Let us first bring some intuition behind the Conditional Value-at-Risk and Granger causality separately and then use this to build CoVaR-NGraCo (Conditional Value-at-Risk-Nonlinear Granger Causality) or NCoVaR for simplicity. In the standard setting we consider two institutions, i and j , whose returns on assets are given by X^i and X^j , respectively. Talking about systemic risk, we set j to be some aggregate variable so that we investigate the relationship between institution i and the system as a whole. Following the original CoVaR literature, let us define VaR_γ as the left γ -quantile of the unconditional returns of a given institution. (In practice γ is chosen from $\{0.01, 0.05, 0.1\}$.) For institution i we have therefore

$$P(X^i \leq \text{VaR}_\gamma^i) = \gamma, \quad (4.1)$$

or equivalently

$$\text{VaR}_\gamma^i = \inf\{x^i : F_{X^i}(x^i) \geq \gamma\}, \quad (4.2)$$

where F_{X^i} is the cumulative distribution function of X^i . (For institution j , the notation is analogous throughout the chapter.) The intuition behind CoVaR is to evaluate VaR_γ of institution j conditional on some event associated with institution i . In particular, Adrian and Brunnermeier (2011) consider two conditioning events, i.e. institution i is at its VaR_γ^i or at its median ($\text{VaR}_{\gamma=0.5}^i = \text{Median}^i$). By comparing the difference between the two, it is possible to estimate the risk contribution of institution i onto j , denoted by ΔCoVaR .

In our study we follow a similar reasoning as Adrian and Brunnermeier (2011), however, we add a (discrete) time dimension. For any period t , let us define the future returns' information set by $\mathcal{G}X_t^i$, and the past and/or current returns' information set by $\mathcal{F}X_t^i$. Following Granger (1969), we say that returns of institution i are Granger causing those of institution j if $\mathcal{F}X_t^i$ contains additional information on $\mathcal{G}X_t^j$ which is not already contained in $\mathcal{F}X_t^j$ alone. We formulate the definition of conditional Granger causality analogously, i.e. we say that returns of institution i are Granger causing those of institution j if, conditional on some past or current

events of those institutions (denoted by $\mathcal{A}(\mathcal{F}X_t^i)$ and $\mathcal{B}(\mathcal{F}X_t^j)$, respectively), $\mathcal{F}X_t^i$ contains additional information on $\mathcal{G}X_t^j$ which is not already contained in $\mathcal{F}X_t^j$ alone.

Given the intuition behind the CoVaR and general Granger causality, we may now turn to NCoVaR. Similarly to Δ CoVaR, we test the difference in Granger causal risk effects from institution i on j , between two conditioning events, i.e. when institution i is and/or was in trouble (or around its VaR_γ^i) and when it is and/or was around the median of its returns. An advantage of allowing institutions to be around (and not exactly at) their VaR_γ or median levels is that we could thereof account for possible nonlinearities in corresponding distributions - something the original methodology could not capture. In particular, we consider a μ -radius ball ($\mu > 0$) centered at VaR_γ or the median. (The following reasoning holds for \mathcal{G} and \mathcal{F} being multivariate, provided that VaR_γ and the medians are taken over the marginals.) We also allow for conditioning on the past and/or current realizations of X_t^j . To formalize this we give the following definition of NCoVaR.

Definition 4.2.1. *Given any stationary bivariate process $\{(X_t^i, X_t^j)\}$, we say that $\{X_t^i\}$ is a nonlinear CoVaR Granger cause of $\{X_t^j\}$ if*

$$P(\|\mathcal{G}X_t^j - \text{VaR}_\gamma^j\| \leq \mu \mid \|\mathcal{F}X_t^i - \text{VaR}_\gamma^i\| \leq \mu, \mathcal{B}(\mathcal{F}X_t^j)) \neq P(\|\mathcal{G}X_t^j - \text{VaR}_\gamma^j\| \leq \mu \mid \|\mathcal{F}X_t^i - \text{Median}^i\| \leq \mu, \mathcal{B}(\mathcal{F}X_t^j)),$$

where $\mu > 0$, $\|\cdot\|$ is the Euclidian distance measure, \mathcal{G} denotes a set of future realizations and \mathcal{F} denotes a set of past and/or current realizations of the corresponding variables and $\mathcal{B}(\cdot)$ reflects some event over the argument.

In this study, we consider two possible scenarios. In the first, we assume that institution j is already in distress, so that potential Granger causal risk effects from institution i do not only induce even higher losses on j but also can clog its recovery. The second scenario is more similar to the traditional risk analysis, where future troubles in institution j come directly from the past problems of institution j . One may thereof reformulate Def. 4.2.1 in the form of two possible scenarios, which we investigate in detail below.

Scenario 1. Given any stationary bivariate process $\{(X_t^i, X_t^j)\}$, we say that $\{X_t^i\}$ is a nonlinear CoVaR Granger cause of $\{X_t^j\}$ in tail if

$$P(\|\mathcal{G}X_t^j - \text{VaR}_\gamma^j\| \leq \mu \mid \|\mathcal{F}X_t^i - \text{VaR}_\gamma^i\| \leq \mu, \|\mathcal{F}X_t^j - \text{VaR}_\gamma^j\| \leq \mu) \neq \\ P(\|\mathcal{G}X_t^j - \text{VaR}_\gamma^j\| \leq \mu \mid \|\mathcal{F}X_t^i - \text{Median}^i\| \leq \mu, \|\mathcal{F}X_t^j - \text{VaR}_\gamma^j\| \leq \mu),$$

where $\mu > 0$, $\|\cdot\|$ is the Euclidian distance measure, \mathcal{G} denotes a set of future realizations and \mathcal{F} denotes a set of past and/or current realizations of the corresponding variables.

Scenario 2. Given any stationary bivariate process $\{(X_t^i, X_t^j)\}$, we say that $\{X_t^i\}$ is a nonlinear CoVaR Granger cause of $\{X_t^j\}$ in median if

$$P(\|\mathcal{G}X_t^j - \text{VaR}_\gamma^j\| \leq \mu \mid \|\mathcal{F}X_t^i - \text{VaR}_\gamma^i\| \leq \mu, \|\mathcal{F}X_t^j - \text{Median}^j\| \leq \mu) \neq \\ P(\|\mathcal{G}X_t^j - \text{VaR}_\gamma^j\| \leq \mu \mid \|\mathcal{F}X_t^i - \text{Median}^i\| \leq \mu, \|\mathcal{F}X_t^j - \text{Median}^j\| \leq \mu),$$

where $\mu > 0$, $\|\cdot\|$ is the Euclidian distance measure, \mathcal{G} denotes a set of future realizations and \mathcal{F} denotes a set of past and/or current realizations of the corresponding variables.

In practice it is impossible to condition on the infinite sets of future or past realizations of variables of interest. Therefore, we reformulate \mathcal{G} and \mathcal{F} as finite sets of future periods or lags, respectively. We limit ourselves to the canonical setting where $\mathcal{G}X_t^j = X_{t+1}^j$, as it is most commonly used in practical Granger causality testing, however, our reasoning holds for any $\mathcal{G}X_t^j = X_{t+k}^j$, $1 \leq k < \infty$. Similarly, we replace $\mathcal{F}X_t^i$ and $\mathcal{F}X_t^j$ by $X_{t,l_i}^i = \{X_{t-l_i+1}^i, \dots, X_t^i\}$ and $X_{t,l_j}^j = \{X_{t-l_j+1}^j, \dots, X_t^j\}$, where $l_i \geq 1$ and $l_j \geq 1$ denote the number of lags of a corresponding variable.

In Granger causality testing, the goal is to find evidence against the null hypothesis of no causality, which according to Def. 4.2.1 is represented by equivalence in conditional probability. We assume that process $\{(X_t^i, X_t^j)\}$ is strictly stationary. In that case, the null hypothesis is a statement about the invariant distribution evaluated at conditional VaR_γ levels of the $(l_i + l_j + 1)$ -dimensional vector $W_t = (Z_t, X_{t,l_i}^i, X_{t,l_j}^j)$, where we substitute $Z_t = X_{t+1}^j$. (For clarity

purposes and to bring forward the fact that we consider the invariant distribution of W_t , we drop the time index, so that $W = (Z, X^i, X^j)$.) Formally, the null hypothesis from Scenarios 1 and 2 can be rewritten as

$$f_{Z, X^i, X^j}(z_\gamma | x_\gamma^i, x_*^j) = f_{Z, X^i, X^j}(z_\gamma | x_m^i, x_*^j), \quad (4.3)$$

where $z_\gamma = \text{VaR}_\gamma^Z$, $x_\gamma^i = \text{VaR}_\gamma^i$, $x_m^i = \text{Median}^i$ and $*$ distinguishes between Scenario 1 and 2 as $x_*^j = \text{VaR}_\gamma^j$ or $x_m^j = \text{Median}^j$, respectively. It is helpful to restate the problem in terms of ratios of joint densities evaluated at given quantiles, as under the null the density of Z evaluated around its VaR_γ level and conditional on specific events in X^i and X^j is equal to the same density conditional on the different set of events in X^i and X^j . Therefore, the joint probability density function, together with its marginals must satisfy

$$\frac{f_{Z, X^i, X^j}(z_\gamma, x_\gamma^i, x_*^j)}{f_{X^i, X^j}(x_\gamma^i, x_*^j)} = \frac{f_{Z, X^i, X^j}(z_\gamma, x_m^i, x_*^j)}{f_{X^i, X^j}(x_m^i, x_*^j)}. \quad (4.4)$$

Since Eq. (4.4) holds for any quantile of the vector (Z, X^i, X^j) in the support of Z, X^i, X^j , Eq. (4.4) might be equivalently rewritten as

$$\frac{f_{Z, X^i, X^j}(z_\gamma, x_\gamma^i, x_*^j)}{f_{X^i, X^j}(x_m^i, x_*^j)} = \frac{f_{X^i, X^j}(x_\gamma^i, x_*^j)}{f_{X^i, X^j}(x_m^i, x_*^j)} \frac{f_{Z, X^i, X^j}(z_\gamma, x_m^i, x_*^j)}{f_{X^i, X^j}(x_m^i, x_*^j)}. \quad (4.5)$$

Analogously to Baeck and Brock (1992) or Hiemstra and Jones (1994), a natural methodology to assess Eq. (4.5) comes from the test for conditional independence. However, as showed by Diks and Panchenko (2005) and Diks and Panchenko (2006), these tests can severely over-reject in Granger causal setting, because its dependence on the conditional variance. Diks and Panchenko (2006) propose to add a positive weight function $g(z_\gamma, x_m^i, x_*^j)$ and, given that the

null should hold in the support of the joint densities, it might be equivalently written as

$$\tau_g \equiv \left(\frac{f_{Z, X^i, X^j}(z_\gamma, x_\gamma^i, x_\gamma^j)}{f_{X^i, X^j}(x_m^i, x_*^j)} - \frac{f_{X^i, X^j}(x_\gamma^i, x_\gamma^j)}{f_{X^i, X^j}(x_m^i, x_*^j)} \frac{f_{Z, X^i, X^j}(z_\gamma, x_m^i, x_*^j)}{f_{X^i, X^j}(x_m^i, x_*^j)} \right) g(z_\gamma, x_m^i, x_*^j) = 0. \quad (4.6)$$

Diks and Panchenko (2006) discuss several possibilities of choosing $g(z_\gamma, x_m^i, x_*^j)$. In this study we focus on $g(z_\gamma, x_m^i, x_*^j) = f_{X^i, X^j}(x_m^i, x_*^j)^2$, as the estimator of τ_g has a corresponding U-statistic representation, bringing the desired asymptotic normality properties for weakly dependent data. Substituting into Eq. (4.6), one finds that

$$\tau = f_{Z, X^i, X^j}(z_\gamma, x_\gamma^i, x_\gamma^j) f_{X^i, X^j}(x_m^i, x_*^j) - f_{X^i, X^j}(x_\gamma^i, x_\gamma^j) f_{Z, X^i, X^j}(z_\gamma, x_m^i, x_*^j). \quad (4.7)$$

To evaluate the data-driven representation of τ , we rely on kernel methods. In particular, we consider the local density estimator

$$\hat{f}_W(w) = \frac{\varepsilon^{-d_W}}{n} \sum_{k=1}^n K\left(\frac{w - w_k}{\varepsilon}\right), \quad (4.8)$$

where n is the sample size, ε is the bandwidth parameter (similar to μ from the Def. 4.2.1), d reflects the dimensionality of a given vector W and $K(\cdot)$ is a bounded Borel function $\mathbb{R}^{d_W} \rightarrow \mathbb{R}$ satisfying

$$\int |K(t)| dt < \infty, \quad \int K(t) dt = 1 \quad \text{and} \quad |tK(t)| \rightarrow 0 \quad \text{as} \quad |t| \rightarrow \infty. \quad (4.9)$$

In practice, $K(\cdot)$ is often chosen to be a probability density function (Wand and Jones, 1995). In order to guarantee the consistency of the pointwise density estimators, we assume that the bandwidth parameter ε comes from the sequence ε_n , which is slowly decreasing with the sample size, i.e.

$$\varepsilon_n \rightarrow 0 \quad \text{and} \quad n\varepsilon_n \rightarrow \infty \quad \text{as} \quad n \rightarrow \infty. \quad (4.10)$$

Parzen (1962) shows that under conditions (4.9) and (4.10) and provided that f is continuous at w , the estimate of density f at a given point w is consistent.

Given a given bandwidth ε , a natural estimator for τ is

$$T_n(\varepsilon) = C \sum_{k=1}^n \sum_{p=1}^n \left[K \left(\frac{(z_\gamma, x_\gamma^i, x_\gamma^j)^T - (z_k, x_k^i, x_k^j)^T}{\varepsilon} \right) K \left(\frac{(x_m^i, x_m^j)^T - (x_p^i, x_p^j)^T}{\varepsilon} \right) - K \left(\frac{(x_\gamma^i, x_\gamma^j)^T - (x_k^i, x_k^j)^T}{\varepsilon} \right) K \left(\frac{(z_\gamma, x_m^i, x_m^j)^T - (z_p, x_p^i, x_p^j)^T}{\varepsilon} \right) \right], \quad (4.11)$$

where ε is the bandwidth and

$$C = \frac{\varepsilon^{-dz-2d_{X^i}-2d_{X^j}}}{n^2}. \quad (4.12)$$

(We sum over two indices as it allows to calculate the variance of $T_n(\varepsilon)$ explicitly.) The asymptotic distribution of the test statistic can be derived from the behavior of the properties of the second order U-statistic, as described by Serfling (1980) and van der Vaart (1998).

Theorem 4.2.1. *Under the conditions described by Eqs. (4.9) and (4.10), for a given set of VaR_γ levels and given bandwidth parameter sequence ε_n , test statistic $T_n(\varepsilon_n)$ satisfies:*

$$\sqrt{n} \frac{(T_n(\varepsilon_n) - \tau)}{S_n} \xrightarrow{d} \mathcal{N}(0, 1),$$

where S_n is a heteroskedasticity and autocorrelation consistent estimator of the asymptotic standard deviation of $\sqrt{n}(T_n(\varepsilon_n) - \tau)$.

The proof of Theorem 4.2.1 can be found in Appendix 4.A. As argued by Diks and Panchenko (2006), although the test statistic is not positive definite, the one-sided version of the test, i.e. rejecting on larger values, turns out to yield better performance.

In this study we choose γ to be 0.05 as it is the most commonly applied VaR significance level. We calculate VaR_γ from the empirical quantile function (Jones, 1992). Following the literature on nonparametric Granger causality testing (Hiemstra and Jones, 1994; Diks and Panchenko, 2006) we take the square kernel function.¹ The square kernel form of the estimator

¹The asymptotic properties of the test statistic are, however, robust to any kernel specification, provided that it

in Eq. (4.8), can be rewritten as

$$\hat{f}_W^{SQ}(w) = \frac{(2\varepsilon)^{-dw}}{n-1} \sum_{k=1}^n I(\|w - w_k\| < \varepsilon), \quad (4.13)$$

where $I(\|w - w_k\| < \varepsilon)$ is the indicator function taking values 1 for any $\|w - w_k\| < \varepsilon$ and zero otherwise, and $\|\cdot\|$ is the supremum norm over all the dimensions.

4.2.1 Optimal bandwidth

Although the asymptotic normality of the test statistic holds for an arbitrary decreasing sequence of bandwidths as long as it satisfies condition from Eq. (4.10), it influences the power of the test to a great extent (Silverman, 1998). Therefore, in order to improve the performance of the test, we calculate the optimal size of the bandwidth explicitly. Following Wand and Jones (1995) and Silverman (1998), the optimal bandwidth minimizes the Mean Squared Error (MSE) of $T_n(\varepsilon_n)$, which may be decomposed into the sum of variance and squared bias of $T_n(\varepsilon_n)$. In our inference it is worthwhile to point out that the optimal bandwidth values of $T_n(\varepsilon_n)$ do not violate the consistency properties of any of the density estimators.

Corollary 4.2.1. *Under the conditions given by Eqs. (4.9) and (4.10), the MSE-optimal sequence of bandwidths of $T_n(\varepsilon_n)$ guarantees consistency of any of the pointwise density estimators contributing to $T_n(\varepsilon_n)$.*

The proof of Corollary 4.2.1 is given in Appendix 4.B. In fact, the MSE optimum rate of convergence of the bandwidth of $T_n(\varepsilon_n)$ is slightly faster than that of individual density estimators, but still much slower than n^{-1} . This is caused by the increased variance of a product of two estimators compared to their individual variances. Therefore, in order to balance this effect in the MSE, the sequence of optimal bandwidths of $T_n(\varepsilon_n)$ should decrease at a slightly faster rate as $n \rightarrow \infty$, but never as fast as n^{-1} . In testing for systemic risk this proves to be of

satisfies conditions (4.9) and (4.10).

large importance as with a bandwidth parameter decreasing just slightly with the sample size we are still able to capture the majority of returns which are left to VaR_γ .

In evaluating the optimal bandwidth value we rely on Monte Carlo methods. Correcting for the weak dependency, we apply the autocorrelation consistent estimator for the variance of $T_n(\varepsilon)$, as proposed in Newey and West (1987). It might be verified that for a given bandwidth ε , the bias of $T_n(\varepsilon)$ may be calculated from the Taylor expansion around any point as

$$\begin{aligned}
 E[T_n(\varepsilon)] - \tau &= \frac{1}{2} \kappa_2 \varepsilon^2 [f_{Z, X^i, X^j}(z_r, x_r^i, x_r^j) \nabla^2 f_{X^i, X^j}(x_s^i, x_s^j) \\
 &\quad + f_{X^i, X^j}(x_s^i, x_s^j) \nabla^2 f_{Z, X^i, X^j}(z_r, x_r^i, x_r^j) \\
 &\quad - f_{X^i, X^j}(x_r^i, x_r^j) \nabla^2 f_{Z, X^i, X^j}(z_r, x_s^i, x_s^j) \\
 &\quad - f_{Z, X^i, X^j}(z_r, x_s^i, x_s^j) \nabla^2 f_{X^i, X^j}(x_r^i, x_r^j)] + o(\varepsilon^2),
 \end{aligned} \tag{4.14}$$

where κ_2 is the second moment of the kernel and $\nabla^2 f_W(w)$ is the trace of the second derivative of density evaluated at point w . Up to the error of order $o(\varepsilon^2)$, Eq. (4.14) has a plug-in estimator, which can be easily calculated using kernel methods (Wand and Jones, 1995).

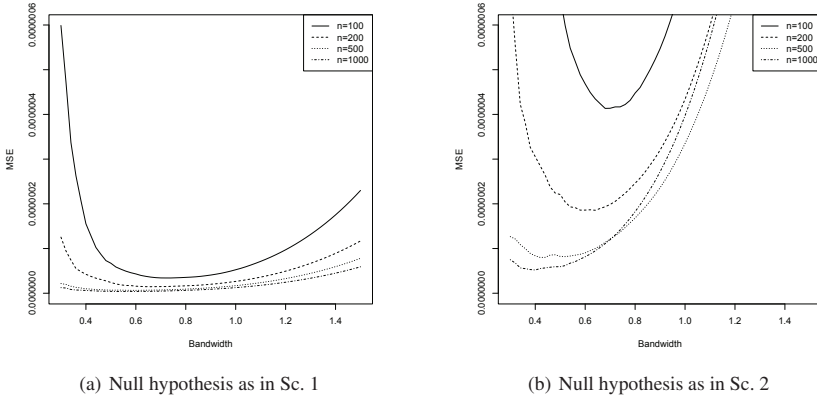
4.3 Numerical simulations

To give an example of the optimal bandwidth value, we perform a numerical experiment on the same bivariate process as considered by Jeong et al. (2012), i.e.

$$\begin{aligned}
 x_t^i &= 1 + \frac{1}{2} x_{t-1}^i + r_{1,t} \\
 x_t^j &= \frac{1}{2} x_{t-1}^j + c (x_{t-1}^i)^2 + r_{2,t},
 \end{aligned} \tag{4.15}$$

where $r_{1,t}$ and $r_{2,t}$ independent standard normal variables. The biggest advantage of the process in Eq. (4.15) is its tuning parameter on Granger causality, c . Clearly, if $c = 0$ the model corresponds to the hypothetical scenario of no Granger causality from X_t^i to X_t^j . The larger the parameter c becomes, the stronger the Granger causal effect, which we thus may control for

Figure 4.1: MSE of the test statistic for bandwidth values in the range $[0.3, 1.5]$ and for different sample sizes, aggregated over 1000 simulations.



explicitly.

We perform 1000 simulations of normalized data of process given by Eq. (4.15) for different sample sizes and evaluate the MSE of the test statistic for different bandwidth values within the range $[0.3, 1.5]$.² For practical reasons, we take lags of order 1 for both variables. The results for two scenarios of Granger causality are presented in Fig. 4.1 and the optimal bandwidths are reported in Table 4.1.

It is straightforward to notice the differences of the MSE curves between two settings. Firstly, for the same sample size and ε , Scenario 2 demonstrates larger MSE than in Scenario 1. Secondly, in Scenario 1 the MSE curve becomes flatter, whereas in Scenario 2 the visible U-shape is preserved as the sample size increases. These, in fact, are direct consequences of the curvature of the true distribution around particular quantiles. Scenario 1 is driven by the tail dependence, where the curvature is relatively flat. On the contrary, Scenario 2 represents the relation between the tail and the median, where the distribution is typically more bell-shaped or simply steeper. This, in fact, shows up in the steepness and in the relative size of the MSE curve. As expected, the minimum of the MSE curves is decreasing with the sample size in both

²We apply the standard score normalization.

scenarios (see Table 4.1).

Table 4.1: Optimal bandwidth values for test statistic evaluated for the process given by Eq. (4.15) for different sample sizes and for two scenarios. The values represent means over 1000 simulations.

	$n = 100$	$n = 200$	$n = 500$	$n = 1000$
ε^* (Sc.1)	0.74	0.66	0.6	0.52
ε^* (Sc.2)	0.68	0.64	0.48	0.44

The reported optimal bandwidth values represent the radius around the VaR_γ which is being considered in the NCoVaR. One may readily observe that Scenario 1 has slightly larger optimal bandwidths than Scenario 2. We may view this as a result of scarcity of data in tails compared with that around the median. Extracting information from tails requires, on average, slightly larger windows in comparison to the region near the median (Caers and Maes, 1998).

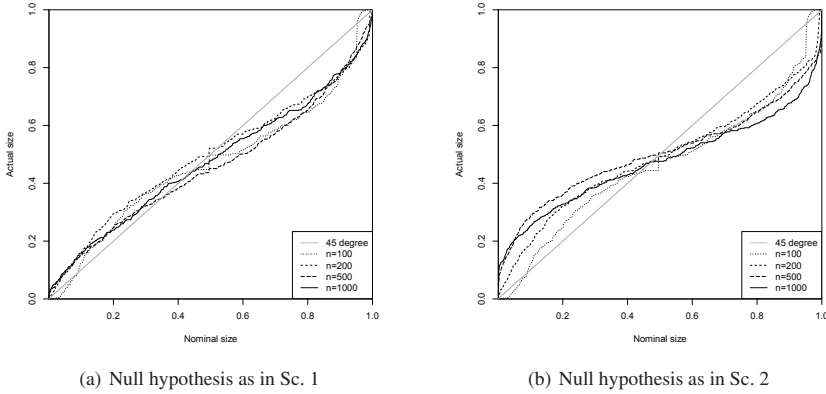
Because the MSE of the test statistics might be calculated explicitly, bootstrapping optimal bandwidths is a powerful technique which might be applied to any data set without assuming an underlying process structure. We recognize, however, that it might take a lot of computational time. For very large samples we suggest taking bins of 0.02 or 0.05 in order to make it computationally less demanding. Our simulations confirm that the power of the test is preserved in the range $[\varepsilon^* - 0.05, \varepsilon^* + 0.05]$.

4.3.1 Performance of the NCoVaR test

We perform two experiments to evaluate the practical side of the test. In both we rely on Monte Carlo methods on the example of the process in Eq. (4.15).³ In the first one, we assess the distribution of the test statistic under the null, evaluated for different sample sizes for 500 runs. In the second experiment, we estimate the power of the test. Given that the null hypothesis is violated ($c > 0$), we estimate rejection rates for different nominal significance levels. We

³One may expect that the numerical size distortions and power of the NCoVaR test would depend on the exact process specification. Eq. (4.15) offers a simple testing environment, which has been already applied in the quantile testing literature (Jeong et al., 2012). We therefore leave the assessment of the NCoVaR numerical performance on other processes for future investigation.

Figure 4.2: Size-size diagram of the NCoVaR test for the process from Eq. (4.15) for different sample sizes over 500 simulations.



summarize the results from both experiments in the size-size plots and size-adjusted power diagrams. The former plots the actual against nominal cumulative rejection rates under the null, showing the size distortions. The latter shows the power of the test corrected for the possible size bias, plotting the observed cumulative rejection rates under the alternative (actual power) against observed rejection rates under the null (actual size). Ideally, the power function should be 1 for any significance level larger than 0, however, in practice we would like to observe an increase in the slope at the origin as the sample size grows. Fig. 4.2 shows the size-size diagrams whereas the size-adjusted power plots are presented in Figs 4.3-4.5.

Fig. 4.2 suggests that the nominal size distortions are larger in Scenario 2 than in Scenario 1. Additionally, the size-size curves are flatter in Scenario 1 whereas they are more wavy in Scenario 2. In fact, this is similar to the pattern observed in the MSE (see Fig. 4.1) and might be largely attributed to the curvature of the true probability density function around particular quantiles.

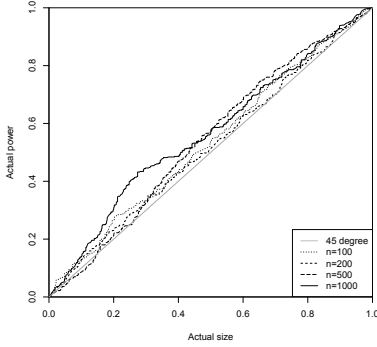
One can readily observe from Figs 4.3-4.5 that the size-adjusted power of the test increases with the sample size and with the strength of Granger causality. Nevertheless, there are two

main patterns emerging from the numerical analysis which deserve to be pointed out.

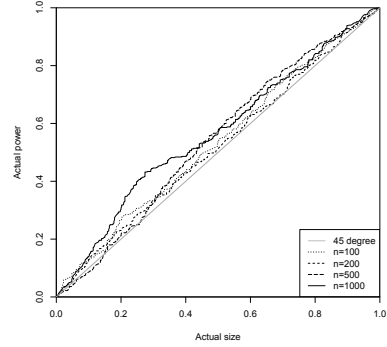
Firstly, for relatively smaller size the power of the test is higher for Scenario 1 than for Scenario 2. This is again the result of model dynamics, where the underlying relation on variable j , i.e. $(X_{t+1}^j \approx \text{VaR}_\gamma^j | X_t^j \approx \text{Median}^j)$ is more rare to observe on the process given by Eq. (4.15). Practically speaking, as the sample size gets larger this effect is hampered.

Secondly, the size-adjusted power is almost negligible for very small Granger causality and short time series. Clearly, one should blame the relative scarcity of observations around quantiles for this discomfort. In order to apply the test to shorter data sets, we propose two solutions to overcome this issue. The first comprises different kernel specifications. The square kernel takes into account only observations which are ε -close to the quantile, leaving out many possibly informative data points. Replacing the kernel by a smoother one, like Gaussian or logistic, should therefore correct for this effect. The second possible solution lies in improving the precision of the density estimators. In the standard kernel estimators (like square kernel estimators applied here) the bias is of order ε^2 (Wand and Jones, 1995). Making the bias smaller should decrease the disinformative effect of the observations around a given quantile so that keeping the sample size fixed we get relatively better representation of the true Granger causal relation, which translates into improved test performance. One may consider Data Sharpening (DS) as being potentially attractive bias reduction method in our setting. Following Hall and Minnotte (2002), the idea behind DS is to slightly perturb the original data set in order to obtain desirable estimator properties (here it is the reduced bias). Diks and Wolski (2013) show that, besides reducing the estimator bias, DS does not affect other asymptotic properties of the test statistic in a similar Granger causality setting. Therefore, it seems to be a straightforward extension to NCoVaR for shorter samples.

Figure 4.3: Size-adjusted power for the NCoVaR test for the process given by Eq. (4.15) for $c = 0.05$ for different sample sizes over 500 simulations.

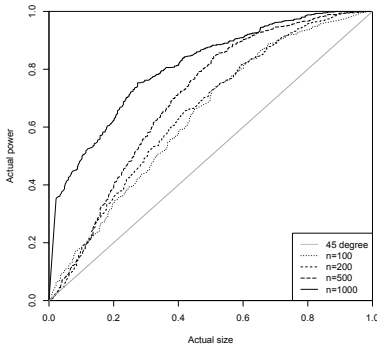


(a) Null hypothesis as in Sc. 1

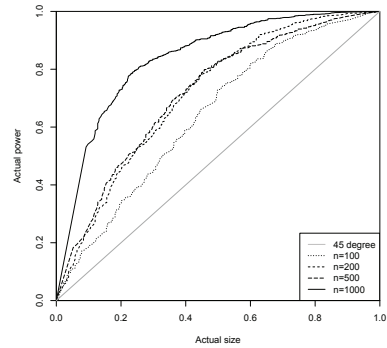


(b) Null hypothesis as in Sc. 2

Figure 4.4: Size-adjusted power for the NCoVaR test for the process given by Eq. (4.15) for $c = 0.25$ for different sample sizes over 500 simulations.

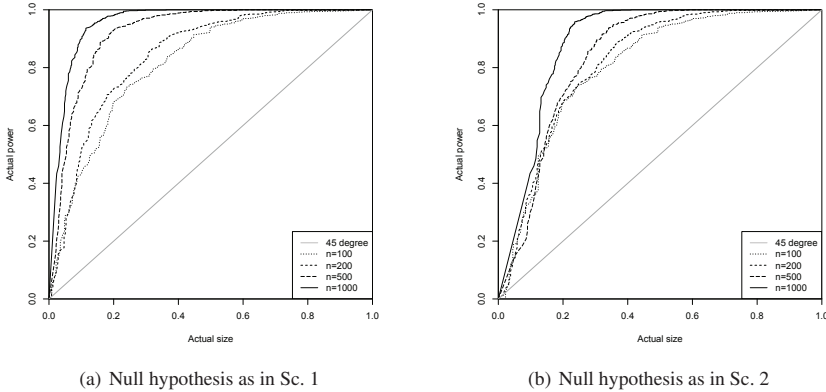


(a) Null hypothesis as in Sc. 1



(b) Null hypothesis as in Sc. 2

Figure 4.5: Size-adjusted power for the NCoVaR test for the process given by Eq. (4.15) for $c = 0.4$ for different sample sizes over 500 simulations.



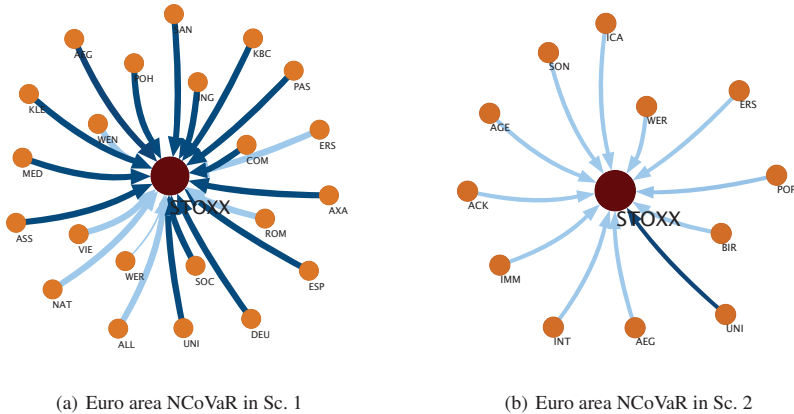
4.4 Assessing financial systemic risk

In our analysis we focus on the NCoVaR of individual institutions on the overall systemic risk. Therefore, we set j to represent the system variable and i individual financial institutions.

We approximate the returns on assets by equity returns and take into account financial institutions publicly traded within the euro zone. In order to make the analysis more transparent we focus on companies which constitute the Euro STOXX Financial Index in years 2000-2012. Our sample thus covers the Great Recession in Europe (2008-12), the financial crisis (2007-2009) and the sovereign debt crisis (2010-2012). In total we collect daily equity returns for 48 companies (3 financial, 13 insurance, 23 banks and 9 real estate) and one aggregate index. For each variable we have 3390 observations. The list of companies, together with the country of origin and their sector can be found in Appendix 4.C. The data have been obtained from the DataStream.

All time series are stationary at the 1% significance level, according to both the Phillips-Perron and Augmented Dickey-Fuller specifications (Phillips and Perron, 1988; Fuller, 1995). We run the pairwise tests against the null of no NCoVaR between each company and system

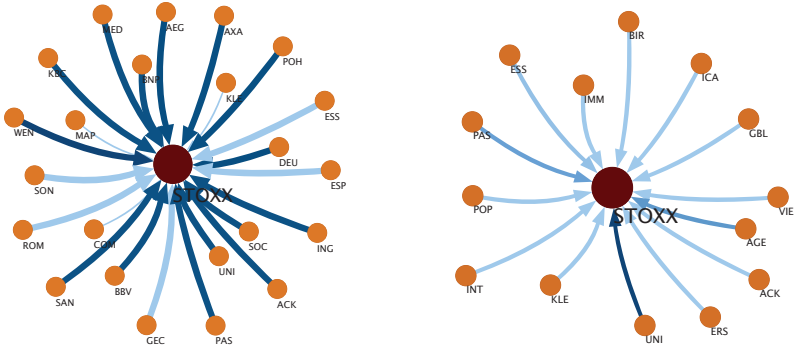
Figure 4.6: NCoVaR between euro area individual financial companies and system variable for raw data.



variable. In order to make sure that all the Granger causal relations are nonlinear, we run the same test specification on VAR-filtered residuals also. In each run the number of lags is taken according to the Schwarz-Bayes Information Criterion of the VAR specification, and the optimal bandwidth value is approximated by bootstrap. As a robustness check, we also correct for possible causality in second moments, as suggested in Francis et al. (2010), by running NCoVaR test on residuals from Dynamics Conditional Correlation GARCH model (Engle, 2002).

The detailed results can be found in Appendix 4.C (Tables 4.C.2, 4.C.3 and 4.C.3), however, for presentational clarity we refer to the star-graphs, which show the NCoVaR between each company and the system as a whole. The center of the star-graph represents the system variable and the satellite nodes correspond to individual institutions. The width of the arrows represents the inverse of the statistical significance level of NCoVaR (the stronger the NCoVaR effect, the wider (and darker) the arrow). Fig. 4.6 shows the results for the raw returns, Fig. 4.7 depicts the VAR-filtered returns and Fig. 4.8 refers to the GARCH residuals. Considering that at least one NCoVaR relation denotes a systemically important institution, our analysis suggests that out of 48 companies 33 might be so described. The group consists of 3 financial services companies,

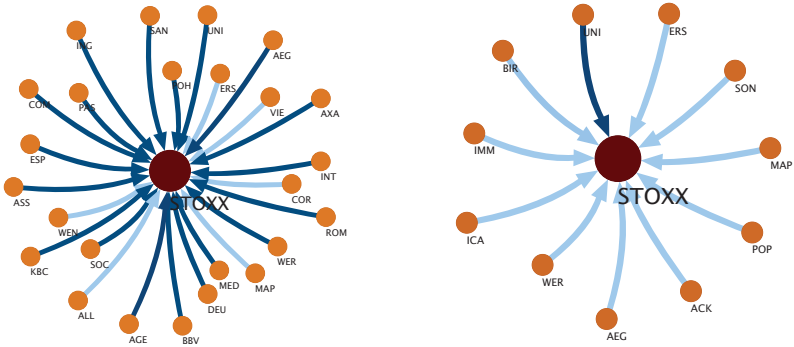
Figure 4.7: NCoVaR between euro area individual financial companies and system variable for VAR-filtered data.



(a) Euro area NCoVaR in Sc. 1

(b) Euro area NCoVaR in Sc. 2

Figure 4.8: NCoVaR between euro area individual financial companies and system variable for GARCH-filtered data.



(a) Euro area NCoVaR in Sc. 1

(b) Euro area NCoVaR in Sc. 2

6 insurance firms, 19 banks and 5 real estate companies. In fact, all of the financial services companies in our sample prove to be systemically important.

There are two main patterns emerging from our analysis. Firstly, there are fewer systemically risky institutions in Scenario 2. Secondly, NCoVaR in Scenario 1 is on average stronger than in Scenario 2. These findings hold for the original as well as the VAR- and GARCH-filtered data. Interestingly, our study suggests that only a few financial institutions pose a serious *ex ante* threat to the systemic risk in the euro area, whereas, given that the system is already in trouble, there are more institutions which hamper its recovery. This result confirms a common view in the literature on macroprudential supervision (Acharya, 2009) that the relative preventive costs are smaller than those after the crisis has already erupted.

The analysis confirms the nonlinear structure of the institutional contribution to the systemic risk. Filtering out the linear relations and second moment spillover effects does not remove the co-risk relations among individual companies and system as a whole. Interestingly, after filtering we observe some new co-risk relations emerging. To illustrate this better let us consider ACK (Ackermans & Van Haaren). The raw returns do not show any NCoVaR, however, after linear filtering it poses a very strong threat to the system's recovery (see Table 4.C.3 in Appendix 4.C which shows a test statistic of order 6.351 in Scenario 1) and after GARCH filtering it has a weak *ex ante* effect on the system's risk (test statistic of order 1.329 in Scenario 2). One may speculate that there are some strong purely nonlinear and second moment co-risk effects from ACK on the system variable, which are being partly offset by their linear equivalents. In other words, under normal circumstances ACK does not seem to be an important systemic risk contributor. However, in abnormal times, like a crisis, it reveals its systemic importance.

There is one more finding which we believe is worth pointing out. We confront our results with the official list of Global Systemically Important Banks (G-SIBs), published by the Financial Stability Board (FSB) in 2011.⁴ The FSB recognizes 11 G-SIBs in the euro area. Our sample covers 8 of them, i.e. Banco Bilbao Vizcaya Argentaria (BBV), Banco Santander

⁴The G-SIBs list is being frequently updated. In our comparison we consider the most recent version of the list, published on November 11th, 2013.

(SAN), BNP Paribas (BNP), Commerzbank (COM), Deutsche Bank (DEU), Societe Generale (SOC), UniCredit (Uni) and ING Bank (ING), as a part of the ING Groep. For all of them we confirm their G-SIB status in at least one NCoVaR setting.

4.5 Conclusions and discussion

Conditional Value-at-Risk-Nonlinear Granger Causality, or NCoVaR, is a new methodology of assessing co-risk relations, designed to capture the possible nonlinear Granger causal effects. Our approach distinguishes between two possible scenarios. In the first one, we test what is the role of individual institutions in hampering the recovery of others, given that they are already in distress. In the second scenario, we assess the contribution of individuals to the others' troubles. We derive the regular asymptotic properties of the NCoVaR test for both scenarios and we confirm them numerically.

We apply our methodology to assess the systemic importance of financial institutions in the euro area. Our findings suggest that (i) only a few financial institutions pose a serious *ex ante* threat to the systemic risk, whereas, given that the system is already in trouble, there are more institutions which hamper its recovery and (ii) there are intriguing nonlinear structures in its systemic risk profile.

Our study suggests that the most systemically risky institution in our sample is UNI (UniCredit), an Italian bank. In all settings it demonstrates a very strong NCoVaR relation to the system. In 2011 it was recognized by the FSB as G-SIB. This analysis confirms its systematic importance, also revealing its nonlinear nature. Interestingly, there are two more companies which demonstrate very strong NCoVaR in 5 out of 6 settings, i.e. ERS (Erste Group Bank), an Austrian bank, and AEG (Aegon), a Dutch insurer. Only the latter was recognized by the FSB to be potentially systemically important, with no official view on the former. However, the former was recognized as a systemically important bank for the Austrian financial sector (von Kruechten et al., 2009). Our results point to potential systemic importance of Erste Group Bank

in the entire euro area.

NCoVaR might be of great use for macroprudential policy, however, it has to be tested on other samples and in other periods. It reveals some intriguing phenomena in the co-risk relations. In order to understand these better, a tempting idea is to investigate the underlying nonlinear structures analytically in models of the aggregate economy. Such settings would allow to capture not only the contribution of individual institutions to systemic risk but also how individual companies are affected by aggregate disturbances. One may also apply NCoVaR as a mapping tool and bring the risk analysis to the network level.

Appendix 4.A Asymptotic properties of test statistic (Theorem 4.2.1)

We first deal with the properties for the independent sample and consider the dependency later. By symmetrization with respect to two indices, the test statistic in Eq. (4.11) has a corresponding U-statistic representation of the form

$$T_n(\varepsilon_n) \equiv T_n(\varepsilon) = \frac{1}{\binom{n}{2}} \sum_{k=1}^n \sum_{p \leq k} \tilde{K}(W_k, W_p), \quad (4.16)$$

with $W_k = (Z_k, X_{k,l_1}^i, X_{k,l_2}^j)$, $k = 1, \dots, n$ and kernel given by

$$\begin{aligned} \tilde{K}(W_k, W_p) = & \frac{\varepsilon^{-d_Z - 2d_{X^i} - 2d_{X^j}} (n-1)}{2n} \left[K_k(z_\gamma, x_\gamma^i, x_\gamma^j) K_p(x_m^i, x_m^j) \right. \\ & - K_k(x_\gamma^i, x_\gamma^j) K_p(z_\gamma, x_m^i, x_m^j) + K_p(z_\gamma, x_\gamma^i, x_\gamma^j) K_k(x_m^i, x_m^j) \\ & \left. - K_p(x_\gamma^i, x_\gamma^j) K_k(z_\gamma, x_m^i, x_m^j) \right], \end{aligned}$$

where for clarity we denote $K_k(w) = K((w - w_k)/\varepsilon)$ and d_Z , d_{X^i} and d_{X^j} are general representations of the dimensionality of \mathcal{G} and \mathcal{F} operators for particular variables. It is worth to remind here that subscript n in the test statistic refers to its sequence.

The asymptotic properties of the sequence of test statistic can be derived by the projection method (van der Vaart, 1998). From the Hájek's projection lemma we know that the projection of $T_n(\varepsilon) - \tau$ on the set of all function of the form $\sum_{k=1}^n \kappa_k(W_k)$ is given by

$$\hat{T}_n(\varepsilon) = \sum_{k=1}^n E[(T_n(\varepsilon) - \tau) | W_k] = \frac{2}{n} \sum_{k=1}^n \tilde{K}_1(w_k), \quad (4.17)$$

where

$$\tilde{K}_1(w_k) = E_{W_p} \left[\tilde{K}(w_k, W_p) \right] - \tau. \quad (4.18)$$

Projection $\hat{T}_n(\varepsilon)$ is mean zero sequence with variance $4/n \text{Var}(\tilde{K}_1(W_1))$. By the Central Limit Theorem, one may verify that $\sqrt{n} \hat{T}_n(\varepsilon)$ converges in distribution to the normal law with

mean 0 and variance given by $4\text{Var}(\tilde{K}_1(W_1))$.

Provided that $\text{Var}(\hat{T}_n(\varepsilon)) \rightarrow \text{Var}(T_n(\varepsilon))$ as $n \rightarrow \infty$, by Slutsky's lemma, we now observe that for a given ε and given quantiles of any independent finite-variance process $(Z_t, X_{t,l_i}^i, X_{t,l_j}^j)$, the sequence $\sqrt{n} \left(T_n(\varepsilon) - \tau - \hat{T}_n(\varepsilon) \right)$ converges in probability to zero as $n \rightarrow \infty$. What follows, the sequence $\sqrt{n} (T_n(\varepsilon) - \tau)$ converges in distribution to $\mathcal{N}(0, \sigma^2)$, where

$$\sigma^2 = 4\zeta_1, \quad (4.19)$$

with $\zeta_1 = \text{Cov} \left(\tilde{K}(W_1, W_2), \tilde{K}(W_1, W_2') \right) = \text{Var}(\tilde{K}_1(W_1))$.

Appendix 4.A.1 Dependence

Following the reasoning from Denker and Keller (1983), the above asymptotic normality properties of the test statistic, $T_n(\varepsilon)$, hold for a weakly dependent process if we take into account the covariance between estimators of particular vectors in the asymptotic variance σ^2 ,

$$\sigma^2 = 4 \left[\zeta_1 + 2 \sum_{t=2}^n \text{Cov} \left(\tilde{K}_1(W_1), \tilde{K}_1(W_t) \right) \right]. \quad (4.20)$$

According to the kernel specification, the estimator for $\tilde{K}_1(W_k)$ is given by

$$\hat{K}_1(W_k) = \frac{(2\varepsilon)^{-dz-2d_{X^i}-2d_{X^j}}}{n} \sum_{p=1}^n \tilde{K}(W_k, W_p).$$

The Newey and West (1987) heteroskedasticity and autocorrelation consistent estimator of σ^2 is

$$S_n^2 = \sum_{b=1}^B R_b \omega_b, \quad (4.21)$$

where B is equal to the floor of $n^{1/4}$, R_b is the sample covariance function of $\hat{K}_1(W_b)$ given by

$$R_b = \frac{1}{n-b} \sum_{a=1}^{n-b} (\hat{K}_1(W_a) - T_n(\varepsilon)) (\hat{K}_1(W_{a+b}) - T_n(\varepsilon)), \quad (4.22)$$

and ω_b is the weight function of the form

$$\omega_b = \begin{cases} 1, & \text{if } b = 1 \\ 2 - \frac{2(b-1)}{\tau}, & \text{if } b > 1. \end{cases} \quad (4.23)$$

For any finite-variance process $(Z_t, X_{t,l_i}^i, X_{t,l_j}^j)$, it follows from Denker and Keller (1983) that

$$\sqrt{n} \frac{(T_n(\varepsilon) - \tau)}{S_n} \xrightarrow{d} \mathcal{N}(0, 1), \quad (4.24)$$

which completes the proof of Theorem 4.2.1.

Appendix 4.B Optimal bandwidth sequence (Corollary 4.2.1)

For a given bandwidth ε , the MSE of the test statistic might be rewritten as as sum of variance and squared bias (Wand and Jones, 1995), i.e.

$$\text{MSE}[T_n(\varepsilon)] = \text{Var}(T_n(\varepsilon)) + \text{Bias}(T_n(\varepsilon))^2, \quad (4.25)$$

where $\text{Bias}(T_n(\varepsilon))$ can be calculated explicitly from the Taylor expansion as in Eq. (4.14) and variance of the test statistic might be represented as $4S_n^2/n$ from Appendix 4.A.1. Asymptotic covariance terms tend to zero as $n \rightarrow \infty$ so that under the null one might find that the asymptotic

variance of $T_n(\varepsilon)$ might be decomposed into

$$\begin{aligned}
 \text{Var}(T_n(\varepsilon)) &= \text{Var}(\hat{f}_{Z, X^i, X^j}(z_\gamma, x_\gamma^i, x_\gamma^j))\text{Var}(\hat{f}_{X^i, X^j}(x_m^i, x_*^j)) \\
 &+ \text{Var}(\hat{f}_{X^i, X^j}(x_\gamma^i, x_\gamma^j))\text{Var}(\hat{f}_{Z, X^i, X^j}(z_\gamma, x_m^i, x_*^j)) \\
 &+ \text{Var}(\hat{f}_{X^i, X^j}(x_m^i, x_*^j))E[\hat{f}_{Z, X^i, X^j}(z_\gamma, x_\gamma^i, x_\gamma^j)]^2 \\
 &+ \text{Var}(\hat{f}_{Z, X^i, X^j}(z_\gamma, x_\gamma^i, x_\gamma^j))E[\hat{f}_{X^i, X^j}(x_m^i, x_*^j)]^2 \\
 &+ \text{Var}(\hat{f}_{X^i, X^j}(x_\gamma^i, x_\gamma^j))E[\hat{f}_{Z, X^i, X^j}(z_\gamma, x_m^i, x_*^j)]^2 \\
 &+ \text{Var}(\hat{f}_{Z, X^i, X^j}(z_\gamma, x_m^i, x_*^j))E[\hat{f}_{X^i, X^j}(x_\gamma^i, x_\gamma^j)]^2 + o(1).
 \end{aligned}$$

One may find that the variance and bias of the individual density estimators are $o(n^{-1}\varepsilon^{-dw})$ and $o(\varepsilon^{-2})$, respectively (Silverman, 1998). Therefore, the dominant terms in the asymptotic variance are of order $o(n^{-1}\varepsilon^{-dz-d_{X^i}-d_{X^j}-4})$.

Taking the first order conditions of the MSE of individual density estimators, one finds that the optimum rate of convergence of bandwidth parameter is $n^{-1/(dw+4)}$. Doing the same for our test statistic, we find that this rate is $n^{-1/(dz+d_{X^i}+d_{X^j})}$. Therefore, for any finite dimension, the MSE-optimal rate of convergence of the test statistic's bandwidth is slightly faster than those of individual density estimators but never as fast as n^{-1} which would violate condition imposed by Eq. (4.10). Provided that the optimum rate of convergence of the individual estimators is sufficient for the consistency (Silverman, 1998), the optimum rate of $T_n(\varepsilon_n)$ guarantees consistency as well.

Appendix 4.C Data description and results

The Euro STOXX Financials Index consists originally of 61 entities. However, only 48 of them cover years 2000-2012 (see Table 4.C.1). For all of them we collect daily equity prices and calculate their log returns accordingly. Data comes from the DataStream and covers period 01/01/2000 till 12/31/2012.

Table 4.C.1: List of all entities used in the empirical analysis.

	Company name/Index	Symbol	Sector	Country
1	Euro STOXX Financials	STOXX	Aggregate	Aggregate
2	Ackermans & Van Haaren	ACK	Financial Services	BE
3	Aegon	AEG	Insurance	NL
4	Ageas	AGE	Insurance	NL
5	Allianz	ALL	Insurance	DE
6	Assicurazioni Generali	ASS	Insurance	IT
7	AXA	AXA	Insurance	FR
8	Bank Of Ireland	BIR	Banks	IR
9	Bankinter	BAN	Banks	ES
10	Banca Monte Dei Paschi	PAS	Banks	IT
11	Banca Popolare Di Milano	MIL	Banks	IT
12	Banca Popolare Di Sondrio	SON	Banks	IT
13	Banca Popolare Emilia Romagna	ROM	Banks	IT
14	BBV Argentaria	BBV	Banks	ES
15	Banco Comr. Portugues	POR	Banks	PT
16	Banco Espirito Santo	ESS	Banks	PT
17	Banco Popolare	POP	Banks	IT
18	Banco Popular Espanol	ESP	Banks	ES
19	Banco Santander	SAN	Banks	ES
20	BNP Paribas	BNP	Banks	FR
21	CNP Assurances	CNP	Insurance	FR
22	Cofinimmo	COF	Real Estate	BE
23	Commerzbank	COM	Banks	DE
24	Corio	COR	Real Estate	NL
25	Deutsche Bank	DEU	Banks	DE
26	Erste Group Bank	ERS	Banks	AT
27	Fonciere Des Regions	FON	Real Estate	FR
28	Gecina	GEC	Real Estate	FR
29	GBL New	GBL	Financial Services	BE
30	Societe Generale	SOC	Banks	FR
31	Hannover Ruck.	HAN	Insurance	DE
32	ICADE	ICA	Real Estate	FR
33	Immofinanz	IMM	Real Estate	AT
34	ING Groep	ING	Insurance	NL
35	Intesa Sanpaolo	INT	Banks	IT
36	KBC Group	KBC	Banks	BE
37	Klepierre	KLE	Real Estate	FR
38	Mapfre	MAP	Insurance	ES
39	Mediobanca	MED	Banks	IT
40	Muenchener Ruck.	MUE	Insurance	DE
41	Natixis	NAT	Banks	FR
42	Pohjola Pankki	POH	Banks	FI
43	Sampo	SAM	Insurance	FI
44	SCOR	SCO	Insurance	FR
45	Unibail-Rodamco	ROD	Real Estate	FR
46	UniCredit	UNI	Banks	IT
47	Vienna Insurance Group	VIE	Insurance	AT
48	Wendel	WEN	Financial Services	FR
49	Wereldhave	WER	Real Estate	NL

APPENDIX 4.C. DATA DESCRIPTION AND RESULTS

Table 4.C.2: NCoVaR from institution i on the system risk in two scenarios in period 01/01/2000 till 12/31/2012 for raw returns. Lags determines the optimal number of lags from the VAR specification using the Schwarz-Bayes Information Criterion. Optimal epsilon values calculated from bootstrap. T-val represents the test statistic of NCoVaR from Eq.(4.11). (*),(**), (***) denotes one-sided p-value statistical significance at 10%, 5% and 1%, respectively.

	Institution i	System variable	Lags	Scenario 1		Scenario 2		
				Opt. ϵ	T-val	Lags	Opt. ϵ	T-val
1	ACK	STOXX	1	0.54	0.913	1	0.44	1.329*
2	AEG	STOXX	1	0.28	6.359***	1	0.4	2.855***
3	AGE	STOXX	4	0.6	0.213	4	0.6	3.516***
4	ALL	STOXX	1	0.38	1.45	1	0.2	-6.365
5	ASS	STOXX	1	0.24	6.351***	1	0.22	-6.359
6	AXA	STOXX	1	0.26	6.351***	1	0.2	-3.035
7	BIR	STOXX	2	0.6	0.585	2	0.6	2.611***
8	BAN	STOXX	1	0.38	-1.262	1	0.22	-3.446
9	PAS	STOXX	1	0.32	6.351***	1	0.2	-4.382
10	MIL	STOXX	1	0.24	-6.369	1	0.2	-4.477
11	SON	STOXX	1	0.3	-3.854	1	0.3	4.448***
12	ROM	STOXX	1	0.48	1.571*	1	0.26	-5.833
13	BBV	STOXX	1	0.38	-1.081	1	0.2	-6.361
14	POR	STOXX	1	0.26	-6.351	1	0.3	-5.378
15	ESS	STOXX	1	0.26	-6.357	1	0.2	-5.44
16	POP	STOXX	1	0.32	-1.551	1	0.26	5.359***
17	ESP	STOXX	1	0.26	6.351***	1	0.2	-6.364
18	SAN	STOXX	1	0.2	6.351***	1	0.42	0.673
19	BNP	STOXX	1	0.34	-1.947	1	0.2	-3.121
20	CNP	STOXX	1	0.3	0.78	1	0.2	-6.367
21	COF	STOXX	1	0.2	-6.351	1	0.2	-6.365
22	COM	STOXX	1	0.26	6.351***	1	0.2	-6.36
23	COR	STOXX	1	0.58	0.411	1	0.22	-6.375
24	DEU	STOXX	1	0.24	6.352***	1	0.24	-6.371
25	ERS	STOXX	1	0.38	1.576*	1	0.26	3.295***
26	FON	STOXX	1	0.26	-6.357	1	0.32	0.052
27	GEC	STOXX	1	0.24	-6.362	1	0.32	-0.389
28	GBL	STOXX	1	0.3	-6.359	1	0.42	0.088
29	SOC	STOXX	1	0.24	6.351***	1	0.22	-6.364
30	HAN	STOXX	1	0.32	-0.645	1	0.2	-3.911
31	ICA	STOXX	1	0.24	-1.474	1	0.34	3.846***
32	IMM	STOXX	2	0.6	-0.316	2	0.6	2.684***
33	ING	STOXX	1	0.26	6.351***	1	0.22	-3.734
34	INT	STOXX	1	0.28	-2.919	1	0.28	4.994***
35	KBC	STOXX	1	0.26	6.352***	1	0.2	-6.359
36	KLE	STOXX	1	0.24	6.351***	1	0.52	-0.315
37	MAP	STOXX	1	0.26	-0.444	1	0.22	-6.371
38	MED	STOXX	1	0.26	6.351***	1	0.2	-6.361
39	MUE	STOXX	1	0.28	-6.351	1	0.2	-6.357
40	NAT	STOXX	1	0.26	1.417*	1	0.28	-4.937
41	POH	STOXX	1	0.28	6.352***	1	0.24	-3.925
42	SAM	STOXX	1	0.6	-0.026	1	0.28	-4.666
43	SCO	STOXX	1	0.32	-2.487	1	0.2	-6.365
44	ROD	STOXX	1	0.52	0.239	1	0.24	-6.377
45	UNI	STOXX	1	0.28	6.352***	1	0.3	5.886***
46	VIE	STOXX	1	0.32	3.747***	1	0.2	-6.362
47	WEN	STOXX	1	0.24	4.733***	1	0.22	-6.364
48	WER	STOXX	1	0.3	3.34***	1	0.44	1.572*

CHAPTER 4. EXPLORING NONLINEARITIES IN FINANCIAL SYSTEMIC RISK

Table 4.C.3: NCoVaR from institution i on the system risk in two scenarios in period 01/01/2000 till 12/31/2012 for VAR-filtered returns. Lags determines the optimal number of lags from the VAR specification using the Schwarz-Bayes Information Criterion. Optimal epsilon values calculated from bootstrap. T-val represents the test statistic of NCoVaR from Eq.(4.11). (*),(**), (***) denotes one-sided p-value statistical significance at 10%, 5% and 1%, respectively.

	Institution i	System variable	Lags	Scenario 1		Scenario 2		
				Opt. ϵ	T-val	Lags	Opt. ϵ	T-val
1	ACK	STOXX	1	0.26	6.351***	1	0.44	1.417*
2	AEG	STOXX	1	0.24	6.355***	1	0.2	-6.36
3	AGE	STOXX	1	0.38	0.983	1	0.24	5.506***
4	ALL	STOXX	1	0.36	0.247	1	0.2	-6.363
5	ASS	STOXX	1	0.3	-2.589	1	0.22	-6.361
6	AXA	STOXX	1	0.32	6.353***	1	0.2	-6.358
7	BIR	STOXX	1	0.22	-6.362	1	0.32	4.329***
8	BAN	STOXX	1	0.28	-2.625	1	0.2	-6.36
9	PAS	STOXX	1	0.34	6.35***	1	0.3	5.469***
10	MIL	STOXX	1	0.34	-5.756	1	0.2	-3.407
11	SON	STOXX	1	0.36	4.458***	1	0.2	-5.44
12	ROM	STOXX	1	0.34	2.491***	1	0.22	-5.627
13	BBV	STOXX	1	0.2	6.35***	1	0.2	-6.36
14	POR	STOXX	1	0.32	-6.36	1	0.24	-6.336
15	ESS	STOXX	1	0.36	1.998***	1	0.24	5.202***
16	POP	STOXX	1	0.34	-3.085	1	0.3	4.667***
17	ESP	STOXX	1	0.36	2.102**	1	0.26	-6.375
18	SAN	STOXX	1	0.2	6.351***	1	0.44	-0.077
19	BNP	STOXX	1	0.22	6.35***	1	0.2	-6.356
20	CNP	STOXX	1	0.34	-0.525	1	0.2	-6.374
21	COF	STOXX	1	0.34	-2.344	1	0.2	-6.363
22	COM	STOXX	1	0.4	3.693***	1	0.2	-6.354
23	COR	STOXX	1	0.6	0.627	1	0.24	-6.38
24	DEU	STOXX	1	0.32	6.355***	1	0.22	-6.366
25	ERS	STOXX	1	0.36	0.493	1	0.34	1.903**
26	FON	STOXX	1	0.3	-6.355	1	0.3	0.633
27	GEC	STOXX	1	0.34	4.518***	1	0.46	-0.53
28	GBL	STOXX	1	0.26	-2.699	1	0.38	3.204***
29	SOC	STOXX	1	0.2	6.351***	1	0.2	-6.36
30	HAN	STOXX	1	0.32	-6.35	1	0.2	-2.984
31	ICA	STOXX	1	0.32	0.502	1	0.32	4.231***
32	IMM	STOXX	1	0.58	0.962	1	0.32	4.231***
33	ING	STOXX	1	0.26	6.35***	1	0.24	-6.363
34	INT	STOXX	1	0.34	-4.345	1	0.28	3.23***
35	KBC	STOXX	1	0.24	6.351***	1	0.2	-6.358
36	KLE	STOXX	1	0.36	2.593***	1	0.36	3.947***
37	MAP	STOXX	1	0.48	1.519*	1	0.2	-6.364
38	MED	STOXX	1	0.2	6.35***	1	0.2	-6.364
39	MUE	STOXX	1	0.24	-6.35	1	0.2	-6.354
40	NAT	STOXX	1	0.28	1.093	1	0.3	-4.97
41	POH	STOXX	1	0.32	6.353***	1	0.44	0.314
42	SAM	STOXX	1	0.52	-0.857	1	0.2	-6.362
43	SCO	STOXX	1	0.32	0.759	1	0.2	-6.364
44	ROD	STOXX	1	0.5	-0.401	1	0.24	-6.38
45	UNI	STOXX	1	0.3	6.351***	1	0.28	5.673***
46	VIE	STOXX	1	0.34	0.611	1	0.3	2.273***
47	WEN	STOXX	1	0.26	6.368***	1	0.2	-6.359
48	WER	STOXX	1	0.32	-2.424	1	0.22	-6.37

APPENDIX 4.C. DATA DESCRIPTION AND RESULTS

Table 4.C.3: NCoVaR from institution i on the system risk in two scenarios in period 01/01/2000 till 12/31/2012 for GARCH-filtered returns. Lags determines the number of lags used in the test. Optimal epsilon values calculated from bootstrap. T-val represents the test statistic of NCoVaR from Eq.(4.11). (*),(**), (***) denotes one-sided p-value statistical significance at 10%, 5% and 1%, respectively.

	Institution i	System variable	Lags	Scenario 1		Scenario 2		
				Opt. ϵ	T-val	Lags	Opt. ϵ	T-val
1	ACK	STOXX	1	0.54	0.913	1	0.44	1.329*
2	AEG	STOXX	1	0.28	6.359***	1	0.4	2.855***
3	AGE	STOXX	1	0.25	6.362***	1	0.23	-4.646
4	ALL	STOXX	1	0.38	1.45*	1	0.21	-6.366
5	ASS	STOXX	1	0.23	6.351***	1	0.21	-6.359
6	AXA	STOXX	1	0.26	6.351***	1	0.2	-3.035
7	BIR	STOXX	1	0.25	-4.733	1	0.21	5.319***
8	BAN	STOXX	1	0.38	-1.262	1	0.23	-3.61
9	PAS	STOXX	1	0.32	6.351***	1	0.2	-4.382
10	MIL	STOXX	1	0.24	-6.369	1	0.21	-4.477
11	SON	STOXX	1	0.3	-3.854	1	0.31	4.083***
12	ROM	STOXX	1	0.27	6.351***	1	0.27	-5.883
13	BBV	STOXX	1	0.25	6.352***	1	0.21	-6.362
14	POR	STOXX	1	0.26	-6.351	1	0.3	-5.378
15	ESS	STOXX	1	0.26	-6.357	1	0.21	-5.771
16	POP	STOXX	1	0.29	-4.344	1	0.27	4.958***
17	ESP	STOXX	1	0.27	6.351***	1	0.21	-6.364
18	SAN	STOXX	1	0.21	6.351***	1	0.43	0.418
19	BNP	STOXX	1	0.34	-1.947	1	0.2	-3.121
20	CNP	STOXX	1	0.31	0.531	1	0.21	-6.37
21	COF	STOXX	1	0.2	-6.351	1	0.21	-6.365
22	COM	STOXX	1	0.26	6.351***	1	0.2	-6.36
23	COR	STOXX	1	0.25	2.84***	1	0.21	-6.371
24	DEU	STOXX	1	0.23	6.352***	1	0.23	-6.37
25	ERS	STOXX	1	0.38	1.576*	1	0.27	3.141***
26	FON	STOXX	1	0.26	-6.357	1	0.32	0.052
27	GEC	STOXX	1	0.24	-6.362	1	0.33	-0.46
28	GBL	STOXX	1	0.3	-6.359	1	0.42	0.088
29	SOC	STOXX	1	0.24	6.351***	1	0.22	-6.364
30	HAN	STOXX	1	0.27	-3.068	1	0.2	-3.911
31	ICA	STOXX	1	0.25	-3.335	1	0.34	3.846***
32	IMM	STOXX	1	0.22	-6.361	1	0.35	4.183***
33	ING	STOXX	1	0.26	6.351***	1	0.22	-3.734
34	INT	STOXX	1	0.26	6.352***	1	0.2	-6.359
35	KBC	STOXX	1	0.24	6.351***	1	0.52	-0.315
36	KLE	STOXX	1	0.26	-0.444	1	0.22	-6.371
37	MAP	STOXX	1	0.25	1.432*	1	0.28	4.994***
38	MED	STOXX	1	0.27	6.351***	1	0.2	-6.361
39	MUE	STOXX	1	0.29	-6.351	1	0.2	-6.357
40	NAT	STOXX	1	0.27	0.34	1	0.28	-4.937
41	POH	STOXX	1	0.29	6.352***	1	0.24	-3.925
42	SAM	STOXX	1	0.59	-0.147	1	0.2	-6.362
43	SCO	STOXX	1	0.32	-2.487	1	0.2	-6.365
44	ROD	STOXX	1	0.52	0.239	1	0.23	-6.375
45	UNI	STOXX	1	0.28	6.352***	1	0.3	5.886***
46	VIE	STOXX	1	0.32	3.747***	1	0.2	-6.362
47	WEN	STOXX	1	0.25	4.869***	1	0.22	-6.364
48	WER	STOXX	1	0.25	6.352***	1	0.45	1.411*

Chapter 5

Do Safe Havens Make Asset Markets Safer?

5.1 Introduction

Over the past half-decade, since the onset of the global financial crisis, asset prices and capital flows have gyrated and their movements have differed strongly across different groups of countries. In particular, some countries (let us call them “safe havens”) experienced strong inflows into sovereign bond markets while others found their most liquid markets drying up. Has the presence of these safe haven flows changed the resilience of the global financial network that was buffeted by repeated shocks since 2007? In this chapter we present some stylized facts on the role of safe havens in spreading or containing contagion.

A rapidly expanding literature has documented contagion across asset prices and, in particular, between sovereign and bank debt. Several authors have provided evidence of cross-country contagion in long-term sovereign bond yields (Basurto et al., 2010; Gilmore et al., 2010) or sovereign Credit Default Swap (CDS) spreads (Caporin et al., 2012) for euro area countries or a broader sample of European countries, the US, and Japan. While there is some concern that strong sovereign-sovereign correlations simply reflect correlations in fundamental finan-

cial factors, especially short-term interest rates (Manganelli and Wolswijk, 2007), Mody (2009) has shown that 2007 was a turning point in sovereign-sovereign correlations with increasing differentiation according to credit risk. In addition to sovereign-sovereign correlations, several authors have also documented sovereign-bank contagion. After bank bailout episodes and financial rescue packages in the euro area, the correlation between bank and sovereign CDS spreads increased significantly (Acharya et al., 2011), and bank and sovereign CDS spreads' sensitivity to a global risk factor became more similar (Ejlsing and Lemke, 2011). Also outside these financial rescue episodes, Merton et al. (2013) show rising correlations between sovereign and bank CDS spreads. By estimating correlations, this literature has essentially mapped the shape of the network of asset prices, whether around periods of stress or over longer time spans. To our knowledge, the literature has not yet analyzed how this network's shape affects contagion once a shock enters this network.

Several authors have shown network measures to be significant correlates of banking system and general financial system stress. Minoiu et al. (2013) found rising interconnectedness (measured as clustering coefficients and degree centrality) in the global network of cross-border banking exposures from the BIS locational statistics to be significant predictors of systemic banking crises. So were degree and betweenness centrality in a bank-level network of syndicated loans (Caballero, 2012). At the same time, increased connectivity in the same network fostered trade (Hale et al., 2013). While the previous papers related mainly to the pre-crisis period, Chinazzi et al. (2013) found that degree centrality in a network of cross-country debt and equity exposures was a significant predictor of the drop growth and stock market volatility during the crisis. The measures these authors used were country-level measures of a country's position in the network. While these are useful to predict crises or trade in any particular country, they do not explain the dynamics of contagion from a crisis. In contrast, here we do not attempt to predict a crisis or any other shock but, contingent on a shock occurring somewhere, we trace how contagion travels through global asset prices.

Blending elements of the literatures on asset price contagion and exposure networks, we

examine how the shape of the global network of asset price co-movements has been conducive (or not) to the spread of contagion. We hone in on a particular group of countries with unique characteristics, the safe haven countries, and their role in amplifying or slowing the spread of contagion across borders and asset classes. In particular, we find important differences in sovereign-bank feedback loops between safe haven and non-safe haven countries. This distinction comes out more clearly in our sample than in those of previous authors because we deliberately expand it to include many emerging markets (50 sovereigns) and individual banks (331 banks). To achieve this larger sample, we rely on sovereign bond yields and bank equity prices, which in many countries are more liquid than CDS spreads. By using individual bank data, we are able to distinguish sovereign-bank correlations between more and less systemic banks which are too big to fail to different degrees.

The existing literature on safe havens has defined safe haven assets as hedges of returns on reference portfolios during times of financial stress or rising risk aversion. This literature has examined exchange rates (Beck and Rahbari, 2008; Habib and Stracca, 2012; Ranaldo and Söderlind, 2010), gold (Baur and McDermott, 2010), or sovereign bonds (Hartmann et al., 2004) as hedges against stock market risk. To our knowledge, the literature has not defined safe haven status based on the potential for sovereign bonds to serve as hedges against individual banking risk. Since ours is a network of sovereign bond yields and individual bank equity returns, we prefer a definition of safe havens relevant to our data instead of one that is exogenous to our data set. However, our definition, as we show below, does not deviate more from common usage than other definitions in the literature or definitions used by financial market participants.

Our definition of a safe haven country explicitly treats sovereign bonds as possible safe haven assets when banks (not the stock market more generally) are under stress: safe havens are those countries where bank equity prices and sovereign bond yields move strongly in tandem. If bank equity prices and sovereign bond yields were purely driven by country-level credit risk, one would have expected the opposite: if credit risk rises, sovereign bond yields increase and bank equity prices fall. In contrast, where credit risk is of negligible concern, i.e. in safe

havens, expectations about future growth and monetary policy become predominant: an improving growth outlook raises bank equity prices and the expectation of tightening monetary policy which, in turn, puts pressure on sovereign bond yields.

Hence, by definition, safe havens are countries without sovereign-bank feedback loops that amplify shocks to both banks and sovereigns. For example, contagion from a global shock that simultaneously raises bond yields and reduces bank equity prices in a safe haven could trigger the expectation of a monetary policy response in the safe haven that would raise bank equity prices and reduce sovereign bond yields. In contrast, outside safe havens, a similar shock could trigger concerns about credit risk and set in motion self-fulfilling bank-sovereign feedback loops. It turns out, however, that in our network, this benign property of safe havens is offset by a less benign one. In particular, safe havens tend to have stronger sovereign-sovereign and stronger bank-bank correlations than non-safe havens. As a result, if a shock arrives in safe havens, they can propagate shocks to other countries faster than non-safe havens. Which of the two effects dominates depends on the nature of the shock and the nature of the broader network. In our sample, we find that, on balance, safe havens amplify shocks (although to varying degrees depending on the shock).

In the next section, we describe our data, followed by our definition of safe havens and their properties in Section 5.3. Section 5.4 describes the global network structure of sovereign bond yields and bank equity returns. In Sections 5.5 and 5.6, we document some stylized facts of feedback loops in shock propagation. In Section 5.7, we examine the role of the two characteristics of safe havens in amplifying or dampening shock propagation. Several of these facts raise intriguing questions, summarized in Section 5.8, that are left for further research.

5.2 Data

We use daily changes in 5-year bond yields of 50 sovereigns and daily log changes in bank equity prices of 331 individual banks using Bloomberg data.¹ Because of limited data availability

¹The results are broadly robust to including the smaller sample using 10-year bond yields.

in the 1990s, the time span for our network of global bank equity prices and sovereign bond yields comprises 2000-2013. The full sample is divided into four subsamples: years 2000-2006 (Great Moderation), 2007-2009 (Subprime Crisis), 2010-2012 (Sovereign Debt Crisis), 2013 (Emerging Markets (EM) Stress).² We adjust the daily data for time zones and exchange rate changes.

For each bank-bank, sovereign-sovereign, and bank-sovereign pair, we calculate bilateral Pearson correlation coefficients between bank equity price log changes and sovereign bond yield changes over each of our four subperiods. Ideally, we would have used measures that explicitly incorporate causality, e.g. Granger (1969) causality or spillover coefficients as in Diebold and Yilmaz (2011), but the estimations necessary to derive these measures would typically have constrained our sample size. Therefore, here we begin by focusing on simple correlations.³ To eliminate spurious correlations, we set the correlations between sovereigns and banks outside their countries to zero.⁴

We call our network $G(V, E)$ a representation of a set of nodes $V = \{v_1, v_2, \dots, v_n\}$, connected by a set of edges $E \subset V \times V$. For now, the strength of the edge between two adjacent nodes is determined by our Pearson correlation coefficient. Formally, we may represent a network G in a matrix form, denote it by $A_{n \times n}$, where all diagonal elements are equal to zero, i.e. the relation between the same assets is irrelevant, and elements a_{ij} represent the correlation between assets i and j . Since we use time adjusted data, matrix A is not symmetric, making the network directed, i.e. $a_{ij} \neq a_{ji}$ for some i and j . Formally, if we denote the number of sovereigns by n^s and number of banking sectors by n^b , one may rewrite the complete network as a block matrix $B_{(n^s+n^b) \times (n^s+n^b)}$, where two diagonal blocks represent the individual networks

²The majority of the euro zone countries provide sovereign bond yields ranging back to 1994. Therefore, we consider this as a special case and we devote Box 1 to analyze the situation in the Economic and Monetary Union (EMU) individually over the years 1994-2012.

³In principle, shocks can of course also jump from equity and bank stock prices to interbank money markets or foreign exchange markets. We will consider these asset classes in future research.

⁴While this does mean that, e.g. the correlation between the Greek sovereign bond yield and a French bank's equity price is eliminated by assumption, it also avoids many spurious correlations, e.g. between the Finnish sovereign and Argentinian banks. Here, to avoid the many spurious correlations at the cost of eliminating some valid ones, we remove all sovereign-bank correlations except those within each country.

and the remaining blocks are zeros except for the case when the sovereign and banks refer to the same country.

Fig. 5.1 shows the distribution of correlations over our four subperiods. The bulk of them are inside the 95% confidence interval $[-0.2, +0.2]$ for Pearson correlation coefficients and, hence, statistically insignificant. This is especially the case for bank-sovereign correlations where less than one-fifth of the correlations are statistically significantly negative or positive. Sovereign-sovereign correlations are stronger than bank-bank and, even more so, bank-sovereign correlations (the distribution of sovereign correlations is further to the right and has fatter tails than that of bank-bank or bank-sovereign correlations). In addition, even if not visible in Fig. 5.1, within-country bank-bank correlations are stronger than cross-country bank-bank correlations.

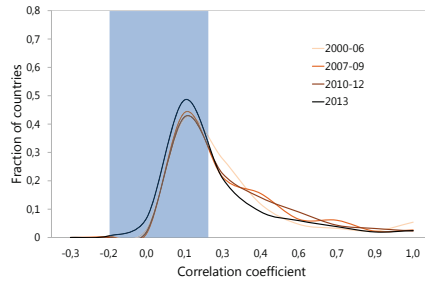
Negative correlations between sovereigns and other sovereigns or banks and other banks are rare. Negative correlations among sovereigns are confined to a few country pairs.⁵

In aggregate, correlations strengthened in 2007-09 but by 2013 had fallen back to pre-crisis levels (the right tail of all three distributions moved sharply out in 2007-09 but has since moved back inwards). The number of strong bank-sovereign correlations has shrunk even below pre-crisis levels.

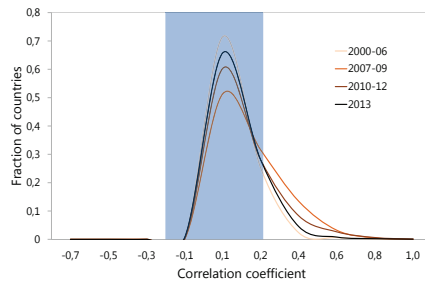
The global financial crisis and the subsequent euro area crisis triggered increasing clustering of sovereign and bank asset prices. Fig. 5.2 shows the distribution of clustering coefficients (loosely speaking, the share of “friends” that are also “friends” with each other) for bank equity prices (Fagiolo, 2007). The distribution shifted sharply to the right as banks equity prices became strongly correlated globally in 2007-09 and regionally in 2010-12. Since then, there has been some reversion towards 2000-06 norms. Similarly, clustering of sovereign bond yields increased sharply in 2007-09 and even more so in 2010-12. The rightward shift of the distribu-

⁵In 2000-06, they include the US against 14 EU countries and Switzerland. In 2007-09, they include Japan against several advanced and emerging market commodity producers (Australia, New Zealand, Canada, Mexico, Brazil, South Africa, Turkey), the US against Switzerland, and Colombia against several Asian and European emerging markets (China, Singapore, Malaysia, Indonesia, Slovakia, Ukraine). In 2010-12, there is only one negative correlation, between Japan and the US. In 2013, negative correlations are between the US and large emerging markets (Brazil, Turkey, Hungary, Mexico, Thailand, Malaysia, Philippines, South Africa) and/or commodity producers (Australia, New Zealand, Norway) and global financial centers (Japan, Switzerland, UK).

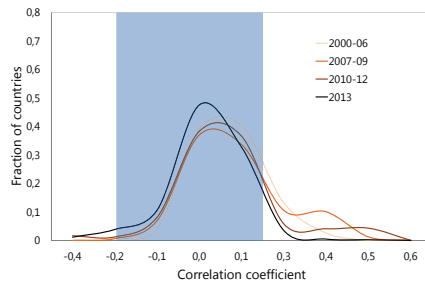
Figure 5.1: Distribution of correlations at the sovereign, bank and sovereign-bank levels.



(a) 5-year sovereign bond yields



(b) Daily bank equity returns



(c) Daily bank equity returns and 5-year sovereign bond yields

tion 2007-09 reflects a stronger clustering especially in Asia whereas the rightward shift of the distribution in 2010-12 reflects especially a stronger clustering in Europe.

5.3 Defining safe havens

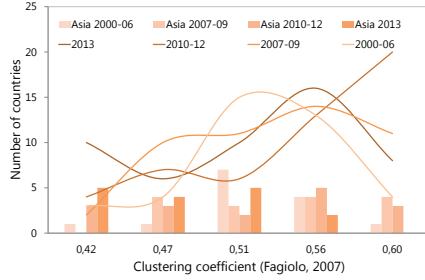
We distinguish between “safe havens” and “non-safe havens” on the basis of their correlation between sovereign bond yields and bank equity prices. For advanced countries, including the US, the positive correlation between sovereign bond yields and prices of riskier assets has been documented by Bauer and Rudebusch (2013) and Pandl (2013). In contrast, for emerging markets, Drainville et al. (2011) show a negative correlation between bond yields and bank equity prices and speculate that this reflects strongly correlated risk premia of EM assets.

Here, we also base our definition on the correlation between sovereign bond yields as potentially the safe assets, and individual bank equity returns as the riskier assets. Specifically, we define countries as “safe havens” if daily changes of sovereign bond yields and bank equity prices are significantly positively correlated (correlation > 0.2). The rationale is as follows.

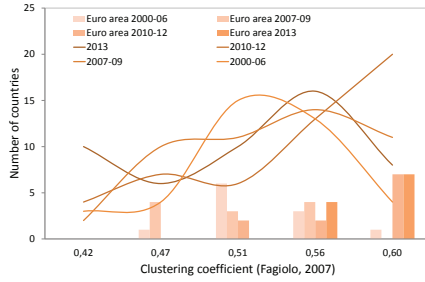
Long term sovereign bond yields can be broadly decomposed into two components: (i) expectations of average future short-term interest rates and (ii) a premium that investors require for bearing the (e.g., credit, liquidity) risk of a long-term bond investment. The expectations component (i) is driven by inflation expectations and expectations of future real rates of return, which depend on future economic growth. The risk premium component (ii) is determined by the degree of uncertainty about these future developments and by the degree of investors’ risk aversion. Similarly, bank equity prices can be decomposed into a component that reflects expectations of future profitability and a risk premium.

During a downturn, a pessimistic economic outlook drives down bank equity prices; the expectation of a loosening monetary policy response drives down sovereign bond yields. This is our expectations component (i). Separately, rising risk aversion during a downturn induces investors to turn away from riskier assets to safer ones. This reduces yields on safe assets and

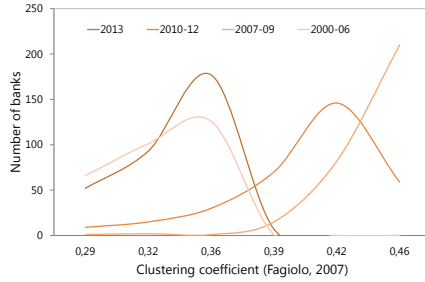
Figure 5.2: Distribution of clustering coefficients for bank equity price correlations and sovereign bond yield correlations. Source: Fagiolo (2007).



(a) Sovereign bond yield (all and Asia)



(b) Sovereign bond yield (all and euro area)



(c) Bank equity returns

raises yields (i.e. reduces prices) of riskier assets. This is our risk premium component (ii).

In safe haven countries, sovereign bonds are considered safe assets. Hence, both effects generate a positive correlation between bank equity prices and sovereign bond yields.

In contrast, in non-safe haven countries, sovereign bonds are not considered a “safe asset” to which investors will turn when risk aversion rises. As global risk aversion rises, therefore, investors will move out of both sovereign bonds and bank equity, sovereign bond yields will rise while bank equity prices fall, and, for a given economic outlook, a negative correlation between sovereign bond yields and bank equity prices will emerge. Since expectations about economic outlook and risk aversion drive the correlation between sovereign bond yields and bank equity prices into opposite directions, the sign of overall correlation is ambiguous.

Safe havens thus defined vary over time (Table 5.1). Japan, Germany, Finland and the United States have been considered safe havens throughout our sample period. But some euro area countries such as Austria, Belgium, France, Italy, Portugal, and Spain lost their safe haven status during the European crisis. Other countries gained safe haven status as the global financial crisis unfolded, including some commodity exporting countries such as Canada and Australia. As European economies are crawling out of recession against the backdrop of public and private deleveraging and as emerging markets are slowing down, many European economies and oil producers with close trade and financial links to emerging markets lost their safe haven status, including Australia, Canada, the UK, and Switzerland. How does our definition compare with other definitions of safe havens? Fig. 5.3 shows our list of safe havens against two other definitions:

- countries with AAA ratings from S&P, Fitch, and Moody’s (similar to definition used in the International Monetary Fund (2012));
- “negative-beta” countries whose sovereign bond yields are negatively correlated with global (here, S&P500) equity prices, a commonly used definition among financial market analysts.

In particular, Luxembourg, Singapore, and some Northern European countries are not identified

5.4. MAPPING THE NETWORK OF SOVEREIGN BOND YIELDS AND BANK EQUITY

Table 5.1: “Safe havens”: positive co-movement between sovereign yields and bank equity.

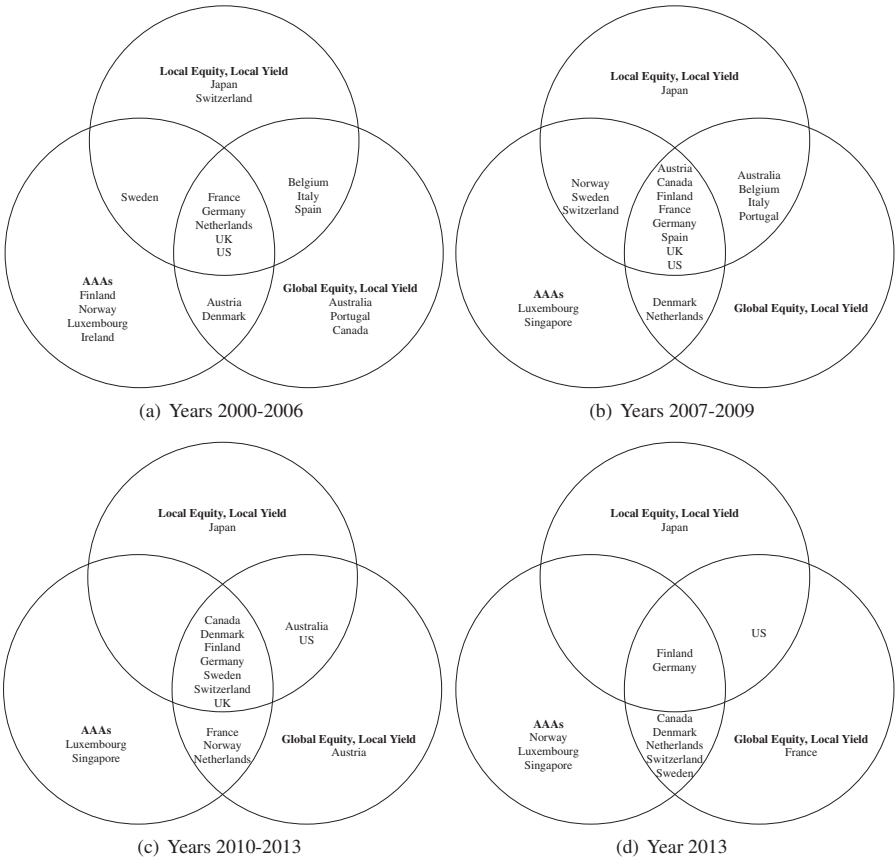
	2000-2006	2007-2009	2010-2012	2013
Austria	0	1	0	0
Australia	0	1	1	0
Belgium	1	1	0	0
Canada	0	1	1	0
Denmark	0	0	1	0
Finland	0	1	1	1
France	1	1	0	0
Germany	1	1	1	1
Italy	1	1	0	0
Japan	1	1	1	1
Norway	0	1	0	0
Netherlands	1	0	0	0
Portugal	0	1	0	0
Spain	1	1	0	0
Sweden	1	1	1	0
Switzerland	1	1	1	0
UK	1	1	1	0
US	1	1	1	1

as safe havens by our definition even though they are either AAA-rated or can be considered “negative beta” countries. Note that all these countries have fixed exchange rate regimes. Our expectations channel that distinguishes safe havens from others, and which works through expected monetary policy changes, would therefore be not expected to be strong.

5.4 Mapping the network of sovereign bond yields and bank equity

For visual clarity, we can only show a subset of all the pairwise links. As discussed above, sovereign-sovereign correlations tend to be much higher than bank-bank correlations and bank-sovereign correlations tend to be the weakest. To make sure that at least some links in each data set are represented, we select the strongest 10 percent of sovereign-sovereign, bank-bank, and

Figure 5.3: Safe havens by three definitions.



5.4. MAPPING THE NETWORK OF SOVEREIGN BOND YIELDS AND BANK EQUITY

bank-sovereign links.⁶

Fig. 5.A.1 shows the characteristics of sovereign-sovereign interconnectedness. Pre-crisis (2000-06), Singapore and, less directly, Korea were the main “bridges” between Emerging Asian and European sovereign bond yields. Since the crisis, sovereign bond yields in Emerging Asia have been drawn into the group of advanced country sovereign bond yield correlations.

European sovereign bond yields remain the most closely intertwined, despite some recent weakening of links with some of the periphery. In contrast, correlations with North American sovereign bond yields have weakened since the pre-crisis period.

Fig. 5.A.2 shows the characteristics of bank-bank interconnectedness. Pre-crisis (2000-06), there were few strong bank-bank correlations and they were confined to individual regions, Europe in particular, or individual countries. The global financial crisis (2007-09) tightened these disparate pre-crisis groups into one knot of cross-border correlations between bank equity returns. One Singaporean bank tied this tight global cluster to Asian-Pacific banks. Since then (2010-12) only the European cluster remains tightly intertwined (see also Box 1 for euro area countries) whereas other countries’ bank equity prices have drifted out of the dense global cluster.

In our sample, bank links across countries are significantly smaller than bank links within a country. There are only a few strong cross-border correlations outside Europe: in Asia-Pacific (Singapore and Australia) and North America (Canada and the US).

During the European crisis (2010-12), Asian and Latin American banks decoupled from banks in other advanced economies. In Asia, two cross-country bank clusters remained strong: one including individual Australian, Singaporean, Korean, and Malaysian banks and another including banks in Hong Kong and China. In Europe, banks in Greece and Cyprus separated from the main European cluster.

In Figs 5.A.3-5.A.5, for individual country groups, we parse the network for cross-country chains of correlations between banks and sovereigns.

⁶Due to space constraints, we only show the networks up to 2010-12. However, networks for 2013 do not materially differ from those for 2010-12.

- Emerging Asia:⁷ Significant (negative) correlations between sovereign bond yields and bank equity prices were present within each country. In contrast to these within-country correlations, cross-country correlations between sovereign bond markets and banking sector were relatively weak prior to the global financial crisis. This suggests that sovereign-bank feedback loops have played an important role in propagating shocks in emerging Asian economies domestically (despite the size and relative impact compare to the European countries) whereas cross-country contagion through either sovereign or banking channels was more muted. At the height of the global financial crisis (2007-09), sovereign-bank linkages strengthened in almost all the emerging Asian economies: shocks from European banks were transmitted through Singapore to other Asian banks which further propagated them to Asian sovereigns. During the subsequent euro area crisis (2010-12), sovereign-bank links weakened again whereas bank-bank links tightened, especially with European banks. Singaporean banks have continued to be the cross-continental “bridge”, affecting directly Thailand, China, Malaysia, and indirectly India and Indonesia. In contrast, Philippine and (some) Korean banks decoupled from the rest.
- Emerging Europe: Like emerging Asian countries, Turkey, Poland and Hungary (less in magnitude than the other two) bank equity prices were highly correlated with their own countries’ sovereign bond yields. During the global financial crisis (2007-09), both sovereign-bank links and cross-country banking sector linkages strengthened, contributing to stronger contagion. Stress in Turkey’s tightly-linked banking sector could potentially affect both sovereigns and banks in Poland and Hungary through the banking channel. During the subsequent European crisis (2010-12), Turkey decoupled from Poland and Hungary which remain together in a tightly interconnected cluster. Turkey and Romania developed into two highly correlated within-country groups.
- GIIPS and Cyprus: Unlike in emerging Asian and European countries, sovereign-bank interconnections were weak prior to the global financial crisis. During the global financial

⁷Malaysia, Thailand, Indonesia, Philippines, China, India.

5.4. MAPPING THE NETWORK OF SOVEREIGN BOND YIELDS AND BANK EQUITY

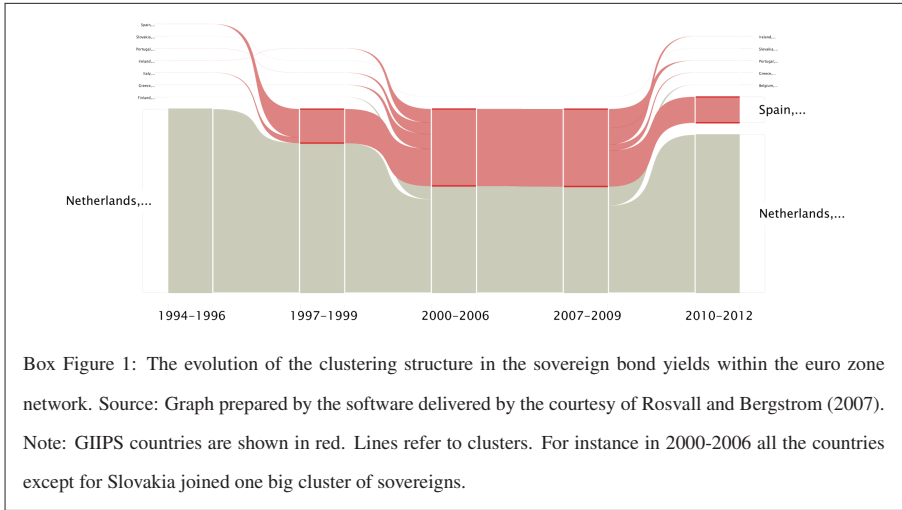
crisis (2007-09), Spanish and Italian banks began to be highly correlated with core European banks whereas the Greek and Cypriot banks formed a separate group of strong bank-bank correlation. Sovereign-bank linkages remained quite weak, however. As the European crisis deepened (2010-12), aside from higher interconnectedness of global banks, sovereign-bank inter-linkages also strengthened. Take Spain for example. Stress in bank 6 would have first affected Spain's sovereign which then could have propagated it most strongly to banks 2, 4 and 7 which, in turn, are highly correlated with Austrian and Italian banks, etc. Similar contagion chains can be drawn for Belgium, Portugal, Italy and Austria. Greece and Cyprus remained decoupled from the other European banks and sovereigns during this period.

Box 1. Clustering and declustering of the euro zone community in sovereign bond yields.

Since data is available from 1994, we construct a time line of the evolution of the network of the Economic Monetary Union (EMU) sovereign bond yields in 12 EMU and later euro area countries (Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Netherlands, Portugal, Slovakia and Spain). We can detect a changing core community in Europe over time. By a community, we understand a part of the global network where the interconnectedness is relatively higher than to the rest of the network. In order to distinguish communities we apply the random-walk algorithm developed by Rosvall and Bergstrom (2007).

In 1994-1996 the core of the EMU: Austria, Belgium, France, Germany and the Netherlands, built a separate cluster, visibly distinct from all remaining countries. In 1997-1999, Italy and Spain joined the core EMU cluster, and Greece, Ireland, Portugal, Finland joined it in 2000-2006. As might be expected from its late membership in the euro area in 2009, Slovakia did not join the community.

In 2010-2012 the core cluster partly dissolved. Belgium, Greece, Ireland and Portugal separated completely, whereas Italy and Spain joined a common cluster.



5.5 Modeling shock propagation

To investigate how shocks are propagated in this network we adapt a standard model from the disease spreading literature, developed by Jammazi and Aloui (2012). Every period each node propagates the cumulative shock it has received to all adjacent nodes. The impact of the shock is weighted by the strength of the link between the nodes. To keep it simple, we make two simplifying assumptions in our use of Jammazi and Aloui (2012). Firstly, we assume that nodes are neutral, i.e. our nodes cannot stop shock propagation, so that the propagation depends on the network structure only. Secondly, we do not put any boundaries on nodes' absorptive capacity, i.e. our nodes cannot slow down shock propagation, as it could hamper the actual cascade effect observed in financial markets. The details of the shock propagation mechanism are described in Appendix 5.B.

Our shock propagation exercise inherently assumes some degree of causality: a shock is “triggered” in one country and “passed on” to others. While the correlations themselves are agnostic on the direction of causality, we posit that causality is unlikely to run from small

entities to large entities. For example, the 93 percent correlation between changes of the 5-year sovereign bond yields of Ireland and Germany in 2010-12 is more likely to reflect the Irish sovereign bond market responding to shocks in Germany than *vice versa*. To capture this discrepancy when the source market is much smaller than the destination market, we scale the correlation between the two entities down proportionately to the relative size: We weight each correlation by the relative size of the source's and destination's total assets (for banks) or government debt (for sovereigns), capping the weight at one. (In future research, we aim to determine the direction of causality of the correlation in a less *ad hoc* manner, e.g. by using Diebold and Yilmaz (2011) spillover coefficients or including Granger (1969) causality measures.)

We simulate two types of shocks, one in each of our markets: a sovereign bond yield shock and a bank equity price shock. The initial shock is assumed to be a 1 percent increase in either sovereign bond yields or in daily bank equity prices. For example, in the first step, the source country's sovereign bond yield is increased by 1 percent. All the adjacent countries' (destinations') sovereign bond yields are then impacted by their (weighted) correlations with the source country sovereign bond yield. Also, the local banks' equity prices are affected by their correlation with their home sovereign bond yield. In the second step, the destination countries themselves become the countries of origins of the next round of shocks: each of them propagates the shock they received in the previous round to all their partner countries. The mechanism repeats step after step and in each step we calculate the cumulative effects of shock propagation in all the countries. We simulate shocks in three subsets of countries: a random shock in any country of the network, a simultaneous shock in all the GIIPS countries (Greece, Ireland, Italy, Portugal, and Spain), or a simultaneous shock in all the Fragile Five countries (Indonesia, India, Brazil, Turkey, South Africa).

Two more caveats are in order. Firstly, by assumption, there is nothing in our experiment that stops shock propagation; in practice, of course, policy steps would (and did) contain shock propagation. Of course, these policy interventions are also implicit in our estimated correla-

tions. Nevertheless, for now we interpret our results as counterfactuals that may have occurred, had modest additional shocks happened and/or had there been no additional policy measures. Secondly, our experiment does not say anything about the speed of contagion from shocks. Since almost all sovereigns bond yields and all bank equity prices have at least some correlation (even if small) and we do not exclude any by assumption. Therefore, the network is complete, i.e. a shock in any one part of the network will immediately travel to all other parts of the network. Instead of speed of contagion, our results are indicative of the size of the impact and the amplification over time of an initial shock on each country and on average. Although the steps have no time dimension, they show the path along which a shock travels around the network. Therefore, in our results below, we retain the notion of distinct steps for illustrative purposes.

5.6 Feedback loops in shock propagation

Feedback loops, even along the relatively weak sovereign-bank correlations in our data set, spread a shock from one asset class into another, where it can then proliferate and return to the initial asset class. We test the effect of feedback loops in sovereign bond contagion by comparing shock propagation under two scenarios: the actual network of sovereign-sovereign and bank-sovereign correlations and a counterfactual network where we assume all bank-sovereign links are zero, i.e. a counterfactual network in which feedback loops are not possible. In our counterfactual network without bank-sovereign links, a sovereign bond yield shock would not travel into the banking system at all and *vice versa*.

5.6.1 Sovereign bond yield shock

Fig. 5.4 shows the results of a sovereign bond shock in the GIIPS, the Fragile Five, or any country. Each line displays the average impact on sovereign bond yields of a 1 percent bond yield shock in any country (bottom panel), the GIIPS (top right panel), or the Fragile Five (top

left panel). We measure the impact relative to the impact under a baseline scenario: the baseline scenario is one where we assume all bank-sovereign feedback loops are zero. On average, feedback loops have amplified the impact of sovereign bond yield shocks (all curves are above 1). However, the strength of feedback loops varies across countries and over time depending on the source of the shock.⁸

On average, the amplification of feedback loops in 2013 is broadly similar for sovereign bond yield shocks in the Fragile Five and in the GIIPS (the curves in the top left chart are about level with those in the top right chart). We speculate that a sovereign bond shock propagates strongly in the highly interconnected sovereign bond yield network. It thus reaches countries with strong bank-sovereign feedback loops quickly and strongly, independent of the source of shocks. (This contrasts with a bank shock, see Section 5.6.2.) Not surprisingly, shocks in an average single country propagate less fast than shocks originating in all of the GIIPS or Fragile Five (the curve in the bottom chart is below those in the top right and top left charts).

For both Fragile Five and GIIPS sovereign shocks, the amplification by feedback loops has strengthened in 2013 compared with 2010-12 (the curve for 2013 is above that for 2010-12). This reflects a strengthening of bank-sovereign correlations in both sets of countries over these two periods. In the case of the GIIPS, for which earlier data are available, feedback loops are now only little stronger than pre-crisis.⁹

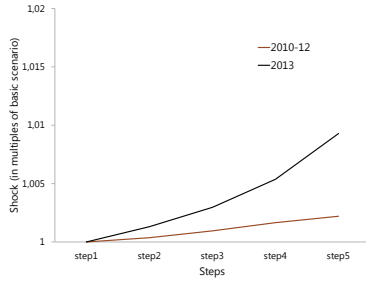
The weak impact of feedback loops in the case of a GIIPS shock in 2010-12 raises the possibility of an interesting interpretation. Shock propagation among asset prices may have been short-circuited by policy interventions, of which there were many in Europe during 2010-12. In 2013, when there were fewer policy actions targeted at dampening GIIPS shocks, a similar GIIPS shock would have propagated more strongly.

This contrasts with a shock in any single country which triggers broadly unchanged feed-

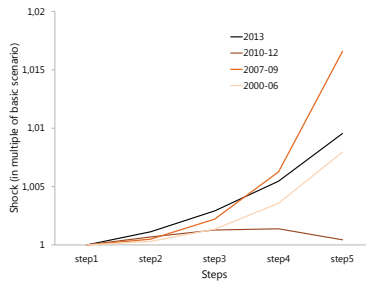
⁸The statistical significance of the differences in shock propagation between various settings and years depend on the size of the initial shock and the number of steps the shock has traveled across the network. Therefore, for presentational clarity, we do not report them in the main body of the text. The exercise aims to illustrate the general patterns of shock propagation in different years with different network structures.

⁹Data is not available for India and Turkey from mid-/late-2001, for Indonesia from 2003, and for Brazil from 2007.

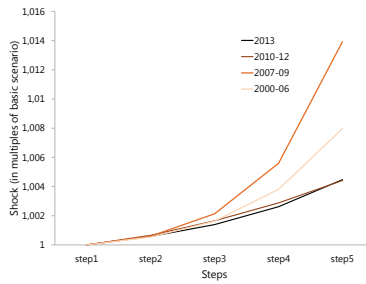
Figure 5.4: Average impact of bond market shock on global sovereign bond yields, with feedback loops (in multiples of average impact of same shock on sovereign bond yields without feedback loops).



(a) Shock propagation to Fragile Five



(b) Shock propagation to GIIPS

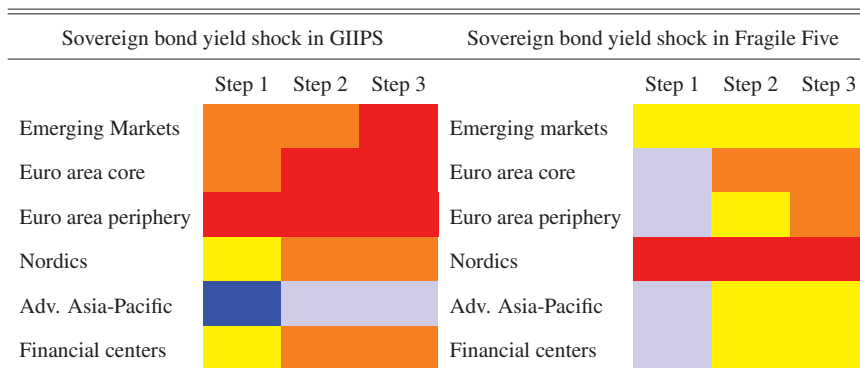


(c) Shock propagation to any country

back loops. The correlation of the Fragile Five and GIIPS sovereign bond yields with their banks' equity prices (already well above the sample average) increased more than for the average country between 2010-12 and 2013. As a result, these correlations intensified feedback loops from shocks originating from the GIIPS and Fragile Five.

Table 5.2 traces how feedback loops amplified sovereign bond yield shocks in the GIIPS or the Fragile Five in 2010-12. A redder tone indicates greater amplification by feedback loops. For example, feedback loops would have intensified the impact of a sovereign bond yield shock in the GIIPS initially (Step 1) more strongly (light orange) to the euro area core than the Nordics. Over time, feedback loops would have also amplified contagion to the Nordics (light orange in Step 2). In contrast, feedback loops would have strongly amplified contagion to the Nordics from a sovereign bond yield shock in the Fragile Five, mainly because strong bank-bank correlations with banks in the Fragile Five would have transmitted the shock to the Nordic sovereigns. In the euro area periphery, where bank-bank correlations with Fragile Five were less strong, feedback loops from a Fragile Five sovereign shock would have built more slowly over time.

Table 5.2: Average strength of feedback loops after sovereign bond yield shock by region in years 2010-12.



Note: Red = increase in sovereign bond yield in the highest quintile of each step; decline in bank equity price in the lowest quintile of each step. Orange = increase in sovereign bond yield in the second highest quintile of each step; decline in bank equity price in the second lowest quintile of each step. Yellow = increase in sovereign bond yield in the third highest quintile of each step; decline in bank equity price in the third lowest quintile of each step. Light blue = increase in sovereign bond yield in the second lowest quintile of each step; decline in bank equity price in the second highest quintile of each step. Dark blue = increase in sovereign bond yield in the lowest quintile of each step; decline in bank equity price in the highest quintile of each step.

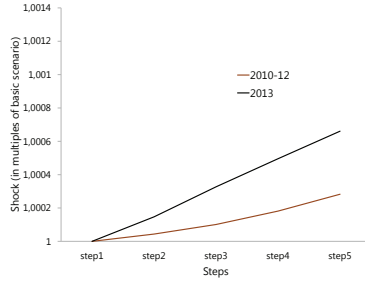
5.6.2 Bank equity price shock

We repeat the exercise but this time for a shock to bank equity prices in the GIIPS, the Fragile Five, or any country (Fig. 5.5). In general, feedback loops matter less for the propagation of bank shocks than for sovereign bond yield shocks (the scale of the vertical axis of Fig. 5.5 below is smaller than that of Fig. 5.4). This presumably reflects the fact that bank-sovereign correlations are generally weaker than sovereign-sovereign correlations and hence dampen the transmission of shocks from loosely interconnected bank equity prices to highly interconnected sovereign bond yields.

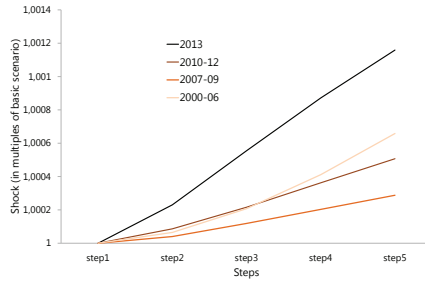
There are some notable differences to the propagation of bond shocks that deserve highlighting. In 2013, feedback loops amplify bank shocks in the GIIPS more than Fragile Five

5.6. FEEDBACK LOOPS IN SHOCK PROPAGATION

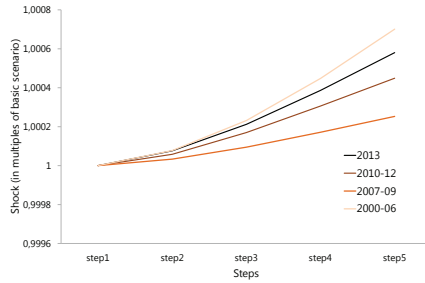
Figure 5.5: Average impact of bank equity shock on bank equity prices, with feedback loops (in multiples of average impact of same shock on bank equity prices without feedback loops).



(a) Shock propagation to Fragile Five



(b) Shock propagation to GIIPS



(c) Shock propagation to any country

shocks. In contrast to the sovereign bond yield shock, the greatest shock propagation occurs when the shock reaches the highly interconnected sovereign bond yield network. The entry point for a bank shock into the sovereign bond network is through bank-sovereign correlations. On average, bank-sovereign correlations in the GIIPS are twice as strong as those in the Fragile Five.¹⁰ As a result, bank shocks originating in the GIIPS are transmitted more strongly than bank shocks in the Fragile Five into the highly interconnected sovereign bond network. From there, shocks spread rapidly.

Also in contrast to sovereign bond shocks, feedback loops amplify bank shocks in the GIIPS or in any country more strongly now (2013) than they did at the height of the global financial crisis (2007-09). The reason for strengthening feedback loops after bank shocks is the shrinking number of safe havens. At the height of the financial crisis, bank equity price shocks in Italy and Spain triggered a decline in yields (probably as a result of a monetary policy response to financial system disruptions) that generated benign spillovers in the sovereign-sovereign network. In the next section, we explore the role of safe havens in more detail.

5.7 The role of safe havens in shock propagation

Our next exercise is focused on safe havens as defined above. In our sample, safe havens have two characteristics, one by definition and one by coincidence. Firstly, by our definition, safe havens display strong positive correlations between sovereign bond yields and bank equity prices. Secondly, by coincidence, they also display strong sovereign-sovereign and bank-bank correlations. This is the case not only for strongly correlated European sovereign bonds but also for non-European safe haven bonds. Even sovereign bond yields for non-European safe havens are, on average, correlated 50 percent more strongly with other sovereign bond yields than non-safe havens. In principle, the first characteristic is stabilizing to the network, whereas the second characteristic is destabilizing. A positive correlation between domestic banks and

¹⁰This difference is due both to stronger unweighted correlations and to higher weights (i.e. higher bank assets relative to sovereign debt) in the GIIPS than in the Fragile Five.

their sovereign (the first characteristic) dampens the impact of a foreign shock that, by itself, would spill over into a spike in sovereign bond yields and a drop in bank equity prices. A strong positive correlation with other sovereigns and banks, however (the second characteristic) generates strong transmission of any shock that arrives in a safe haven. Fig. 5.6 shows the different distributions of sovereign-sovereign, bank-bank and bank-sovereign correlations for safe havens and non-safe havens.

5.7.1 Sovereign bond yield shock

To distill the unique role of safe havens, we need to construct a “no-safe havens” counterfactual network that we can compare against our actual network. For our “no-safe havens” counterfactual network, we replace all the safe havens’ correlations with average correlations of non-safe haven countries (for bank-sovereign links alone, or in a separate experiment for sovereign-sovereign, bank-bank, and bank-sovereign links) as if they were the average non-safe haven country. Then we repeat the shock propagation exercises and compare with the results for the actual network.

Figs 5.7 and 5.8 show the role of safe havens in the propagation of sovereign bond shocks in the Fragile Five and the GIIPS. We measure the impact of a shock in a network without safe havens (one in which all safe haven correlations have been replaced with non-safe haven average correlations (continuous line)) against a baseline of the actual network of correlations. A line below 1 indicates that shocks propagate more strongly in a network with safe havens than in one without safe havens: the destabilizing effect of safe havens’ first characteristic predominates.

To distil separately the stabilizing effect of safe havens, we compare the same baseline of actual correlations against another counterfactual (dotted line) in which only bank-sovereign correlations of safe havens have been replaced with average non-safe haven correlations but all sovereign-sovereign and bank-bank correlations remain actual correlations. A dotted line above 1 indicates that the bank-sovereign links of safe havens dampen the propagation of shocks. In all our scenarios, shocks eventually propagate faster in networks with safe havens than without

Figure 5.6: Distribution of bilateral correlations for safe havens and non-safe havens.

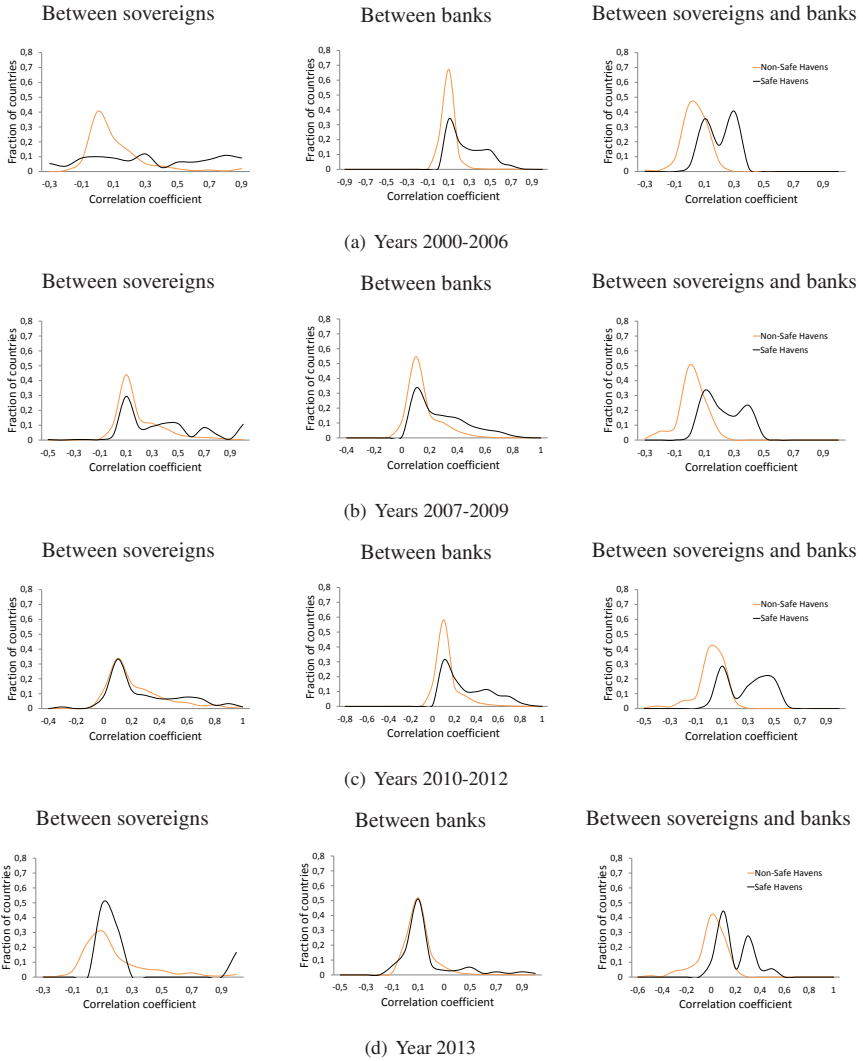
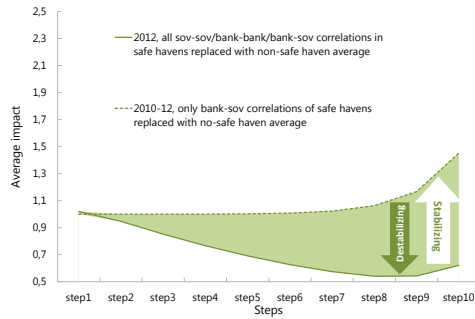
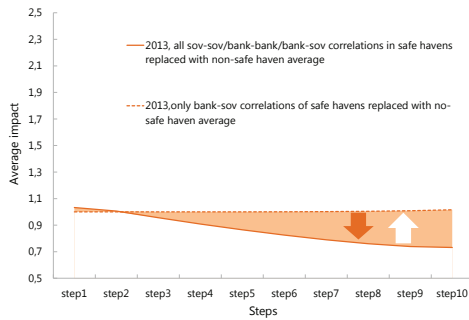


Figure 5.7: Average impact of Fragile Five sovereign bond shock without safe havens (in multiples of average impact of Fragile File bond market shock in the actual network of correlations).



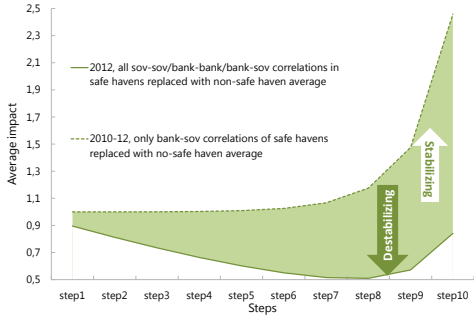
(a) Fragile Five sovereign bond shock in 2010-12



(b) Fragile Five sovereign bond shock in 2013

Note: The upper limit of the band indicates the impact when only bank-sovereign correlations are replaced for safe havens by sample averages; the lower limit indicates the impact when all correlations for safe havens are replaced by sample averages.

Figure 5.8: Average impact of GIIPS sovereign bond shock without safe havens (in multiples of average impact of GIIPS market shock in the actual network of correlations).



(a) GIIPS sovereign bond shock in 2010-12



(b) GIIPS sovereign bond shock in 2013

Note: The upper limit of the band indicates the impact when only bank-sovereign correlations are replaced for safe havens by sample averages; the lower limit indicates the impact when all correlations for safe havens are replaced by sample averages.

safe havens (the continuous lines are eventually below 1). Not surprisingly, the larger group of safe havens in 2010-12 than in 2013 results in stronger effects in 2010-12 than in 2013.

The stabilizing effects of safe havens take time to gather momentum after a sovereign bond shock. A sovereign bond shock spreads rapidly and strongly across the highly interconnected sovereign bond network. In contrast, the stabilizing bank-sovereign effect in safe havens only operates once a shock hits either a safe haven banking system or a safe haven sovereign.

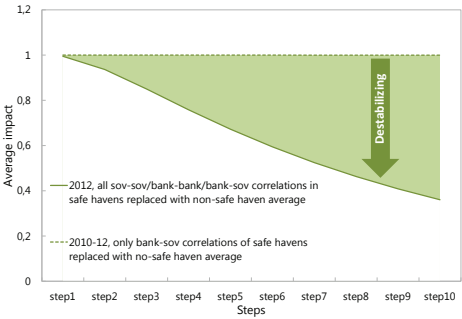
The stabilizing effect of safe havens depends on the origin of the shock. For example, in 2010-12, the stabilizing effect emerged more strongly and faster if the shock originated in the GIIPS than in the Fragile Five. Because GIIPS sovereign bond yields were on average one-third more strongly correlated with safe haven sovereign bond yields than Fragile Five sovereign bond yields, a sovereign shock originating in the GIIPS reached safe havens more strongly. This also triggered stronger stabilizing bank-sovereign links in safe havens.

5.7.2 Bank equity price shock

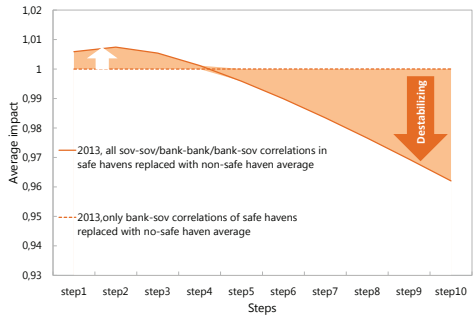
In Fig. 5.9, we conduct the same experiment for a bank equity shock in the Fragile Five countries. Again, a continuous line below 1 indicates that the presence of safe havens amplifies the propagation of shocks. The stabilizing effect of safe havens is too small to be noticeable in the chart because the origin of the shock (the Fragile Five) is weakly correlated with safe haven banks or sovereigns. However, as the shock reaches into the sovereign bond yield network, it is strongly amplified by the presence of safe havens.

In 2013, bank-bank correlations between Fragile Five banks and safe havens, on average, doubled compared with 2010-12 whereas sovereign-sovereign correlation between Fragile Five and safe havens halved. As a result, the stabilizing effects of safe havens were triggered more strongly in the initial phases of shock propagation but were later superseded by the destabilizing effects.

Figure 5.9: Average impact of Fragile Five bank equity shock on bank equity prices without safe havens (in multiples of average impact of Fragile File bank equity shock on bank equity prices in the actual network of correlations).



(a) Fragile Five bank equity shock in 2010-12



(b) Fragile Five bank equity shock in 2013

The upper limit of the band indicates the impact when only bank-sovereign correlations are replaced for safe havens by sample averages; the lower limit indicates the impact when all correlations for safe havens are replaced by sample averages.

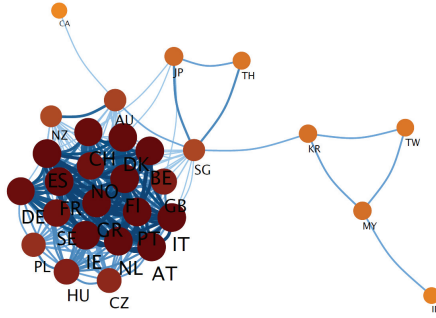
5.8 Conclusions and issues for further research

Our results thus far highlight a few stylized facts. We show how competing features of safe havens (highly interconnected sovereign bond yields versus stabilizing bank-sovereign links) combine to accelerate shock propagation in global bond and bank equity prices. We also show how feedback loops amplify especially shocks in the highly interconnected sovereign bond yield network. We speculate that these feedback loops may have been short-circuited by policy measures to contain contagion from GIIPS sovereign bond stress during the euro area crisis of 2010-12.

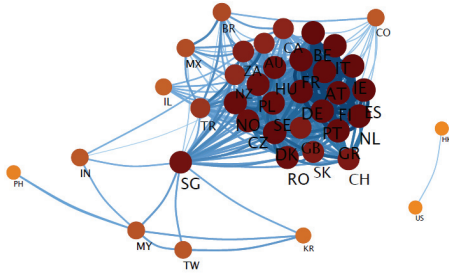
Our results raise some intriguing follow-on questions for further research. Firstly, the role of safe havens probably changes depending on their “neighborhood” in the network. Safe havens in deeply interconnected Europe may well play a different role than safe havens in Asia. Secondly, although we speculate in some instances about policies, their role is not directly addressed in this chapter. It is likely that announced policies altered the shape of the correlation network and drastically change shock propagation.

Appendix 5.A Network graphs

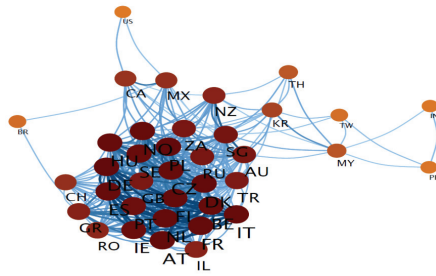
Figure 5.A.1: Sovereign interconnectedness.



(a) Years 2000-2006

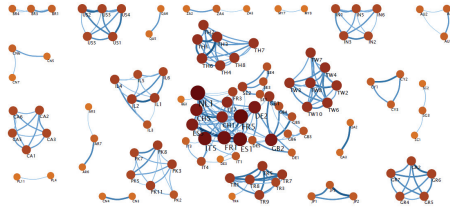


(b) Years 2007-2009

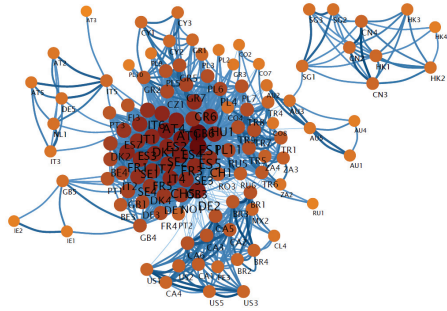


(c) Years 2010-2012

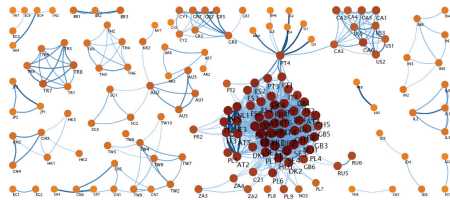
Figure 5.A.2: Bank interconnectedness.



(a) Years 2000-2006

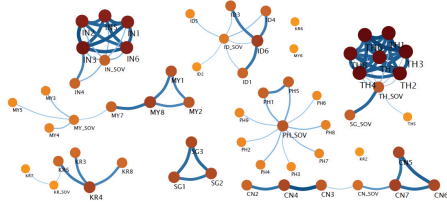


(b) Years 2007-2009

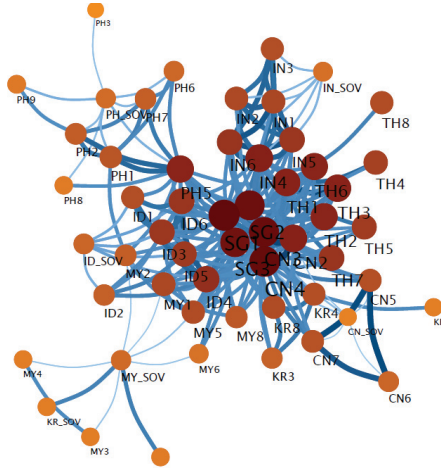


(c) Years 2010-2012

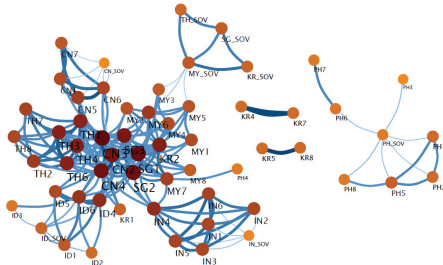
Figure 5.A.3: Sovereign-bank correlations in Emerging Asia.



(a) Years 2000-2006



(b) Years 2007-2009



(c) Years 2010-2012

Figure 5.A.4: Sovereign-bank correlations in Emerging Europe.

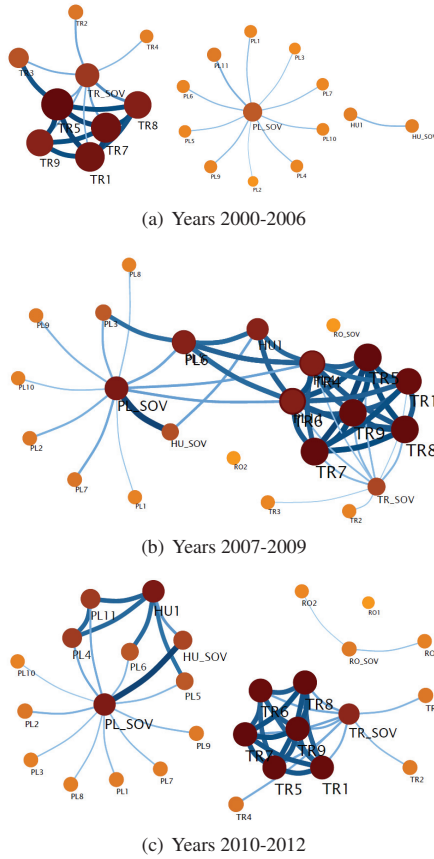
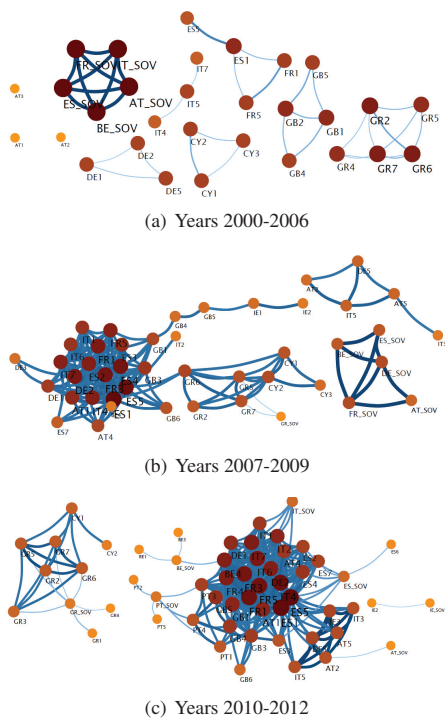


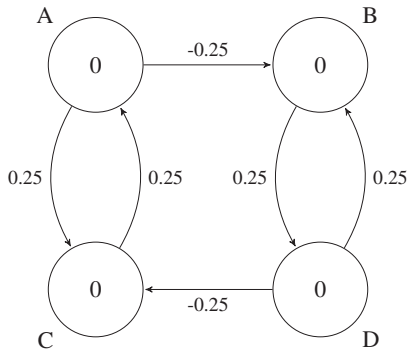
Figure 5.A.5: Sovereign-bank correlations in GIIPS and Cyprus.



Appendix 5.B Shock propagation mechanism

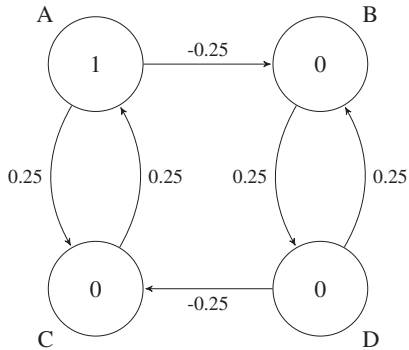
For illustration purposes, imagine a very simple network structure, consisting of 4 nodes connected by links of weights -0.25 and 0.25 in the following way

Figure 5.B.6: System setup.



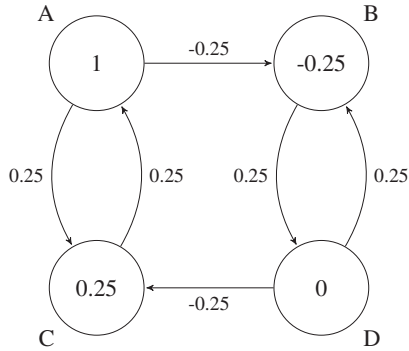
Before the shock, none of the nodes is affected so that all of them are 0. Imagine now, that in step 1 node A is hit by a shock of magnitude one.

Figure 5.B.7: Shock in node A.



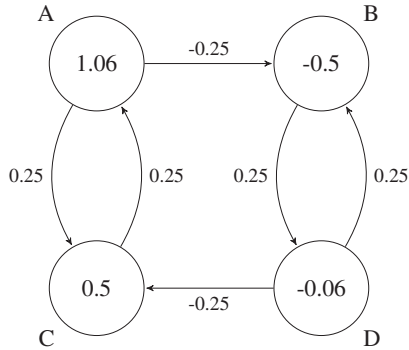
The node is now a source of the shock to the adjacent nodes B and C, propagating 25% of its initial magnitude with an appropriate sign.

Figure 5.B.8: Shock propagation (step 1).



In the second step, there are three sources of the shock, i.e. nodes A, B and C each, propagating 25% of the initial shock accumulated. Node A would therefore propagate 0.25 to the adjacent nodes B and C again. Node B would which would propagate 0.0625 to adjacent node D, and node C would propagate 0.0625 to node A. At the end of the second step, the network looks the following:

Figure 5.B.9: Shock propagation (step 2).



The process repeats itself for 10 steps. In each of them we calculate the cumulative shock in each of the nodes.

Chapter 6

Summary

In this thesis we aim at exploring the captivating and highly nonlinear profile of the modern world and assess its relevance in monetary policy conduct and macroprudential supervision. In particular, we focus on three different aspects of possible nonlinearities, i.e. as arising from (i) heterogeneous and boundedly rational expectations, (ii) probability distribution irregularities and (iii) complex network structures in the globalized economy.

We propose formal practical tools for central bankers and financial authorities to assess nonlinear structures among various institutions and system as a whole. In times of very non-standard policy actions, these tools might prove to be of great importance as they may reveal existing nonlinear relations and dependencies which standard econometric models cannot capture.

Highlighting the detailed outcomes from individual chapters, in Chapter 2 we investigate the possible irregularities arising from the presence of boundedly rational agents in the economy. We show that, in the setting with an active banking sector and extrapolative heuristics, the range of determinate policy instruments is narrowed. In fact this might have significant consequences for the real world. Pfajfar and Zakelj (2011) suggest that the fraction of extrapolative agents might be as high as 30%, which is even larger than in our setting. Given the fact that the estimated Taylor rule parameters vary usually in the region of $(0,1)$ for the output gap weight and of $(1,2)$ for the inflation weight (Taylor, 1999; Woodford, 2003), this may suggest that the

system is very close to indeterminacy, if not indeterminate already, which is the consequence of financial intermediation. This, in fact, is an interesting topic for further detailed investigation.

Chapters 3 and 4 propose formal methodologies to assess the influence of nonlinear causal structures in time series. We correct for these nonlinearities by assuming no underlying parameter structure but we test for the equivalence in conditional probability distribution instead. A clear advantage of such an approach lies in its generality. Chapter 3 reveals the intriguing relationships in the US grain market. Besides highlighting the role of nonlinearities, we discover a dual role of weather forecasts. Firstly, they seem to drive the causal relation from wheat to corn in the pairwise setting as they serve as a common factor, i.e. they affect both variables at the same time. Secondly, they are masking the causal relations from corn to beans and from beans to wheat in the system setting. Correcting for the common factor, we reveal the nonlinear Granger casual relations in the US grain market, suggesting that the causality runs from bigger, i.e. deeper and more liquid, to smaller markets. Chapter 4 tests the co-risk relations in the euro area financial sector. The results suggest that (i) only a few financial institutions pose a serious *ex ante* threat to the systemic risk, whereas, given that the system is already in trouble, there are more institutions which hamper its recovery and (ii) there are intriguing nonlinear structures in the euro zone systemic risk profile.

Chapter 5 treats nonlinearities from a network's perspective. Interestingly, our model highlights a few stylized facts observed over the past decade. We show how competing features of safe havens (highly interconnected sovereign bond yields versus stabilizing bank-sovereign links) combine to accelerate shock propagation in global bond and bank equity prices. We also show how feedback loops amplify shocks in the highly interconnected sovereign bond yield network. We speculate that these feedback loops may have been short-circuited by policy measures to contain contagion from GIIPS sovereign bond stress during the euro area crisis of 2010-12, supporting the actions of the European Commission, ECB and IMF.

A view to the future

The general conclusion arising from this thesis underpins the relevance of the nonlinear dynamics in economics and econometrics. The complexity of the real world has proven to play an important role in economic constructions since the very first models. As pointed out by Blinder (2013), this complexity has increased substantially, being one of the core reasons for the financial malaise during the crisis in the years 2007-2009. Therefore, even though the role of a model is to simplify the real world, it can cause severe consequences if it simplifies the reality by too much.

This thesis provides the tools and general directions on how to incorporate more complex structures into the existing economic methodologies. One should never claim that the ideas contained on these pages are ultimate as they arose as an answer to the recent financial crisis. With the technological advance and rapidly changing global environment, it is just a matter of time that these tools would not be enough. The same as the telescopes evolved satisfying the constantly rising curiosity about the mysteries of the deep universe, the methodologies describing complex and highly nonlinear economic structures should change in line with the advances in their fields.

Even though it is not possible to predict the stance of economic modeling in ten years from now, heterogeneous agents and network models constitute a very promising direction. With the support from behavioral economics and natural sciences, they create a natural horizon for future study of the main questions raised in this thesis.

The common conclusion, arising probably not only from this, but from many theses around the world, is to ask the right questions. This ultimately drives a researcher towards right answers, and right answers are likely to improve the world. But when a right questions is asked, a researcher can apply the question-specific methodology and work therefore more efficiently. This thesis advertises nonparametric statistics and econometrics, proposing novel approaches of assessing the role of nonlinear dynamics in the financial world. Although designed through a prism of a policy maker, the advances of these pages can be viewed as question-specific

methodologies so that to a large extent they rely on the ability to ask right questions.

The beginning of *modern monetary policy*, advertised in the Introduction, signalled the importance of macroprudential supervision and regulation. Hopefully, it will also encourage economists to look outside the box and to ask the right questions. For such researchers, the ideas contained in this thesis offer a powerful set of tools on how to capture the complexity around us.

Bibliography

- ABRAMSON, I. S. (1982): “On bandwidth variation in kernel estimates – A square root law,” *The Annals of Statistics*, 10, 217–1223.
- (1984): “Adaptive density flattening – A metric distortion principle for combating bias in nearest neighbor methods,” *The Annals of Statistics*, 12, 880–886.
- ACHARYA, V. (2009): “A theory of systemic risk and design of prudential bank regulation,” *Journal of Financial Stability*, 5, 224–255.
- ACHARYA, V., L. PEDERSEN, T. PHILIPPON, AND M. RICHARDSON (2010): “Measuring systemic risk,” Working Paper 1002, New York University.
- ACHARYA, V. V., I. DRECHSLER, AND P. SCHNABL (2011): “A Pyrrhic victory? – Bank bailouts and sovereign credit risk,” Working Paper 17136, National Bureau of Economic Research.
- ADRIAN, T. AND M. K. BRUNNERMEIER (2011): “CoVaR,” Working Paper 17454, FED New York.
- ASSENZA, T., P. HEEMEIJER, C. HOMMES, AND D. MASSARO (2011): “Individual expectations and aggregate macro behavior,” CeNDEF Working Paper 11-01, University of Amsterdam.
- BAECK, E. AND W. BROCK (1992): “A nonparametric test for independence of a multivariate time series,” Working Paper 9204, University of Wisconsin – Madison.

BIBLIOGRAPHY

- BAI, Z., W.-K. WONG, AND B. ZHANG (2010): “Multivariate linear and nonlinear causality tests,” *Mathematics and Computers in Simulation*, 81, 5–17.
- BASURTO, M. A. S., C. CACERES, AND V. GUZZO (2010): “Sovereign spreads: Global risk aversion, contagion or fundamentals?” Working Paper 120, International Monetary Fund.
- BAUER, M. D. AND G. D. RUDEBUSCH (2013): “What caused the decline in long-term yields?” Economic Letters, Federal Reserve Bank of San Francisco.
- BAUR, D. G. AND T. K. MCDERMOTT (2010): “Is gold a safe haven? International evidence,” *Journal of Banking and Finance*, 34, 1886–1898.
- BECK, R. AND E. RAHBARI (2008): “Optimal reserve composition in the presence of sudden stops: The euro and the dollar as safe haven currencies,” Working Paper 916, European Central Bank.
- BEKIROS, S. AND C. DIKS (2008a): “The nonlinear dynamic relationship of exchange rates: Parametric and nonparametric,” *Journal of Macroeconomics*, 30, 1641–1650.
- (2008b): “The relationship between crude oil spot and futures prices: Cointegration, linear and nonlinear causality,” *Energy Economics*, 30, 2673–2685.
- BERNANKE, B. S. AND M. GERTLER (1989): “Agency costs, net worth, and business fluctuations,” *The American Economic Review*, 79, 14–31.
- BERNANKE, B. S., M. GERTLER, AND S. GILCHRIST (1999): “The financial accelerator in a quantitative business cycle framework,” in *Handbook of Macroeconomics*, ed. by J. B. Taylor and M. Woodford, New York: Elsevier, 1341–1393.
- BERNANKE, B. S. AND M. WOODFORD (1997): “Inflation forecasts and monetary policy,” *Journal of Money, Credit and Banking*, 29, 653–684.
- BLANCHARD, O. J. AND C. M. KAHN (1980): “The solution of linear difference models under rational expectations,” *Econometrica*, 48, 1305–1311.

- BLINDER, A. S. (2013): *After the Music Stopped: The Financial Crisis, the Response, and the Work Ahead*, New York: Penguin Press HC.
- BRANCH, W. A. (2004): “The theory of rationally heterogeneous expectations: Evidence from survey data on inflation expectations,” *The Economic Journal*, 114, 592–621.
- BRANCH, W. A. AND B. MCGOUGH (2009): “A new Keynesian model with heterogeneous expectations,” *Journal of Economic Dynamics and Control*, 33, 1036–1051.
- (2010): “Dynamic predictor selection in a new Keynesian model with heterogeneous expectations,” *Journal of Economic Dynamics and Control*, 34, 1492–1508.
- BROCK, W. A. AND C. H. HOMMES (1997): “A rational route to randomness,” *Econometrica*, 65, 1059–1096.
- BRUNNERMEIER, M. (2009): “Deciphering the liquidity and credit crunch 2007-2008,” *Journal of Economic Perspectives*, 23, 77–100.
- BRUNNERMEIER, M., A. CROCKET, C. GOODHART, A. PERSSAUD, AND H. SHIN (2009): *The Fundamental Principles of Financial Regulation: Geneva Reports on the World Economy*, London: Centre for Economic Policy Research.
- BUITER, W. (2009): “The unfortunate uselessness of most state of the art academic monetary economics,” *Financial Times*, March 3.
- BULLARD, J. AND K. MITRA (2002): “Learning about monetary policy rules,” *Journal of Monetary Economics*, 49, 1105–1129.
- CABALLERO, J. (2012): “Banking crises and financial integration,” Research Department Publications 4816, Inter-American Development Bank.
- CAERS, J. AND M. A. MAES (1998): “Identifying tails, bounds and end-points of random variables,” *Structural Safety*, 20, 1–23.

BIBLIOGRAPHY

- CALVO, G. A. (1983): “Staggered prices in a utility-maximizing framework,” *Journal of Monetary Economics*, 12, 383–398.
- CAPORIN, M., L. PELIZZON, F. RAVAZZOLO, AND R. RIGOBON (2012): “Measuring sovereign contagion in Europe,” Working Paper 05, Norges Bank.
- CARRECK, N. AND D. CHRISTIAN (1997): “A study of grain dormancy and viability in spring barley,” *European Journal of Agronomy*, 6, 155–161.
- CARROLL, C. D. (2003): “Macroeconomic expectations of households and professional forecasters,” *Quarterly Journal of Economics*, 118, 269–298.
- CARTWRIGHT, N. (2007): *Hunting Causes and Using Them: Approaches in Philosophy and Economics*, Cambridge: Cambridge University Press.
- CASARES, MIGUEL AND POUTINEAU, J.-C. (2010): “Short-run and long-run effects of banking in a new Keynesian model,” *The B.E. Journal of Macroeconomics*, 11, 1–34.
- CHINAZZI, M. AND G. FAGIOLO (2013): “Systemic risk, contagion, and financial networks: A survey,” LEM Working Paper 2013/08, Sant’Anna School of Advanced Studies.
- CHINAZZI, M., G. FAGIOLO, J. A. REYES, AND S. SCHIAVO (2013): “Post-mortem examination of the international financial network,” *Journal of Economic Dynamics and Control*, 37, 1692–1713.
- CHOI, E. AND P. HALL (1999): “Data sharpening as a prelude to density estimation,” *Biometrika*, 86, 941–947.
- CHRISTOFFEL, K., G. COENEN, AND A. WARNE (2010): “Forecasting with DSGE models,” Working Paper 1185, European Central Bank.
- CLARIDA, R., J. GALI, AND M. GERTLER (1999): “The science of monetary policy: A new Keynesian perspective,” *Journal of Economic Literature*, 37, 1661–1707.

- CONLISK, J. (1996): "Why bounded rationality?" *Journal of Economic Literature*, 34, 669–700.
- DATZ, G. (2013): "The narrative of complexity in the crisis of finance: Epistemological challenge and macroprudential policy response," *New Political Economy*, 18, 459–479.
- DENKER, M. AND G. KELLER (1983): "On U -statistics and v. Mises' statistics for weakly dependent processes," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 64, 505–522.
- DIEBOLD, F. X. AND K. YILMAZ (2011): "On the network topology of variance decompositions: Measuring the connectedness of financial firms," Working Paper 17490, National Bureau of Economic Research.
- DIKS, C. AND V. PANCHENKO (2005): "A note on the Hiemstra-Jones test for Granger non-causality," *Studies in Nonlinear Dynamics and Econometrics*, 9, 1–7, art. 4.
- (2006): "A new statistic and practical guidelines for nonparametric Granger causality testing," *Journal of Economic Dynamics and Control*, 30, 1647–1669.
- DIKS, C. AND M. WOLSKI (2013): "Nonlinear Granger causality: Guidelines for multivariate analysis," CeNDEF Working Paper 13-15, University of Amsterdam.
- DIXIT, A. K. AND J. E. STIGLITZ (1977): "Monopolistic competition and optimum product diversity," *The American Economic Review*, 67, 297–308.
- DRAINVILLE, B., E. KHATCHADOURIAN, AND M. TORREY (2011): "The case for balanced investing in Emerging Markets," *Leadership Series (Investment insights)*.
- EDGE, R. M. AND R. S. GÜRKAYNAK (2010): "How useful are estimated DSGE model forecasts for central bankers?" Discussion Paper 8158, Centre for Economic Policy Research.
- EJSING, J. AND W. LEMKE (2011): "The Janus-headed salvation: Sovereign and bank credit risk premia during 2008-2009," *Economics Letters*, 110, 28–31.

BIBLIOGRAPHY

- ENGLE, R. (2002): "Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models," *Journal of Business & Economic Statistics*, 20, 339–350.
- EVANS, G. W. AND S. HONKAPOHJA (2001): *Learning and Expectations in Macroeconomics*, Princeton: Princeton University Press.
- EVANS, G. W. AND B. MCGOUGH (2005): "Monetary policy, indeterminacy and learning," *Journal of Economic Dynamics and Control*, 29, 1809–1840.
- FAGIOLO, G. (2007): "Clustering in complex directed networks," *Physical Review E*, 76, 2–8.
- FRANCIS, B. B., M. MOUGOUA, AND V. PANCHENKO (2010): "Is there a symmetric nonlinear causal relationship between large and small firms?" *Journal of Empirical Finance*, 17, 23–38.
- FRIEDMAN, M. (1953): *Essays in Positive Economics*, Chicago: Chicago University Press.
- FRYDMAN, R. AND M. D. GOLDBERG (2007): *Imperfect Knowledge Economics: Exchange Rates and Risk*, Princeton: Princeton University Press.
- FULLER, W. A. (1995): *Introduction to Statistical Time Series (Wiley Series in Probability and Mathematical Statistics)*, New York: Wiley-Interscience.
- GAO, B. AND R.-E. REN (2013): "The topology of a causal network for the Chinese financial system," *Physica A: Statistical Mechanics and its Applications*, 392, 2965–2976.
- GARCIA, N. E. (2011): "DSGE macroeconomic models: A critique," *Economie Appliquee*, 64, 149–171.
- GILBERT, C. L. (2010): "How to understand high food prices," *Journal of Agricultural Economics*, 61, 398–425.

- GILMORE, C. G., B. M. LUCEY, AND M. W. BOSCIA (2010): "Comovements in government bond markets: A minimum spanning tree analysis," *Physica A: Statistical Mechanics and its Applications*, 389, 4875–4886.
- GOODFRIEND, M. (2005): "Narrow money, broad money, and the transmission of monetary policy," in *Models and Monetary Policy: Research in the Tradition of Dale Henderson, Richard Porter, and Peter Tinsley*, ed. by J. Faust, A. Orphanides, and A. Reifschneider, Washington, DC: Board of Governors of the Federal Reserve System.
- GOODFRIEND, M. AND B. T. MCCALLUM (2007): "Banking and interest rates in monetary policy analysis: A quantitative exploration," *Journal of Monetary Economics*, 54, 1480–1507.
- GRAMMATIKOS, T. AND R. VERMEULEN (2012): "Transmission of the financial and sovereign debt crises to the EMU: Stock prices, CDS spreads and exchange rates," *Journal of International Money and Finance*, 31, 517–533.
- GRANGER, C. W. J. (1969): "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica*, 37, 424–438.
- GRANOVSKY, B. L. AND H. G. MLLER (1991): "Optimizing kernel methods: A unifying variational principle," *International Statistical Review*, 59, 373–388.
- GUO, S., C. LADROUE, AND J. FENG (2010): "Granger causality: Theory and applications," in *Frontiers in Computational and Systems Biology*, ed. by J. Feng, W. Fu, and F. Sun, London: Springer, vol. 15 of *Computational Biology*, 83–111.
- HABIB, M. H. AND L. STRACCA (2012): "Getting beyond carry trade: What makes a safe haven currency?" *Journal of International Economics*, 87, 50–64.
- HALE, G., C. CANDELARIA, J. CABALLERO, AND S. BORISOV (2013): "Bank linkages and international trade," Working Paper 2013-14, Federal Reserve Bank of San Francisco.

BIBLIOGRAPHY

- HALL, P. AND M. C. MINNOTTE (2002): “High order data sharpening for density estimation,” *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 64, 141–157.
- HARTMANN, P., S. STRAETMANS, AND C. DE VRIES (2004): “Asset market linkages in crisis periods,” *Review of Economics and Statistics*, 86, 313–326.
- HE, Z. AND A. KRISHNAMURTHY (2012): “A macroeconomic framework for quantifying systemic risk,” Working Paper 233, National Bank of Belgium.
- HIEMSTRA, C. AND J. D. JONES (1994): “Testing for linear and nonlinear Granger causality in the stock price – volume relation,” *Journal of Finance*, 49, 1639–1664.
- HOMMES, C. (2011): “The heterogeneous expectations hypothesis: Some evidence from the lab,” *Journal of Economic Dynamics and Control*, 35, 1–24.
- (2013): *Behavioral Rationality and Heterogeneous Expectations in Complex Economic Systems*, Cambridge: Cambridge University Press.
- HOMMES, C., J. SONNEMANS, J. TUINSTRA, AND H. VAN DE VELDE (2005): “Coordination of expectations in asset pricing experiments,” *Review of Financial Studies*, 18, 955–980.
- HUANG, X., H. ZHOU, AND H. ZHU (2010): “Systemic risk contributions,” Working Paper 60, Bank for International Settlements.
- INTERNATIONAL MONETARY FUND (2012): *Global Financial Stability Report*, Washington, DC: International Monetary Fund.
- (2013): *World Economic Outlook: Hopes, Realities, and Risks*, Washington, DC: International Monetary Fund.
- JAMMAZI, R. AND C. ALOUI (2012): “Crude oil price forecasting: Experimental evidence from wavelet decomposition and neural network modeling,” *Energy Economics*, 34, 828–841.

- JEONG, K., W. K. HÄRDLE, AND S. SONG (2012): “Nonparametric test for causality in quantile,” *Econometric Theory*, 28, 861–887.
- JONES, M. C. (1992): “Estimating densities, quantiles, quantile densities and density quantiles,” *Annals of the Institute of Statistical Mathematics*, 44, 721–727.
- KIYOTAKI, N. AND J. MOORE (1997): “Credit cycles,” *The Journal of Political Economy*, 105, 211–248.
- KRUGMAN, P. (2009): “How did economists get it wrong?” *New York Times*, September 6, MM36.
- KUPIEC, P. H. (2002): “Stress testing in a value at risk framework,” in *Risk Management: Value at Risk and Beyond*, ed. by M. A. H. Dempster, Cambridge: Cambridge University Press, 76–99.
- KYDLAND, F. E. AND E. C. PRESCOTT (1982): “Time to build and aggregate fluctuations,” *Econometrica*, 50, 1345–1370.
- LEVINE, R. (1997): “Financial development and economic growth views and agenda,” *Journal of Economic Literature*, 35, 688–726.
- LUCAS, R. E. (1972): “Expectations and the neutrality of money,” *Journal of Economic Theory*, 4, 103–124.
- MANGANELLI, S. AND G. WOLSWIJK (2007): “Market discipline, financial integration and fiscal rules: what drives spreads in the euro area government bond market?” Working Paper 745, European Central Bank.
- MANKIW, N. G. AND R. REIS (2007): “Sticky prices in general equilibrium,” *Journal of European Economic Association*, 5, 603–613.

BIBLIOGRAPHY

- MANKIW, N. G., R. REIS, AND J. WOLFERS (2003): “Disagreement about inflation expectations,” in *NBER Macroeconomics Annual*, ed. by M. Gertler and K. Rogoff, Cambridge: MIT Press, December 2002, 209–248.
- MASSARO, D. (2013): “Heterogeneous expectations in monetary DSGE models,” *Journal of Economic Dynamics and Control*, 37, 680–692.
- MERTON, R. C., M. BILLIO, M. GETMANSKY, D. GRAY, A. W. LO, AND L. PELIZZON (2013): “On a new approach for analyzing and managing macrofinancial risks,” *Financial Analysts Journal*, 69.
- MINOIU, C., C. KANG, V. SUBRAHMANIAN, AND A. BEREÄ (2013): “Does financial connectedness predict crises?” Working Paper WP/13/267, International Monetary Fund.
- MODY, A. (2009): “From Bear Stearns to Anglo Irish: How eurozone sovereign spreads related to financial sector vulnerability,” Working Paper WP/09/108, International Monetary Fund.
- MUTH, J. F. (1961): “Rational expectations and the theory of price movements,” *Econometrica*, 29, 315–335.
- NEWKEY, W. AND K. D. WEST (1987): “A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix,” *Econometrica*, 55, 703–08.
- PANDL, Z. (2013): “The bond risk premium,” *Columbia Management*.
- PARZEN, E. (1962): “On estimation of a probability density function and mode,” *The Annals of Mathematical Statistics*, 33, 1065–1076.
- PFAJFAR, D. AND B. ZAKELJ (2011): “Inflation expectations and monetary policy design: Evidence from the laboratory,” CentER Working Paper 2011-091, Tilburg University.
- PHILLIPS, P. C. B. AND P. PERRON (1988): “Testing for a unit root in time series regression,” *Biometrika*, 75, 335–346.

- POPP, M. P., C. R. DILLON, AND T. C. KEISLING (2003): "Economic and weather influences on soybean planting strategies on heavy soils," *Agricultural Systems*, 76, 969–984.
- POWELL, J. L. AND T. M. STOKER (1996): "Optimal bandwidth choice for density-weighted averages," *Journal of Econometrics*, 75, 219–316.
- RANALDO, A. AND P. SÖDERLIND (2010): "Safe haven currencies," *Review of Finance*, 14, 385–407.
- ROSVALL, M. AND C. T. BERGSTROM (2007): "Maps of random walks on complex networks reveal community structure," *Proceedings of the National Academy of Sciences of the USA*, 195, 1118–1123.
- ROTEMBERG, J. J. AND M. WOODFORD (1997): "An optimization based econometric framework for the evaluation of monetary policy," *NBER Macroeconomics Annual*, 12, 297–346.
- SAMIUDDIN, M. AND G. M. EL-SAYYAD (1990): "On nonparametric kernel density estimates," *Biometrika*, 77, 865–874.
- SARI, R., S. HAMMOUDEH, C.-L. CHANG, AND M. MCALEER (2012): "Causality between market liquidity and depth for energy and grains," *Energy Economics*, 34, 1683–1692.
- SCOTT, D. W. (1992): *Multivariate Density Estimation (Wiley Series in Probability and Mathematical Statistics)*, New York: Wiley-Interscience.
- SERFLING, R. J. (1980): *Approximation Theorems of Mathematical Statistics*, New York: Wiley.
- SHI, S. (1999): "Search , inflation and capital accumulation," *Journal of Monetary Economics*, 44, 81–103.
- SILVERMAN, B. W. (1998): *Density Estimation for Statistics and Data Analysis (Chapman & Hall/CRC Monographs on Statistics & Applied Probability)*, London: Chapman and Hall/CRC.

BIBLIOGRAPHY

- SIMON, H. A. (1955): "A behavioral model of rational choice," *Quarterly Journal of Economics*, 69, 99–118.
- (1957): *Models of Man: Social and Rational : Mathematical Essays on Rational Human Behavior in a Social Setting*, New York: Garland Publishing.
- SOROS, G. (2003): *The Alchemy of Finance*, New York: Wiley.
- STANISLAWSKA, E. AND E. TOMCZYK (2010): *Inflation Expectations: A View from Various Perspectives*, Warsaw: Warsaw School of Economics Publishing.
- TABE, H. (2010): *The Unravelling of Structured Investment Vehicles: How Liquidity Leaked Through SIVs: Lessons in Risk Management and Regulatory Oversight*, New York: Thoth Capital.
- TALEB, N. N. (2010): *The Black Swan: The Impact of the Highly Improbable*, New York: Random House Trade Paperbacks.
- TAYLOR, J. B. (1999): "A historical analysis of monetary policy rules," in *Monetary Policy Rules*, Cambridge: National Bureau of Economic Research, Inc.
- TOVAR, C. (2008): "DSGE models and central banks," Working Paper 258, Bank for International Settlements.
- TVERSKY, A. AND D. KAHNEMAN (1974): "Judgment under uncertainty: Heuristics and biases," *Science*, 185, 1124–1131.
- VAN DER VAART, A. (1998): *Asymptotic Statistics*, Cambridge: Cambridge University Press.
- VON KRUECHTEN, V., E. ROMANOVA, AND M. SCHMAUS (2009): "Austria-based Erste Group Bank AG 'A' rating affirmed on high systemic importance," Discussion Paper 2013/08, Standard & Poor's Rating Direct.
- WALSH, C. E. (2010): *Monetary Theory and Policy*, Cambridge/London: MIT Press, 3rd ed.

- WAND, M. P. AND M. C. JONES (1995): *Kernel Smoothing (Chapman & Hall/CRC Monographs on Statistics & Applied Probability)*, London: Chapman and Hall/CRC.
- WOLSKI, M. (2013a): "Exploring nonlinearities in financial systemic risk," CeNDEF Working Paper 13-14, University of Amsterdam.
- (2013b): "Monetary policy, banking and heterogeneous expectations," Working Paper 136, National Bank of Poland.
- WOODFORD, M. (1994): "Determinacy of equilibrium under alternative policy regimes," *Economic Theory*, 4, 323–326.
- (2003): *Interest and Prices*, Princeton: Princeton University Press.
- XIAOHUA, D. AND K. SHIYING (2012): "International early warning nonlinear model for securities systemic risk based on RBFNN using UDM and BVM," *Procedia Engineering*, 29, 1378–1386.

Samenvatting (Summary in Dutch)

In dit proefschrift willen we het boeiende en sterk niet-lineaire profiel van de moderne financiële wereld en zijn relevantie in het voeren van monetair beleid en macro-prudentieel toezicht verkennen. In het bijzonder richten we ons op drie verschillende aspecten van mogelijke niet-lineariteiten, te weten degenen die voortvloeien uit (i) heterogene en begrensd rationele (boundedly rational) verwachtingen, (ii) kansverdeling onregelmatigheden en (iii) complexe netwerkstructuren in de geglobaliseerde economie.

We opperen formele praktische instrumenten voor centrale bankiers en financiële autoriteiten ter beoordeling van niet-lineaire structuren tussen verschillende instellingen en het systeem als geheel. In tijden van zeer niet-standaard beleidsmaatregelen, kunnen deze instrumenten van groot belang blijken te zijn, aangezien zij aanwezige niet-lineaire relaties en afhankelijkheden kunnen laten zien die standaard econometrische modellen niet kunnen vatten.

Aandacht gevend aan de gedetailleerde resultaten van afzonderlijke hoofdstukken, onderzoeken we in Hoofdstuk 2 de mogelijke onregelmatigheden die voortvloeien uit de aanwezigheid van begrensd rationele agenten in de economie. We laten zien dat, in een omgeving met een actieve bankensector en extrapolerende agenten, het bereik van gedetermineerde beleidsinstrumenten wordt versmald. In feite kan dit aanzienlijke gevolgen voor de reële wereld hebben. Pfajfar en Zakelj (2011) suggereren dat de fractie van extrapolerende agenten zo groot als 30% kan zijn, zelfs groter dan in ons model. Gezien het feit dat de geschatte Taylor-regel parameters meestal in het interval $(0, 1)$ variëren voor het output-gap gewicht en $(1, 2)$ voor het inflatiegewicht (Taylor, 1999; Woodford, 2003), kan dit erop wijzen dat het systeem zich zeer dicht bij

onbepaaldheid bevindt, zo niet onbepaald is, ten gevolge van van financiële bemiddeling. Dit is een interessant onderwerp voor verder gedetailleerd onderzoek.

Hoofdstukken 3 en 4 introduceren formele methoden ter beoordeling van de invloed van niet-lineaire structuren in de tijdreeksanalyse. We corrigeren voor deze niet-lineariteiten door aan te nemen dat er geen onderliggende parameter-structuur is, maar we toetsen in plaats daarvan de gelijkwaardigheid in voorwaardelijke kansverdeling. Een duidelijk voordeel van een dergelijke aanpak ligt in zijn algemeenheid en intuïtieve gevolgtrekking. Hoofdstuk 3 toont de intrigerende relaties in de Amerikaanse graanmarkt aan. Naast het aantonen van de rol van niet-lineariteiten, ontdekken we een dubbele rol van weersvoorspellingen. Ten eerste lijken zij in de paarsgewijze toets de causale relatie van tarwe tot maïs te veroorzaken, waarin ze een gemeenschappelijke factor vormen. Ten tweede maskeren zij de causale relaties tussen maïs en bonen en tussen bonen en tarwe in de multivariate context. Corrigerend voor de gemeenschappelijke factor, onthullen we de niet-lineaire Granger causaliteitsrelaties in de Amerikaanse graanmarkt, die suggereert dat de causaliteit loopt van grotere, namelijk diepere en meer liquide, tot kleinere markten. Hoofdstuk 4 toetst de co-risico relaties in de financiële sector binnen het eurogebied. De resultaten suggereren dat (i) slechts een paar financiële instellingen een ernstige ex ante bedreiging voor de systeemrisico's vormt, terwijl, aannemende dat het systeem al in de problemen is, er meer instellingen zijn die het herstel belemmeren en (ii) er intrigerende niet-lineaire structuren zijn in het risicoprofiel van de eurozone.

Hoofdstuk 5 behandelt niet-lineariteiten vanuit het oogpunt van een netwerk. Interessant is dat ons model wijst op een aantal gestileerde feiten die gedurende de afgelopen tien jaar waargenomen zijn. We tonen aan hoe concurrerende kenmerken van veilige havens (sterk met elkaar verbonden rente op staatsobligaties tegenover stabiliserende banksoevereine koppelingen) combineren om schokvoortplanting in wereldwijde obligatie- en bank aandelenkoersen te versnellen. We laten ook zien hoe feedback loops schokken versterken in het sterk gekoppelde netwerk van opbrengsten van overheidsobligaties. We speculeren dat deze feedback loops kortgesloten kunnen zijn geweest door middel van beleidsmaatregelen om besmetting van GIIPS

staatsobligatie-stress tijdens de eurozone crisis van 2010-2012 te voorkomen, de acties van de Europese Commissie, ECB en IMF ondersteunend.

Een blik op de toekomst

De algemene conclusie uit dit proefschrift ligt ten grondslag aan de relevantie van de niet-lineaire dynamica in economie en financiën. De complexiteit van de echte wereld heeft een belangrijke rol gespeeld in economische constructies sinds de allereerste modellen. Zoals Blinder (2013) opmerkt is deze complexiteit substantieel toegenomen, en is het een van de hoofdredenen voor de financiële malaise gedurende de crisis in de jaren 2007-2009. Hoewel de rol van een model is om de werkelijkheid te vereenvoudigen, kunnen de gevolgen groot zijn als een model de werkelijkheid te veel vereenvoudigt.

Dit proefschrift verschaft middelen en algemene suggesties om complexere structuren in bestaande economische methodologieën te verwerken. We zullen echter niet beweren dat de ideeën uiteengezet op deze bladzijden het laatste woord zijn, omdat ze als een antwoord zijn gekomen op de recente financiële crisis. Met technologische vooruitgang en een snel veranderende globale omgeving is het slechts een kwestie van tijd eer deze hulpmiddelen tekort schieten. Net zoals telescopen zijn geëvolueerd om aan de alsmear toenemende nieuwsgierigheid omtrent de mysteries van het diepe heelal tegemoet te komen, zullen de methodologieën die complexe en sterk niet-lineaire economische structuren beschrijven zich moeten aanpassen aan de vooruitgang binnen hun onderzoeksvelden.

Hoewel het onmogelijk is om de stand van zaken in economisch modelleren over tien jaar te voorspellen, vormen heterogene-agentmodellen en netwerkmodellen een veelbelovende richting. Met hulp van gedragseconomie en de natuurwetenschappen creëren zij een natuurlijke horizon voor toekomstige bestudering van de hoofdvragen die in dit proefschrift zijn gesteld.

De gebruikelijk conclusie, die waarschijnlijk niet alleen uit dit proefschrift kan worden getrokken, maar uit vele proefschriften wereldwijd, is dat de juiste vragen gesteld moeten wor-

den, en dat de juiste vragen de wereld kunnen verbeteren. Pas als de juiste vraag gesteld is kan een onderzoeker vraag-specifieke methodologie toepassen en aldus efficiënter werken. Dit proefschrift staat het gebruik van niet-parametrische statistiek en econometrie voor, en stelt verschillende aanpakken voor het bepalen van de rol van niet-lineaire dynamica in de financiële wereld voor. Hoewel ontworpen vanuit de visie van een beleidsmaker, kunnen de vorderingen die hier zijn beschreven worden gezien als vraag-specifieke methodologieën, zodat zij in grote mate steunen op het vermogen om de juiste vragen te stellen.

Het begin van *modern monetair beleid*, waarvan de relevantie is aangegeven in de Inleiding (Introduction), gaf het belang aan van macro-prudentiële toezicht en regulatie. Hopelijk zal dit economen ook aanmoedigen om van de gebaande paden af te wijken en de juiste vragen te stellen. Voor dergelijke onderzoekers bieden de ideeën uit dit proefschrift een rijke verzameling hulpmiddelen om de complexiteit om ons heen te vatten.