REVEALING FINANCIAL INTERACTIONS

DISSERTATION

to obtain the degree of Doctor at the Maastricht University,
on the authority of the Rector Magnificus, Prof.dr. Rianne M. Letschert
in accordance with the decision of the Board of Deans,
to be defended in public
on Wednesday, 5th of July 2017, at 12:00 hours.

by

Hang Sun
Supervisor
Prof. dr. Jaap W. B. Bos

Co-supervisor
Dr. Paulo J. M. Rodrigues

Members of the Degree Committee
Prof. dr. Peter Schotman, Chairman
Prof. dr. Franz Palm
Prof. dr. Mike Tsionas (Lancaster University)
Prof. dr. Stefan Zeisberger (Radboud University Nijmegen)
This thesis is financially supported by the China Scholarship Council.
## Contents

**Preface** vii  

**Acknowledgments** xi  

1 **Introduction** 1  
   1.1 Interactions between Trader Behaviors and Commodity Prices 3  
   1.2 Modeling Dynamics of Financial Interconnections 5  

2 **Destabilizing or Passive?** 7  
   2.1 Introduction 7  
   2.2 Hypothesis Development and Empirical Strategy 13  
   2.3 Data and Empirical Results 26  
   2.4 Conclusion 34  
   2.A Adjusting the MHW Estimation with IID Reported Positions 36  

3 **In the Nick of Time** 39  
   3.1 Introduction 39  
   3.2 Identification 41  
   3.3 Methodology 44
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4</td>
<td>Baseline Model</td>
<td>50</td>
</tr>
<tr>
<td>3.5</td>
<td>Models with Extra Controls</td>
<td>64</td>
</tr>
<tr>
<td>3.6</td>
<td>Conclusion</td>
<td>67</td>
</tr>
<tr>
<td>3.A</td>
<td>Monte Carlo Analyses</td>
<td>68</td>
</tr>
<tr>
<td>4</td>
<td>Crisis-Contingent Dynamics of Connectedness</td>
<td>73</td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>73</td>
</tr>
<tr>
<td>4.2</td>
<td>A Not-So-Technical Overview of the Threshold Tripod Model</td>
<td>76</td>
</tr>
<tr>
<td>4.3</td>
<td>SVAR and T-SVAR Models</td>
<td>80</td>
</tr>
<tr>
<td>4.4</td>
<td>From Structural Models to Spatial Models</td>
<td>93</td>
</tr>
<tr>
<td>4.5</td>
<td>Network Analyses</td>
<td>107</td>
</tr>
<tr>
<td>4.6</td>
<td>An Application: Financial Contagion in the Eurozone</td>
<td>110</td>
</tr>
<tr>
<td>4.7</td>
<td>Conclusions</td>
<td>127</td>
</tr>
<tr>
<td>4.A</td>
<td>Mathematical Proofs</td>
<td>129</td>
</tr>
<tr>
<td>5</td>
<td>Conclusion</td>
<td>131</td>
</tr>
<tr>
<td>6</td>
<td>Research Impact</td>
<td>135</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>145</td>
</tr>
<tr>
<td></td>
<td>Curriculum Vitae</td>
<td>147</td>
</tr>
</tbody>
</table>
Listing of figures

2.1.1 S&P GSCI and Estimated CIT Net Long Positions ............... 9
2.2.1 CIT Net Long Positions Estimations .............................. 17

3.4.1 CIT Net Long Positions Estimations ............................... 53
3.4.2 State Probabilities of Baseline Model with Different Number of States ............................. 55
3.4.3 Impulse-Response Functions of the Restricted Baseline Model .......... 60
3.4.4 Impulse-Response Functions of the Unrestricted Baseline Model ........ 62
3.4.5 Impulse-Response Functions of the Restricted Baseline Model with 13 Lags .......... 63
3.5.1 Impulse-Response Functions of the Model with VIX and Term Spread Controlled .............................. 66

4.6.1 Stock Market Indices .............................................. 113
4.6.2 iTraxx Europe Values .............................................. 114
4.6.3 Values of the Threshold Function ................................. 116
4.6.4 Goolge Trends Indices .............................................. 118
4.6.5 An Overview of the Connectedness among the Five-Country System .... 122
The illustration of illustrious virtue is required below,
And the dread majesty is on high.
Heaven is not readily to be relied on;
It is not easy to be king.
Yin’s rightful heir to the heavenly seat,
Was not permitted to possess the kingdom.

Chinese civilization starts from the exploration of the law of Heaven\textsuperscript{1}, a supreme divinity that ancient Chinese people believed to rule the nature and human society. Heaven was believed to decide whether it would be rainy tomorrow, whether the yield of millet would be high for the next harvest, and even whether the king could continue ruling his kingdom. But the will of Heaven is just so unpredictable. Rulers were always concerned with the decisions of Heaven. What if Heaven decides that the king cannot rule any longer in the next morning, just like the rule of Shang dynasty – also known as "Yin" – was overthrown at the dawn of 20th January, 1046 BCE?\textsuperscript{2} Rulers at that time could do little but hymn with fear:

Heaven is not readily to be relied on;  
It is not easy to be king.

Earliest intellectual class in the Chinese society arose from the diviners who tried to understand the law of Heaven, nature, and humanity by communicating with Heaven with ox scapulae and turtle plastrons. They would write\textsuperscript{3} down their questions to Heaven with logograms on bones, heat the bones until cracked, judge the will of Heaven according to the pits, and then write the answer of Heaven down on the bones again with logograms. These logograms are later known as Chinese characters. Based on their judge of the will of Heaven, diviners were able to offer professional advice to their kings in all the aspects relevant to the current society.

But after hundreds of years’ practice of bone-cracking, people gradually found that pits on oracle bones could not predict the decision of Heaven significantly. New methodologies

\textsuperscript{1}During the Shang dynasty (c.1600-1046 BCE), the supreme ruler of the universe is more often called “Shangdi”, or the “Supreme Deity”. But later the conception is conflated with Heaven. I do not differentiate these conceptions in this short article.

\textsuperscript{2}The date is in Julian calendar. This is the most commonly accepted date of the historic battle called “Battle of Muye.” Other opinions about this specific date also exist.

\textsuperscript{3}It is worth noting that the original meaning of the word “write” is also “to incise, to engrave.”
must be created. After the historic failure for the House of Shang to maintain Heaven’s favor and rule, the new ruler of China, King Wu of Zhou, announced his people’s novel understanding of Heaven:

Heaven sees as my people see; Heaven hears as my people hear.\(^4\)

This is the first thought in the history of China to relate to the will of Heaven to the behaviors of the humanity. Hundreds of years later, following the idea of King Wu of Zhou, Confucianists raised an even stronger proposition:

What Heaven has conferred is called The [Human] Nature.\(^5\)

Nowadays, we know clearly that the answer to such a question as “How is the weather decided?” is only to a small extent related to the human nature, despite there is global warming and other environmental problems caused by human activities. But if we ask “How is the yield of millet decided?”, the answer will be more related to human behaviors, since farmers’ operations are affected by many social factors such as the price of millet as well as the risk. Moreover, whether a government can rule well or not is to an even more extent an issue of humanity. If there were really a “Heaven” that decided the patterns of the human society, the best tool to study the law of this “Heaven” we know today is social science, which is eventually a collection of different disciplines of scientific studies on the human nature.

But it is never easy to understand humans despite that we are humans ourselves. When humans gather and form the society, it becomes even more difficult to study the society than individuals. We still know so little about the human society. Even the particular question that how people allocate scarce goods in the society is so poorly understood. Worse, even the much smaller sub-question that how people exchange their goods and share risks in markets is little known. But we are still confident. At least, compared with our predecessors thousands of years ago who could only rely on ox scapulae and turtle plastrons, nowadays we have scientific theories and quantitative methods to study the society.

When I am writing this thesis, I often think of those professional writers on oracle bones. They wrote with knife and ink, whilst I write with computer and \LaTeX. But they and I all

write to understand more about the nature and the human society. It is hard to know whether they had ever thought that the logograms they wrote would still be used and their works still be read after millennia. But I am sure many people indeed had thought about why they were writing. So wrote a poet call Cao Pi (187–226 CE), who later became the first emperor of the state of Wei in the Three Kingdoms period of Chinese history:

Writing, it is the great work that frames a state, and the magnificent deed that never decays.⁶

⁶“Lun Wen (Essay on Literature)”, *Dian Lun (Normative Disquisitions)*.
In Chinese traditions, five things are the most respected, namely Heaven, Earth, parents, the monarch, and teachers. Let me acknowledge support for this thesis according to this order as well.

First and foremost are Heaven and Earth, which represent the nature and the natural environment that nurtures the human civilization. Well, I ought to thank the nature for the very existence of the cellular life and the human civilization, which are the greatest things in the entire universe; but no matter humans are grateful for the nature or not, nothing changes. Our greatest nature is just there.

I can never be too grateful for my parents, not only for the PhD, but for everything in my life. My brother and my father’s mother is always there supporting me, too. I love you all, and I am really sorry for leaving home for so many years and could not stay with you. My father’s father, and my mother’s parents, I wish you would be happy to see this thesis in Heaven.

Next, according to the tradition I should thank my monarch. Unfortunately, People’s Republic of China is a republic, so there is no longer a monarch to thank. I could thank the state, but such thanks are not much more substantial than thanks to the nature. But I sincerely appreciate the financial support from the China Scholarship Council (CSC), which affiliates to the PRC government. Thanks to everybody who works for the CSC, to taxpayers who fund the CSC, and to those who decided to establish the CSC to sponsor students for pursuing their academic dreams overseas.

The most crucial people for this thesis are my supervisors. Jaap, you used to say that you and Paulo were my academic parents. Indeed, without you and Paulo, there would not be this thesis, just like there would not be I without my parents. But you have done more than helping me writing this thesis, you are also my life mentor. The word “supervisor” in Chinese is composed by two morphemes “guide” and “teacher”, and you are undoubtedly the greatest supervisor, guide, and educator for me. Paulo, I always appreciate all your guidance and advice, too. They are invaluable to me, and I will never forget them. Thank
you.

After Heaven, Earth, parents, the monarch, and teachers, I have to admit there are too many people for me to fully acknowledge. I can try to name a small part of them:

Reading committee members, Peter Schotman, Franz Palm, Mike Tsionas, and Stefan Zeisberger, thank you for your time and efforts in reading and evaluating this thesis. You precious comments are truly appreciated.

My co-author, Zhuo Li, without your work, this thesis could not have been completed. Moreover, I will never forget your help during my master period and after that. Thank you again, and this time it is for the PhD years.

I always feel so lucky to have met all the nice colleagues in Maastricht. Gaby and Omar, you are always the kindest people I feel I can rely on when I am lonely. You are like my sister and brother in my mind. Runliang and Ming, jiàoyóu! I firmly believe you will become the best PhDs. Shusen, you are my true friend. Discussing with you is the most enjoyable thing for me in these years. Alessandro, you are the best office-mate, and the best tutor, ever. Rogier Quaedvlieg, I owe you for the help and guidance regarding the job market as well as being part of my defence committee. Pomme and Rogier Holtermans, I really enjoy the time to work in the same office with you. Iman, I enjoy chatting with you. I always feel that we have so much to share. Of course, I will not forget your efforts in organizing all the PhD colloquium things. They help me a lot. Mike, thank you, too, for organizing those great colloquiums. Clarissa, Judy, and Wiebke, our trip to Philadelphia and New York is the greatest memory during my PhD time. Leonard and Oana, thank you heartily for being so caring when everything in the Netherlands was new and strange to me. Inka, Patrick, and Tobias, it is fun to stay with you. Aderajew, I believe it will be a very great thing to collaborate with you.

Dennis Bams, Jeroen Derwall, Piet Eichholtz, Stefanie Kleimeier, Thomas Post, Peter Schotman, Paul Smeets, and Stefan Straetmans, I sincerely appreciate all the favor, advice, and feedback you gave me, no matter they are about economics or any other thing in life.

Special thanks go to the ladies of the secretariat, Carina, Cécile, Els, and Francien. It is you and your work that make everyday life in the Finance Department as happy as everyday life in a warm family!

I also want to say much thanks to Alain Hecq, you really helped me much in econometrics. Nalan Bastürk, thanks for attending my defense. Martien Lamers and Irina Stanga, you are the people who guided me into the area of banking. I cherish the memory in Groningen a lot. Ines Wilms, we will definitely do the best job together!
This list of acknowledgements will soon become too long if I name every friend I got to know during my PhD years, not to mention those old friends who have always been supporting me. I have to save some pages for the chapters. This thesis is dedicated to you all.
In 1999, Ahmed Zewail was awarded the Nobel Prize in Chemistry for his work on femtochemistry, which reveals mechanisms of chemical reactions on extremely short timescales. Using a rapid “ultrafast laser” technique, he managed to observe how reactions take place step by step cross $10^{-15}$ seconds, also known as femtoseconds.

For financial economists, financial markets are like test tubes where traders, assets, and information react day by day. These “reactions” discover prices, and match demand with supply. They can also have real consequences: for example, in the middle of 2008, the S&P GSCI, a price index of major commodity futures contracts over the world, almost doubled last year, and then rapidly plumbed to a lowest value since 2005 in half a year. This immense
INTRODUCTION

volatility involves many necessities including energy and food. It does not only destabilizes the economy, but can also even be life-threatening – the poorest people spend roughly three-quarters of their incomes on staple foods (Hertel et al., 2007). More recently, a massive crisis has swept over Europe. Although the crisis originated from only a few countries such as Ireland and Greece, it contaminated the entire Europe and even countries outside Europe. Markets panic, banks fail, and workers lose jobs...

Naturally, financial economists want to understand how these financial “reactions” take place. But researchers of empirical finance often face significant identification problems when they try to extract a financial mechanism from the data. For example, financial market variables often response to shocks so promptly that many causal relationships will be lost among the observations, unless researchers can find some exceptionally good instrumental variables or a satisfactory high-frequency data set. But instrumental variables are not always available, and high-frequency data sets are usually either confidential or expensive, if they even exist. Can an economist still identify causality in financial markets even in this situation?

Certainly: this thesis discusses a series of issues concerning the revelation of interactions among traders and markets with relatively low-frequency data. Just like femtochemists who use a laser to reveal detailed processes of chemical reactions, I try to extract as much information about the financial mechanisms as is possible with limited data. Chapters 2 and 3 focus on the formation of the volatility in recent commodity markets. Through a series of tests, Chapter 2 shows that the large price volatility in most major commodity markets can be attributed to traders’ interactions following the positive feedback mechanism similar to De Long et al. (1990). Chapter 3 tries to look into this issue in a model-free manner. Using a novel identification strategy for a structural vector-autoregression (SVAR) model, this chapter figures out the simultaneous equilibrium between commodity index trader flows and crude oil futures returns. Chapter 4 further develops the econometric model in Chapter 3 and establishes a unified dynamic model that combines SVAR, spatial models, and network models together. This model can be used to analyze the dynamics of the interconnections among financial entities. As an example, I show what we can learn by applying the model to the contagion during recent Eurozone crises.
The contribution of this thesis is both empirical and methodological. One the one hand, this thesis sheds new light on two important recent financial phenomena that have significant real consequences, namely the price volatility in commodity markets and the contagion during recent Eurozone crisis. On the other hand, this thesis develops some novel econometric models. They can reveal information from relatively low-frequency data that traditional methodologies cannot do.

The themes of this thesis are further discussed in the remainder of this chapter with an overview of each chapter.

1.1 Interactions between Trader Behaviors and Commodity Prices

The cause for the significant increase in the volatility of major commodities poses a puzzle that has been heavily debated in recent years. Since 2009, many explanations have been proposed. Most naturally, many papers have argued that it was the shifts of physical demand and supply such as the energy demand of emerging economies, and biofuel production that drove the price pattern. However, people doubt that fundamentals were the sole factors that led to the large-scale price volatilities. They noticed, ‘Commodity markets have become more like financial markets in terms of the motivations and strategies of participants’ (Domanski and Heath, 2007). An epitome of the so-called “financialization” of commodity markets is the massive inflow of long-only Commodity Index Traders (CIT’s), who invest in commodity markets by duplicating the weights of a commodity price index. Since the flows of CITs to a large extent coincide with the fluctuations of commodities, many suspect that CITs played an important role in deciding commodity prices in recent years.

Chapter 2 empirically investigates 35 major commodity markets and shows that the recent price pattern can be attributed the positive feedback mechanism similar to the model in De Long et al. (1990, hereinafter DSSW). We identify investors who follow the trading strategies hypothesized in the DSSW model and show that such trading strategies lead to price patterns predicted by the DSSW model. Specifically, this chapter identifies two of the three classes of investors hypothesized by the DSSW model in commodity markets, namely the passive investors and the positive feedback traders. The third class of investors in the
INTRODUCTION

DSSW model is the “rational speculators” and naturally omnipresent in every market. We demonstrate that CITs did want to curtail their long demand when commodity prices went up and had a stabilizing effect on prices, similar with what the passive investors do in the DSSW model. Utilizing the public Commitments of Traders weekly report released by the United States Commodity Futures Trading Commission (CFTC), we find the ratio of long positions to total open interest of the positions of “CFTC nonreportable traders” can be a good proxy for positive feedback traders’ activities. These non-reportable traders are generally smaller traders and not required to report their weekly positions to the CFTC. We find they become more likely to take long positions rather than short positions when the return increased, making them victims of by the rational arbitrageurs as the return will subsequently fall back. The collaboration of index traders, positive feedback traders, and rational speculators drives commodity prices away from the fundamentals.

While Chapter 2 examines the interactions between traders and commodity prices, it relies on the hypotheses of the DSSW model and does not do the examination in a direct way. Is it possible to directly figure out how much traders respond to return shocks as well as the other way around without assuming any model? The major difficulty in doing so is that the available data on index trading are reported weekly, whereas commodity futures markets are generally so liquid that responses to shocks can regarded as instantaneous compared with the weekly data. Due to this endogeneity problem, it is not possible to estimate the instantaneous bidirectional influences between traders and commodity returns with traditional econometric models.

Chapter 3 tries to solve this problem with a novel methodology based on a new identification strategy for an SVAR model. Once an SVAR model is identified, the instantaneous impulse-response relationships among the variables considered will become clear. In order to identify the SVAR, we assume the data are generated by a heteroskedastic process and capture this heteroskedasticity with Markov-switching volatility states. This assumption provides us with new information about the structure of an SVAR model. In the context of the two-variable system, as long as we can further find a new variable that interacts (simultaneously) with one existing variable whilst not interacting (simultaneously) with the other, the model becomes fully identified. Focusing on the interrelationship between crude oil futures returns and CIT flows in particular, we confirm
the positive instantaneous responses of crude oil excess returns to CIT inflow shocks. In addition, our approach allows us to demonstrate that the response of crude oil excess returns to CIT inflow shocks decays at an exceptionally low speed: the response remains significant in several following weeks, suggesting that the information contained in the position changes of CITs is absorbed very slowly. In fact, we believe this methodology can be applied to many similar issues besides commodity markets considered in this chapter.

1.2 Modeling Dynamics of Financial Interconnections

Economic entities are often interconnected and these interconnections can have significant economic consequences. Particularly, the interconnections among financial entities, whether they are asset markets, firms, financial institutions, or economies, have drawn much attention since the 2008 financial crisis. A growing number of methodologies have been proposed to empirically measure financial connectedness. However, most of these existing literature only tries to depict a static state of financial connectedness and seldom attempts to model the dynamic evolution of the connectedness of a financial system.

Chapter 4 of this thesis proposes a novel and systematic way to empirically depict and analyze the connectedness among economic or financial entities in a given regime and its transitions across different regimes. Thus, the dynamics of connectedness are captured in a regime-switching model. With this model, we can figure out how the interconnections in an economic or financial system are decided by some threshold variables that we are interested in, and make comparisons among the interconnections in different regimes.

The model is particularly useful for applications where it is as important to study transitions of connectedness as the connectedness itself. As an illustration of my model, I show how it sheds some new light on the topic of financial contagions in the EU, which has become a hotly discussed topic since the start of the Eurozone crisis in late 2009.
INTRODUCTION
Destabilizing or Passive?

2.1 Introduction

This paper analyses the impact of index traders on the equilibrium prices of commodity futures. This question arises since the growth in importance of commodity markets as financial investment vehicles in recent years has been contemporaneous to the significant increase in the volatility of commodity indices. This fact is shown in Figure 2.1.1. The upper panel of Figure 2.1.1 shows the price process of the S&P GSCI index, a major commodity price index. It can be seen that commodity prices were subject to several boom and bust cycles in recent years. A possible explanation for this pattern, that naturally arises, would be a shift in the demand and supply for commodities. One example would be the
shift in energy demand of emerging economies, and biofuel production that drove the price pattern. Papers that provide analysis along this line include Hamilton (2009), Mueller (2011), and Kilian and Murphy (2014), only to name a few. However, the results in these studies show that fundamentals alone are not the sole reason for the observed patterns in prices.

A further reason put forward for the increasing price volatility is the flow of investment capital into and out of commodity markets. As noticed, e.g., in Domanski and Heath (2007) “Commodity markets have become more like financial markets in terms of the motivations and strategies of participants.” This flow is depicted in the lower panel of Figure 2.1.1, which shows the estimated long positions of index traders during the period between 2006 and 2014. A quick inspection of the movements shown makes it clear why the connection between index trading and market volatility has been pointed out in media outlets and has been the target of investigation in academic papers.

The most prominent proponent of this 'financialization' theory is Mr. Michael W. Masters, a hedge fund portfolio manager. He argued in front of the United States Commodity Futures Trading Commission (CFTC)(Masters, 2009) that the inflow of money into commodity index funds caused the price spike observed in 2008. Following the terminology of Irwin and Sanders (2012), we denote the claim made by Mr. Masters as the “Masters Hypothesis” hereafter.

Our goal is to shed some further light on the role that commodity index funds play in the futures markets. To do this we follow the theoretical model of De Long et al. (1990) (hereafter DSSW) and assume that the futures market is populated by three investor categories. These are positive feedback traders, passive investors, and rational speculators. We show, after controlling for further effects, that CIT positions are negatively affected by past returns and that they positively predict future returns. This finding resembles the passive investor category in the DSSW model best. We therefore conclude that the flows caused by index traders are unlikely to be a destabilizing factor in the price finding process that caused the increase in price volatility.

Our next step is to identify trading positions that resemble the positive feedback traders in
INTRODUCTION

Figure 2.1.1: S&P GSCI and Estimated CIT Net Long Positions
“CIT” stands for Commodity Index Traders. Their net long positions were estimated with the Rectified Masters–Hamilton–Wu Algorithm.
the DSSW model. To do that we analyze the flows of the non-reportable traders collected by the Commodity Futures Trading Commission (CFTC). The CFTC demands all the clearing members, futures commission merchants, and foreign brokers should report daily to them and we can therefore identify the trading positions of numerous trader categories. The reason to concentrate on the trading positions of non-reportable traders is that they are generally smaller in scale and, as has been shown in, e.g., Abraham and Ikenberry (1994), Lee et al. (1998), Bange (1994), smaller investors are more likely to follow simplistic trading strategies as postulated positive feedback traders would. Indeed, in our empirical applications we find that the change in positions of non-reportable traders are positively related to commodity returns and negatively to future returns. A characteristic that resembles the positive feedback traders in DSSW.

Combining both results, we argue that, although the financialization of commodity markets could have contributed to the increasing price volatility, commodity index traders are unlikely to be the source of this destabilization.

We have two caveats here. First, we are certainly not arguing that the non-reportable traders are all positive feedback traders and none of the reportable traders are sophisticated non-feedback traders. Given the dominating proportion of reportable traders’ open interest in the market, the bulk of these positive feedback traders should be concealed in the reportable section of the COT report. And second, we did not forget commercial hedgers, which are a separate type of investor in commodity futures markets. We simply did not differentiate financial or commercial traders or “speculators” and “hedgers”. Actually, nothing prevents commercial hedgers from being positive feedback traders, since commercial hedgers also desire minimum hedging cost as well as minimum commodity price risk. It is not unlikely for a farmer to decide to sell his or her harvest as quickly as possible if the market turned bear and he or she believed that the pessimistic trend were going to continue.

Certainly, so as to portray a panorama of the stage where positive feedback traders met passive investors, we have another important task. That is, we need to gain more knowledge

---

1 According to the CFTC, “[t]he aggregate of all traders’ positions reported to the Commission usually represents 70 to 90 percent of the total open interest in any given market.”
about the CITs in two directions: How CITs’ activities affected commodity prices, and how their activities were, in turn, affected by market conditions, including price trends. The first question is at the same time the continuation of the debate on the Masters Hypothesis.

The first part, which is the core argument of Masters Hypothesis, has received intensive study recently. Besides the informal evidences shown by Masters (2009) at the testimony, Singleton (2014) was the first formal attempt to discover some trace of the predictive power of CIT positions on commodity prices. Employing some simple regressions, he immediately found the changes of CIT net long positions significantly predicted following crude oil futures returns. But it did not take a long time before contrary evidences appeared. Irwin and Sanders (2012) argued that the mapping algorithm employed by Singleton (2014), which was once again proposed by Masters (2009), and hence his results, were problematic. Hamilton and Wu (2015) improved the Masters Algorithm somehow (we will denote this improved algorithm as the Masters–Hamilton–Wu, MHW algorithm), but still found the relationship discovered by Singleton (2014) between CIT positions and commodity futures returns insignificant, using to-date data.

However, when we look into the impacts of CITs with a much wider horizon of 35 different commodity futures as a whole, we still find commodity returns across markets are predicted by CITs flows, in contrast with Hamilton and Wu (2015) but consistent with the DSSW model. Besides the results, we also contribute to the topic methodologically. Employing the algorithm invented by Chow and Lin (1971), we further improve the MHW estimation by rectifying the estimated values with the monthly reporting Index Investment Data (IID).

Compared with the topic of how CITs’ activities affected commodity prices, there are few studies about how their positions were decided by market conditions. Cheng et al. (2014) is the only paper we know till now studying the determination of CIT positions. They find that VIX changes can significantly predict CITs’ positions changes. We hope to go beyond this, trying to answer whether CITs behaved like the passive investor class in the DSSW model or not. The DSSW model hypothesizes that passive investors decide their positions negatively correlated with the departure of prices from the values of underlying assets. Unfortunately, it is nearly impossible to know how much exactly the values of physical

---

2 Before June 2010, the IID reports were announced every quarter instead of every month, though.
commodities are – otherwise there would not have been any debate on the topic we are discussing in this paper. Therefore we have to bypass any potential changes in the real values of commodities and test a simpler hypothesis which says positive innovations in commodity futures prices lead to negative changes in CIT net long positions. However, in reality portfolios of CITs, who are usually large funds, are affected by many factors including the macro situations. Since our data contain the observations during 2008 financial crisis, these macro shocks will be especially important (Etula, 2013). We manage to make these macro effects properly controlled, and finally find that CITs are cutting their net long positions when contract prices were going up. This observation docks at the assumption of the DSSW model and implies that CITs were actually not a destabilizing power, instead of the other way around. Positive feedback trading, as we argue hereinbefore, is a reasonable candidate for the responsible party of recent commodity price volatilities, although the conditions for it to exert influences are likely to be provided by CITs’ activities.

Existing hypotheses for the channel through which the financial players, or so-called “speculators”\(^3\), in commodity markets are able to influence the level of commodity price can be roughly divided into three groups, all of which basically feature the a “speculator–destabilize” narrative: First, due to the capital constraints and hence the limits of arbitrage, shifts of contrast of the capital strength of the short side, typified by commercial hedgers, and the long side, typified by the CITs, will be reflected in prices. Acharya et al. (2013) exhibit a formal model and empirical results on how producers’ anxious hedging demand is alleviated by increased speculation activities. Mou (2010) provides some immediate evidences by showing CITs (used to) suffer significant loss when they have to go through the Goldman Roll and their counter-parties are financially incapable to fully neutralize the massive opening and closing of CIT positions. Second, some literature emphasizes information frictions. Information frictions may take the form of asymmetric information sets, and may also take the form of symmetric but differently comprehended information sets, also known as “agree to disagree”. Singleton (2014) supports this view with the correlation between the front-month NYMEX WTI futures

\(^3\)There have been no consistent definitions of terms “speculator” and “speculation” in existing literature. We therefore try to refrain from these terms as much as possible in the main body of this paper and we do not name any trader as a “speculator”, nor any trading as “speculative trading” because of the serious ambiguity of these words. Fattouh et al. (2013) has a special section discussing the pragmatics of them.
HYPOTHESIS DEVELOPMENT AND EMPIRICAL STRATEGY

prices and the cross-sectional dispersion of forecasts of oil prices one-year ahead by the professionals, surveyed by Consensus Economics. And third, Basak and Pavlova (2013) construct a theoretical model rooted in fund management structures, arguing that the common practice of institutional investors taking indices as benchmarks of their performance will also lead to an impact of institutional investors’ entering on commodity prices. The first two types of literature all regard the “speculators” as a homogeneous group who trade in the same way. Basak and Pavlova (2013) are among the earliest attempts to recognize different subgroups of “speculators”. However, in our knowledge there is still no empirical evidence supporting Basak and Pavlova (2013).

Our findings in this paper altogether add to the literature with a novel angle in understanding the roles that different classes of financial investors played in commodity markets, which is quite distinctive from the existing “speculator–destabilize” narratives.

2.2 HYPOTHESIS DEVELOPMENT AND EMPIRICAL STRATEGY

2.2.1 INVESTOR CATEGORIES

Our analysis is based on the theoretical framework put forward in De Long et al. (1990). We assume that the futures market for commodities is populated by three classes of investors: (1) Positive feedback traders, (2) passive investors, and (3) rational speculators.

De Long et al. (1990) assume a three period model with an underlying asset that pays a risky cash flow at the end of period three. The investors base the demand for the underlying asset on different information sets. The positive feedback traders buy the asset in case there is a price increase and sell otherwise. Passive investors’ demand is a function of the difference between the fundamental value of the asset and its current market price. Rational speculators’ demand function has two components. First, they also react to price differences between fundamental value and the current price of the asset. Second, they seek to take advantage of the behavior of the positive feedback traders by anticipating the increased demand due to price movements. De Long et al. (1990) show in their analysis that an increase in the number of rational speculators or positive feedback traders does lead
to a divergence of market prices from fundamental value and to an increase of price volatility.

Naturally, the model is to stylized to be taken directly to the data. We use it as guidance to make predictions on what we should observe in terms of demand functions for the different investor classes assumed in the model. The first modification we need to make is that the demand functions of the investors are not as simple as postulated in the theoretical model, i.e., do not depend only on the price movements of the underlying. We therefore control for other factors that have been shown in the literature to have an impact on fund flows. Second, we are not able to observe the fundamental value of a commodity, and therefore need to work with a modified demand function for the passive investor class. We argue that the probability for the market price to be higher than the fundamental value is increasing in the returns of that commodity, and vice versa. This brings us to our first two testable hypothesis.

1st and 2nd Hypothesis: Regressing the change in CIT positions on returns and control variables results in the coefficient for the returns to be negative for passive investors and positive for positive feedback traders.

In a second step we use the assumption that any deviation of fundamental value gets corrected in a short period of time. This implies that if the demand of an investor is driven by price differences from fundamental value it should be positively related to future returns. For example, if prices are below fundamental value the passive investor has positive demand. If prices revert to fundamental value, that means that the return is positive in the next period. This observation leads us to the second testable hypothesis.

3rd and 4th Hypothesis: The demand of passive investors positively predict future returns and the demand for positive feedback traders negatively predict future returns.

Taken together, we are able to classify the trading positions of commodity index traders and nonreportable traders into either passive investor or feedback trader, investment classes postulated in the paper by De Long et al. (1990).
2.2.2 Behaviors of CITs and Their Impacts on Commodity Markets

We will begin our empirical studies with CITs, the center of the Masters Hypothesis. In this paper we want to study the behaviors of CITs and their impacts on commodity prices immediately, so the foremost problem for us to solve is a proper estimation of CIT positions.

Estimating CIT Positions

Currently the sole accurate data source for CIT positions in non-agricultural commodity markets is the monthly CFTC IID report. Unluckily, the frequency of this report is too low to satisfy the need of many studies including ours. Meanwhile, the CFTC also announce the positions of CITs in major United States agricultural futures markets weekly, known as the Commodity Index Trader Supplement (hereafter CITS) of the Commitments of Traders report. These data on agricultural markets provided us with the necessary information to infer the CITs positions in those non-agricultural markets. This is the basic idea behind the Masters Algorithm.

Masters Algorithm assumes that CITs invested in commodity futures according to two most popular commodity indices, namely the S& GSCI and the DJ UBSCI. In the original version of the algorithm (Masters, 2009), Masters notices that the soybean oil contract in the S&P GSCI was not a component of the DJ AIG (later renamed as DJ UBSCI) and the cocoa, Kansas wheat and feed cattle contracts in the DJ AIG were not components of the S&P GSCI at that time. Since CIT positions in these contracts are known from the CFTC, and the compositions of the two indices are also announced to the public every year by the makers of the indices, one can easily figure out the money invested by CITs according to these two indices by dividing the known positions on index-exclusive agricultural commodities with their respective weights in the indices. Finally adding the amount of money invested according to each index together, one gets an estimation of total CIT positions. Singleton (2014) also employs this approach in obtaining his CIT positions data. However, this original version of this algorithm is documented by Irwin and Sanders (2012) to have various problems.
In response to the criticism of Irwin and Sanders (2012), Hamilton and Wu (2015) intelligently observe that “in fact there is no need to restrict the inference as Masters did only to commodities that appear in one of the indexes but not the other”. For any two arbitrary agricultural commodities $i$ and $j$ there exists a Masters-type estimate of the CIT holdings of any commodity contract $m$:

$$
\hat{X}_{mt}^{[i,j]} = \begin{pmatrix}
\delta^G_{mi} & \delta^D_{mi} \\
\delta^G_{mj} & \delta^D_{mj}
\end{pmatrix}
\begin{pmatrix}
\delta^G_{it} & \delta^D_{it} \\
\delta^G_{jt} & \delta^D_{jt}
\end{pmatrix}^{-1}
\begin{pmatrix}
X_{it} \\
X_{jt}
\end{pmatrix},
$$

where $\delta^G_{it}$ and $\delta^D_{it}$ are respectively the designated weights of contracts $i$ and $j$ in SP GSCI and DJ UBSCI, and $X$ denotes dollar positions. This estimation based on two arbitrary known contracts can be easily extended to a linear projection type estimator in the same way as we extend linear equation systems to linear regressions with OLS estimators, as is shown in details in Hamilton and Wu (2015). In this paper, we also use the linear projection version of the MHW algorithm based on all the 13 agricultural contracts contained in the weekly CITS report.

The new version of the algorithm improved the quality of estimations significantly, but after comparing with the accurate IID monthly data, we found there were still a gap to fill. The MHW estimations and the IID data are displayed in Figure 2.2.1 with dashed and double dashed lines respectively. The two lines generally keep similar trends but there remain a spread between them.

However, it does not mean that the MHW Algorithm is useless now. The MHW estimations and the IID data still have a Pearson correlation coefficient of 0.79 and a Spearman correlation coefficient of 0.83. MHW is still the only feasible way to obtain weekly estimations of CIT positions in non-agricultural markets. It is better to combine the information from both sources together; and we find Chow and Lin (1971) have proposed an ideal way to do this job.

The algorithm proposed by Chow and Lin (1971) has become a standard approach to secure monthly GDP estimations based on seasonally released GDP statistics and monthly released series like industrial added values (see, e.g., Cuche and Hess, 1999). The Chow-Lin
Figure 2.2.1: CIT Net Long Positions Estimations

Dashed, double dashed, and solid lines represent estimations of CIT net long positions in billion US$ estimated with the MHW Algorithm, reported by the CFTC IID (Index Investment Data) report, and rectified values.
Algorithm has two steps: In the first step the series with high frequency is projected onto the space of the series with lower frequency, and then the regression error is allocated optically to the projected values. The details of our rectification of the MHW estimations can be found in Appendix 2.A. We plot the rectified values in Figure 2.2.1 with the solid line. We can see for the figure that the rectified values keep a similar pattern with original MHW values and are consistent with the IID values. In fact, the rectified and original values share both Pearson and Spearman correlation coefficients of 0.97.

Modeling CIT Positions

In this experiment we try to test whether CITs’ behaviors resembled the hypothesized “passive investors” class in the DSSW model. That is, commodity futures returns negatively predict CIT net long positions growth. In the DSSW model passive investors are assumed to decide their demands to be $-a(p - \Phi)$, where $a$ is a constant coefficient, $p$ is the price of the asset, and $\Phi$ is a fundamental shock. Unfortunately we are not able to observe the exact values of fundamental shocks in commodity markets, so we have to simply ignore this term and try to test a reduced relationship that commodity futures returns negatively predict CIT net long positions growth. The returns are defined as the full collateralized returns of the front contracts. Hereinafter, we will abbreviate commodity futures returns and CIT net long positions growth as ComRet and CITPosGr respectively. Our dropping of the fundamental term is reasonable not only because we are also going to control many variables that are closely correlated to the macroeconomic environment and therefore the demands and supplies of commodities, but also because the observations are on a weekly basis, whereby the slow-moving fundamentals of commodities should not change much each week.

We examine 35 major commodity markets as a whole. These markets include the most important commodity futures markets in the United States, as are studied by Tang and Xiong (2012) except pork bellies due to the liquidity problem, and several metal contracts traded in London and got included in S&P GSCI or DJ UBSCI, plus an obsolete contract known as unleaded gas, which used to be a component of the indices. A detailed list of these contracts can be found in Table 2.2.1.
Table 2.2.1: Contracts and Weights as of 03-12-2013

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Energy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CL</td>
<td>WTI Crude Oil</td>
<td>NYM</td>
<td>.2568</td>
<td>.1038</td>
<td>.1311</td>
<td>.5339</td>
<td>Yes</td>
</tr>
<tr>
<td>HO</td>
<td>Heating Oil</td>
<td>NYM</td>
<td>.0652</td>
<td>.0393</td>
<td>.0439</td>
<td>.5932</td>
<td>Yes</td>
</tr>
<tr>
<td>HU</td>
<td>Unleaded Gas</td>
<td>NYM</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.2765</td>
<td>No</td>
</tr>
<tr>
<td>NG</td>
<td>Natural Gas</td>
<td>NYM</td>
<td>.0284</td>
<td>.1389</td>
<td>.1192</td>
<td>.2157</td>
<td>Yes</td>
</tr>
<tr>
<td>RB</td>
<td>RBOB Gasoline</td>
<td>NYM</td>
<td>.0583</td>
<td>.0372</td>
<td>.0410</td>
<td>.4330</td>
<td>Yes</td>
</tr>
<tr>
<td>LCO</td>
<td>Brent Crude Oil</td>
<td>ICE-UK</td>
<td>.2348</td>
<td>.0647</td>
<td>.0951</td>
<td>.6209</td>
<td>No</td>
</tr>
<tr>
<td>LGO</td>
<td>Gasoil</td>
<td>ICE-UK</td>
<td>.0889</td>
<td>.0000</td>
<td>.0159</td>
<td>.5906</td>
<td>No</td>
</tr>
<tr>
<td><strong>Grains</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BO</td>
<td>Soybean Oil</td>
<td>CBT</td>
<td>.0000</td>
<td>.0242</td>
<td>.0151</td>
<td>.5979</td>
<td>Yes</td>
</tr>
<tr>
<td>C</td>
<td>Corn</td>
<td>CBT</td>
<td>.0321</td>
<td>.0488</td>
<td>.0458</td>
<td>.3136</td>
<td>Yes</td>
</tr>
<tr>
<td>O</td>
<td>Oats</td>
<td>CBT</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.2438</td>
<td>Yes</td>
</tr>
<tr>
<td>RR</td>
<td>Rough Rice</td>
<td>CBT</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.2528</td>
<td>Yes</td>
</tr>
<tr>
<td>S</td>
<td>Soybeans</td>
<td>CBT</td>
<td>.0258</td>
<td>.0568</td>
<td>.0512</td>
<td>.4573</td>
<td>Yes</td>
</tr>
<tr>
<td>SM</td>
<td>Siybean Meal</td>
<td>CBT</td>
<td>.0139</td>
<td>.0381</td>
<td>.0338</td>
<td>.2706</td>
<td>Yes</td>
</tr>
<tr>
<td>W</td>
<td>Chicago Wheat</td>
<td>CBT</td>
<td>.0319</td>
<td>.0336</td>
<td>.0333</td>
<td>.3083</td>
<td>Yes</td>
</tr>
<tr>
<td>KW</td>
<td>Kansas Wheat</td>
<td>KBT</td>
<td>.0068</td>
<td>.0128</td>
<td>.0112</td>
<td>.3455</td>
<td>Yes</td>
</tr>
<tr>
<td>MWE</td>
<td>Minneapolis Wheat</td>
<td>MGE</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.2609</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Softs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Hypothesis Development and Empirical Strategy**
### Table 2.2.1 continues.

<table>
<thead>
<tr>
<th>CC</th>
<th>Cocoa</th>
<th>ICE-US</th>
<th>.0028</th>
<th>.0000</th>
<th>.0005</th>
<th>.2235</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT</td>
<td>Cotton #2</td>
<td>ICE-US</td>
<td>.0100</td>
<td>.0202</td>
<td>.0183</td>
<td>.1962</td>
<td>Yes</td>
</tr>
<tr>
<td>KC</td>
<td>Coffee</td>
<td>ICE-US</td>
<td>.0047</td>
<td>.0197</td>
<td>.0170</td>
<td>.3526</td>
<td>Yes</td>
</tr>
<tr>
<td>OJ</td>
<td>Orange Juice</td>
<td>ICE-US</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.1434</td>
<td>Yes</td>
</tr>
<tr>
<td>SB</td>
<td>Sugar #11</td>
<td>ICE-US</td>
<td>.0139</td>
<td>.0381</td>
<td>.0338</td>
<td>.1880</td>
<td>Yes</td>
</tr>
<tr>
<td>LB</td>
<td>Lumber</td>
<td>CME</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.1390</td>
<td>Yes</td>
</tr>
</tbody>
</table>

**Livestock**

<p>| FC  | Feeder Cattle | CME | .0059 | .0000 | .0011 | .1877 | Yes  |
| LH  | Lean Hogs    | CME | .0166 | .0215 | .0207 | .0449 | Yes  |
| LC  | Live Cattle  | CME | .0297 | .0364 | .0352 | .3336 | Yes  |</p>
<table>
<thead>
<tr>
<th>Metals</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>GC</td>
<td>Gold</td>
<td>CMX</td>
<td>0.0227</td>
<td>0.0883</td>
<td>0.0766</td>
</tr>
<tr>
<td>SI</td>
<td>Silver</td>
<td>CMX</td>
<td>0.0031</td>
<td>0.0271</td>
<td>0.0228</td>
</tr>
<tr>
<td>HG</td>
<td>New York Copper</td>
<td>NYM</td>
<td>0.0000</td>
<td>0.0689</td>
<td>0.0566</td>
</tr>
<tr>
<td>PA</td>
<td>Palladium</td>
<td>NYM</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>PL</td>
<td>Platinum</td>
<td>NYM</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>MAL</td>
<td>Aluminum</td>
<td>LME</td>
<td>0.0181</td>
<td>0.0450</td>
<td>0.0402</td>
</tr>
<tr>
<td>MCU</td>
<td>London Copper</td>
<td>LME</td>
<td>0.0298</td>
<td>0.0000</td>
<td>0.0053</td>
</tr>
<tr>
<td>MNI</td>
<td>Nickel</td>
<td>LME</td>
<td>0.0045</td>
<td>0.0192</td>
<td>0.0166</td>
</tr>
<tr>
<td>MPB</td>
<td>Lead</td>
<td>LME</td>
<td>0.0041</td>
<td>0.0000</td>
<td>0.0007</td>
</tr>
<tr>
<td>MZN</td>
<td>Zinc</td>
<td>LME</td>
<td>0.0050</td>
<td>0.0256</td>
<td>0.0219</td>
</tr>
</tbody>
</table>

This table lists the 35 different commodity futures contracts used by this paper, together with their demonstrative dollar weights as of 3rd December, 2013. Tickers for contracts traded in United Kingdom follows the handbooks of S&P GSCI Index Methodology. Weights are rounded to four digits. Market average weights are the dollar weights averaged across S&P GSCI and DJ UBSCI by their estimated market shares. These weights are then adjusted by their historical return correlations, and the results are put in the respective column. Note that these adjusted weights do not sum to one.
We can notice from Table 2.2.1 that the weights allocation of each contract in the two contracts is very unbalanced. As an example, on 3rd December, 2013, the WTI crude oil contract has a dollar weight of 0.36 in S&P GSCI and 0.10 in DJ UBSCI, but the weights of lean hogs are only about 0.02 in both indices. Furthermore, some contracts are not even included in the indices. This sharp contrast gives us an ideal way to test the hypothesis more efficiently, that is, we expect the price changes in contracts of larger weights have a significantly larger negative impact on CIT positions in contracts. We therefore calculate the market average weights of each contract by weighted-averaging the weights designated by two indices based on respective estimated market shares, and test whether this average weight has an interactive effect together with ComRet. The variable is abbreviated to MktWt. In this way, our design for testing this hypothesis can be written as:

\[ \text{CITPosGr}_t = \alpha_1 \text{CITPosGr}_{t-1} + \alpha_2 \text{ComRet}_{i,t-1} + \alpha_3 \text{MktWt}_{i,t-1} \\
+ \alpha_4 \text{ComRet}_{i,t-1} \times \text{MktWt}_{i,t-1} + \alpha \text{Controls}_{t-1} + \epsilon_{i,t}, \quad (2.2) \]

where \( t \) represents the week index of observations, \( i \) represents different contracts, Controls are the controlling variables which are to be introduced afterwards, and \( \epsilon \) is the error term.

However, we are now facing a serious problem in this experiment design. We know that many commodity contracts are close substitute goods or complimentary goods with each other, and therefore their price movements are correlated to a large extent. For instance, prices of RBOB and WTI crude oil co-move by nature, This means one percent change in the price of a RBOB futures contract is also expected to have large impact on CIT positions despite its small weight, only because there will almost surely be a shift in the price of WTI crude, whenever there is a change in the RBOB price. Thus, RBOB price shocks will have an influence that is not proportional to the nominal weight of RBOB.

In response to the the natural correlation between contracts, we improve our experimental design. We adjust the simple market average weights in the following way: Let \( \mathbf{w}_t \) be the vector containing the market average weights of the contracts at time \( t \), and \( \mathbf{r}_{i,t} \) and \( \mathbf{r}_{j,t} \) be respectively vectors of full collateralized returns of contracts \( i \) and \( j \neq i \). For the regression \( \mathbf{r}_{j,t} = \mathbf{a}_{ij} \mathbf{r}_{i,t} + \epsilon_{i,t} \), we have \( \mathbf{a}_{ij} \) as the OLS estimation of \( \mathbf{a}_{ij} \). Square matrix \( \mathbf{A} \) is defined as \( (\hat{\mathbf{a}}_{ij})_{n \times n} \), where \( \hat{\mathbf{a}}_{ij} \) for any \( i \), and \( n \) is the number of different contracts, and then our
HYPOTHESIS DEVELOPMENT AND EMPIRICAL STRATEGY

adjusted average weights $w_t \triangleq Aw_t$. We denote the adjusted weights as AdjWt hereinafter. We test the Hypothesis 1 again using the adjusted weights again, i.e.,

$$\text{CITPosGr}_t = \alpha_1 \text{CITPosGr}_{t-1} + \alpha_2 \text{ComRet}_{i,t-1} + \alpha_3 \text{AdjWt}_{i,t-1} + \alpha_4 \text{ComRet}_{i,t-1} \times \text{AdjWt}_{i,t-1} + \alpha_{\text{Controls}_{i,t-1}} + \epsilon_{i,t}, \quad (2.3)$$

The purpose for our controlling variables is to control the macro market conditions. They include CDS spread of major investment banks (CDSSprd), major investment bank returns (IBRet)$^4$, difference of VIX (DiffVIX), LIBOR, term spread (TermSprd), and TED spread (TEDSprd), and S&P 500 returns (SPRet).

Previous studies have laid considerable emphasis on the determining effect of market risk and investors’ own sensitivity to risks on the portfolio allocation of institutional investors. As we mentioned, the pioneering empirical study on the decision of CIT positions, Cheng et al. (2014), focus on the VIX and its interacting effect with CDS spread. Etula (2013) also model how financial constraints of investors affect their willingness to invest in the risky commodity markets. Even pension funds, usually known as more risk averse, also have different risk-taking behaviors based on the macro market conditions and their financial constraints (see, e.g., Mohan, 2014). Market conditions are especially important because our sample period contains the crisis period. After all, our controlling variables are also related to the fundamentals of commodities, since market conditions are themselves reflections of underlying economy. For example, term spread is known to be a leading indicator of the entire economy.

IMPLICATION OF CHANGES IN CIT POSITION

In the DSSW model, passive investors are indispensable in keeping a constant risk-bearing capacity. They thereby help stabilize the market. So there comes our Hypothesis 2: Growth rates of CIT net long positions positively predict commodity returns.

$^4$The so-called ‘major investment banks’ above refer to Bear Stearns, Citigroup, Credit Suisse, Goldman Sachs, HSBC, JP Morgan, Merrill Lynch, and Morgan Stanley. The two variables related to these investment bank, i.e., CDS and IBRet, are the market value weighted average.
Hamilton and Wu (2015) test a similar hypothesis. They directly test for each commodity contract whether CIT positions changes have significant predictive power for the returns in the following week, controlling one lag of the returns. In contrast, in this paper we test the hypothesis in the panorama of 35 contracts, and the test can be specified as:

\[
\text{ComRet}_{it} = \beta_1 \text{ComRet}_{i,t-1} + \beta_2 \text{CITPosGr}_{t-1} + \beta_3 \text{SPGSRet}_{t-1} + \epsilon_{it},
\]

where \(i\) stands for different contracts, and SPGSRet is short for S&P GSCI returns\(^5\). We include this variable for controlling the global conditions of commodity markets.

We then try to test the robustness by additionally controlling the returns of other assets, namely returns of S&P 500 (SPRet), MSCI Emerging Markets Index (EMRet), U.S. Dollar Index (USDRet), and J.P. Morgan U.S. Aggregate Bond Index (BondRet).

### 2.2.3 Nonreportable Traders

We focus on the behaviors of positive feedback traders and their purport to the markets in this section. Naturally, the first question we have to answer is whether there really exists such a type of investors or not. After that, we need to further figure out whether the existence of positive feedback traders did destabilize commodity prices significantly. We begin on the existence.

#### Modeling the Change in Nonreportable Trader Holdings

As we have discussed in the introductory section, it is not possible for us to exhibit a roster of each positive feedback traders in the markets as we do not have access to the trading records of the traders. That being said, we can still endeavor to expose their trails. We turn our glares to the nonreportable traders, who belong to non of the three categories defined as clearing members, futures commission merchants, and foreign brokers and therefore tend to be smaller in size and more possible to take the naïve positive feedback strategy.

\(^5\) Be careful not to confuse SPGSRet with the S&P 500 return, which is abbreviated as SPRet.
Both the long and short positions on each commodity contract are reported by the COT report, so the most intuitive indicator of nonreportable traders’ strategies should be their net long positions. But it is not the case. The simple net long interests of nonreportable traders are likely to be contaminated by the capital flows into and off commodity markets instead of pure reflection of nonreportable commodity traders’ true opinions towards the markets. In other words, commodity markets are always known to be with considerably high market entry barriers, and for this reason an increase in nonreportable net long interests probably merely indicates that more investors have stridden over the barrier of entry.

Based on the above considerations, we alternatively represent nonreportable traders’ strategies with the long-to-total open interests ratio, or in other words, the percentage of long interests in total open interests. In such an approach, the effect of investors flows will be ruled out, allowing us to isolate the shifts in nonreportable traders’ behaviors. The variable will be abbreviated as PctLong in the following parts of the paper.

The construction of the test is similar with the regressions used for testing Hypothesis 1 in Subsection 2.2.2. However, there is no need to cover the weights of contracts now, for that the COT report releases nonreportable positions separately for each contract and these positions are mutually independent. So the regress will be:

\[
PctLong_{t} = \gamma_1 PctLong_{t-1} + \gamma_2 ComRet_{i,t} + \gamma Controls_{t-1} + \epsilon_{i,t}, \tag{2.5}
\]

where Controls are still the same as the set of controlling variables in the test for Hypothesis 1 and all the other symbols remain the same meanings. Furthermore, note that PctLong is not differentiated, which follows the assumption of the DSSW model; and for this very reason we regress on ComRet, instead of ComRet_{i,t-1}. We expect a significantly positive \( \gamma_2 \) if nonreportable traders did follow a positive feedback strategy.

Besides, comparing with the tests in Section 2.2.2, we drop 9 of the 35 contracts which are traded outside US in testing the hypotheses in this section because the COT report does not cover them. The list of dropped contracts can be found in Table 2.2.1.
Impact of the Change in Noreportable Trader Holdings on Returns

Provided that there existed positive feedback traders, the DSSW model would predict that they had been followed by negative returns as prices must regress to fundamental levels and the naïve positive feedback traders would lose money to the rational traders who had already foreseen and decided to take advantage of the positive feedback demands.

This hypothesis can be easily tested with

\[
\text{ComRet}_t = \delta_1 \text{ComRet}_{t-1} + \delta_2 \text{PctLong}_{t-1} + \delta_3 \text{SPGSRet}_{t-1} + \epsilon_t, \quad (2.6)
\]

and similar with the test for Hypothesis 2 in Subsection 2.2.2, we also redo the regression after adding the asset returns as controlling variables.

2.3 Data and Empirical Results

2.3.1 Data Sources and Summary Statistics

Our dataset range from 17th January, 2006 to 3rd December, 2013, containing 412 weekly observations. CIT and nonreportable positions are obtained from or deducted based on the COT, CITS, andIID reports released by the CFTC. Investment bank market values and returns are obtained from the CRSP database. And all other variables are from Datastream.

The descriptive statistics for the variables can be found in Table 2.3.1. It is worth noting that the magnitude of AdjWt is approximately one order larger than MktWt because we do not scale the sum of AdjWt to one.

2.3.2 Negative Feedback Trading Pattern of CITs

The results of the tests for the CITs’ negative feedback tradings are shown in Table 2.3.2. The regression is estimated with contract fixed effects, and all the following regressions in
Table 2.3.1: Descriptive Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Min.</th>
<th>L. Hinge</th>
<th>Median</th>
<th>H. Hinge</th>
<th>Max.</th>
<th>Mean</th>
<th>Std.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ComRet</td>
<td>-0.330</td>
<td>-0.0230</td>
<td>0.00310</td>
<td>0.0261</td>
<td>0.396</td>
<td>0.00195</td>
<td>0.0453</td>
</tr>
<tr>
<td>MktWt</td>
<td>0</td>
<td>0</td>
<td>0.0166</td>
<td>0.0379</td>
<td>0.317</td>
<td>0.0286</td>
<td>0.0445</td>
</tr>
<tr>
<td>AdjWt</td>
<td>0.00898</td>
<td>0.244</td>
<td>0.325</td>
<td>0.474</td>
<td>0.723</td>
<td>0.352</td>
<td>0.155</td>
</tr>
<tr>
<td>PctLong</td>
<td>0.176</td>
<td>0.427</td>
<td>0.532</td>
<td>0.644</td>
<td>0.958</td>
<td>0.536</td>
<td>0.137</td>
</tr>
<tr>
<td>CITPosGr</td>
<td>-0.148</td>
<td>-0.0161</td>
<td>0.00193</td>
<td>0.0195</td>
<td>0.188</td>
<td>0.00200</td>
<td>0.0373</td>
</tr>
<tr>
<td>TermSprd</td>
<td>-0.598</td>
<td>1.12</td>
<td>2.13</td>
<td>2.93</td>
<td>3.77</td>
<td>1.92</td>
<td>1.25</td>
</tr>
<tr>
<td>TEDSprd</td>
<td>-0.422</td>
<td>0.488</td>
<td>0.788</td>
<td>1.80</td>
<td>5.98</td>
<td>1.36</td>
<td>1.43</td>
</tr>
<tr>
<td>LIBOR</td>
<td>0.00500</td>
<td>1.23</td>
<td>4.93</td>
<td>5.75</td>
<td>6.78</td>
<td>3.95</td>
<td>2.18</td>
</tr>
<tr>
<td>DiffVIX</td>
<td>-14.8</td>
<td>-1.14</td>
<td>-0.104</td>
<td>1.09</td>
<td>16.2</td>
<td>0.0175</td>
<td>2.77</td>
</tr>
<tr>
<td>CDSSSprd</td>
<td>23.3</td>
<td>79.3</td>
<td>133</td>
<td>166</td>
<td>296</td>
<td>126</td>
<td>63.9</td>
</tr>
<tr>
<td>SPRet</td>
<td>-0.119</td>
<td>-0.00776</td>
<td>0.00168</td>
<td>0.0107</td>
<td>0.0614</td>
<td>0.000571</td>
<td>0.0173</td>
</tr>
<tr>
<td>EMRet</td>
<td>-0.167</td>
<td>-0.0132</td>
<td>0.00354</td>
<td>0.0178</td>
<td>0.115</td>
<td>7.79e-5</td>
<td>0.0326</td>
</tr>
<tr>
<td>BondRet</td>
<td>-0.0191</td>
<td>-0.00242</td>
<td>0.000159</td>
<td>0.00245</td>
<td>0.0147</td>
<td>4.21e-5</td>
<td>0.00409</td>
</tr>
<tr>
<td>USDRet</td>
<td>-0.05</td>
<td>-0.00595</td>
<td>-3.80e-4</td>
<td>0.00592</td>
<td>0.0359</td>
<td>-2.00e-4</td>
<td>0.009722</td>
</tr>
<tr>
<td>SPGCRet</td>
<td>-0.112</td>
<td>-0.0172</td>
<td>0.00282</td>
<td>0.0194</td>
<td>0.161</td>
<td>0.00142</td>
<td>0.0342</td>
</tr>
</tbody>
</table>

This table reports descriptive statistics for the variables used in this paper. L. Hinge and H. Hinge mean respectively the lower and higher quartiles. The first four lines are the pooled statistics of the weekly observations for all the 35 contracts as is listed in Table 2.2.1 from 17-01-2006 to 03-12-2013, which sum to 14,420 observations, and the rest lines are the time series statistics of the 412 weekly observations in the same period. All numbers are precised to three significant numbers.
this paper are also estimated in the same way.

We start from the controlling variables. The most striking pattern in the block of controls is the significance: Not a single variable except LIBOR in the nominal weights setting can be rejected to have non-zero predictive power for CIT net long positions growth, demonstrating that the inflows and outflows of CITs had a very close relationship with macro market situations. This is another direct reflection of the traits of CITs as financial investors. Additionally, the choice of the measure of market weights do not influence the coefficients of controlling variables much, except LIBOR, while its signs keep negative in spite of the change in the significant level.

The coefficients of IB CDS spread and differences of VIX are quite consistent with the story in Cheng et al. (2014), though our settings are slightly different from theirs at several points. Distress of the financial investors impel them to reallocate their assets to less risky highlands. A similar story goes with the negative coefficients of TED spreads, another indicator of perceived credit risk. The positive coefficients of IB returns are also intuitive, since higher IB returns indicate the financial sector are doing better. When funding cost, as is reflected by LIBOR, is higher, it is natural that CITs will hold less long interests in commodity futures. The negative coefficients of term spreads are not quite intuitive, which are possibly driven by CITs' hedge demands against inflation. Finally, the negative predictive power of S&P 500 returns are probably reflections of the inter-market money flows.

The key to understanding the core result is the coefficients of the interaction terms, i.e. $\text{ComRet}_{t-1} \times \text{AdjWt}_{t-1}$. The coefficients of the individual term ComRet are negative, though not significantly different from zero when using adjusted weights; however, they stand for nothing more than the the impacts on CIT positions of contracts with zero weights, or in other words, contracts outside commodity indices. The predictive power of returns of a non-zero weighted contract are given by the sum of the coefficients of this individual term and the interaction term, which are found to be always negative. These results are in good accordance with the predictions of Hypothesis 1: CITs trade in a negative feedback way in response to commodity price shifts and this negative feedback is even
**DATA AND EMPIRICAL RESULTS**

Table 2.3.2: Predicting CIT Net Long Positions with Commodity Returns

<table>
<thead>
<tr>
<th></th>
<th>Nominal Weights</th>
<th>Adjusted Weights</th>
</tr>
</thead>
<tbody>
<tr>
<td>CITPosGr&lt;sub&gt;t-1&lt;/sub&gt;</td>
<td>1.639e-1</td>
<td>131.96***</td>
</tr>
<tr>
<td>ComRet&lt;sub&gt;t-1&lt;/sub&gt;</td>
<td>-1.269e-2</td>
<td>-10.05***</td>
</tr>
<tr>
<td>MktWt&lt;sub&gt;t-1&lt;/sub&gt;</td>
<td>2.949e-3</td>
<td>6.83***</td>
</tr>
<tr>
<td>ComRet&lt;sub&gt;t-1&lt;/sub&gt; × MktWt&lt;sub&gt;t-1&lt;/sub&gt;</td>
<td>-2.849e-2</td>
<td>-2.08*</td>
</tr>
<tr>
<td>AdjWt&lt;sub&gt;t-1&lt;/sub&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ComRet&lt;sub&gt;t-1&lt;/sub&gt; × AdjWt&lt;sub&gt;t-1&lt;/sub&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDSSprd&lt;sub&gt;t-1&lt;/sub&gt;</td>
<td>-1.515e-5</td>
<td>-3.07**</td>
</tr>
<tr>
<td>DiffVIX&lt;sub&gt;t-1&lt;/sub&gt;</td>
<td>-7.062e-4</td>
<td>-26.21***</td>
</tr>
<tr>
<td>IBRet&lt;sub&gt;t-1&lt;/sub&gt;</td>
<td>1.741e-1</td>
<td>41.83***</td>
</tr>
<tr>
<td>LIBOR&lt;sub&gt;t-1&lt;/sub&gt;</td>
<td>-1.230e-4</td>
<td>-0.63</td>
</tr>
<tr>
<td>TermSprd&lt;sub&gt;t-1&lt;/sub&gt;</td>
<td>-1.809e-3</td>
<td>-11.47***</td>
</tr>
<tr>
<td>TEDSprd&lt;sub&gt;t-1&lt;/sub&gt;</td>
<td>-1.120e-3</td>
<td>-16.10***</td>
</tr>
<tr>
<td>SPRet&lt;sub&gt;t-1&lt;/sub&gt;</td>
<td>-1.263e-1</td>
<td>-72.46***</td>
</tr>
</tbody>
</table>

| Multiple R²          | 0.0491          | 0.0553           |

This table reports the results of Regressions 2.2 and 2.3 in Subsection 2.2.2 for testing the trading behaviors of CITs. “Nominal Weights” and “Adjusted Weights” stand for the regressions using MktWt and AdjWt respectively. The dependent variable is CITPosGr<sub>t</sub>. Regressions were estimated with contract fixed effects. The data cover 35 contracts and 410 effective time points, and 14,350 observations in total.

Nota bene, we report Wald Z-statistics instead of t-type statistics as is programmed in the plm package for R, whereby the computations were done. Coefficients are precised to four significant figures and the Wald Z-statistics rounded to two digits. Significant levels: *** for 0–0.001, ** for 0.001–0.01, * for 0.01–0.05, and . for 0.05–0.1.
Table 2.3.3: CITs Influences on Commodity Returns

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ComRet(_{t-1})</td>
<td>-0.01456</td>
<td>-766.47***</td>
</tr>
<tr>
<td>CITPosGr(_{t-1})</td>
<td>0.05902</td>
<td>6.80***</td>
</tr>
<tr>
<td>SPGCRet(_{t-1})</td>
<td>-0.0667</td>
<td>-25.38***</td>
</tr>
<tr>
<td>SPRet(_{t-1})</td>
<td>-0.0667</td>
<td>-25.38***</td>
</tr>
<tr>
<td>EMRRet(_{t-1})</td>
<td>-0.8557</td>
<td>-40.62***</td>
</tr>
<tr>
<td>USDRet(_{t-1})</td>
<td>-0.8998</td>
<td>-8.98***</td>
</tr>
</tbody>
</table>

This table reports the results in Subsection 2.2.2 for testing the impacts of CITs’ net long positions growth on commodity returns. We additionally include returns of other financial assets in (2). The dependent variable is ComRet\(_{t}\). Regressions are estimated with contract fixed effects. The data cover 35 contracts and 410 effective time points, and 14,350 observations in total. Coefficients are precised to four significant figures and the Wald Z-statistics rounded to two digits. Significant levels: *** for 0–0.001, ** for 0.001–0.01, * for 0.01–0.05, and . for 0.05–0.1.

larger for those contracts with larger weights in commodity indices, which we call “giants”.

2.3.3 Influences of CITs on Commodity Prices

We report the results of the tests for Hypothesis 2 in this subsection. The regression coefficients can be found in Table 2.3.3.

We can see from Table 2.3.3 that decreases of CITs’ demand indeed predict negative commodity futures returns in the following weeks. Along with the finding in the previous subsection that CITs traded in a negatively feedback way, CITs are now confirmed to be the
“passive investors” hypothesized in the DSSW model which help stabilize the market—though we do not call them as stabilizer because they statistically intermediated the mean reversion of commodity prices but because they help the whole market to keep a constant risk-bearing capacity. The outcome remains the same after controlling returns of some other assets as in (2).

In these group of experiments we also control the lagged returns of S&P GSCI since we are observing the 35 different markets at the same time. Foreseeably, returns of S&P GSCI take coefficients of the same signs as individual lagged returns. We then tried to add the returns of S&P500, MSCI Emerging Market Index, US Bond Index, and US Dollar Index into the regression. Albeit the new controlling variables but the MSCI EM Index share significant predictive power, the coefficients of the CIT net long positions growth do not change much. However, the standard deviance of the estimations becomes larger, as is implied by the absolute values of the Z-values, but the significance levels still remain high.

Besides the story about the impacts of CIT flows on the levels of commodity prices, our results in this and the previous subsection have a very interesting connotation on the relationship among different commodity markets as well. Noticing the quite unbalanced allocation of weights across different commodity contracts, it can be expected that a price shock in the “giant” markets such as crude oil affects CIT demands far more than in the “dwarf” markets such as coffee, and consequently gains much larger influences on all commodity markets. This is especially interesting because an increase in the crude oil price will be followed by a decrease in the CIT demands and then a decrease in the prices of all the markets, which forms a crowding-out effect, or “see-saw effect” as CITs needed to re-balance their dollar positions in different markets. The price correlations among intra-index commodity contracts can thus diminish, which is not considered in the style investing theory (Barberis and Shleifer, 2003) and makes an interesting footnote to the empirical findings of Cheng et al. (2014).
Table 2.3.4: Predicting Proportions of Nonreportable Long Interests with Commodity Returns

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PctlLong_{t,-1}</td>
<td>9.176e-1</td>
<td>1205.51***</td>
</tr>
<tr>
<td>ComRet_{t}</td>
<td>1.526e-1</td>
<td>177.26***</td>
</tr>
<tr>
<td>CDSSprd_{t,-1}</td>
<td>-4.998e-5</td>
<td>-1.26</td>
</tr>
<tr>
<td>DiffVIX_{t,-1}</td>
<td>-4.933e-4</td>
<td>-2.91**</td>
</tr>
<tr>
<td>IBRet_{t,-1}</td>
<td>3.493e-2</td>
<td>0.71</td>
</tr>
<tr>
<td>LIBOR_{t,-1}</td>
<td>-1.671e-3</td>
<td>-4.99***</td>
</tr>
<tr>
<td>TermSprd_{t,-1}</td>
<td>-3.913e-4</td>
<td>-0.23</td>
</tr>
<tr>
<td>TEDSprd_{t,-1}</td>
<td>1.823e-3</td>
<td>5.35***</td>
</tr>
<tr>
<td>SPRet_{t,-1}</td>
<td>-6.407e-2</td>
<td>-3.37***</td>
</tr>
<tr>
<td>Multiple R^2</td>
<td>0.9642</td>
<td></td>
</tr>
</tbody>
</table>

This table reports the results of Regressions 2.5 in Subsection 2.2.3 for testing the trading behaviors of nonreportable traders. Regressions are estimated with contract fixed effects. The data cover 26 contracts and for each contract there are at least 326 and at most 390 weekly observations, which sum up to 9,880 observations in total. Coefficients are precised to four significant figures and the Wald Z-statistics rounded to two digits. Significant levels: *** for 0–0.001, ** for 0.001–0.01, * for 0.01–0.05, and . for 0.05–0.1.

2.3.4 Positive Feedback Trading Pattern of Nonreportable Traders and Its Impacts

The outcomes of the test for Hypothesis 3 are reported in Table 2.3.4.

The significantly positive coefficient of commodity returns indicate that the nonreportable traders collectively show positive feedback trading behaviors indeed. When commodity prices increase, these nonreportable traders will prefer long to short interests more. Positive feedback traders as are defined in the DSSW model thereby prove to exist in commodity markets.
DATA AND EMPIRICAL RESULTS

Table 2.3.5: Predicting Commodity Returns with Nonreportable Long Proportions

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ComRet(_{t-1})</td>
<td>-0.01952</td>
<td>-14.62***</td>
<td>-0.02958</td>
<td>-17.87***</td>
</tr>
<tr>
<td>PctLong(_{t-1})</td>
<td>-0.003552</td>
<td>-1.89</td>
<td>-0.006102</td>
<td>-6.50***</td>
</tr>
<tr>
<td>SPGCRet(_{t-1})</td>
<td>-0.01857</td>
<td>-10.54***</td>
<td>-0.1310</td>
<td>-35.44***</td>
</tr>
<tr>
<td>SPRet(_{t-1})</td>
<td>0.04399</td>
<td>2.66**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EMRet(_{t-1})</td>
<td>0.01937</td>
<td>1.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>USDRet(_{t-1})</td>
<td>-0.8136</td>
<td>-21.41***</td>
<td>-0.8156</td>
<td>-18.75***</td>
</tr>
<tr>
<td>BondRet(_{t-1})</td>
<td>-0.8156</td>
<td>-18.75***</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Multiple R\(^2\) 0.0011 0.0266

This table reports the results in Subsection 2.2.3 for testing the influences of nonreportable long interests proportions on commodity returns. We additionally include returns of other financial assets in (2). The dependent variable is ComRet\(_t\). Regressions are estimated with contract fixed effects. 26 contracts and for each contract there are at least 326 and at most 390 weekly observations, which sum up to 9,881 observations in total. Coefficients are precised to four significant figures and the Wald Z-statistics rounded to two digits. Significant levels: *** for 0–0.001, ** for 0.001–0.01, * for 0.01–0.05, and . for 0.05–0.1.

Comparing with Table 2.3.2, a couple of the coefficients of the controlling variables, namely CDSSprd and TermSprd become insignificant. The coefficients of the remaining variables reverse signs except DiffVIX, hinting the sharp contrast between the strategies of CITs and the candidates for positive feedback traders. Nevertheless, the difference of VIX still predicts nonreportable long interests proportions and CIT net long positions in the same direction, which echoes with Cheng et al. (2014) in the important role of market volatility for commodity investors’ decisions.

Following this, we then report the results for the tests in the opposite direction in Table 2.3.5.
As expected, nonreportable long interests proportions prove to have negative predicting power for the following commodity returns. However, the coefficient in the simple regression, which is reported in Group (1), is significantly different from zero merely with a $p$-value of a little larger than 0.05 (0.05811, to be more precise), due to the large estimation variance. This possibly corresponds to the fact that doubtless fact that every nonreportable trader is not positive feedback trader and hence PctLong is only a proxy of positive feedback trader existence to a certain extent. But anyway, the $p$-value here is still not quite large, and we only have an opportunity of less than 6% to have made a wrong judgement here.

In fact, after we add the four returns of other assets into the regression, the significance level of the coefficient of PctLong turns larger. And comparing with Table 2.3.3, the coefficients of SPRet, EMRet, USDRet, and BondRet remain unchanged to a large extent.

By now, we have confirmed Hypotheses 3 and 4. Combining them together, we find that these exist a group of positive feedback traders in commodity markets, who tend to hold more long positions following higher commodity returns but are then taken advantage of by rational arbitrageurs and lose money. With the concomitance of passive traders who traded in the opposite way, the DSSW model forecasts a deviance of prices from their fundamental values. Surely, the presence of positive feedback traders per se is not able to destabilize the markets either, but if there were not their naïve positive feedback tradings, there would not be opportunities for rational traders to make money by utilizing them to drive prices off fundamental levels, either. We believe this is one of the contributors to the recent commodity market trends.

2.4 Conclusion

In this paper we find that the recent price patterns of commodity futures markets can be partly attributed to the concurrence of trading strategies of positive feedback traders and commodity index traders, and the theoretical model about positive feedback tradings proposed by De Long et al. (1990) fit the historical market data well. We show that the CFTC nonreportable long to total open interests ratios can be a decent proxy for the positive feedback tradings as nonreportable traders prefer more long short positions, and
such a strategy empirically proves to be unwise because it will be taken advantage of by rational arbitrageurs and significant losses will follow. As a byproduct of rational arbitrageurs’ utilization of the positive feedback strategy, commodity prices will also depart from fundamental levels. Furthermore, expected price effects of positive feedback tradings rely on the presence of passive investors who decide demands negatively related with prices to keep a constant risk-bearing capacity of the economy; and this is also supported by empirical data. We discover that commodity index investors rightly traded in this way and accordingly their changes of demands are followed by commodity return changes in the same direction.

Noticing the concurrence of CIT capital flows and commodity price trends, there has been suspect that it was the CITs who destabilized the markets. Moreover, many hope to test CITs’ impacts by using CIT dollar positions to forecast commodity returns. We argue that the market is a little more sophisticated than this. Even though CIT positions positively predict commodity prices, we are still be at risk to jump to the conclusion that CITs are the ones to blame, especially when one fails to fully understand the determination of CITs’ behaviors. In fact, we would rather call CITs a stabilizing power due to their negative feedback strategies; however, they are still the necessary condition for others to destabilize the markets. Furthermore, positive feedback traders do not either destabilize the markets by themselves. The destabilized price pattern is a collaboration of all the three classes of traders, none of whom is dispensable.

Finally, we basically limit our study within the relationship between traders’ behaviors and the levels of prices. But rightly as aforementioned, our findings also imply that an increase in the prices of commodities with larger weight in commodity indices would also be followed by larger decreases in the prices of all the markets, thus changing the correlation of prices between different commodity contracts. We do not dive into this phenomenon more deeply in this paper, but this is any way a very interesting point.
2.A Adjusting the MHW Estimation with IID Reported Positions

Let \( y \) be a column vector of \( T \) dimensions representing the true values of index investor positions, and \( x \) be another related series of \( T \) observations, or the MHW estimations. We have a linear model

\[
y = x\beta + u,
\]

(2.7)

where \( u \) is a covariance-stationary stochastic process uncorrelated with \( x \) and has zero mean and a covariance matrix \( V \), and \( \beta \) is the regression coefficients.

However, not all elements in \( y \) can be observed. Actually, what we observe is a series with lower frequency, \( y^* \), or the IID reported index investor positions. Denote the transformation matrix as \( C \), and \( y^* = Cy \). Premultiplying Eq. (2.7) by \( C \), we obtain

\[
y^* = Cx\beta + Cu = x^*\beta + u^*,
\]

(2.8)

where \( x^* = Cx \) and \( u^* = Cu \).

Eq. (2.8) can be estimated with GLS or other feasible tools.

The best linear unbiased estimator (BLUE) of \( y \) is found to be

\[
\hat{y} = x\hat{\beta} + (V^*[V^*]^{-1})\hat{u},
\]

(2.9)

where

\[
V^{**} = E[u^*(u^*)'] = E[CVC'],
\]

(2.10)

\[
V^* = E[u^*(u^*)'],
\]

(2.11)

\[
\hat{\beta} = [(x^*)'(V^{**})^{-1}x^*]^{-1}x^*(V^{**})^{-1}y^*
\]

\[
= \beta + [(x^*)'(V^{**})^{-1}x^*]^{-1}x^*(V^{**})^{-1}u^*,
\]

(2.12)

and

\[
\hat{u}^* = y^* - x^*\hat{\beta}.
\]

(2.13)
ADJUSTING THE MHW ESTIMATION WITH IID REPORTED POSITIONS

The estimator in Eq. (2.9) has two parts. The first part, $x\hat{\beta}$ can be obtained easily, hence the key point is the estimation of the second part, $(V^\ast [V^{\ast\ast}]^{-1})\hat{\mu}$, wherein $(V^\ast [V^{\ast\ast}]^{-1})$ is still unknown.

We thus assume $u$ is subject to a first-order autoregression. Denote in the $t$-th week the residual in Eq. (2.7) as $u_t$, then it follows

$$u_t = au_{t-1} + \varepsilon_t, \quad (2.14)$$

where $a$ is the autoregressive coefficient, and $\varepsilon_t$ is a white noise series with $E[\varepsilon_t \varepsilon_s] = \delta_{ts}\sigma^2$.

Then the autocovariances of $u_t$ are

$$V = E[uu'] = \begin{pmatrix} 1 & a & a^2 & \cdots & a^{T-1} \\ a & 1 & a & \cdots & a^{T-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a^{T-1} & a^{T-2} & a^{T-3} & \cdots & 1 \end{pmatrix} \frac{\sigma^2}{1 - a^2} \cong A \frac{\sigma^2}{1 - a^2}. \quad (2.15)$$

$A$ can be estimated from known information. Then we have

$$V^\ast (V^{\ast\ast})^{-1} = VC'(CVC')^{-1} = AC'(CAC')^{-1}. \quad (2.16)$$

With Eq. (2.16), we are able to obtain $\hat{y}$. 

37
DESTABILIZING OR PASSIVE?
3.1 Introduction

In empirical analyses, we often observe data at a frequency that is not high enough to capture the speed of the data generating processes behind the causal effects we are interested in. In this paper, we look into this problem by proposing a novel identification strategy based on a structural vector-autoregression (SVAR) model.

We subsequently demonstrate how our method can shed light on the relationship between commodity futures prices and commodity index trader flows. There are several reasons why this relationship is hotly debated and appropriate for our purpose. One reason is because
commodity index traders (hereafter CITs) are seen by some as “speculators,” a group of passive investors who are believed to have large-scale and overwhelmingly long-only positions in the markets (Masters, 2009). Another reason is because this relationship is typically studied using weekly data, since these are readily available, even though the impact of trades on returns is generally assumed to materialize almost instantaneously.

One way of examining the impact of CITs on commodity prices is through an indirect experiment design (see, e.g., Tang and Xiong, 2012). Another way is by examining the correlation between CIT positions and commodity prices directly (e.g., Irwin and Sanders, 2012, Singleton, 2014, Hamilton and Wu, 2015). However, there is little agreement among authors. Singleton (2014) finds the 13-week changes of CIT positions can positively predict excess returns of crude oil futures. Irwin and Sanders (2012) and Hamilton and Wu (2015) find no evidence of a positive, significant impact of index trading on commodity futures prices.

One possible explanation for these inconsistent findings is the mismatch between the frequency of data employed and the speed of the response being studied. Since the available index trading data are reported weekly, all of the existing studies that attempt to examine the influence of index trading directly can only observe the delayed interaction between index trading and futures markets after one week. Indeed, if we regress returns directly on the “instantaneous” changes of index trader positions, there will be an obvious reverse causality problem: index traders in commodity markets will adjust their positions over time depending on market conditions, even if they maintain a fixed portfolio based on the commodity index they follow. Indeed, using a proprietary database of the U.S. Commodity Futures Trading Commission (CFTC), Cheng et al. (2014) find that CIT positions are significantly affected by the market stress indicated by the VIX.

If we had daily or even intra-day data, the relationship between prices and trader flows would be easily observed. Using a daily dataset of commodity-linked notes, Henderson et al. (2014) show these notes reflect the participation of financial investors in commodity markets and contain information about future returns of referenced futures contracts; moreover, the information dissolves largely with one week. But can we still capture the effects found by Henderson et al. (2014) if we use our method and apply it to weekly data?
Indeed, in line with Singleton (2014) and Henderson et al. (2014), we also find positive instantaneous responses of crude oil excess returns to CIT inflow shocks. In addition, our approach allows us to demonstrate that the response of crude oil excess returns to CIT inflow shocks decays at an exceptionally low speed. The response remains significant in several following weeks, suggesting that the information contained in the position changes of CITs is absorbed very slowly. Hence, we complement Singleton (2014) who studies the effects of CIT inflow after a week, we have reasons to believe that his findings only show part of a larger impact as revealed by this paper.

The remainder of this paper continues as follows. In Section 3.3, we introduce a method for local identification using heteroskedasticity, before arriving at a more general identification strategy. Section 3.4 contains our baseline results as well as a description of the data employed in our empirical analysis. We describe the mechanism behind our impulse response analyses and introduce some further robustness tests. In Section 3.5, we enrich our model with additional controls based on the existing literature. We conclude in Section 3.6.

### 3.2 Identification

In principle, instantaneous impulse responses can be described by an SVAR model. Compared with reduced-form vector-autoregression (VAR) models, in an SVAR model the residual covariance already has more structure since some variables are allowed to respond to shocks instantly. However, since we often believe that data do not contain enough information for the full identification of the structure of shocks, we typically impose some additional constraints on the structure of the VAR model.¹

The most common form of constraints are triangular constraints. They require the variables in the SVAR to be ordered as a chain where shocks to any variable can only affect the variables ranked after the shocked variable directly. This is often a very strong requirement, since there does not exist such an order for many economic systems. Other, similar identification strategies face similar problems in deciding which variables, either in the

---

¹ For a review of common SVAR identification strategies, see Kilian (2013).
short or long run, are immune to which shocks.

This problem is to a large extent solved by a new class of identification strategies using sign restrictions, first proposed by Faust (2015). With sign restrictions, we constrain the direction of the responses to shocks, but not their timing. For example, a negative shock to the money supply will lead to an increase in the federal funds rate. In practice, however, this approach still suffers from two issues. First, for many economic systems, we do not have enough prior knowledge to establish convincing sign restrictions. For instance, the relationship between CIT flows and crude oil futures returns is still debated and it is therefore not possible to assert whether a certain response should be positive or negative. Second, unlike traditional short or long-run restrictions, the feasible shock structures identified by sign restrictions are not unique. As a result, in order to arrive at a single (set of) possible impulse response functions, we need to introduce additional restrictions, which are often difficult to interpret based on economic theory.

In this paper, we try to fully identify the structure of our SVAR model with a novel strategy based on heteroskedasticity. We follow Lanne et al. (2010) and Lütkepohl and Netšunajev (2014) and identify the SVAR model by assuming the data are generated by a heteroskedastic process which can be captured by Markov-switching volatility states.

The idea to identify an SVAR through heteroskedasticity is pioneered by Rigobon (2003). He assumes that if we believe the underlying economic mechanism behind an SVAR model is constant over the observed period, then we can try to split the observations into sub-periods, and the shape of impulse-response functions (IRFs) should remain the same across these sub-periods. This testable assumption provides us with new information about the structure of an SVAR model.

Inspired by Rigobon (2003), Lanne and Lütkepohl (2008) notice that heteroskedasticity can be utilized in many general cases for SVAR specification tests. While in Rigobon (2003) the observed series are simply split at a pre-determined breakpoint that is often not so easy to select, Lanne et al. (2010) show they can also be split according to a Markov Switching VAR (MSVAR) model. In this way, we can obtain a unique pattern of impulse-response functions, and do not need any further identification method like sign restrictions. Lütkepohl and Netšunajev (2014) built on their work to establish a series of
standard test procedures.

Although Lanne and Lütkepohl (2008), Lanne et al. (2010), and Lütkepohl and Netšunajev (2014) have established SVAR model specification tests, drawing inference from SVAR models often remains problematic. As an example, consider an SVAR system with two variables A and B. For this system, we can draw four impulse-response plots showing the response of A and B to structural shocks on A and B. But we still do not know which plot represents the response of A to shocks on A, and which represents its response to shocks on B - a problem we will describe in more detail in the following section. Only in a few cases can this problem be solved by prior economic knowledge, e.g., while demand shocks are transient, supply shocks have long-term effects.

This paper solves this shock-labelling issue by making heteroskedastic models part of a general-purpose SVAR identification strategy. In the context of the two-variable system, if we introduce a new variable that interacts (simultaneously) with one existing variable whilst not interacting (simultaneously) with the other, shocks can be labelled at once. Hereinafter, we call this newly introduced variable a “catalyst.” Essentially, the catalyst works because its relationship with the two variables we care about is asymmetric, and therefore informs us which variable is which. A catalyst is usually much easier to find than an instrumental variable. While an instrumental variable has to be uncorrelated with the error term, about which we have little information to justify the choice of the instrumental variable, an catalyst in our model only needs to be uncorrelated with one of the variable we are interested in.  

In this paper, we use crude oil inventory levels as a catalyst. Obviously, these interact instantly with oil prices. But since index traders are passive traders who keep proportional positions in a blanket of commodities, it is unlikely that they would adjust their positions according to real time inventory data, if they ever had them.  

---

2 Our method can be easily generalized to the cases with more variables: for a system with K variables, K − 1 additional catalysts enable us to tell the shocks apart.

3 Meanwhile, except for the Commodity Futures Trading Commission’s weekly report, there is no other source for the inventory holders to learn index traders’ positions, let alone react to them. As we shall show later, using inventory levels as the catalyst is also statistically justifiable, since they pass the model specification tests proposed by Lanne et al. (2010) and Lütkepohl and Netšunajev (2014).
3.3 Methodology

We start by closely tracing the path laid out by Lanne et al. (2010) and Lanne and Lütkepohl (2008). We argue that their approach enables us to “locally” identify the impact of shocks. Subsequently we move towards a more general identification procedure that allows us to fully identify the impact of the same shocks.

3.3.1 Local identification through heteroskedasticity

Consider a typical stable $k$-dimensional VAR($p$) model:

$$y_t = c + \sum_{i=1}^{p} A_i y_{t-i} + v_t,$$  \hspace{1cm} (3.1)

where $y_t = (y_{1t}, y_{2t}, \ldots, y_{kt})'$ are the observations of $k$ endogenous variables at time $t$, $c$ is the intercept, $\{A_i\}_{i=1}^{p}$ are the coefficient matrices, and $v_t$ is a vector white noise error term that has zero mean and a covariance matrix $\Sigma$. $v_t$ represents a mixture of different shocks and can thus be decomposed as $v_t \equiv B\varepsilon_t$, where $\varepsilon_t$ has a mean of zero and a unit covariance matrix $I_k$. Now $\varepsilon_t$ represents the orthogonalized shocks that are mutually independent. Since the matrix $B$ contains information about how the independent shocks are reconstructed in $v_t$, $\Sigma$ can also be decomposed as the inner product of $B$, $\Sigma = BB'$. Hereafter we call $B$ the “structural matrix.”

Rewriting the error term in the reduced-form VAR model with orthogonal errors, we obtain the SVAR($p$) model:

$$y_t = c + \sum_{i=1}^{p} A_i y_{t-i} + B\varepsilon_t.$$  \hspace{1cm} (3.2)

Compared with the reduced-form VAR model, the structural matrix $B$ is now the extra unknown parameter for an SVAR model. Given the constraint $\Sigma = BB'$, the estimation of $B$ is equivalent to finding a proper decomposition of the reduced-form covariance matrix $\Sigma$. 44
METHODOLOGY

However, there are infinitely many ways to decompose $\Sigma$. So we need additional information to estimate the true $B$ that generates the data.

For this additional information, we need to understand the economic process we are investigating. For example, if we are certain that some variables are immune to orthogonal instantaneous shocks, we can implement zero restrictions. An effective set of zero restrictions must have at least $k(k - 1)$ restrictions, which correspond to at least $k(k - 1)$ elements in $B$ to be set to zero a priori. Sometimes, the hierarchy among shocks is fairly obvious. For example, at times we may expect a certain shock to have an impact on the entire system instantaneously, whereas another has an impact on the entire system but for one variable instantaneously, and so forth, until the final shock that can only have an impact on a single variable instantaneously. This then corresponds to a triangular matrix $B$ that satisfies the conditions for the zero restrictions. However, in practice ranking all the shocks in a strict order is often not possible. For example, we simply do not know the order of shocks between CIT positions and oil futures prices, which is indeed the very reason we are interested in examining their relationship.

In this paper we try to identify an SVAR model with much weaker assumptions based on heteroskedasticity, which proves to contain important additional information (Rigobon, 2003). Our method allows a pair of variables in an SVAR system to interact freely and does not require any sign restrictions, thus enabling us to observe the intricate short-term dynamics among variables in the system that are previously considered unobservable, including the case in this paper, i.e., the instantaneous relationship between CIT flows and crude oil prices.

Instead of the previously mentioned rather strong assumptions, for our approach we have to make a relatively mild assumption: that the data are generated by a heteroskedastic process. Hence, although the variables we are interested in influence one another in the same way during the entire sample period, the strength of the shocks that reflect their

---

4 Obviously, if $B$ is a decomposition such that $\Sigma = BB'$, then $BQ$ is also a decomposition, where $Q$ is an orthogonal matrix, such that $\Sigma = BQQ'B'$.

5 Another popular strand of the literature tries to identify a VAR model with sign restrictions. This requires weaker assumptions. However, in our case we still do not have information to support sign restrictions, which also introduce other problems (e.g., the loss of the uniqueness of the solution; see Fry and Pagan, 2011).
dynamics varies.

We start with the simplest situation, in which volatility suddenly jumps from one level to the other at a single breakpoint time $T$, which we assume - for now - is known a priori. We then get a heteroskedastic reduced-form VAR model with a jump in volatility similar to Lanne and Lütkepohl (2008):

$$y_t = c + \sum_{i=1}^{p} A_i y_{t-i} + u_t,$$

(3.3)

where $E[u_t] = 0$, $E[u_t u_t'] = \Sigma_1 \cdot 1_{(t < T)} + \Sigma_2 \cdot 1_{(t \geq T)}$, and the remainder stays the same as in equation (3.1). Since the relationship between the variables is assumed to be stable, the coefficients of the model are constant. The identification of orthogonal shocks is still equivalent to finding the proper decomposition of the reduced-form residual covariance matrices. But the breakpoint $T$ divides the model into two states, $t < T$ and $t \geq T$, with different volatilities. We call them State 1 and State 2 respectively.

Lanne and Lütkepohl (2008) shows that the assumption that the relationship among the variables we are interested in remains stable across different volatility states leads to the local identification of the model in equation (3.1). We call it a local identification because the possible decompositions of the reduced-form residual covariance matrices are still not unique.

In line with the standard SVAR model, the IRFs have the same shapes across different states. Therefore, if $B$ is a proper decomposition of $\Sigma_1$ for State 1, then the decomposition of $\Sigma_2$ must be in the form of $BA^{1/2}$, where $\Lambda$ is a nonnegative-definite diagonal matrix, in order to keep the shapes of IRFs unchanged across states. Indeed, for a pair of symmetric matrices $\Sigma_1$ and $\Sigma_2$, if matrices $B$ and $\tilde{B}$ both satisfy $\Sigma_1 = BB' = \tilde{B}B'$ and $\Sigma_2 = B\Lambda B' = \tilde{B}\Lambda\tilde{B}'$, where $\Lambda$ is a nonnegative-definite diagonal matrix whose diagonal elements are pairwise unequal, then there must exist one and only one orthogonal matrix $Q$ so that $\tilde{B} = BQ$ and $Q = TR$, where $T$ is a column-switching elementary matrix and $R$ is a diagonal matrix whose diagonal elements are either 1 or -1.\footnote{Its proof can be found in many textbooks. Lanne and Lütkepohl (2008) also provide a proof in their appendix.} Interestingly, this
decomposition can also be generalized to more than two states, so in fact we can have as many breakpoints as we want. If there are $S$ states, and the covariance matrix of the $i$-th state $\Sigma_i$ can be decomposed as $\Sigma_i = B \Lambda_i B'$, then the sufficient condition of the local uniqueness of $B$ up to sign and permutation of columns is the existence of $j \in \{1, \ldots, S\}$ such that the diagonal elements of $\Lambda_j$ are not pairwise equal (Lanne et al., 2010).

So far, we have closely followed Lanne and Lütkepohl (2008) and Lanne et al. (2010), who mainly develop their approach for local identification as benchmarks for tests of established identification strategies. In order to fully identify the model, we still need to solve four outstanding problems. First, the matrix $R$ allows us to arbitrarily switch the signs of each column in $B$. Second, the local uniqueness of $B$ only exists for non-trivial cases where the diagonal elements of $\Lambda$ are all unequal in at least one state. Third, the matrix $T$ allows us to arbitrarily swap different columns of matrix $B$. Fourth, the model specification outlined so far (Rigobon, 2003) requires a choice of breakpoint a priori, the determination of which is often problematic itself. We solve each of these four problems in the next subsection.

3.3.2 Towards full identification

As it turns out, providing a solution to the first two problems discussed above is rather trivial, as we can directly follow earlier work. The second set of problems, however, requires somewhat more work.

The first problem, of arbitrarily switching signs, is indeed trivial. Technically, any locally identified orthogonal shock can be either positive or negative, since the signs of the columns of $B$ are arbitrary. But we can simply assume that the instantaneous response of a variable to a shock originating from itself is always positive, following Lanne and Lütkepohl (2008) and Lanne et al. (2010). For example, CIT positions always increase instantaneously when a CIT inflow shock takes place. As a result, we require the diagonal parameters of $B$ to be positive, thus effectively determining the signs of the columns of $B$.

For the second problem, the semi-identification, we can follow in the footsteps of Lanne et al. (2010) and Lütkepohl and Netšunajev (2014). They have proposed a series of Wald tests for testing the equality among the diagonal elements of $\Lambda$. 

47
The third problem, of arbitrarily swapping columns of $B$, is however less trivial. Since we have not imposed any further restrictions on $B$, we do not know to which shock any particular IRF belongs, even if we have obtained all IRFs. As a result, we have at most quasi-identification: sometimes orthogonal shocks can be labelled based on prior knowledge, but more often this is simply not possible.

We provide a solution to this problem by making a minimal concession. Note that for a $3 \times 3$ sized matrix $B$, if two elements in $B$ that are neither in the same row nor the same column are known, the whole matrix is almost surely identified. Based on this, we introduce a new “catalyst” variable that interacts with one of the existing variables but does not interact with the other variable(s). Assume the system in equation (3.3) is composed of $y_t = (y_{1t}, y_{2t}, y_{3t})'$ and $y_{3t}$ only interacts simultaneously with $y_{1t}$ and not with $y_{2t}$. Then we know the structural matrix $B$ must have the form

$$B = \begin{pmatrix}
* & * & * \\
* & * & o \\
* & o & *
\end{pmatrix},$$

where the asterisks represent free parameters to be estimated. We can see that the columns of $B$ are now identified by the two zeros.

For a three-variable system, a set of traditional triangular restrictions requires three elements in $B$ to be zero, while our method requires two. One less restriction is far more than a marginal improvement. First, even for a three-variable case, losing one restriction can make the SVAR model applicable to far more economic systems—while triangular restrictions correspond to a strictly recursive causal chain, which is uncommon in reality, the difficulty of building a system that meets our requirements is comparable to finding an instrumental variable, which is related to one variable and unrelated to the other. Second, when the dimensionality of the system increases, we only require $k - 1$ restrictions, in contrast with the previous $k(k - 1)/2$ restrictions, hence our advantage over the existing approach grows at a quadratic rate.

Our approach also compares favorably to a traditional instrumental variables approach. An instrumental variable is required to be uncorrelated with the error term, i.e., all potential
omitted variables. Since we can never be completely sure what variables are omitted, finding a good instrumental variable is far from trivial. The requirements for our “catalyst”, however, are much weaker, as the latter is only required to be uncorrelated with one of the two variables in our analysis.

As is the case with an instrumental variable, the fact that we choose a catalyst based on our prior knowledge puts us at the risk of having a misspecified model. However, a second advantage of the identification strategy we follow here is that it allows us to test for possible misspecification of the structural matrix in an intuitive and straightforward manner. Lanne et al. (2010) and Lütkepohl and Netšunajev (2014) have already shown that any restriction on the structural matrix becomes over-binding if the true model is in reality unrestricted. This allows to build a simple likelihood-ratio test: if the likelihood of the restricted model is significantly lower than the unrestricted model, we know that our restrictions are very likely to be misspecified.

The fourth issue is the choice of the breakpoint. Lanne et al. (2010) show that to circumvent this issue, we can introduce a Markovian Regime-Switching VAR (MSVAR) model with two or more different volatility states. We only introduce the simplest case with two states here, since the approach easily generalizes to more states. Let a Markov chain \( \{s_t\}_{t=0,\pm 1,\pm 2,\ldots} \), \( s_t \in \{0, 1\} \) denote the underlying state of the system at time \( t \), with a transition probability matrix

\[
Q = \begin{pmatrix} q_{00} & q_{01} \\ q_{10} & q_{11} \end{pmatrix},
\]

where \( q_{ij} \equiv \Pr[s_t = j | s_{t-1} = i] \), \( i, j \in 0, 1 \). Then the model can be specified as

\[
y_t = c + \sum_{i=1}^{p} A_i y_{t-i} + u_t, \tag{3.4}
\]

where \( \mathbb{E}[u_t] = 0 \), and \( \mathbb{E}[u_t u'_t | s_t] = \Sigma_u \). At each time \( t \), the system can be in either state with certain probabilities and the probability distribution can be estimated from the data. There is no need to identify the states ex ante any more. The identification of this model is equivalent to finding the proper decomposition of \( \Sigma_o \) and \( \Sigma_1 \) and thus can be achieved in the same way as for equation (3.3).
These properties make the MSVAR model particularly suited to represent financial market variables, since financial markets are often described with ARCH-type models and the MSVAR model with different volatility states shares similar features. The estimation of the MSVAR parameters and the identification of the structure can be achieved simultaneously through maximum likelihood estimation (Lanne et al., 2010). A brief Monte Carlo analysis of the estimators can be found in Appendix 3.A. Since we are mainly interested in the case where the true data generating process has a higher-frequency structure than the observations, we also construct an example to examine the performance of our model in this situation.

3.4 Baseline Model

To demonstrate our approach, we now turn to the relationship between CIT positions and commodity futures excess returns. We start by introducing our data, before we turn to our model setup and results.

3.4.1 Data

The empirical analysis in this paper will begin with the baseline model containing the three most important variables: West Texas Intermediate (WTI) futures excess returns, CIT position changes, and crude oil inventory shifts. For these variables, we can obtain weekly observations for the period starting January 17th, 2006 and ending December 31st, 2012. We use the weekly changes of US crude oil inventory as the catalyst variable. First, inventory is a critical factor in commodity pricing. Inventory adjustments can directly affect the balance of demand and supply and thereby the spot price. Moreover, inventory is closely correlated with the convenience yield (Gorton et al., 2008), which is an important pricing factor of futures. Second, as index traders are believed to be passive traders tracking commodity indices, they are not likely adjust their positions according to inventory information promptly. The other way round, the only data source that an ordinary

---

7 The computer program used in this paper builds on the MATLAB code published together with Lütkepohl and Netšunajev (2014) and is available upon request.
inventory holder can get access to is the weekly report published by the Commodity Futures Trading Commission (CFTC), which makes it reasonable to assume that it takes at least one week for the inventory holders to react towards CIT flows, if they do. Since a catalyst is expected to interact with only one of the other two variables we are concerned, inventory shifts constitute an ideal catalyst.

US crude oil inventory data and WTI futures price data can be easily found at the US Energy Information Administration (EIA) website and Datastream. Futures excess returns are calculated in the same way as Singleton (2014), and we use the log excess returns of the front contracts in the baseline model. Correspondingly, the changes of inventories are also represented by the weekly log difference.

The remaining variable, CIT positions, is not readily available. The “Commodity Index Trader Supplement” (CITS) report and the “Index Investment Data” (IID) report released by the CFTC are the sole reliable data sources about the commitment of CITs, but neither of them can be used as is: although the former is announced weekly, it only contains information about agricultural commodity markets; the latter does contain information for the crude oil market, but it is released monthly. For these reasons CIT positions in the WTI market can be only estimated based on the available data. The standard practice in the existing literature (Singleton, 2014, Hamilton and Wu, 2015) is based on a mapping algorithm proposed by Masters (2009). Masters (2009) considers two major commodity indices, S&P GSCI and DJ-UBSCI and finds that CIT positions in any non-agricultural commodity can be deduced from the information released in the CITS reports and the weights of the two commodity indices, after noticing each of the two indices contains a characteristic agricultural contract that is not shared by the other. Hamilton and Wu (2015) further notice that the reliance of Masters (2009) on characteristic contracts is not necessary, and the position data in every agricultural contract made public by the CITS reports can be utilized for the estimation of CITs’ engagement in a non-agricultural contract. Hereafter we refer to the estimation through the algorithm of Hamilton and Wu (2015) as the Masters-Hamilton-Wu (MHW) model.

The MHW model is not without critics, though. Irwin and Sanders (2012) argue that there

---

8 Before June, 2010, the IID report was even released on a seasonal base.
exists a considerable discrepancy between the MHW estimated values and the official monthly values in the IID reports. We confirm their findings for the WTI market, as is shown by the spread between the dashed line and the double dashed line in Figure 2.2.1. Nevertheless, despite the spread between the level values, we find the trends of the two lines are to a large extent consistent, which implies that apart from the spread, the MHW estimation is close to the true values. We therefore try to improve the MHW estimation by combining the information in both the CITS and the IID reports. We employ the interpolation algorithm devised by Chow and Lin (1971) and rectify the MHW estimation with the IID reported values while keeping the general trend. The result is shown by the solid line in Figure 2.2.1. We then use the log growth of the rectified estimation of CIT positions in the baseline model.

For simplicity, we abbreviate the WTI front futures contract excess returns, inventory growth, and CIT position growth respectively as ER1, INVENTORY, and CITPos.

3.4.2 Model Setup

To further specify our MSVAR model, we need to fix two parameters: the lag order and the number of states. Taking into account that the market variables we are studying are expected to respond to new information very quickly, we use one lag in the baseline model.⁹

<table>
<thead>
<tr>
<th>Number of States</th>
<th>AIC</th>
<th>SC</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>-6577.22</td>
<td>-6483.82</td>
</tr>
<tr>
<td>3</td>
<td>-6608.75</td>
<td>-6488.11</td>
</tr>
<tr>
<td>4</td>
<td>-6608.31</td>
<td>-6452.65</td>
</tr>
</tbody>
</table>

This table shows the AIC and Schwartz Criterion (SC) values for the SVAR model containing WTI front futures contract excess returns, inventory growth, and CIT position growth with two, three, and four different volatility states. The three variables are respectively abbreviated as ER1, INVENTORY, and CITPos.

⁹Indeed, increasing the lag order results in substantial increases in AIC values.
Figure 3.4.1: CIT Net Long Positions Estimations

Dashed, double dashed, and solid lines represent estimations of CIT net long positions in billion US$ estimated with the MHW algorithm, reported by the CFTC IID (Index Investment Data) report, and corrected values.
Next, we have to determine the number of states. As can be seen in Table 3.4.1, both AIC and Schwartz Criterion (SC) values point to three states. The existence of three states becomes more intuitive once we have a look at the estimated probability distribution for different states, as shown in Figures 3.4.2a–3.4.2c. With three different volatility states (Figure 3.4.2a), we obtain a clear division between the three sample periods. Judging from the residual covariances reported in Table 3.4.2, the volatilities are roughly the lowest in State 1 and the highest in State 3. This pattern is consistent with our impression about the recent crude oil market, as the most volatile State 3 roughly corresponds to the spectacular bust after mid-2008, and the market also saw large volatilities in the first half year of 2008 and during the rally after 2010, which is included in the distribution of State 2. For comparison, consider the results with two states, shown in Figure 3.4.2c. The state probability distribution lacks an obvious pattern, and neither of the two states remains stable for a long time.\footnote{In fact, the state distribution shown in Figure 3.4.2c appears only to be generated by splitting State 2 in the 3-state case into two states.}
Figure 3.4.2: State Probabilities of Baseline Model with Different Number of States

(a) 3 states (preferred specification)

(b) 2 states

(c) 4 states

This figure shows the estimated probability of the MSVAR system with three volatility states, which contains WTI front futures contract excess returns, inventory growth, and CIT position growth, to be respectively in States 1, 2, 3 and 4.
This table shows the estimated covariance matrices and their eigen values of the non-orthogonal residuals in the reduced-form VAR model containing WTI front futures contract excess returns, inventory growth, and CIT position growth with three states. The three variables are respectively abbreviated as ER1, Inventory, and CITPos.

Since the identification of the baseline model relies on the assumption that Inventory and CITPos do not respond to the shocks to each other instantly, we test whether the model is misspecified. The unrestricted model has 33 free parameters and its log likelihood is 3336.09. When we restrict the instantaneous responses between Inventory and CITPos to zero, the number of free parameters drops to 31 and the log likelihood drops to 3335.38. With a p-value of 0.7011, the likelihood ratio test cannot reject the validity of our restrictions in the baseline model.

Before we start analyzing the baseline model, we also have to test whether the diagonal elements in the estimated $\Lambda$’s are different. The estimated values of the $\Lambda$’s of both the unrestricted model and the model with restrictions on the instantaneous responses between Inventory and CITPos are reported in Table 3.4.3, where $\lambda_{ij}$ denotes the $j$-th diagonal element in the matrix $\Lambda_i$ of State $i$. We see that the estimations of the $\lambda$’s do not change much after we impose the restrictions. For each state, no pair of $\lambda$’s in the restricted model appears to be equal. The Wald tests in Table 3.4.4 provide the details: the p-value of
**Table 3.4.3:** Estimated $\Lambda$’s in Baseline Model with 3 States

<table>
<thead>
<tr>
<th></th>
<th>Unrestricted Model</th>
<th></th>
<th>Restricted Model</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimation</td>
<td>SD</td>
<td>Estimation</td>
<td>SD</td>
</tr>
<tr>
<td>$\lambda_{21}$</td>
<td>0.840 316</td>
<td>0.200 677</td>
<td>0.851 057</td>
<td>0.275 694</td>
</tr>
<tr>
<td>$\lambda_{22}$</td>
<td>1.529 337</td>
<td>0.829 916</td>
<td>1.512 752</td>
<td>0.317 398</td>
</tr>
<tr>
<td>$\lambda_{23}$</td>
<td>7.830 264</td>
<td>3.563 535</td>
<td>7.822 529</td>
<td>1.616 984</td>
</tr>
<tr>
<td>$\lambda_{31}$</td>
<td>6.974 877</td>
<td>2.915 683</td>
<td>6.930 944</td>
<td>3.758 024</td>
</tr>
<tr>
<td>$\lambda_{32}$</td>
<td>1.300 846</td>
<td>0.747 950</td>
<td>1.310 487</td>
<td>0.515 381</td>
</tr>
<tr>
<td>$\lambda_{33}$</td>
<td>13.348 555</td>
<td>6.852 705</td>
<td>13.277 863</td>
<td>6.675 058</td>
</tr>
</tbody>
</table>

This table shows the estimated diagonal elements of $\Lambda_i$ of State $i$ of the unrestricted and restricted baseline model, where $\lambda_{ij}$ denotes the $j$-th diagonal element in $\Lambda_i$.

**Table 3.4.4:** Tests of the Equality among $\lambda$’s

<table>
<thead>
<tr>
<th>Null Hypothesis</th>
<th>Wald Statistic</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_{21} = \lambda_{22}$</td>
<td>1.574</td>
<td>0.057 75</td>
</tr>
<tr>
<td>$\lambda_{21} = \lambda_{23}$</td>
<td>4.250</td>
<td>0.000 01</td>
</tr>
<tr>
<td>$\lambda_{22} = \lambda_{23}$</td>
<td>3.829</td>
<td>0.000 06</td>
</tr>
<tr>
<td>$\lambda_{31} = \lambda_{32}$</td>
<td>1.482</td>
<td>0.069 21</td>
</tr>
<tr>
<td>$\lambda_{31} = \lambda_{33}$</td>
<td>0.829</td>
<td>0.203 68</td>
</tr>
<tr>
<td>$\lambda_{31} = \lambda_{32}$</td>
<td>1.788</td>
<td>0.036 93</td>
</tr>
</tbody>
</table>

This table shows the Wald statistics and respective p-values of the tests of the pairwise equality among the diagonal elements of $\Lambda_2$ and $\Lambda_3$ of State 2 and 3 of the restricted baseline model. $\lambda_{ij}$ denotes the $j$-th diagonal element in $\Lambda_i$.  

57
the equality test for the first diagonal elements of $\Lambda$, at State 2 slightly exceeds 0.05, while the tests for the other two pairs at this state are highly significant. However, the statistics for the inequality of the $\lambda$’s at State 3 cannot provide us with strong enough evidence against the null hypothesis, as $\lambda_{31}$ is not significantly different from $\lambda_{33}$, mainly driven by the large standard deviation of the estimations. Nonetheless, since a $\Lambda$ with all different diagonal elements is only required in one among all the states for the model to be identified, we can proceed with our restricted baseline model based on the inequalities of the $\lambda$’s for State 2.

3.4.3 A shock to CIT inflows and other impulse responses

Once we obtain the proper structural matrix through the decomposition of the covariance matrices at different states, we can identify three orthogonal shocks from the non-orthogonal reduced-form residuals. The impulse-response functions for our restricted baseline model are shown in Figure 3.4.3.

We observe the instantaneous responses and the responses in the following nine weeks. Each column in Figure 3.4.3 represents a specific orthogonal shock as identified by the estimation of the structural matrix, namely the shock of WTI futures excess returns, inventory shifts, and CIT inflows. Each row represents the instantaneous and continued responses of a certain variable, ER1, Inventory, and CITPos ordered from top to bottom. The solid line denotes the true IRF and the dashed lines mark off the 68% confidence intervals. Comparing the magnitude of each of the responses is not straightforward, as these vary across states.

We first consider the three IRFs on the diagonal of Figure 3.4.3. Recall that we assume that all orthogonal shocks are positive shocks, such that all the variables respond positively to the shocks to themselves at the beginning. Therefore the three IRFs on the diagonal all start with positive values. The impact of an excess return shock is transient and only significant within the same week, which is typical for a financial asset. In contrast, the impact of both the inventory and the CIT inflow shock remains significant for several weeks: interestingly, a positive shock in the inflow of index investors can persist for weeks. Since CITs are mainly institutional investors, this is probably a result of the fact that the funding of institutional
investors is mostly affected by slow-moving factors such as the business cycle. The relatively slow decay of the inventory shock is less surprising, and likely related to its adjustment costs (Jin, 2013).

The assumption that inventory and CIT do not react instantaneously to each other, and thus the zero restrictions on the respective positions in the structural matrix is the key condition for the identification. It enables us to impose the necessary zero restrictions on two positions in the structural matrix that stand for the instantaneous interaction between Inventory and CITPos. This assumption is reflected in the responses of Inventory to CIT inflow shocks and CITPos to inventory shocks, which are both zero in the starting week. Although by no means the consequence of the (initial) zero restriction, the response of Inventory to CIT inflow shocks remains insignificant and close to zero for the whole observed period, indicating that inventory holders ignore the CITs in the futures market almost completely. This appears to reflect the insensitivity of participants in the physical market to financial investors’ behavior. Finally, the IRF of CITPos to inventory shocks is not significant either, which is in line with the view of CITs as passive investors.

Our SVAR model can also reveal some interesting aspects about the relationship between inventory and crude oil futures excess returns. While it is well known that the inventory level of commodities is an important determinant of convenience yields (Pindyck, 1994), the precise mechanism is typically rather opaque. Gorton et al. (2008) predict a negative relationship between futures excess returns and commodity inventories. In line with their theory, the subplot in the first row, the second column in Figure 3.4.3 finds a negative response of ER1 to positive inventory shocks, which is only significant in the first week. In what is perhaps the consequence of the release of inventories to stabilize prices, responses of Inventory to excess return shocks turn out to be negative.

Now that we have discussed how each of our variables of interest behaves when subject to a shock, it is time to focus on our main point: what happens to crude oil futures excess returns after a shock to CIT inflows? In answering this question, we follow in the footsteps of, e.g., Singleton (2014), Hamilton and Wu (2015) and Irwin and Sanders (2012). Although both focus on long-term CIT flows, they find contradicting evidence as to how
**Figure 3.4.3:** Impulse-Response Functions of the Restricted Baseline Model

This figure shows the impulse-response functions of the restricted baseline model containing WTI front futures contract excess returns, inventory growth, and CIT position growth with 3 states. Each column represents a specific orthogonal shock as is identified by the estimation of the structural matrix, and each row represents the instantaneous and continued responses of a certain variable within 9 weeks. The solid line denotes the true IRF and the dashed lines mark off the 68% confidence intervals.
these flows affect crude oil: Irwin and Sanders (2012) and Hamilton and Wu (2015) find no effect, whereas Singleton (2014) does.

Our results are in line with Singleton (2014), and rather strong: as Figure 3.4.3 shows, the response of ER1 - excess returns - to CIT inflow shocks is positive and persists for many weeks following the initial shock. Apparently, the information contained in CIT flows is absorbed by the market very slowly, even compared with the findings of Henderson et al. (2014). Summing up, our analysis so far shows that there is indeed an instantaneous response of ER1 to CIT inflow shocks, and capturing it does not take away from a longer-run impact.

### 3.4.4 Additional Tests

Recall that we impose two restrictions on the structural matrix in order to fully identify the baseline model. Although we cannot reject these restrictions statistically, one may still wonder how these restrictions affect the results. We can now compare the IRFs from the restricted baseline model with those based on the unrestricted model. The unrestricted IRFs are shown in Figure 3.4.4.

Note that the order of the $\lambda$’s for the restricted and the unrestricted models is the same. Hence we can label the three columns of Figure 3.4.4 in the identical order as for the restricted case. Comparing Figure 3.4.4 with Figure 3.4.3, we find most of the IRFs remain similar. Of course for the response of ER1 to inventory shocks, we have effectively introduced a new channel in our fully identified model, via the CIT positions. Indeed, this is where the IRFs from both models are markedly different. An inventory shock now also has a positive instantaneous impact on CIT positions, reflecting investors’ reaction to physical traders’ activities. Although the instantaneous impact of a CIT inflow shock on INVENTORY is only marginally significantly positive at the 68% level, the overall picture from both models is therefore different.

Given that the effects of CIT flows can persist for a long time, a logical next step is to investigate how our model fares once we extend our horizon. To do so, we increase the
**Figure 3.4.4:** Impulse-Response Functions of the Unrestricted Baseline Model

This figure shows the impulse-response functions of the unrestricted baseline model containing WTI front futures contract excess returns, inventory growth, and CIT position growth with three states. Each column represents a specific orthogonal shock as is semi-identified by the estimation of the structural matrix, and each row represents the instantaneous and continued responses of a certain variable within nine weeks. The labels of the shocks is attached through the comparison with the restricted model. The solid line denotes the true IRF and the dashed lines mark off the 68% confidence intervals.
**Figure 3.4.5:** Impulse-Response Functions of the Restricted Baseline Model with 13 Lags

This figure shows the impulse-response functions of the restricted baseline model containing WTI front futures contract excess returns, inventory growth, and CIT position growth with three states and 13 lags. Each column represents a specific orthogonal shock as is identified by the estimation of the structural matrix, and each row represents the instantaneous and continued responses of a certain variable within 52 weeks. The solid line denotes the true IRF and the dashed lines mark off the 68% confidence intervals.

Once we review the IRFs over very long periods, it obviously becomes less easy to find a consistent sign, since we can be less sure to assess the impact of a single shock. Indeed, the response of ER1 to inventory shocks is no longer purely positive (or negative, for that matter) over such a long horizon. Nevertheless, the response of INVENTORY to shocks in the futures market is still largely negative, even we extend the lag number to 13. In addition, the impact of CIT inflow shocks on ER1 is still persistently positive, and the impact of
excess return shocks on CIT Position changes also remains positive, though the former fades much more quickly than the latter.

3.5 Models with Extra Controls

It is possible that the relationships identified in the baseline model vary in different financial environments, especially since our sample contains the 2008 crisis. We therefore estimate some models in which the VIX and the US term spread are included as extra control variables in our model.

We choose these two variables because they are likely to have a direct influence on both crude oil futures excess returns and CIT positions. The yield curve of the US treasury bond is widely used as a leading indicator of economic growth (Dotsey, 1998), which may affect both oil prices and CITs funding. The VIX is often used as an indicator of investors’ fear (e.g. Whaley, 2000), and has proven to helpful in predicting returns of many assets, including crude oil futures (see Chevallier and Benoit, 2013). Finally, CITs are also known to actively adjust their positions according to the risks represented by VIX (Cheng et al., 2014). We use the log difference of the VIX (VIX) and the difference of the 10Y-3M term spread of US treasury bonds (TermSprd).

In the models in this section, we therefore regard VIX and TermSpr as exogenous variables. Let \( y_t \) be the endogenous vector \((E_{1t}', \text{Inventory}_{1t}', \text{CITPos}_{1t}')\)', and \( x_t \) be the exogenous vector \((\text{VIX}_{1t}', \text{TermSprd}_{1t}')\) or \((\text{VIX}_{1t}', \text{TermSprd}_{1t}')\)', then the reduced-form model is expressed as:

\[
y_t = c + \sum_{i=1}^{p} A_i y_{t-i} + \sum_{i=1}^{p} \Phi_i x_{t-i} + u_t, \quad (3.5)
\]

where \( u_t \) still captures the non-orthogonal residuals with Markov-Switching volatility regimes. This reduced-form model is identified in the same way as the baseline model by restricting the instantaneous mutual responses between inventory and CIT positions. To be consistent with the previous results, we again employ three volatility states and one lag. In order to save space, we only report the IRFs of the model with both the VIX and
MODELS WITH EXTRA CONTROLS

TERMSPRD included as controls in Figure 3.5.1.\textsuperscript{11} Despite the difference in scales, there is no difference in the significance and signs between this figure and Figure 3.4.3: the impulse-responses with extra controls change little compared with the baseline model, again proving that our baseline model is rather robust.

Table 3.5.1: Coefficients and Wald Statistics of Exogeneous Variables

<table>
<thead>
<tr>
<th>Model</th>
<th>Endogeneous Var.</th>
<th>TERMSPRD\textsubscript{t-1}</th>
<th>VIX\textsubscript{t-1}</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>ER\textsubscript{t}</td>
<td>0.020154</td>
<td>1.384</td>
</tr>
<tr>
<td></td>
<td>INVENTORY\textsubscript{t}</td>
<td>0.000509</td>
<td>0.489</td>
</tr>
<tr>
<td></td>
<td>CITPOS\textsubscript{t}</td>
<td>0.004385</td>
<td>1.006</td>
</tr>
<tr>
<td>(2)</td>
<td>ER\textsubscript{t}</td>
<td>0.019229</td>
<td>1.375</td>
</tr>
<tr>
<td></td>
<td>INVENTORY\textsubscript{t}</td>
<td>0.000601</td>
<td>0.616</td>
</tr>
<tr>
<td></td>
<td>CITPOS\textsubscript{t}</td>
<td>0.003978</td>
<td>0.952</td>
</tr>
</tbody>
</table>

This table shows the coefficients of the exogenous variables and respective Wald statistics.

Table 3.5.1 shows the coefficients of the exogenous variables.\textsuperscript{12} The TERMSPRD is indeed a good predictor for our endogenous variables. Since a higher term spread signals higher growth, its positive predictive power for the crude oil futures excess returns is to be expected. The VIX negatively predicts CITPOS\textsubscript{t}, in line with the arguments in Cheng et al. (2014) that index traders follow the convective flow and chase the less risky assets. Although the coefficients for the VIX in the prediction of ER\textsubscript{t} are positive, contradicting findings in the literature (e.g. Saria et al., 2011), they are not very significant.\textsuperscript{13}

Summing up, our results appear to robust to both changes in the lag structure and the

\textsuperscript{11}The full results are available upon request.

\textsuperscript{12}We also show Wald statistics, which generally do not have very significant p-values except the prediction of ER\textsubscript{t} with TERMSPRD. However, due to the highly nonlinear nature of the MSVAR model, Wald tests are in fact not that reliable.

\textsuperscript{13}Further evidence in this direction can be gathered from the insignificant Wald statistics for the relationship between the VIX and INVENTORY.
Figure 3.5.1: Impulse-Response Functions of the Model with VIX and Term Spread Controlled

This figure shows the impulse-response functions of the model containing WTI futures contract excess returns, inventory growth, and CIT position growth as endogenous variables, and VIX and term spread of US treasury bonds as exogenous variables with three states and one lag. Each column represents a specific orthogonal shock as is identified by the estimation of the structural matrix, and each row represents the instantaneous and continued responses of a certain variable within nine weeks. The solid line denotes the true IRF and the dashed lines mark off the 68% confidence intervals.
CONCLUSION

inclusion of additional controls.

3.6 Conclusion

In this paper we design a novel identification strategy for structural vector autoregression models, which enables us to capture (some of) the short-run dynamics between variables that we observe at a low(er) frequency. We show how our method can shed light on the hotly-debated issue on the relationship between commodity prices and commodity index trader flows with weekly data.

Our SVAR identification strategy relies on an underlying Markov-Switching VAR model with two or more volatility states and the assumption that the impulse-response functions have the same shape across different states. If we are interested in the relationship between two variables, the mutual instantaneous impulse-responses between them can be identified with the inclusion of a single additional “catalyst” variable, which only interacts instantaneously with one of two variables and not with the other. Likewise, we can identify a system with more variables, with the inclusion of other catalyst variables.

Compared with traditional SVAR identification strategies such as triangular restrictions or sign restrictions, the restriction conditions of our strategy are much easier to satisfy, and the impulse-responses resulting from estimating the model are uniquely identified. In fact, the impulse responses from our approach can be regarded as the cumulative responses generated by the real data-generating process at a higher frequency. Hence, our methodology can help solve many identification problems for situations where data are available at a relatively low frequency, for example for financial markets, but also for many macroeconomic analyses.

As an illustration, we demonstrate how our methodology can help us understand the relationship between crude oil futures prices and commodity index trader flows. Using the estimated index trader positions and other weekly data, we find that an index trader inflow shock can cause a positive response of the WTI futures excess returns instantly, and the response remains significant in several following weeks. Commodity index traders
themselves also positively respond to shocks of crude oil futures excess returns instantly, though the responses will not last for long and are only significant in the immediate week. These findings are robust even after we add some extra control variables that are likely to affect oil prices and index traders into our model. These findings nicely complement the work done with higher frequency data by Singleton (2014).

3.A Monte Carlo Analyses

Concerning the estimation of the model, we present a brief Monte Carlo analysis here. Consider the model in equation 3.4. Following Droumaguet (2012), we let

\[ s_t = \{0, 1\}, \]

\[ p = 1, \]

\[ c = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \]

\[ A_1 = \begin{pmatrix} 0.6 & 0 \\ 0 & 0.9 \end{pmatrix}, \]

\[ \Sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 2.236 \end{pmatrix}, \]

\[ \Sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & 2.236 \end{pmatrix}, \]

and

\[ Q = \begin{pmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{pmatrix}. \]

Obviously, with these parameters, the correct decomposition of \( \Sigma_0 \) and \( \Sigma_1 \) must be

\[ \Sigma_0 = BB' = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{5} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{5} \end{pmatrix} \approx \begin{pmatrix} 1 & 0 \\ 0 & 2.236 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 2.236 \end{pmatrix}, \]
and
\[
\Sigma_1 = BAB' = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{5} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{5} \end{pmatrix}.
\]

The initial values are \(y_0 = (0, 0)'\) and \(s_0 = 0\). The first 300 simulated observations are dropped to warrant the randomness, and the following 300 observations are kept for the Monte Carlo experiments. We repeat 500 simulations in total, and the results are reported in Table 3.A.1.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Component/Element</th>
<th>True Value</th>
<th>Mean</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>(c)</td>
<td>(c_1)</td>
<td>0.000</td>
<td>0.000295</td>
<td>0.003877</td>
</tr>
<tr>
<td></td>
<td>(c_2)</td>
<td>0.000</td>
<td>-0.000102</td>
<td>0.054215</td>
</tr>
<tr>
<td></td>
<td>(a_{11})</td>
<td>0.600</td>
<td>0.586254</td>
<td>0.002059</td>
</tr>
<tr>
<td></td>
<td>(a_{12})</td>
<td>0.000</td>
<td>0.010764</td>
<td>0.026284</td>
</tr>
<tr>
<td></td>
<td>(a_{21})</td>
<td>0.000</td>
<td>0.000374</td>
<td>0.000055</td>
</tr>
<tr>
<td></td>
<td>(a_{22})</td>
<td>0.900</td>
<td>0.888535</td>
<td>0.000782</td>
</tr>
<tr>
<td></td>
<td>(b_{11})</td>
<td>1.000</td>
<td>0.990843</td>
<td>0.013173</td>
</tr>
<tr>
<td></td>
<td>(b_{12})</td>
<td>0.000</td>
<td>0.004290</td>
<td>0.008727</td>
</tr>
<tr>
<td></td>
<td>(b_{21})</td>
<td>0.000</td>
<td>-0.012335</td>
<td>0.289684</td>
</tr>
<tr>
<td></td>
<td>(b_{22})</td>
<td>2.236</td>
<td>2.082039</td>
<td>0.148946</td>
</tr>
<tr>
<td>(A_1)</td>
<td>(\lambda_{11})</td>
<td>1.000</td>
<td>0.996814</td>
<td>0.149508</td>
</tr>
<tr>
<td></td>
<td>(\lambda_{22})</td>
<td>5.000</td>
<td>6.190787</td>
<td>5.000563</td>
</tr>
<tr>
<td>(B)</td>
<td>(q_{00})</td>
<td>0.800</td>
<td>0.750209</td>
<td>0.025488</td>
</tr>
<tr>
<td></td>
<td>(q_{01})</td>
<td>0.200</td>
<td>0.249791</td>
<td>0.025488</td>
</tr>
<tr>
<td></td>
<td>(q_{10})</td>
<td>0.200</td>
<td>0.241747</td>
<td>0.027461</td>
</tr>
<tr>
<td></td>
<td>(q_{11})</td>
<td>0.800</td>
<td>0.758253</td>
<td>0.027461</td>
</tr>
</tbody>
</table>

This table shows the mean and the variance of parameter estimations of the model in equation 3.4 in the Monte Carlo experiments. \(c_1\) and \(c_2\) are respectively the first and second components of \(c\), \(q_{00}\) to \(q_{11}\) follow the same notation in equation 3.3.2, and \(a_{ij}\), \(b_{ij}\) and \(\lambda_{ij}\) in rest rows denote the element of \(A_1\), \(B\), and \(\Lambda\) in the \(i\)-th row and \(j\)-th coloumn. Since \(\Lambda\) is required to be diagonal in the estimation, we only report the diagonal elements.

On top of the estimation of parameters, we are especially interested in the performance of the model when the true data generating process is at a higher frequency than the observations. We present an example here to show how our model performs in this case.
In this analysis, the data generating process generally resembles the one in the previous process, whereas the shocks take place every other time, and observations are taken in between. Assume

\[ y_t = c + A_1 y_{t-1} + u_t, \]

where \( u_t = 0 \) for \( t = 0, 2, 4, \ldots \) and \( u_t | s_t \sim (0, \Sigma_{s_t}) \) for \( t = 1, 3, 5, \ldots \), and \( \tau = (t - 1)/2 \). Similar with the previous example, \( s_\tau, \tau = 0, 1, 2, \ldots \) is a Markov chain whose values are either 0 or 1, and the transition matrix is the same:

\[ Q = \begin{pmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{pmatrix}. \]

And the parameters \( c, A_1, \Sigma_0, \) and \( \Sigma_1 \) take the same values as in the previous analysis. However, here we assume that the data are censored at \( t = 1, 3, 5, \ldots \) and we only observe \( y_t \) at \( t = 0, 2, 4, \ldots \). The model established on the observed data should therefore be

\[ y_{2\tau} = c + A_1 y_{2\tau-1} + u_{2\tau} \]
\[ = c + A_1 (c + A_1 y_{2\tau-2} + u_{2\tau-1}) + u_{2\tau} \]
\[ = c + A_1 c + A_1^2 y_{2\tau-2} + A_1 u_{2\tau-1}. \]

for \( \tau = 0, 1, 2, \ldots \). An observer who is not informed with the data generating process will thus believe the data are generated by an equivalent process

\[ y_{2\tau} = \tilde{c} + \tilde{A}_1 y_{2\tau-2} + \tilde{u}_{2\tau}, \]

\[ (3.6) \]

where \( \tilde{c} = c + A_1 c, \tilde{A}_1 = A_1^2, \) and \( \tilde{u}_{2\tau} = A_1 u_{2\tau-1} \sim (0, \tilde{\Sigma}_{s_t} \equiv A_1 \Sigma_{s_t} A_1^T). \)

We then randomly generate 1,000 observations and then drop every other one, thus obtaining a 500-observation sample. According to equation 3.6, we expect to find the
parameters in the equivalent model to be:

\[
\tilde{c} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},
\]

\[
\tilde{\lambda}_i = \begin{pmatrix} 0.36 & 0 \\ 0 & 0.81 \end{pmatrix},
\]

\[
\tilde{\Sigma}_0 = \begin{pmatrix} 0.36 & 0 \\ 0 & 4.05 \end{pmatrix},
\]

\[
\tilde{\Sigma}_1 = \begin{pmatrix} 0.36 & 0 \\ 0 & 20.25 \end{pmatrix},
\]

and for the observer transition matrix there is also

\[
\tilde{Q} = \begin{pmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{pmatrix}.
\]

As a result, the decomposition of the equivalent covariance matrices should be

\[
\tilde{\Sigma}_0 = \tilde{B}\tilde{B}' = \begin{pmatrix} 0.6 & 0 \\ 0 & \sqrt{4.05} \end{pmatrix} \begin{pmatrix} 0.6 & 0 \\ 0 & \sqrt{4.05} \end{pmatrix} \approx \begin{pmatrix} 0.6 & 0 \\ 0 & 2.012 \end{pmatrix} \begin{pmatrix} 0.6 & 0 \\ 0 & 2.012 \end{pmatrix},
\]

and

\[
\tilde{\Sigma}_1 = \tilde{B}\tilde{\Lambda}\tilde{B}' = \begin{pmatrix} 0.6 & 0 \\ 0 & \sqrt{4.05} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 20 \end{pmatrix} \begin{pmatrix} 0.6 & 0 \\ 0 & \sqrt{4.05} \end{pmatrix}.
\]

We run 500 simulations in total again, and the results are reported in Table 3.A.2. We see the accuracy of the estimations are in any way comparable to the analysis where the SVAR model is consistent with the true data generating process.
This table shows the mean and the variance of parameter estimations of the censored model in the Monte Carlo experiments. \( \tilde{\mathcal{C}} \) and \( \tilde{\mathcal{C}} \) are respectively the first and second components of \( \tilde{\mathcal{C}} \). \( \tilde{a}_{ij} \), \( \tilde{b}_{ij} \), and \( \tilde{\lambda}_{ij} \) denote the element of \( \tilde{\mathcal{A}}_i \), \( \tilde{\mathcal{B}}_i \), and \( \tilde{\Lambda}_i \) in the \( i \)-th row and \( j \)-th column. Since \( \tilde{\Lambda} \) is required to be diagonal in the estimation, we only report the diagonal elements.
Economic entities are often interconnected and these interconnections can have significant economic consequences. Particularly, the interconnections among financial entities, whether they are asset markets, firms, financial institutions, or economies, have drawn much attention since the 2008 financial crisis. A growing number of methodologies have been proposed to empirically measure financial connectedness, e.g., Billio et al.
(2012), Diebold and Yilmaz (2014), Giraitis et al. (2016), and Barigozzi and Brownlees (2016). Following this strand of literature, Scidá (2016) finds that financial connectedness can not only be depicted with network models but also with spatial econometric models, and both approaches are unified by an underlying structural vector auto-regression (structural VAR, SVAR) model. In this paper, I call her model a “SVAR-spatial-network” model, or a “tripod” model, as it is an integration of these three econometric models.

However, most of existing literature only tries to depict a static state of financial connectedness and seldom attempts to model the dynamic evolution of the connectedness of a financial system. As an exception, Giraitis et al. (2016) indeed considers a dynamic model, but unlike most other literature, they assume the connections among the entities are known a priori, which is not realistic for many applications.

As the first contribution of this paper, I propose a novel and systematic way to empirically depict and analyze the connectedness among economic or financial entities in a given regime and its transitions across different regimes. Thus, the dynamics of connectedness are captured in a regime-switching model. With this model, it is possible to figure out how the interconnections in an economic or financial system are determined by some threshold variables that we are interested in, and make comparisons among the interconnections in different regimes.

The model in this paper is rooted in the tripod model of Scidá (2016). Since it contains a regime-switching mechanism that is described by a threshold function, my model may be called a SVAR-spatial-network model with thresholds, or a “threshold tripod” model. As a secondary contribution of this paper, the threshold tripod model improves the original model of Scidá (2016) in various ways, on top of the dynamic regime-switching mechanism. First, I improve the identification strategy used in Scidá (2016) for an SVAR model, which is never “an easy problem to address” in any SVAR literature. In this paper, I solve the identification problem in the SVAR part, which is fundamental to the whole tripod model, with a novel strategy that exploits the normal-mixture structure of the residuals. Second, I show that the tripod model actually provides a benchmark to test a broad class of interesting hypotheses on the spatial weight matrices. Third, the networks implied by the tripod model in this paper are usually complete and often have negative edges, which make them difficult to analyze with traditional techniques. I show that such
INTRODUCTION

networks can still be analyzed by some novel approaches devised recently.

The threshold tripod model is particularly useful for applications where it is as important to study transitions of connectedness as the connectedness itself. A typical example is the empirical modeling of financial contagion: a large body of literature agrees that a definition of contagion should emphasize the change of entity interrelationship between crisis and non-crisis periods, rather than the relationship in each period. As an illustration of my model, I show how it sheds some new light on the topic of financial contagions in the EU, which has become a hotly discussed topic since the start of the Eurozone crisis in late 2009. While many existing studies rely on manual choice of start points and end points of crises, I pick a European credit default swap (CDS) index as the threshold variable and let the threshold variable determine two distinctive regimes where the five selected countries, namely Germany, Italy, Spain, Ireland, and Greece, have different interconnections. Under each regime, the mutual connectedness among the five countries is displayed clearly. Furthermore, I construct statistical tests of two hypotheses. The first test confirms the existence of contagion among the countries; and the second test demonstrates that the contagion is not solely spread through the cross-border interbank channel. The existence of contagion is reconfirmed by global network measures. Moreover, centrality analyses and clustering analyses reveal clearly the distinctive roles played by the selected countries in different regimes.

Besides the contagion among markets in different countries, my model may be applied to profiling the transitions of many other instances of financial interconnections as well. For instance, many studies focus on the interconnections among financial institutions or firms (e.g. Diebold and Yilmaz, 2014, Barigozzi and Brownlees, 2016). With my model, we may clearly see how these interconnections among institutions or firms change in different situations. As another interesting application of my model, we may also try to see how the spillover effect across different asset markets shifts. For example, the volatility transmission among different commodity markets is determined by various macroeconomic and financial factors and known to shift significantly in different environment (Nazlioglu et al.,

---

1 See, e.g., Forbes and Rigobon (2002), Dungey et al. (2005), Chiang et al. (2007), Caporin et al. (2013). For an in-depth discussion, see Section 4.6.1.

2 CDS is a swap agreement, with which the seller insures the buyer against certain reference debts.
It might be interesting to how it evolves over time with the model in this paper. Applications of the threshold tripod model are certainly not limited to the area of finance. In various fields of economics, we encounter issues concerning interconnections of entities. For example, there is a large number of studies of the connections among firms or regions through which knowledge spillover takes place (e.g., Lee, 2006, Maliranta et al., 2009). It is possible to document the transitions of such connections as well.

A not-so-technical overview of the threshold tripod model is presented in the following section. The rest of this paper is organized as follows. Section 4.3 reviews the basics of an SVAR model and shows how the T-SVAR model used in the following sections is identified. Section 4.4 introduces the restricted SAR model and the way to construct tests based on it. Section 4.5 presents the network analyses. Section 4.6 reports the empirical results of the application of the model in the Eurozone markets. Section 4.7 concludes this paper. And in the end, the appendices collect the necessary mathematical proofs.

4.2 A Not-So-Technical Overview of the Threshold Tripod Model

Before I start to introduce the model formally, I present an overview of it in a less technical manner here.

As its name suggests, the SVAR-spatial-network model is composed of three steps. First, the entities of interest form an SVAR model. As we identify the SVAR, the responses of the entities to the structural shocks with clear economic interpretations are revealed. This provides a baseline for the following steps. Second, by adding some extra restrictions to the SVAR model, we get a spatial auto-regression (SAR) model, which shows clearly the interconnections implied by the SVAR structure. Third, we use certain network analytical techniques to further analyze the connectedness revealed by the SAR model. Since the tripod model in this paper has thresholds to model the dynamics of the interconnections, I add threshold functions to both the SVAR and SAR steps, and thereby upgrade them to threshold SVAR (T-SVAR) and threshold SAR (T-SAR) models. I introduce these steps one by one in this section.
4.2.1 The T-SVAR Step

Now it has become a popular practice to start inferring the structure of connectedness of a system from a VAR-type model. Most of the existing literature, such as Billio et al. (2012) and Diebold and Yılmaz (2014), uses reduced-form VARs. In contrast, the tripod model starts from a structural VAR model. Compared to reduced-form VARs, an SVAR allows the variables in the system to react instantaneously to shocks from elsewhere, which is closer to reality if the frequency of observations is not so high that the shocks and responses cannot be separated temporally.

In this paper, I add a threshold function to the SVAR model, thus upgrading it to a T-SVAR model. In the T-SVAR model, regimes are determined by the values of the threshold function that maps one or more exogenous variables to the probabilities of each state. There is no requirement of the form of the threshold function except that its value must be a probability, i.e., a real number ranging from zero to one. Therefore, we can use all kinds of threshold functions of convenience, which can be transient (similar to Hansen, 1997), smooth (see van Dijk et al., 2002, for a survey), or even more complex forms. There is no restriction on the number of threshold variables, either. This feature of my model can in fact be employed to construct a regime indicator based on the shift of variable connectedness. If we abstract from the structural part, the remaining reduced-form VAR part of the model will be similar to the threshold VAR models widely used in empirical studies such as Li and St-Amant (2010) and Afonso et al. (2011).

Compared with reduced-form VARs, there is always an identification problem accompanying SVAR models, as the data do not provide enough information to estimate all the parameters in an SVAR, unless we impose some additional assumptions on the model a priori. With proper prior assumptions, all the parameters of an SVAR can be estimated, and then we call the SVAR “identified.” The literature has proposed a number of identification strategies, but common strategies such as identifications based on recursive restrictions or sign restrictions, either require very strong prior assumptions that are rarely justified in reality, or do not result in unique identifications (see Kilian, 2013, for a review). Scidá (2016) suggests using a machine-learning algorithm to select an optimum set of restrictions, but this methodology requires there not to exist cyclical causalities in the model,
which is also often overly strong in reality.

In my model, the SVAR is identified by assuming that the residuals follow normal-mixture distributions.\(^3\) I assume the simplest case, where the normal-mixture distribution of the reduced-form residuals under each regime follow a mixture of two normal distributions with different covariances and zero mean. This is an even weaker assumption than the regular assumption, since an ordinary normal distribution is a special case of a mixture-normal distribution and mixture distributions can be used to approach random distributions in the real world (e.g. Kon, 1984, on distributions of asset returns). In an SVAR model with normal-mixture distributed residuals, Lanne and Lütkepohl (2010) show that we can achieve a unique local identification without any prior restriction on the interrelationship among the variables. Indeed, this only leads to local rather than global identifications and therefore we cannot know which variable is responding to which shock merely from identification results.\(^4\) But fortunately this problem can still be solved in actual empirical applications. In this paper, I follow Kohonen (2013) and fully identify of the T-SVAR model by referring to some external source of information.

4.2.2 The T-SAR Step

In the tripod model, the connectedness is recorded by the spatial weight matrix in a SAR model. The spatial weight matrix records the inverse distance between each pair of entities in the system, and the inverse distance can either be an inverse of the real geographical distance, or some closeness in the figurative sense, such as the amount of trade. There is in fact no substantial difference between the “inverse distance” and the “connectedness” discussed in this paper. Unlike traditional spatial modeling where the user must designate a spatial weight matrix (e.g. Tonzer, 2015), the spatial weight matrix in the tripod model becomes endogenous and can be estimated from the SVAR identification obtained in the

\(^3\)In order to understand what is a normal-mixture distribution, we can imagine the return of a certain stock follows a different distribution in different situations in the future, and each situation takes place in the future at given probabilities. Since we are not sure which situation will actually take place, the distribution of the return of the stock conditioned on present information will be a mixture of the normal distributions in different situations, or a “normal-mixture” distribution.

\(^4\)This problem is called “label switching” problem by Maciejowska (2010). She offers a detailed proof of it in the appendix.
A NOT-SO-TECHNICAL OVERVIEW OF THE THRESHOLD TRIPOD MODEL

previous SVAR step. This step can be compared with some recent attempts to infer spatial weight matrices from data such as Manresa (2015) and Lam and Souza (2015).

Specifically, the SAR in the tripod model is obtained by imposing some additional linear constraints on the parameters of the SVAR. This step is carried out with a “minimum distance (MD)” approach, which is easy to implement. Since I consider tripod models with thresholds in this paper, a similar MD approach can also be employed to obtain a threshold SAR (T-SAR) model from the T-SVAR estimation.

Moreover, I show that the thresholds in the threshold tripod model enable us to construct tests of many interesting hypotheses. I focus on two types of tests in this paper. First, I test whether the spatial structure contained in an economic system is identical across different regimes. Second, I test whether the difference between spatial structures in different regimes follows a certain pattern. Since spatial weights are a measure of the connectedness of the system, the two tests are actually equivalent to tests of the existence of the transition of connectedness, and the channel through which the transition of the connectedness takes place. Tests on other types of restrictions are also possible. For example, it is possible to test whether a given spatial weight matrix is misspecified or not.

4.2.3 The Network Analysis Step

In this step, we can further explore the entity interconnections revealed by the T-SAR model using some network analysis techniques. If we recall that a network is just defined as a group of entities and the connectedness among them, then the system we study can be regarded as a blend of different networks in different regimes and the transitions among the regimes can be regarded as the transitions among different networks.5

However, the networks generated by the tripod model in this paper are characterized by several unconventional traits, including non-unitary edge weights,6 directed edges, and

5 This kind of network blends is also called “overlapping networks” (Battiston et al., 2013) or “overlay networks” (De Domenico et al., 2013). Kivelä et al. (2014) provides a detailed review of the broader class of multilayer networks.

6 Usually a network is composed by “nodes” and “edges.” Each node represents an entity in the network and an edge represents a connection between the two nodes on its two ends.
often negatively weighted edges. Furthermore, the networks are usually complete, which means every pair of nodes in the network is to some extent connected. These traits make the application of classical network analytical techniques difficult, as many of them are actually devised to handle common incomplete networks with undirected edges with positive unitary weights. Nevertheless, I show that such unconventional networks can still be analyzed by some novel methods that have been devised recently. First, I use the simple strength measure and the PN measure to analyze centrality (see Everett and Borgatti, 2014), i.e., the importance of the role of a certain node in the networks. Second, McAssey and Bijma (2015) propose a novel clustering measure for complete and weighted networks, with which we can measure how close are the neighbors of a node connected mutually and to how close the whole network is interconnected. Finally, Scidá (2016) suggests that when we estimate the SAR model, we automatically obtain a global measure for the influence strength of the whole network, which is called the $\rho$-measure. Scidá (2016) also proposes many other analyses, but most of them can only be applied to networks with only positively weighted edges. Since I try to provide some analyses that still work even when there are negative edges, I switch to a new set of analyses and only keep the $\rho$-measure in this paper.

4.3 SVAR and T-SVAR Models

Before introducing my model formally, I introduce the notations used throughout the remaining parts of this paper. $I_a$ is the identity matrix of order $a$. $O_a$ is the $a \times a$ matrix where all the elements are zeros. Both $AB$ and $A \times B$ are matrix products of matrices $A$ and $B$. $1_a$ is the $a$-dimensional vector of ones, and $0_a$ is the $a$-dimensional vector of zeros. $A \otimes B$ is the Kronecker product of matrices $A$ and $B$. $A \odot B$ is the Hadamard product of $A$ and $B$. Next, $\text{diag}(a_1, a_2, \ldots)$ is the diagonal matrix where diagonal elements are $a_1, a_2, \ldots$, $\text{diag}(X)$ is the diagonal matrix where diagonal elements are respectively the elements of vector $X$, and $\text{diag}(A_1, A_2, \ldots)$ is the block diagonal matrix where $A_1, A_2, \ldots$ are on the main diagonal. $\text{vec}(A)$ is the standard vectorization of matrix $A$.

---

7 This is also known as the Schur product, or entrywise product.
4.3.1 A Review of SVAR Models

In this section, I present a brief review of the SVAR model, which is the foundation of the SVAR-spatial-network “tripod” model.

As the main advantage of an SVAR model, it allows us to capture the instantaneous response of variables to shocks. For instance, if an SVAR system at a daily frequency is composed of two stock market indices, then the SVAR model can reveal how one market responds to a market-specific shock in the other market on the same day. This feature of the SVAR model is especially valuable for financial issues, as many financial variables such as asset prices respond to shocks promptly. Even for economic systems where variables respond more slowly, this can still be very useful, because many data can only be obtained at an even lower frequency.

In the rest part of this paper, I assume $Y_t = (y_{1t}, y_{2t}, \ldots, y_{Kt})'$ is a $K$-dimensional vector representing endogenous variables observed at time $t = 1, 2, \ldots, T$, and $X_t = (x_{1t}, x_{2t}, \ldots, x_{Mt})'$ is a vector with the $M$ dimensions representing exogenous variables observed at time $t$. Additionally, let $\{u_t\}_{t=1}^T$ be the $K$-dimensional error term, then the reduced-form model with $p$ lags studied in this paper can be defined as $Y_t = \phi(Y_{t-1}, Y_{t-2}, \ldots, Y_{t-p}; X_t, u_t)$. Correspondingly, the structural-form model can be specified as $\psi(Y_t, Y_{t-1}, Y_{t-2}, \ldots, Y_{t-p}; X_t, u_t) = 0$.

In the baseline SVAR model, let the function $\phi$ be linear and $u_t$ be a white noise. Then the reduced-form VAR model becomes $Y_t = c + \sum_{i=1}^{p} A_i Y_{t-i} + A_X X_t + u_t$, where $c$ is a $K$-dimensional constant term, $A_1, A_2, \ldots, A_p$ are autoregressive coefficient matrices with size $K \times K$, $A_X$ is the exogenous coefficient matrix with size $K \times M$, and $u_t \sim (\mathbf{0}, \Sigma_u)$. Obviously, this reduced-form model is equivalent to various structural form models. But most SVAR literature considers the simple case, where the function $\psi$ is linear.

---

8Surely, it is possible to make $Y_t$ dependent on lags of $X_t$. In this case, we can stack the lags of $X_t$ in one vector and this definition still applies.
CRISIS-CONTINGENT DYNAMICS OF CONNECTEDNESS

and additively separable, i.e.,

\[ AY_t = A(c + \sum_{i=1}^{p} A_i Y_{t-i} + AX_t + u_t) = 0, \]

where \( A \) is an invertible \( K \times K \) matrix. Hereinafter, I call \( A \) the “structural matrix”. Defining \( \varepsilon_t \equiv Au_t \), the model can be rearranged as

\[ AY_t = Ac + \sum_{i=1}^{p} AA_i Y_{t-i} + AA_X X_t + \varepsilon_t. \] (4.1)

It is easy to see that \( \varepsilon \sim (0, A \Sigma_{\varepsilon} A') \). If we further require \( \varepsilon_t \) to be orthogonal such that its covariance matrix \( \Sigma_{\varepsilon} \) is diagonal, then the model in equation (4.1) can be called an SVAR model. A diagonal \( \Sigma_{\varepsilon} \) indicates that all pairs of components in the error term \( \varepsilon_t \) have zero correlation, therefore the shocks to the economic system represented by the SVAR model are uncorrelated as well. Such shocks are called “structural shocks”. Equation (4.1) is in fact the “A-model” in Lütkepohl (2007) plus some exogenous variables.

In many papers and textbooks about SVAR, the structural matrix \( A \) is chosen to let \( \Sigma_{\varepsilon} \) be an identity matrix \( I_K \). Under this convention, the structural shocks all have the same size, which provides empirical studies with much convenience in comparing the effects of different shocks. However, since in this paper I do not focus on comparing the effects of economic shocks, I choose another way of normalization that is more convenient for further research in the following sections. I require all the diagonal elements of \( A \) to be one, but do not restrict the scale of \( \Sigma_{\varepsilon} \). It is easy to see that my normalization is equivalent to the conventional normalization: if \( A \) is a structural matrix following my rule of normalization, and \( \Sigma_{\varepsilon} = \text{diag}(\sigma_{11}, \sigma_{22}, \ldots, \sigma_{KK}) \), then \( A \times \text{diag}(\sigma_{11}^{1/2}, \sigma_{22}^{1/2}, \ldots, \sigma_{KK}^{1/2}) \) is obviously a structural matrix following the conventional rule of normalization such that \( \Sigma_{\varepsilon} \) is an identity matrix. The most significant advantage of my way of normalization can be shown by rearranging equation (4.1) as

\[ Y_t = GY_t + Ac + \sum_{i=1}^{p} AA_i Y_{t-i} + AA_X X_t + \varepsilon_t, \] (4.2)
where $G \equiv I_K - A$. Since all diagonal elements of $A$ have been set to one, the diagonal elements of $G$ will be zero. This will be the most convenient form to display an SVAR model in the next section, but within this section, I still focus on the specification in the form of equation (4.1).

There are two groups of unknown parameters in the SVAR model in equation (4.1). The first group, $\{c, A_1, A_2, \ldots, A_p, \Sigma_u, \Sigma_u\}$, is inherited from the reduced-form VAR; and the second group, $\{A, \Sigma_e\}$, contains the additional parameters that are exclusive to the structural form model. Now let us assume that we had obtained a proper estimation of the first group of parameters, then we would only need to find a way to get an estimation of $A$ and $\Sigma_e$ to fully identify the structural model.

However, the identification step is never easy. Straightforwardly, if we assume the distribution of $u_t$ is completely decided by $\Sigma_u$ and there is no restriction on the free elements in $A$ and $\Sigma_e$, then the SVAR model is not identifiable. The reason is intuitive: $A$ and $\Sigma_e$ are regulated by the equation

$$\Sigma_e = A \Sigma_u A' \cdot$$

Since $A$ is invertible, this equation can be rewritten as

$$\Sigma_u = A^{-1} \Sigma_e A'^{-1} \cdot$$

(4.3)

Since $\Sigma_u$ is a covariance matrix and therefore symmetric, it has $K(K+1)/2$ free elements. But in the right hand of the equation, $A$ has $K^2 - K$ free elements and $\Sigma_e$ has $K$ free elements. Putting them together, the right side of equation (4.3) requires $K^2$ free elements, whereas the left side can only offer $K(K + 1)/2$. Hence, there will be infinite number of solutions to the equation, and the SVAR model fails to be identified. Lütkepohl (2007) also offers a formal rank condition for the identifiability of the SVAR model discussed here.

The traditional way of solving the problem is to reduce the number of free parameters in the right side of equation (4.3) by imposing additional restrictions. As the most straightforward way to do this, we may simply set the lower- or upper-triangle of $A$ to be 

---

9 The estimation methods of reduced-form VARs can be found in Lütkepohl (2007).
zero, and thereby A and $\Sigma_u$ can be immediately obtained by the means of a Cholesky decomposition. However, this is a very strong set of restrictions, as it assumes a recursive causal chain among the variables. In other words, it assumes we know that among the $K$ variables, variable A is known to be instantly affected by all $K$ structural shocks, variable B by certain $K-1$ structural shock, variable C by certain $K-2$ structural shocks, etc. But in practice, it is really difficult to predicate the such relationships among the variables. Scidá (2016) suggests a machine learning algorithm called “PC algorithm” to find the optimal recursive order of the variables. However, we still have to make sure that the recursive causal chain is really there among the variables before we apply the algorithm. Besides the recursive restrictions, there are other identification strategies that rely on a priori restrictions, and most of them suffer from similar problems in imposing the restrictions, as is echoed by Kilian (2013). Such problems are quite undesirable, as the very purpose of the tripod model is to reveal the connectedness among variables or entities without knowing much about their interrelationship a priori. In this paper, the identification problem is handled in a novel way, by exploiting the structure of $\Sigma_u$ and extracting more information from it, which enables us to obtain a unique identification without imposing strong a priori restrictions on the model. This methodology for the SVAR model is described in the following part, and in Section 4.3.3 I show that the generalized case of the SVAR model with thresholds can be also identified with it.

4.3.2 Identification of an SVAR Model with Normal-Mixture Residuals

Although the main purpose of this paper is to establish a model to describe transitions of connectedness, which is based on an SVAR model with thresholds, in order to illustrate how the identification strategy works, I first introduce how a non-threshold SVAR model with normal-mixture residuals is identified by exploiting the distribution structure of residuals. The idea is pioneered by Lanne and Lütkepohl (2010), who show normal-mixture residuals can be utilized to construct identification tests.

From the previous discussion, we have already seen that an SVAR is identified by finding a proper “structure”, which combines orthogonal structural shocks into the visible

---

10 For details of the PC algorithm, see Spirtes et al. (2000, p.117–119).
reduced-formed innovations. A mixture distribution can be regarded as a mixture of several components of simple distributions. The SVAR model is identified, if we assume that the "structure", i.e., the way that the orthogonal shocks are combined into reduced-form innovations is the same for all the components.

Consider the reduced-form VAR model

$$Y_t = \epsilon + \sum_{i=1}^{p} A_i Y_{t-i} + A_\chi X_t + v_t,$$

where

$$v_t = \begin{cases} v_{1t} \sim N(0, \Sigma_1) \quad \text{with probability } \gamma, \\ v_{2t} \sim N(0, \Sigma_2) \quad \text{with probability } 1 - \gamma, \end{cases}$$

Now the reduced-form error term $v_t$ follows a mixture of two normal distributions with zero mean and different covariance matrices. If $\Sigma_1 = \Sigma_2$, or $\gamma = 0, 1$, then the normal-mixture distribution reduces to a regular normal distribution. The covariance matrix of $v_t$ is also the weighted combination of the covariance of its two components, $\gamma \Sigma_1 + (1 - \gamma) \Sigma_2$. Proposition 1 in Lanne and Lütkepohl (2010) demonstrates that given the normal-mixture distributed $v_t$, there exists a diagonal matrix $\tilde{\Lambda} = \text{diag}(\tilde{\lambda}_1, \tilde{\lambda}_2, \ldots, \tilde{\lambda}_K)$ with $\tilde{\lambda}_i > 0$ for all $i = 1, 2, \ldots, K$ and an invertible matrix $\tilde{\Omega}$ such that $\Sigma_1 = \tilde{\Omega} \tilde{\Lambda} \tilde{\Omega}'$ and $\Sigma_2 = \tilde{\Omega} \tilde{\Lambda} \tilde{\Omega}'$. Furthermore, if every pair of diagonal elements of $\tilde{\Lambda}$ is different, the decomposition $\tilde{\Omega}$ is unique up to the inversion of the signs of all elements in the same column.\footnote{That is to say, if two decompositions $\tilde{\Omega}$ and $\hat{\Omega}$ both satisfy $\Sigma_1 = \tilde{\Omega} \tilde{\Lambda} \tilde{\Omega}' = \hat{\Omega} \hat{\Lambda} \hat{\Omega}'$ and $\Sigma_2 = \tilde{\Omega} \tilde{\Lambda} \tilde{\Omega}' = \hat{\Omega} \hat{\Lambda} \hat{\Omega}'$, then there must exists a diagonal matrix $J$, all diagonal elements of which are either 1 or -1, such that $\tilde{\Omega} = \hat{\Omega} J$.}

Suppose the diagonal elements of $\tilde{\Omega}$ are $(\tilde{\omega}_{11}, \tilde{\omega}_{22}, \ldots, \tilde{\omega}_{KK})$. Due to my rule of normalization, I then further rewrite the decompositions of covariances as $\Sigma_1 = \Omega \Lambda_1 \Omega'$ and $\Sigma_2 = \Omega \Lambda_2 \Omega'$, where $\Omega$ has all-one diagonal elements, $\Lambda_1 = \text{diag}(\tilde{\omega}_{11}^2, \tilde{\omega}_{22}^2, \ldots, \tilde{\omega}_{KK}^2)$ and $\Lambda_2 = \Lambda_1 \tilde{\Lambda}$. The covariance matrix of $v_t$ can then be written as

$$\Sigma_v = \gamma \Omega \Lambda_1 \Omega' + (1 - \gamma) \Omega \Lambda_2 \Omega' = \Omega (\gamma \Lambda_1 + (1 - \gamma) \Lambda_2) \Omega'. \quad (4.4)$$

Suppose that this reduced-form model is equivalent to the SVAR model specified by
CRISIS-CONTINGENT DYNAMICS OF CONNECTEDNESS

equation (4.1) so that
\[ \Sigma_v = A^{-1} \Sigma_c A' \, A^{-1}. \]

Comparing equation (4.5) with (4.4), we find
\[ A = \Omega^{-1}. \]

The SVAR model is identified up to this moment. Indeed, we may still invert the column signs of \( \Omega \) arbitrarily, but this is no problem in practice, as we can always distinguish between a positive shock and a negative shock easily.

The parameters in this SVAR model can be locally estimated by maximum likelihood (ML) estimation. Denote \( \epsilon = \{ \epsilon_t \}_{t=1}^T \). Collecting all the parameters in the model except the mixture weight \( \gamma \) in the vector \( \theta \), then the likelihood function is
\[ L_{SVAR}(\theta, \gamma; \epsilon) = \prod_{t=1}^T L_{SVAR}^t(\theta, \gamma; \epsilon_t) \]
where the one-time density function is
\[ f(\epsilon_t; \theta, \gamma) = \gamma (2\pi)^{-K/2} \det(\Omega \Lambda, \Omega')^{-1/2} \exp \left( -\frac{1}{2} \epsilon_t' \left( \Omega \Lambda, \Omega' \right)^{-1} \epsilon_t \right) \]
\[ + (1 - \gamma) (2\pi)^{-K/2} \det(\Omega \Lambda, \Omega')^{-1/2} \exp \left( -\frac{1}{2} \epsilon_t' \left( \Omega \Lambda, \Omega' \right)^{-1} \epsilon_t \right). \]

However, it should be noticed that we can only obtain local estimation through ML. To illustrate this problem, let us assume \( \hat{\Omega}, \hat{\Lambda}_1 \) and \( \hat{\Lambda}_2 \) are a local ML estimation that maximizes the likelihood function in equation (4.6). If we permute the \( i \)-th and \( j \)-th columns of \( \hat{\Omega} \) by postmultiplying \( \hat{\Omega} \) by the elementary column-switching operator \( E_{ij} \), then \( \hat{\Omega}^* = \Omega E_{ij}, \Lambda_1^* = E_{ij} \hat{\Lambda}_1 E_{ij}' \) and \( \Lambda_2^* = E_{ij} \hat{\Lambda}_1 E_{ij}' \) are also a group of local ML estimation.

Since \( E_{ij} E_{ij}' = I_K \), we have \( \hat{\Omega} \Lambda, \hat{\Lambda}' = \Omega^* \Lambda^*_1 \Omega^* \Lambda^*_2 \) and \( \hat{\Omega} \Lambda, \hat{\Omega}' = \Omega^* \Lambda^*_1 \Omega^* \Lambda^*_2 \). Therefore both groups of estimations maximize the one-time density function in equation (4.7) for all \( t \) and thus the overall likelihood function. This is called the “label switching” problem in Maciejowska (2010). She provides a formal proof of the problem in her appendix. Since \( \Omega = A^{-1} \) represents the simultaneous effects of each structural shock, the arbitrariness of
the column permutation of $\Omega$ implies that we can only know how each structural shock affects the economic system, but do not know which structural shock identified by the SVAR model corresponds to which economic shock in reality.

Following Kohonen (2013), I solve this label switching problem by ordering information contained in some proxy variable in this paper. Kohonen (2013) argues that if we can find a proxy variable that provides information on the ordering of the diagonal elements of $\Sigma_\varepsilon$, then it becomes possible for us to pick the correct global estimation of the SVAR model among the local ML estimations. For example, if $Y_t$ represents the returns of the stock markets in several countries, then the size of return shocks indicated by $\Sigma_\varepsilon$ can be proxied by the magnitude of news in each country. Should we find some variable showing the magnitude of news, the diagonal elements of $\Sigma_\varepsilon$ should have the same order as the magnitude of news. In this way, those local estimations where the diagonal elements of the estimated $\hat{\Sigma}_\varepsilon = \hat{\gamma} \hat{\Lambda}_1 + (1 - \hat{\gamma}) \hat{\Lambda}_2$ is ordered incorrectly will be filtered out.

So far I have shown how to identify an SVAR model with normal-mixture residuals. In following sections, I will show that an SVAR actually contains all the information I need for revealing the connectedness among the entities. However, since this paper aims to describe how the connectedness transitions in different circumstances, we need a model that internalizes the transitions. In this paper, the transitions of connectedness are modeled by a threshold SVAR (T-SVAR) model, which I present in the following section.

### 4.3.3 The T-SVAR Model with Normal-Mixture Residuals

In this part, I show how to construct a T-SVAR model. In the T-SVAR model, the data-generating process is partitioned by thresholds into $N$ different regimes. Under each regime, the system is interconnected in a certain way, which is described by an SVAR model. As all the $N$ regimes are estimated simultaneously, we obtain $N$ different sets of interconnections of the entities or variables we are interested in, and the transitions among them are described by the threshold function. The T-SVAR model is still identified through its normal-mixture residuals. The model introduces a threshold mechanism in a similar way of the threshold VAR models widely used in empirical studies such as Li and St-Amant
CRISIS-CONTINGENT DYNAMICS OF CONNECTEDNESS

(2010) and Afonso et al. (2011), despite it is a structural model.

Since in my model the threshold function is restricted to a number between zero and one, the threshold model has a probabilistic interpretation in economics. In the model without thresholds, the connectedness is fixed for all the time. After the threshold function is introduced, the connectedness is no longer certain, and we can only know that the connectedness is one among several possible states. The probabilities for each possible state are determined by the threshold function, which is a certain reflection of the exogenous threshold variables. The dynamics of the connectedness are thus portrayed by the time-varying probability distributions determined by the threshold variables.

Assume there is an $H$-dimensional vector of exogenous threshold variables $Q_t = (q_{t1}, q_{t2}, \ldots, q_{tH})'$ for time $t = 1, 2, \ldots, T$, and the unnormalized threshold function $\tau^*_\delta : \mathbb{R}^H \rightarrow [0, \infty)^N$, where $\delta$ represents the unknown parameter(s) of $\tau^*_\delta$. I further normalize $\tau^*_\delta$ by the sum of its value and get the normalized threshold function

$$\tau(Q_t; \delta) \equiv \tau_\delta(Q_t) = \frac{\tau^*_\delta(Q_t)}{1^N \tau^*_\delta(Q_t)}. \quad (4.8)$$

This means there are $N$ regimes, and $\tau_\delta$ can map the threshold variables to the probability of the system to be in each regime. For simplicity, in the rest of this paper, when I mention the “threshold function”, I always refer to the normalized threshold function. The components of the threshold function are denoted as $\tau(Q_t; \delta) = (\tau_1(Q_t; \delta)', \tau_2(Q_t; \delta)', \ldots, \tau_N(Q_t; \delta)')'$.

In practice, the threshold function can take any form. For example, the transilient form with one threshold variable is

$$\tau^\text{Tran}_b(Q_t) = \begin{cases} (1, 0)', & Q_t \geq b, \\ (0, 1)', & Q_t < b, \end{cases} \quad (4.9)$$

where $b$ is the breakpoint. This implies that the system has probability one to be in Regime 1 if the threshold variable $Q_t \geq b$, and in Regime 2 otherwise. We can also use the the
SVAR AND T-SVAR MODELS

multivariate logistic form with any number of variable(s) as the threshold function

$$\tau^\text{Logi}_\beta(Q_t) = \begin{pmatrix} 1/(1 + e^{-\beta^\prime Q_t}) \\ 1 - 1/(1 + e^{-\beta^\prime Q_t}) \end{pmatrix} ,$$

(4.10)

where $\beta$ is the vector of scales of the logistic function.\footnote{In actual applications of this multivariate logistic threshold function, the actual threshold variables are usually the desired threshold variables $Q_t$ together with a constant term, say, $Q_t^{\text{Actual}} = (1, Q_t^\prime)$, in order to keep a intercept term.} This threshold function allows the system to switch between two regimes smoothly, and the location and speed of the switch are decided by parameters $\beta$.

Note that according to the definition of the threshold function (4.8), the threshold variables $Q_t$ are given exogenously, but the parameters $\delta$ is unkown a priori. This means before we estimate the T-SVAR model, we only known different values of $Q_t$ correspond to different probability distributions of regimes, but do not know how they correspond. The parameter values are estimated from the data. For instance, in the transilient threshold function (4.9), we only know that when $Q_t \geq b$, the system is at Regime 1, and let the data decide the value of $b$.

With these definitions, we consider the reduced-form and corresponding structural-form threshold model with $N$ regimes and $p$ lags:

$$Y_t = \Phi(Y_{t-1}, Y_{t-2}, \ldots, Y_{t-p}, X_t, r_t) \tau_\delta(Q_t),$$

and

$$\Psi(Y_t, Y_{t-1}, Y_{t-2}, \ldots, Y_{t-p}, X_t, r_t) \tau_\delta(Q_t) = 0_K,$$

(4.11)

where $r_t$ is still the error term, $\Phi(\cdot) = (\phi_1(\cdot), \phi_2(\cdot), \ldots, \phi_N(\cdot))$ and $\Psi(\cdot) = (\psi_1(\cdot), \psi_2(\cdot), \ldots, \psi_N(\cdot))$ are matrices of size $K \times N$ denoting the model conditioned on each regime. Notice that I *do not* impose restrictions on the columns of $\Phi(\cdot)$ and $\Psi(\cdot)$ here, so they can be either mutually equal, partly equal, or completely unequal. In the application part of this paper, I let all columns of them be equal. Certainly, other kinds of restrictions are possible as well.

\footnote{In actual applications of this multivariate logistic threshold function, the actual threshold variables are usually the desired threshold variables $Q_t$ together with a constant term, say, $Q_t^{\text{Actual}} = (1, Q_t^\prime)$, in order to keep a intercept term.}
The T-SVAR model considered in this paper can still be specified by assuming \( \Psi \) is linear and additively separable. Its summation form is:

\[
\sum_{n=1}^{N} A_n Y_t \tau_n(Q_t; \delta) = \\
\sum_{n=1}^{N} \left( A_n c^{(n)} + \sum_{i=1}^{p} A_n A_i^{(n)} Y_{t-i} + A_n A_X^{(n)} X_t + \varepsilon_{nt} \right) \tau_n(Q_t; \delta),
\]  

(4.12)

where \( A_n \) denotes the structural matrix conditioned on regime \( n \), \( \tau_n(Q_t; \delta) \) is the probability for the system to be in regime \( n \), \( c^{(n)} \), \( A_i^{(n)} \), and \( A_X^{(n)} \) are the autoregressive coefficients in regime \( n \), and \( \varepsilon_{nt} = A_n r_t \) are the structural shocks conditioned on regime \( n \).

The key to the identification of the T-SVAR model still rests on the distributional assumption of \( r_t \). In the T-SVAR model I assume the conditional distribution of \( r_t \) is a time-constant normal-mixture distribution, and hence its unconditional distribution becomes a mixture of \( N \) normal-mixture distributions, i.e.

\[
\Sigma_{r,t|Q_t;\delta} = \sum_{n=1}^{N} (\gamma_n \Sigma_{1n} + (1 - \gamma_n) \Sigma_{2n}) \tau_n(Q_t; \delta).
\]  

(4.13)

Recall the results about the covariance decompositions in Section 4.3.2, and then we know there are matrices \( \{\Omega_n\}_{n=1}^{N}, \{\Lambda_{1n}\}_{n=1}^{N}, \{\Lambda_{2n}\}_{n=1}^{N} \) such that \( \Omega_n \Lambda_{1n} \Omega_n' = \Sigma_{1n} \) and \( \Omega_n \Lambda_{2n} \Omega_n' = \Sigma_{2n} \) for all \( n = 1, 2, \ldots, N \) under regular conditions. Equation (4.13) can thereby be rewritten as

\[
\Sigma_{r,t|Q_t;\delta} = \sum_{n=1}^{N} \Omega_n (\gamma_n \Lambda_{1n} + (1 - \gamma_n) \Lambda_{2n}) \Omega_n' \tau_n(Q_t; \delta).
\]  

(4.14)

Denote

\[
\Upsilon_n = \Omega_n (\gamma_n \Lambda_{1n} + (1 - \gamma_n) \Lambda_{2n})^{1/2},
\]  

(4.15)
then equation (4.14) can also be written as

$$
\Sigma_{r,t|Q;\delta} = \sum_{n=1}^{N} \Upsilon_n \Upsilon_n' \tau_n(Q_t;\delta).
$$

Further, as assumed, for all $n$, we have $\epsilon_{nt} = A_n r_t$ conditioned on regime $n$. Therefore the unconditional relationship between the reduced-form and structural-form residuals will be

$$
r_t|Q;\delta = \sum_{n=1}^{N} (A_n^{-1} \epsilon_{nt}) \tau_n(Q_t;\delta),
$$

and

$$
\Sigma_{r,t|Q;\delta} = \sum_{n=1}^{N} (A_n^{-1} \Sigma_{\epsilon}^{(n)} A_n'^{-1}) \tau_n(Q_t;\delta), \tag{4.16}
$$

where $\Sigma_{\epsilon}^{(n)}$ denotes the covariance matrix of $\epsilon_{nt}$. Comparing equations (4.16) and (4.14), we get

$$
A_n = \Omega_n^{-1},
$$

for all $n = 1, 2, \ldots, N$. Meanwhile, we have

$$
A_n^{-1} \Sigma_{\epsilon}^{(n)} = \Upsilon_n I_K^J.
$$

This implies that $\Upsilon_n$ is actually the decomposition of the reduced-form covariance matrices if we require the covariance matrices of the structural shocks $\epsilon_{nt}$ to be unit matrices instead. In fact, $\Upsilon_n$ is the matrix that people are usually most interested in with traditional SVAR analyses, because it represents the instantaneous response of each variable to unit-sized structural shocks.

The T-SVAR model can still be locally estimated with ML. Denote $\varepsilon_t = \{\varepsilon_{nt}\}_{t=1}^{T}$ and $\varepsilon = \{\varepsilon_{nt}\}_{t=1}^{T_n}$. Stacking the mixture weights $(\gamma_1, \gamma_2, \ldots, \gamma_N)$ in vector $\Gamma$, and remaining unknown parameters in the T-SVAR model except $\delta$ in vector $\Theta$, then we get
CRISIS-CONTINGENT DYNAMICS OF CONNECTEDNESS

the likelihood function

$$L^{T-SVAR}(\Theta, \Gamma, \delta; \varepsilon) = \prod_{t=1}^{T} L^{T-SVAR}_t(\Theta, \Gamma, \delta; \varepsilon_t) = \prod_{t=1}^{T} \phi(\varepsilon_t; \Theta, \Gamma, \delta), \quad (4.17)$$

where the one-time density function is

$$\phi(\varepsilon_t; \Theta, \Gamma, \delta) = \sum_{n=1}^{N} \phi_n(\varepsilon_{nt}; \Theta, \Gamma) \tau_n(Q; \delta),$$

and the one-time density function conditioned on regime $n$ is

$$\phi_n(\varepsilon_{nt}; \Theta, \Gamma) = \gamma_n (2\pi)^{-K/2} \det(\Omega_n \Lambda_{1n} \Omega_n')^{-1/2} \exp\left(-\frac{1}{2} \varepsilon_{nt}' (\Omega_n \Lambda_{1n} \Omega_n')^{-1} \varepsilon_{nt}\right) + (1 - \gamma_n) (2\pi)^{-K/2} \det(\Omega_n \Lambda_{2n} \Omega_n')^{-1/2} \exp\left(-\frac{1}{2} \varepsilon_{nt}' (\Omega_n \Lambda_{2n} \Omega_n')^{-1} \varepsilon_{nt}\right).$$

After we obtain local estimations of the T-SVAR model by maximizing the likelihood function, global estimation can still be achieved with the method proposed by Kohonen (2013), which has been described in Section 4.3.2 in this paper.

Finally, before I finish this section, I have a caveat here concerning the computational implementation of the ML estimation. Although in theory it is nothing more than finding a maximum of the log-likelihood function, in practice, direct numerical optimization algorithms work poorly for the SVAR models with normal-mixture residuals no matter whether there are thresholds or not, given the complexity of the likelihood functions. In order to achieve a high-quality optimum, I use the expectation-maximization (EM) method. Maciejowska (2010) offers an in-depth discussion about the technical issues in the ML estimation of these mixture models.

The T-SVAR model is the foundation of the tripod model in this paper. In the following section, I show how to transform a T-SVAR model to a spatial auto-regression model, which displays the connectedness in different regimes in a straightforward manner.
FROM STRUCTURAL MODELS TO SPATIAL MODELS

4.4 From Structural Models to Spatial Models

An SVAR model depicts the instantaneous and lagging responses of the variables to structural shocks, and we want to further understand what interconnections the impulse-response relationship implies. Scidá (2016) demonstrates that an identified SVAR model actually contains a spatial auto-regression (SAR) process, and the spatial weight matrix in a SAR model is a good indicator of the connectedness among the entities of interest. While in traditional spatial auto-regressive models the spatial weights have to be exogenously given, an SVAR model does not impose any structure on the spatial relationship among the variables. Any exogenously given spatial structure thus becomes testable against the SAR process transformed from an SVAR model. Similarly, a T-SVAR model also contains a threshold SAR (T-SAR) process, which implies different spatial weights, and thereby different states of connectedness, in different regimes. Procedures to transform a T-SVAR model to a T-SAR are shown in Section 4.4.1. After we obtain the T-SAR model, we may construct statistical tests of various interesting hypotheses, which is discussed in Section 4.4.2.

4.4.1 The SAR Model

In this section, I start from the SAR model

\[ Y_t = \rho W Y_t + c_W + \sum_{i=1}^{p} B_i Y_{t-i} + B_X X_t + \epsilon_t, \]  

(4.18)

where \( \rho \) is a scalar scaling coefficient, \( W = (w_{ij})_{K \times K} \) is a spatial weight matrix, \( c_W, \{ B_i \}_{i=1}^{p} \) and \( B_X \) are autoregressive coefficients, and \( \epsilon_t \) is still the orthogonal error term with zero mean and a diagonal covariance matrix \( \Sigma_\epsilon \). Following the convention in spatial models, I assume \( W \) is row-standardized to unitary sums and its diagonal elements are all zeros. \( W \) represents the inverse distances, or in other words, spatial connectedness among the entities in an economic system, and the “spatial connectedness” can either be an inverse measure of real geographical distances among the entities, or some figurative connection such as the amount of mutual trade among economies. Specifically, the element \( w_{ij} \) in \( W \) indicates the
connection from the $j$-th entity to the $i$-th entity.\footnote{If the connections are not directional, then $W$ will be symmetric and thus $w_{ij} = w_{ji}$ for all $i \neq j$. But the connections can be directional, too. In fact, the connections considered in this paper are mostly directional.} And since $W$ is standardized by $\rho$, $\rho$ serves as an overall measure of the connectedness intensity of the system.

If we compare equation (4.18) with equation (4.2), it is straightforward to see how the similarity. Since the diagonal elements of $G = I_K - A$ are zeros, equation (4.18) is similar to (4.2) with row-standardizing constraints on the sans-diagonal structural matrix $G$. Certainly, in the SVAR model specified by equation (4.2) $A$ is unknown a priori and in spatial models $W$ is usually given. But in this paper, I estimate a special SAR model with endogenous a spatial weight, and test a given spatial weight matrix against the endogenous one implied by the data. In fact, the relationship between the SAR model with endogenous spatial weights and the SVAR model specified by equation (4.2) has been formally demonstrated by Proposition 3.1 in Scidá (2016). Because in this paper it is very unlikely that the estimation of $G$ would have a row where all the elements are zeros due to the special SVAR identification strategy, I only present a corollary of this proposition below:

**Proposition 1.** Assume the SVAR model (4.2) is identified, and no row in $G$ is fully composed by zeros. Then the SAR model (4.18) is a constrained SVAR model (4.2) with $(K - 1)$ independent linear restrictions on $G$ given by

$$\mathcal{R} G 1_K = o_{K-1},$$

where the $(K - 1) \times K$ matrix

$$\mathcal{R} = \begin{pmatrix} -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & -1 & 1 \end{pmatrix}.$$

This proposition is an immediate corollary of Proposition 3.1 in Scidá (2016).

Based on the finding above, I proceed to the SAR model with thresholds. Consider the
N-regime T-SAR model

\[
\sum_{n=1}^{N} Y_t \tau_n(Q_t; \delta) = \\
\sum_{n=1}^{N} \left( \rho_n W_n Y_t + \varepsilon_n^{(n)} + \sum_{i=1}^{P} B_{n(i)}^{(n)} Y_{t-i} + B_{X}^{(n)} X_t + \varepsilon_{nt} \right) \tau_n(Q_t; \delta),
\]

(4.19)

where \( \rho_n, W_n, \varepsilon_n^{(n)}, B_{n(i)}^{(n)}, \text{ and } B_{X}^{(n)} \) are respectively the scaling coefficient, spatial weight matrix, and autoregressive coefficients in regime \( n \). Still, \( \varepsilon_{nt} \) is the structural error term.

Applying Proposition 1 to each regime in equation (4.12), we have for each regime \( n \),

\[
\mathcal{R}(I_K - A_n) 1_K = o_{K-1}.
\]

Since \( \mathcal{R}I_K 1_K \equiv o_{K-1} \), this can be rewritten as

\[
\mathcal{R}A_n 1_K = o_{K-1},
\]

for all \( n = 1, 2, \ldots, N \), or equivalently in the stacking form

\[
\mathcal{R} \mathcal{A} 1_{NK} = o_{N(K-1)},
\]

where \( \mathcal{R} = I_N \otimes \mathcal{R} \) and \( \mathcal{A} = \text{diag}(A_1, A_2, \ldots, A_N) \). This conclusion turns the estimation of the T-SAR model into a restricted ML estimation program:

\[
\max_{\Theta, \Gamma, \delta, \varepsilon} L^{T-\text{SVAR}}(\Theta, \Gamma, \delta; \varepsilon),
\]

s.t. \( \mathcal{R} \mathcal{A} 1_{NK} = o_{N(K-1)} \).

However, while the constrained ML gives the parameter estimations of the T-SAR model, in practice we have to perform a very complex constrained non-linear optimization procedure; but even the unconstrained non-linear optimization in the T-SVAR model is already technically challenging, let alone the restricted non-linear optimization. Therefore, in this paper, I employ the minimum distance (MD) approach proposed by Scidá (2016) to
obtain the estimation constrained T-SAR model from the unconstrained T-SVAR model.

Let \( \mathbf{A} = (\text{vec}(A_1)', \text{vec}(A_2)', \ldots, \text{vec}(A_N)')' \) collect all the elements of \( \{A_n\}_{n=1}^N \) or \( \mathcal{A} \), and

\[
\mathbf{G} \equiv \begin{pmatrix}
\text{vec}(G_1)'
\text{vec}(G_2)'
\vdots
\text{vec}(G_N)'
\end{pmatrix} = \begin{pmatrix}
\text{vec}(I_K - A_1)'
\text{vec}(I_K - A_2)'
\vdots
\text{vec}(I_K - A_N)'
\end{pmatrix}.
\]

Denote a matrix where \( e_{ij} = 0 \) for all \( i = j \) and \( e_{ij} = 1 \) for all \( i \neq j \) as \( \mathbf{E} = (e_{ij})_{K \times K} \). Let \( \mathbf{S} \) be a \( (K^2 - K) \times K^2 \) row-selection matrix which deletes all the all-zero rows in \( \text{diag}(\text{vec}(\mathbf{E})) \) such that \( \mathbf{S} \times \text{diag}(\text{vec}(\mathbf{E})) \) does not have all-zero rows\(^{14} \). Then we can see the matrix

\[
\mathbf{S} = \mathbf{I}_N \otimes \mathbf{S}
\]

contains all the non-diagonal elements of \( \{A_n\}_{n=1}^N \) and stacks them in the vector \( \mathbf{G} \) by letting

\[
\mathbf{G} = \mathbf{S}\mathbf{G} = \mathbf{S}\mathbf{A}.
\]

In fact, \( \mathbf{G} \) is the parameter vector that collects all the unknown parameters in \( \{A_n\}_{n=1}^N \) or \( \mathcal{A} \) in the T-SVAR model as all the fixed parameters in \( \mathbf{G} \) are removed. Assume we have already obtained an estimation of \( \mathbf{G} \) in the unconstrained T-SVAR model denoted by \( \hat{\mathbf{G}} \), and the squared error matrix of \( \hat{\mathbf{G}} \) is \( \Sigma_{\hat{\mathbf{G}}} \).\(^{15} \) Now the MD estimation of \( \mathbf{G} \) in the constrained T-SAR model is given by minimizing the weighted distance between the constrained \( \mathbf{G} \) and the estimated unconstrained \( \hat{\mathbf{G}} \):

\[
\min_{\mathbf{G}} \quad (\mathbf{G} - \hat{\mathbf{G}})'\Sigma_{\hat{\mathbf{G}}}^{-1}(\mathbf{G} - \hat{\mathbf{G}}),
\]

\[\text{s.t.} \quad \mathcal{R}_N \mathbf{1}_{NK} = \mathbf{0}_{N(K-1)}. \quad (4.20)\]

Note that although the constraint \( \mathcal{R}_N \mathbf{1}_{NK} \) is not explicitly a function of \( \mathbf{G} \), it can actually be rewritten as a linear combination of \( \mathbf{G} \), which I show later in this section. Now the

\(^{14} \)In fact, now we have \( S \times \text{diag}(\text{vec}(\mathbf{E})) = S \).

\(^{15} \)Since \( \hat{\mathbf{G}} \) is actually obtained through the ML estimation in this paper, \( \Sigma_{\hat{\mathbf{G}}} \) can be easily calculated via estimating Fisher’s information of \( \hat{\mathbf{G}} \).
FROM STRUCTURAL MODELS TO SPATIAL MODELS

problem has been turned into the minimization of a quadratic form with linear constraints, which has a closed-form solution and is thus far less burdensome in computation than the constrained ML estimator. Moreover, it has been demonstrated by Property 9 in Gourieroux and Monfort (1989) that the MD estimator is asymptotically equivalent to the constrained ML estimator, which guarantees its efficiency.

In order to find the closed-form solution of the program in (4.20), we first rewrite the constraints $RA_{1NK} = o_{N(K-1)}$ as a linear combination of the parameter vector $G$.

Concerning this, I show the following proposition:

**Proposition 2.** Let $\text{vec}(G_n^{\wedge 1})$, $n = 1, 2, \ldots, N$ be a $K^2$-dimensional vector whose $i$-th component is 1 if the $i$-th component of $\text{vec}(G_n)$ is not zero, and 0 otherwise, and

$$
\mathcal{H}^{\wedge 1} = \begin{pmatrix}
\text{vec}(G_1^{\wedge 1})_{1K-1}' \\
\text{vec}(G_2^{\wedge 1})_{1K-1}' \\
\vdots \\
\text{vec}(G_N^{\wedge 1})_{1K-1}'
\end{pmatrix},
$$

The constraint in the MD program (4.20) can be equivalently written as:

$$
RA_{1NK} = \mathcal{H}G = o_{N(K-1)},
$$

(4.21)

where

$$
\mathcal{H} = (I_N \otimes (1_K \otimes R)^\prime) \otimes \mathcal{H}^{\wedge 1}.
$$

Now that the MD program becomes a classical minimization problem of quadratic forms with linear constraints, the MD estimator is given by its solution, as is described in the following proposition with $\mathcal{U} = o_{N(K-1)}$:

**Proposition 3.** Suppose the estimation $\tilde{G}$ of a certain model is given by the following MD program:

$$
\min_{\mathcal{G}} \quad (\mathcal{G} - \mathcal{G}^*)' \Sigma_{\mathcal{G}}^{-1} (\mathcal{G} - \mathcal{G}^*),
$$

s.t. $\mathcal{H}G = \mathcal{U},$

97
where the $d \times N(K^2 - K)$ matrix $\mathcal{H}$ contains the $d$ distinctive linear constraints imposed on the parameter vector $G$, and $\mathcal{U}$ is the $d$-dimensional vector indicating the values of the linear constraints, then we have

$$
\hat{G} = G^* - \Sigma_G^* \mathcal{H}' (\mathcal{H} \Sigma_G^* \mathcal{H}')^{-1} \mathcal{H} G^* + \Sigma_G^* \mathcal{H}' (\mathcal{H} \Sigma_G^* \mathcal{H}')^{-1} \mathcal{U}.
$$

(4.22)

Particularly, equation (4.22) reduces to

$$
\hat{G}^{T-SAR} = G^* - \Sigma_G^* \mathcal{H}' (\mathcal{H} \Sigma_G^* \mathcal{H}')^{-1} \mathcal{H} G^*,
$$

when the model is a T-SAR model without any additional restrictions and thus $\mathcal{H} = \mathcal{H}$, $\mathcal{U} = 0_{N(K-1)}$.

Since $\hat{G}^{T-SAR}$ contains the estimation of all the unknown parameters of $\{G_n\}_{n=1}^N$ in equation (4.19), the estimations of $\{\rho_n\}_{n=1}^N$ and $\{W_n\}_{n=1}^N$ can be easily obtained by standardizing the rows of $\{\rho_n\}_{n=1}^N$ by their row sums: for all $n = 1, 2, \ldots, N$,

$$
\hat{\rho}_n = \frac{1_K \hat{G}_n 1_K}{K},
$$

and

$$
\hat{W}_n = \hat{\rho}_n^{-1} \hat{G}_n,
$$

where $\hat{\rho}_n$, $\hat{W}_n$, and $\hat{G}_n$ are respectively the estimations of $\rho_n$, $W_n$, and $G_n$ in regime $n$.

$\hat{W}_n$ displays how the variables or entities are interconnected in Regime $n$. Based on these results, we may carry out a class of hypothesis tests to draw clearer conclusions about the interconnections in different regimes.

### 4.4.2 Hypothesis Tests

In this section I present how to construct a broad class of hypothesis tests of additional restrictions in the T-SAR model.

Throughout this section I denote $G_n = \rho_n W_n$ for $n = 1, 2, \ldots, N$, the elements in
FROM STRUCTURAL MODELS TO SPATIAL MODELS

\{G_n\}_{n=1}^N are stacked in \(G\), and the free parameters \(\mathcal{G} = S\overline{G}\) satisfies the restrictions of a T-VAR model so that \(\mathcal{HG} = o_N(K-1)\).

In this paper, I focus on a special class of null hypotheses. Let \(V\) be a given \(K \times K\) matrix with the same row sums and zeros on all diagonal elements, and \(\Xi_1, \Xi_2, \ldots, \Xi_N\) be \(N\) given subsets of \(\mathbb{R}^1\). Then a null hypothesis considered in this paper is always such a proposition: there exists at least one group of coefficients 

\[
(\xi_1, \xi_2, \ldots, \xi_N) \in \Xi_1 \times \Xi_2 \times \cdots \times \Xi_N \subseteq \mathbb{R}^N
\]

such that

\[
\sum_{n=1}^N \xi_n G_n = V.
\]

Such propositions in fact cover a broad class of interesting hypotheses. For example, assume \(N = 2\). If we let \(V = V^{\text{Exo}}\) be a given spatial matrix, \(\Xi_1 \subseteq \mathbb{R}^1\), and \(\Xi_2 = \{0\}\), then the hypothesis equals 

\[H^\text{One-Regime}_0: W_1 = V^{\text{Exo}};\]

if \(V = O_K, \Xi_1 = \mathbb{R}^1, \text{and } \Xi_2 = \{-1\}\), then the hypothesis equals 

\[H^\text{Equal}_0: W_1 = W_2.\]

And if we let \(V = V^{\text{Exo}}, \Xi_1 = \Xi_2 = \mathbb{R}^1 \setminus \{0\}\), then under the null hypotheses, there exist \(\xi_1^{\text{Diff}}\) and \(\xi_2^{\text{Diff}}\) such that 

\[
W_2 = -\frac{\xi_1^{\text{Diff}}}{\xi_2^{\text{Diff}}} \rho_1 \rho_2 W_1 + \frac{1}{\xi_2^{\text{Diff}}} \rho_2 C_w^{\text{Diff}} W_1 + \frac{\rho_2^{\text{Diff}}}{} V^{\text{Exo}},\]

where \(\xi_1^{\text{Diff}} = -\frac{\xi_1^{\text{Diff}}}{\rho_1} \rho_2^{\text{Diff}}\), and \(\xi_2^{\text{Diff}} = \frac{1}{\rho_2^{\text{Diff}}}\). This implies \(W_2\) is a kind of “aggregation” of \(W_1\) and \(C_w^{\text{Diff}}\). I denote this null hypothesis as \(H^{\text{Diff}}_0\).

Taking an unrestricted or less restricted T-SAR model as the model under the alternative hypothesis, the null hypothesis is nested in the alternative hypothesis. Therefore, the null hypotheses can be tested with likelihood ratio (LR) tests against the less restricted model under the alternative hypothesis. Since \(\{G_n\}_{n=1}^N\) are estimated through the MD approach, we have to first obtain the estimations with additional restrictions under the null
hypotheses and then evaluate the likelihood of the estimations under the null hypotheses.

The main problem of the estimation of the restricted model under the null hypothesis is the nonlinearity of the additional restrictions. They appear to be linear but in fact are not, as some $\xi_n$’s are uncertain. This being said, we may still convert it to a linear problem. First rewrite the restrictions (4.23) in the vectorized form. Vectorizing both sides of equation (4.23), we get

$$\xi_1 \text{vec}(W_1) + \xi_2 \text{vec}(W_2) + \cdots + \xi_N \text{vec}(W_N) = \text{vec}(V).$$

Then let $\xi = (\xi_1, \xi_2, \ldots, \xi_N)'$. Deleting all the fixed diagonal elements and applying multiplication rules of block matrices, we obtain

$$(\xi' \otimes I_K) G = V,$$  \hspace{1cm} (4.26)

where $V = S \text{vec}(V)$ is the remaining components after deleting all the diagonal elements of $V$ from $\text{vec}(V)$.

However, we still cannot simply put the additional restrictions in equation (4.26) and the basic restrictions for a T-SAR model in equation (4.21) together and let the restriction matrix under the null hypothesis be the simple combination of the two groups of restrictions. This is because some restrictions in equations (4.26) and (4.21) are equivalent - in equation (4.21) we require the row sums of each $G_n, n = 1, 2, \ldots, N$ to be the same; and when the restrictions in equation (4.26) are satisfied, since $V$ has the same row sums, as long as $(N - 1)$ in the $N G_n$’s have the same row sums, the remaining one must have the same row sums as well. Formally, under the null hypothesis, if $V$ and $G_1, G_2, \ldots, G_{N-1}$ all have the same row sums so that $\mathcal{R} V_{1K} = \mathcal{R} G_1 1_K = \mathcal{R} G_2 1_K = \cdots = \mathcal{R} G_{N-1} 1_K = 0$, we have

$$\mathcal{R} G_N 1_K = \mathcal{R}(V - \xi_1 G_1 - \xi_2 G_2 - \cdots - \xi_{N-1} G_{N-1}) 1_K = 0.$$  

Therefore, we need to remove these redundant restrictions in order to estimate the restricted model under the null hypothesis. Actually, the final restrictions needed for the T-SAR model under the null hypothesis are constructed in the following way:
FROM STRUCTURAL MODELS TO SPATIAL MODELS

Without loss of generality, assume that \( \Xi_N \) is not restricted to \( \{0\} \). Denote

\[
\tilde{\mathcal{H}}^{\Lambda_1} = \begin{pmatrix}
\text{vec}(G_1^{\Lambda_1})_{1_{K-1}} \\
\text{vec}(G_2^{\Lambda_1})_{1_{K-1}} \\
\vdots \\
\text{vec}(G_{N-1}^{\Lambda_1})_{1_{K-1}}
\end{pmatrix},
\]

and

\[
\hat{\mathcal{H}} = (I_{N-1} \otimes (1_{K} \otimes \mathcal{R})) \odot \hat{\mathcal{H}}^{\Lambda_1},
\]

then the final restrictions of the model under the null hypothesis (4.23) are

\[
\mathcal{H}_t G = \mathcal{Y},
\]

where

\[
\mathcal{H}_t = \begin{pmatrix}
\tilde{\mathcal{H}} \\
\xi' \otimes I_K
\end{pmatrix},
\]

and

\[
\mathcal{Y} = \begin{pmatrix}
\mathcal{O}((N-1)(K-1)) \\
\mathcal{V}
\end{pmatrix}.
\]

It should be stressed again that the restrictions in equation (4.27) are not linear, because \( \xi \) are unknown parameters and \( \mathcal{H}_t \) depends on \( \xi \). But we can convert them to linear restrictions. If \( \xi \) is known, then we get an estimation of \( G \) by applying Proposition 3:

\[
\hat{G}^{\text{Ho}}_t = G^* - \Sigma^*_G \mathcal{H}_t' (\mathcal{H}_t \Sigma^*_G \mathcal{H}_t')^{-1} \mathcal{H}_t G^* + \Sigma^*_G \mathcal{H}_t' (\mathcal{H}_t \Sigma^*_G \mathcal{H}_t')^{-1} \mathcal{V},
\]

and \( \hat{G}^{\text{Ho}}_t \) is a function of \( \xi \). The final estimation, \( \hat{G}^{\text{Ho}}_t \), can be obtained by searching for a minimum of the weighted distance between \( \hat{G}^{\text{Ho}}_t \) and \( G^* \), adjusting \( \xi \) within the given space \( \Xi_1 \times \Xi_2 \times \cdots \times \Xi_N \). Hence, the estimation of the T-VAR model under the null hypothesis (4.23) is given as:

\[
\hat{G}^{\text{Ho}}_t = \arg \min_{\hat{G}^{\text{Ho}}_t, \xi \in \Xi_1 \times \Xi_2 \times \cdots \times \Xi_N} (\hat{G}^{\text{Ho}}_t - G^*)' \Sigma^{-1}_G (\hat{G}^{\text{Ho}}_t - G^*).
\]
Now that we have obtained $\hat{G}^H_0$, we try to test the restrictions against a less restricted model under the alternative hypothesis. Suppose the estimation of $G$ under the alternative hypothesis is $\hat{G}^H_1$. In this paper, I assume the model under the alternative hypothesis is also a T-SAR model with less or no additional restrictions in the form of equation (4.23). Therefore, $\hat{G}^H_1$ can be obtained and evaluated in the same way as $\hat{G}^H_0$, for which reason I only discuss $\hat{G}^H_0$ in the following paragraphs.

As mentioned above, I test the null hypothesis with an LR test. In order to construct the LR statistic, we must first evaluate $\hat{G}^H_0$ in the likelihood function. Recalling that the T-SAR model is nothing more than a constrained T-SVAR model, we can express the likelihood function with the same form as $L^{T-SVAR}$ and the same parameter sets $\Theta$, $\Gamma$, and $\Delta$. Thus, the likelihood function

\[
L^{H_0}(\hat{\Theta}^{H_0}, \hat{\Gamma}^{H_0}, \hat{\delta}^{H_0}; \hat{\epsilon}^{H_0}) = L^{T-SVAR}(\hat{\Theta}^{H_0}, \hat{\Gamma}^{H_0}, \hat{\delta}^{H_0}; \hat{\epsilon}^{H_0}),
\]

where $L^{T-SVAR}$ has been given in equation (4.17), $\hat{\Theta}^{H_0}$, $\hat{\Gamma}^{H_0}$, and $\hat{\delta}^{H_0}$ are the estimations of respective parameters under the null hypothesis. The problem is, $\hat{G}^H_0$ is merely a part of $\hat{\Theta}^{H_0}$, and the remaining parameters in $\Theta$, together with $\Gamma$ and $\delta$, are not obtained through the MD estimation. I try to solve this problem by returning to the ML problem

\[
L^{H_0}(\hat{\Theta}^{H_0}, \hat{\Gamma}^{H_0}, \hat{\delta}^{H_0}; \hat{\epsilon}^{H_0}) = \max_{\Theta \setminus G, \Gamma, \delta} L^{H_0}, \tag{4.28}
\]

where the set difference $\Theta \setminus G$ contains the parameters in $\Theta$ but not in $G$.

Given the large number of unknown parameters in the maximization problem (4.28), it is still a rather computationally demanding task.

Fortunately, for three of the most widely used specifications of the reduced-form model, I show that this procedure may be significantly simplified by a generalized least square (GLS) estimator of the autoregressive coefficients. In the first specification, all the autoregressive coefficients of the reduced-form model in different regimes are free to vary. In the second specification, the autoregressive coefficients of the reduced-form model are assumed to be the same across different regimes. And the third specification is a mix of the previous two.
In order to apply the GLS estimator, first let us restore the T-SAR model (4.19) to the reduced-form

\[ Y_t = \sum_{n=1}^{N} \left( c^{(n)} + \sum_{i=1}^{p} A_i^{(n)} Y_{t-i} + A_X^{(n)} X_t \right) \tau_n(Q; \delta) + r_t, \quad (4.29) \]

where the covariance matrix of \( r_t \) has been given by equation (4.14). For simplicity, in the following I denote \( \Sigma_{r,t}^jQ_t; \delta \) as \( \Sigma_{r,t} \). Moreover, let \( C \) collect all the unknown autoregressive coefficients \( \{c^{(n)}, A_i^{(n)}, A_X^{(n)} : n = 1, 2, \ldots, N, i = 1, 2, \ldots, p\} \), and \( \vartheta \) collect all the rest unknown parameters in \( (\Theta \setminus \mathcal{G}) \cup \Gamma \cup \delta \). We can see that \( \vartheta = \{A_{11}, A_{12}, \ldots, A_{1N}, A_{21}, A_{22}, \ldots, A_{2N}, \gamma_1, \gamma_2, \ldots, \gamma_N, \delta \} \).

With these preparations, let us discuss the following three specifications:

**Specification A.** In this specification, all the autoregressive coefficients of the reduced-form model in different regimes are free to vary. That is to say, for any \( n_1, n_2 = 1, 2, \ldots, N \) and \( i = 1, 2, \ldots, p \), we do not require \( c^{(n_1)} = c^{(n_2)} \), \( A_i^{(n_1)} = A_i^{(n_2)} \), or \( A_X^{(n_1)} = A_X^{(n_2)} \).

The reduced-form (4.29) collapses to

\[ Y_t = \sum_{n=1}^{N} C_n \tilde{Z}_t \tau_n(Q; \delta) + r_t, \quad (4.30) \]

where \( C_n = \{c^{(n)}, A_i^{(n)}, A_X^{(n)} : n = 1, 2, \ldots, N, i = 1, 2, \ldots, p\} \) is a stack of autoregressive coefficients, and \( \tilde{Z}_t = (1, Y_{t-1}^\prime, Y_{t-2}^\prime, \ldots, Y_{t-p}^\prime, X_t^\prime)^\prime \). Further supposing \( C^{\text{SpecA}} = (C_1, C_2, \ldots, C_N) \), and

\[ Z_t^{\text{SpecA}} = \tau_\delta(Q; \delta) \tilde{Z}_t = \begin{pmatrix} \tau_1(Q; \delta) \tilde{Z}_t \\ \vdots \\ \tau_N(Q; \delta) \tilde{Z}_t \end{pmatrix}, \]

equation (4.30) can be rewritten as

\[ Y_t = C^{\text{SpecA}} Z_t^{\text{SpecA}} + r_t. \]
Vectorizing both side, we get

\[ Y_t = (Z_t^{SpecA'} \otimes I_K) \text{vec}(C^{SpecA}) + r_t. \]

This is a typical multivariate regression model with time-varying error covariances.

Assuming that \( \vartheta \) is known, we can write the GLS estimator as

\[ \text{vec} \hat{C}^{SpecA}_{\vartheta} = (\Sigma_{Z}^{SpecA})^{-1} \text{vec}(\Sigma_{Y}^{SpecA}), \]

where

\[ \hat{\Sigma}_{Z}^{SpecA} = \frac{1}{T} \sum_{t=1}^{T} (Z_t^{SpecA} Z_t^{SpecA'} \otimes \Sigma_{r,t}^{-1}), \]

and

\[ \hat{\Sigma}_{Y}^{SpecA} = \frac{1}{T} \sum_{t=1}^{T} \Sigma_{r,t}^{-1} Y_t Y_t^{SpecA'}. \]

**Specification B.** In this specification, all the autoregressive coefficients of the reduced-form model are the same across different regimes. Therefore, for any \( n_1, n_2 = 1, 2, \ldots, N \) and \( i = 1, 2, \ldots, p \), we require \( c^{(n_i)} = c^{(n)} = c, A_i^{(n_i)} = A_i^{(n)} = A_i, \) and \( A_X^{(n_i)} = A_X^{(n)} = A_X \).

Then the threshold function can be removed from equation (4.29), yielding

\[ Y_t = c + \sum_{i=1}^{p} A_i Y_{t-i} + A_X X_t + r_t. \tag{4.31} \]

Now, let \( C^{SpecB} = (c, A_1, A_2, \ldots, A_p, A_X) \), and \( Z^{SpecB} = (1, Y_{t-1}', Y_{t-2}', \ldots, Y_{t-p}', X_t')' \), then the model (4.31) can also be written as

\[ Y_t = (Z_t^{SpecB'} \otimes I_K) \text{vec}(C^{SpecB}) + r_t. \]

The GLS estimator follows similarly to specification A:

\[ \text{vec} \hat{C}^{SpecB}_{\vartheta} = (\Sigma_{Z}^{SpecB})^{-1} \text{vec}(\Sigma_{Y}^{SpecB}), \]
where
\[ \hat{\Sigma}_Z^{\text{SpecB}} = \frac{1}{T} \sum_{t=1}^{T} (Z_t^{\text{SpecB}} Z_t^{\text{SpecB}'} \otimes \Sigma_{r,t}^{-1}), \]
and
\[ \hat{\Sigma}_Y^{\text{SpecB}} = \frac{1}{T} \sum_{t=1}^{T} \Sigma_{r,t}^{-1} Y_t^{\text{SpecB}}. \]

**Specification C.** It is also easy to specify the reduced-form model (4.29) as a mixture of specifications A and B. There are many possible specifications in this category; as an example, consider the following case:

Assume that the intercepts \( c^{(1)}, c^{(2)}, \ldots, c^{(n)} \) can freely vary in different regimes, whereas the other autoregressive coefficients are the same across different regimes. This is a typical varying-incept threshold model that is often seen in literature. Let \( A_i^{(n_i)} = A_i^{(n_i)} = A_i \), and \( A_X^{(n_i)} = A_X^{(n_i)} = A_X \). Similar with specifications A and B, let
\[ C^{\text{SpecC}} = (c^{(1)}, c^{(2)}, \ldots, c^{(N)}, A_1, A_2, \ldots, A_p, A_X), \]
and
\[ Z^{\text{SpecC}} = (\tau_1(Q_t; \delta), \tau_2(Q_t; \delta), \ldots, \tau_N(Q_t; \delta), Y_{t-1}', Y_{t-2}', \ldots, Y_{t-p}', X_t'), \]
then we can rewrite the model of this specification in the vectorized form again:
\[ Y_t = (Z_t^{\text{SpecC}'} \otimes I_K) \text{vec}(C^{\text{SpecC}}) + r_t. \]

The GLS estimator goes as always:
\[ \text{vec} \hat{C}_g^{\text{SpecC}} = (\hat{\Sigma}_Z^{\text{SpecC}})^{-1} \text{vec}(\hat{\Sigma}_Y^{\text{SpecC}}), \]
where
\[ \hat{\Sigma}_Z^{\text{SpecC}} = \frac{1}{T} \sum_{t=1}^{T} (Z_t^{\text{SpecC}} Z_t^{\text{SpecC}'} \otimes \Sigma_{r,t}^{-1}), \]
and
\[ \hat{\Sigma}_Y^{\text{SpecC}} = \frac{1}{T} \sum_{t=1}^{T} \Sigma_{r,t}^{-1} Y_t^{\text{SpecC}}. \]

Besides the three specifications discussed above, it is always possible to specify the
reduced-form model (4.29), but then we have to construct the GLS estimators for them ad hoc.

No matter what specification we use, we have now obtained the GLS estimation of vec(\(C\)), which is a function of unknown parameters \(\vartheta\). Patilea and Raïssi (2012) have proven that these GLS estimators are asymptotically consistent. Recalling that vec(\(C\)) is in fact a vector composed of the autoregressive coefficients \(C\), I denote it as \(\hat{C}_\vartheta\). It follows that

\[
L^{H_0}(\hat{\Theta}^{H_0}, \hat{\Gamma}^{H_0}, \hat{\delta}^{H_0}; \hat{\varepsilon}^{H_0})
= L^{H_0}(\hat{G}^{H_0}, \hat{C}^{H_0}, \hat{\vartheta}^{H_0})
= \max_{\vartheta} (\hat{G}^{H_0}, \hat{C}_\vartheta, \vartheta).
\]

As long as we obtain \(L^{H_0}(\hat{\Theta}^{H_0}, \hat{\Gamma}^{H_0}, \hat{\delta}^{H_0}; \hat{\varepsilon}^{H_0})\) and \(L^{H_1}(\hat{\Theta}^{H_1}, \hat{\Gamma}^{H_1}, \hat{\delta}^{H_1}; \hat{\varepsilon}^{H_1})\), the LR statistic is given by

\[
LR = 2(L^{H_1}(\hat{\Theta}^{H_1}, \hat{\Gamma}^{H_1}, \hat{\delta}^{H_1}; \hat{\varepsilon}^{H_1}) - L^{H_0}(\hat{\Theta}^{H_0}, \hat{\Gamma}^{H_0}, \hat{\delta}^{H_0}; \hat{\varepsilon}^{H_0})).
\]

LR can be compared with a \(\chi^2\) distribution. The degree of freedom of the \(\chi^2\) distribution can be calculated by comparing the numbers of restrictions of the null and alternative hypotheses. Taking the T-SAR model without any additional restrictions as the benchmark, the model under the null or alternative hypothesis with restrictions in equation (4.27) has at most \(K^2 - (K - 1)\) more restrictions on \(\mathcal{G}\) than the benchmark. But there are also a certain number of free parameters \(\xi_1, \xi_2, \ldots, \xi_N\), and they must be subtracted from the total number of additional restrictions. For example, under \(H^\text{Equal}_0\), since \(\xi\) is a free parameter, there should be \(K^2 - (K - 1) - 1\) more restrictions on \(\mathcal{G}\) than the benchmark. Following steps of the test do not generate more restrictions; therefore the final number of additional restrictions under a hypothesis in the form of equation (4.23) is \(K^2 - (K - 1)\) minus the number of free parameters in \(\xi_1, \xi_2, \ldots, \xi_N\). The degree of freedom of the \(\chi^2\) distribution follows.
4.5 Network Analyses

While the connectedness among economic entities can be described by the spatial weight matrices in the T-SAR model, there are ways to analyze the connectedness further. Noticing the similarity between a spatial structure and a network structure, networks can be built from the spatial weight matrices estimated in Section 4.4.1. After this, many methods proposed by the large amount of literature on networks can be applied to analyze the connectedness in details. This completes the final part of the tripod model.

Same as Scidá (2016), I take the transposes of the spatial weight matrices \( W'_1, W'_2, \ldots, W'_N \) as the network adjacency matrices \( A_1, A_2, \ldots, A_N \). That is to say, for each regime \( n = 1, 2, \ldots, N \), I assume there is a network defined by the adjacency matrix \( A_n = W'_n \). The transpose in fact arises from the different conventions in different strands of literature: in spatial econometrics, the element \( w_{n,ij} \) on the \( i \)-th row and \( j \)-th column of the spatial weight matrix \( W_n \) usually denotes the influence of the \( j \)-th variable on the \( i \)-th variable, whereas in network studies, the element \( a_{n,ij} \) on the \( i \)-th row and \( j \)-th column of the adjacency matrix \( A_n = (a_{n,ij})_{K \times K} \) usually denotes the influence of the \( i \)-th variable on the \( j \)-th variable. The diagonal elements of \( A_n \) are set to zeros, which means that there is no “loop” connecting a node to itself.\(^{16}\)

Adjacency matrices defined in this way have several significant features. First, they are usually asymmetric, which results in directed networks. Second, they are weighted networks, where the strength of each edge usually does not equal to one. Third, they are usually complete, which means usually every pair of nodes is somehow connected. Fourth, they often have negatively weighted edges. On the contrary, networks studied in traditional network studies typically lack these features. A large number of methods are devised for analyzing networks that are undirected, non-weighted, incomplete, and do not have negative edges. Certainly, there are also many ways to analyze networks with some of these features, but it is indeed not easy to find some method to analyze a network with all four features. In this paper, I show that it is still possible to apply several analytical techniques to these highly featured networks built from spatial weight matrices.

---

\(^{16}\)A network is composed of “nodes” and “edges”. Nodes are the entities involved in the network, and edges are the connections between the nodes.
First, centrality measures can be used to measure how well a node is connected to others. The simplest centrality measure is the degree centrality, which counts the number of neighbors of each node. This measure obviously does not work for a complete network as all nodes in a complete network have the same number of neighbors. Scidá (2016) also propose to use three other types of popular centrality measures, namely closeness centrality, betweenness centrality, and Bonacich power centrality. Among them the closeness centrality and the betweenness centrality do not work for networks with negative edges, because they evaluate the shortest path between each pair of nodes, and the common definition of “shortest path” becomes invalid when there are negative edges, through which the length of the path decreases. Finally, the Bonacich power centrality works only for networks where all edges are simultaneously positive or simultaneously negative.

In this paper, I propose to use two different approaches to measure the centrality of a complete network with both positive and negative edges. The first is intuitive. It generalizes the degree centrality to weighted networks. Since the networks we consider are also directed networks, we have both the “in-degrees” and the “out-degrees”, which respectively measure how much influence a node receives from other nodes, and how much influence a node exerts on other nodes. Formally, for node $k$ in network $n$ we have the in-degree

$$c_{d,in}^{n,k} = \frac{\sum_{i=1}^{K} a_{n,ik}}{K - 1},$$

and the out-degree

$$c_{d,out}^{n,k} = \frac{\sum_{i=1}^{K} a_{n,ki}}{K - 1}.$$
both positive and negative edges, following Bonacich and Lloyd (2004), Everett and Borgatti (2014) devises the PN centrality measure. The PN centrality is basically a generalized version of the Hubbell measures that considers the positive and negative edges separately and then sums both sides together. Formally, for network n, we have the in-PN measure

\[ C_{n,\text{in}}^{\text{PN}} = \left( I_K - \frac{1}{4(K-1)^2} A_n A_n' \right)^{-1} \left( I_K + \frac{1}{2(n-1)} A_n \right) 1_K, \]

and the out-PN measure

\[ C_{n,\text{out}}^{\text{PN}} = \left( I_K - \frac{1}{4(K-1)^2} A_n A_n' \right)^{-1} \left( I_K + \frac{1}{2(n-1)} A'_n \right) 1_K, \]

where both \( C_{n,\text{in}}^{\text{PN}} \) and \( C_{n,\text{out}}^{\text{PN}} \) are K-dimensional vectors containing the PN centralities for all the nodes \( 1, 2, \ldots, K \).

Furthermore, because the adjacency matrix \( A_n \), \( n = 1, 2, \ldots, N \) is always column-standardized, we will always get the same in-degree and in-PN centralities for all the nodes. Therefore, in this paper, we only consider the out-degree and out-PN centrality measures.

Besides centrality measures, the clustering analysis can also reveal some important aspects of a network. Different from the centralities, clustering is both a local analysis of the position of each node in the network and a global analysis of the overall connectedness of the entire network. For each node \( i = 1, 2, \ldots, K \), we have a clustering coefficient for the node. While the centrality measures how well the node is connected to other nodes, the clustering coefficient of the node measures how well its neighbors are mutually connected. If the neighbors of a node are closely connected, then the node has a larger clustering coefficient, and vice versa. Moreover, if we take the mean of the clustering coefficients of all the nodes, we get the average clustering coefficient for the whole network, which is a good measure of the overall connectedness of the whole network.

However, traditional measures of clustering coefficient do not work for a complete network. Since all the nodes in a complete network are connected, the clustering
coefficients of all the nodes will reach their maximum value, one. Therefore, the clustering coefficients for a complete network must be constructed in a different way.

In this paper, I use the clustering coefficient recently constructed by McAssey and Bijma (2015), which takes the edge weights into consideration and therefore works well for complete networks. Formally, for \( k = 1, 2, \ldots, K \) and \( n = 1, 2, \ldots, N \), the clustering coefficient of node \( i \) in network \( n \) is given as

\[
C_{nk} = \frac{(A_n + A_n')_{kk}^k}{2((A_n + A_n')(I_k - I_k)(A_n + A_n'))_{kk}},
\]

where the subscript \( kk \) denotes the \( k \)-th diagonal element of the matrix. Further, the average clustering coefficient for the whole network is

\[
\bar{C} = \frac{\sum_{k=1}^{K} C_k}{K}.
\]

Finally, Scidá (2016) suggests that the estimations of the scaling parameters in the T-SAR model, \( \hat{\rho}_n \), \( n = 1, 2, \ldots, N \) are themselves natural indicators of the intensity of the interactions among the variables in regime \( n \). This is another global measure of the overall connectedness among the variables.

With these network analytical techniques, the SVAR-spatial-network “tripod” model is completed. In the following section, I show how to apply this model to study contagion during recent Eurozone crises.

### 4.6 An Application: Financial Contagion in the Eurozone

#### 4.6.1 Introduction

In this section, I present an application of my model, which is on contagion among the stock markets in the Eurozone.\(^{18}\)

---

\(^{18}\)The application is mostly coded in R, together with a little C. I thank the authors of R and the R packages listed below: actuar, combinat, doParallel, fields, foreach, ggnetwork,
The Eurozone markets have been in turmoil for nearly a decade. At first, they were influenced considerably by the 2008 Financial Crisis; and not after long, the European countries found that they were in trouble again due to debt problems in a handful of countries such as Ireland and Greece. During this period, the contagion among international asset markets received much attention from both the academia and the public.

However, there is no consensus on the definition of contagion. I identify three categories of definitions. First, recent literature tend to define contagion as the intensification of shock transmissions among markets during turbulent periods. Especially, many studies model the contagion under this definition empirically by examining the structural breaks of bivariate correlations of assets, including Forbes and Rigobon (2002), Caporale et al. (2005), Chiang et al. (2007), and Caporin et al. (2013). Second, some authors anchor the definition of contagion on economic fundamentals, and define contagion as a shock transmission in excess of what can be explained by fundamentals. Dungey and Martin (2007) is an example of this category of literature. Third, some earlier literature, e.g. Edwards (1998), does not strictly differentiate contagion and spillovers. They do not assume that the shocks are transmitted in a different way during crisis periods. This is a very broad definition and rarely used in recent literature.

In this paper, I follow the first definition of contagion. I define contagion as the increase of connectedness among markets in the turbulent periods. However, my methodology is distinguished from most existing literature in this category, as I do not consider the statistical correlations among the markets; rather, I focus on the connectedness inferred from the contemporaneous impulse-response relationships with the threshold tripod model.19 My measure has several clear features. First, my model is based on contemporaneous impulse-response relationships, which show a clear causal map of the propagation of shocks. It is widely known that correlation is different from causality; so when we observe a correlation between two markets, we cannot know what causes the correlation. But in my model, it is easy to observe how the markets are connected by

19In fact, the correlation can be a measure of connectedness as well. For example, the weighted gene co-expression network analysis (WGCNA), which has been widely applied in bioinformatics, builds networks on correlations. It might be interesting to investigate market correlations from the network angle, too.
mutual shocks and responses. Related to this, while correlation analyses suffer from problems due to heteroskedasticity and must be corrected (see Forbes and Rigobon, 2001), my model accommodates the heteroskedasticity easily. Second, network analysis techniques can efficiently summarize the interrelationship among the markets as well as the unique roles played by each market in both tranquil and turbulent periods. Moreover, it is easy to see how the market interrelationship changes in different periods from a network angle, which provides us with clues for the channels of contagion. Third, my model allows the data to decide the crisis and non-crisis states of the market interrelationship. I also use a smooth threshold function, which does not draw a clear-cut separation between crisis and non-crisis periods. Therefore, I do not need to define the beginnings and ends of crises manually, thus making my measure of contagion more flexible and less arbitrary.

Following Kohonen (2013), I select five representative countries to study the contagion in Eurozone stock markets, namely Greece (GR), Germany (DE), Italy (IT), Spain (ES), and Ireland (IE). I let the data determine the contagion and non-contagion regimes according to the iTraxx Europe, a European credit default swap (CDS) index. With the threshold tripod model, I confirm the existence of contagion among the stock markets in these countries in recent years, and reject the hypothesis that the contagion is solely conducted by the cross-border interbank channel. Further analyses reveal the distinctive roles played by the five countries in both the contagion and the non-contagion regimes clearly.

4.6.2 Data and Model Setup

I study stock market logarithm returns in the five representative countries. The market indices are respectively the Athex Composite, the DAX 30, the FTSE MIB, the IBEX 35, and the Ireland SE Overall, all of which are obtained from Datastream. The observation period ranges from January 2nd, 2006 to May 26th, 2016, therefore there are 2,714 observations in total. For computational reasons, I multiply all the log returns by 100.

Figure 4.6.1 shows the trends of the data in the observed years. Note that all the data have been rescaled by their initial values. It is easy to notice that these markets are to a large extent

\[1^3\text{Higher iTraxx suggests higher CDS spreads for European financial assets, and therefore higher default risk in European markets.}\]
extent interrelated. Especially, all the countries saw a large simultaneous downturn during the 2008 crisis, and a relatively less significant one during the recent European debt crisis. Moreover, the German market and the Greek market show extremely different trends, which hint the different roles played by the two countries.

**Figure 4.6.1: Stock Market Indices**

This figure reports daily closing values from January 2nd, 2006 to May 26th, 2015 of the representative stock market indices in five Eurozone countries. All the data are rescaled by their initial values.

I assume there are two different regimes, and the regime with generally higher market connectedness is called “Regime 2”, and the other “Regime 1.” Contagion can be thus viewed as the difference of the market connectedness between the two regimes. I then choose the iTraxx Europe index as the threshold variable, which is retrieved from
Datastream. It measures the overall default risk in Europe. In fact, default risk is the most common crisis indicator used in contagion literature, e.g., Kalbaska and Gątkowski (2012), Manasse and Zavalloni (2013), among others. It is also used by De Bruyckere et al. (2013) in their study of the Eurozone contagion. Their historical trends are plotted in Figure 4.6.2. For the threshold function, I use the multivariate logistic function described in equation (4.10)

\[ \tau_1^{Logi}(iTraxx_t; \beta) = \frac{1}{1 + e^{-\beta_0 - \beta_1 iTraxx_t}}, \]  

and

\[ \tau_2^{Logi}(iTraxx_t; \beta) = 1 - \tau_1(iTraxx_t; \beta), \]

which allows the system to transition between each other smoothly. The location and speed of the transition are decided by the coefficients of the multivariate logistic function \( \beta_0 \) and \( \beta_1 \). Keeping other parameters constant, the larger the absolute value of \( \beta_1 \) is, the faster the system switches between the regimes.

**Figure 4.6.2: iTraxx Europe Values**

This figure reports daily closing values of iTraxx Europe from January 2nd, 2006 to May 26th, 2015.

All the autoregressive coefficients in the reduced-form VAR are assumed to be constant. This means the relationship between different markets are the same across different regimes barring the different instantaneous effects in the same day when shocks take place. This
assumption can be supported by the so-called “non-crisis-contingent”\footnote{This is the name given by Forbes and Rigobon (2001) to a class of economic theories. There is no relationship between this name and the title of this paper at all, as they refer to different things.} category of contagion theories in Forbes and Rigobon (2001), who argue that “any large cross-market correlations after a shock are a continuation of linkages that existed before the crisis,” such as trade. In fact, in this application part I test whether the Eurozone contagion relies on cross-border interbank market linkages. The lag order is set to one to rule out potential autocorrelation effects.

It should be noted that I do not include any factor that predicts the returns in this application. The tripod model framework definitely allows me to do so, as I always allow exogenous variables in the model description in previous sections.

4.6.3 The Local Identification

I first estimate the two-regime T-SVAR model of market log returns with constant reduced-form autoregressive coefficients, a multivariate logistic threshold function, and one lag. Without any additional information, I can achieve the local identifications. But due to the “label-switching” problem, these local identifications of $\Omega_1$ and $\Omega_2$ do not make economic sense at this moment, because we can always arbitrarily permute the columns of the local estimations of them.\footnote{For the meaning of the parameters, see Section 4.3.3.} Nevertheless, since we do not need the global identification to obtain the estimations of the parameters in the threshold functions and the mixture weights, we can still know now that the estimations of the mixture weights under Regimes 1 and 2 are respectively $\hat{\gamma}_1 = 0.310$ and $\hat{\gamma}_1 = 0.118$, and the estimations of the threshold function parameters in equation (4.32) are respectively $\hat{\beta}_0 = 2.703$, and $\hat{\beta}_1 = -0.037$, where the variable with hats are the estimations of respective parameters without hats. We can see that the coefficient for iTraxx, $\hat{\beta}_i$, is negative. This implies that an increase in iTraxx will lead to a higher probability for the system to be in Regime 2, where the markets have closer interconnections as I show in the following sections. The trend of the threshold function can be viewed more clearly in Figure 4.6.3. This figure reveals two major clusters of high-interconnection periods: the 2008 crisis, and the following Eurozone
crisis. Interestingly, during the period starting from late 2015, there is likely to be a recurrence of the Eurozone crisis, though not as significant as the previous two crises.

**Figure 4.6.3**: Values of the Threshold Function

This figure shows the probability for the system to be in Regime 1, which has generally closer market interconnections from January 2nd, 2006 to May 26th, 2015.

4.6.4 **From Local Identification to Global Identification**

Now that we have obtained local identification, we may proceed to the global identification of the T-SVAR model.

Based on the theoretical model of King and Wadhwani (1990), Kohonen (2013) suggests to use the magnitude of news in each market, which is extracted from Google Trends Indices (GTIs), as the proxy variable to attain the correct permutation of the columns of covariance matrix decompositions $\hat{\Omega}_1$ and $\hat{\Omega}_2$. The idea is simple: the GTI of a certain term indicates the relative volumes of weekly queries of the term on Google. Therefore, a more volatile GTI of a market-related term will imply stronger news in this period, and stronger news will lead to stronger innovations in asset returns, as are described by the diagonal elements of $\Sigma_1^{(i)} = \gamma_1 \Lambda_{11} + (1 - \gamma_1) \Lambda_{21}$ and $\Sigma_2^{(i)} = \gamma_1 \Lambda_{12} + (1 - \gamma_1) \Lambda_{22}$.

Once the correct $\hat{\Lambda}_{11}$, $\hat{\Lambda}_{21}$, $\hat{\Lambda}_{12}$ and $\hat{\Lambda}_{22}$ are pinned down, the column permutations of $\hat{\Omega}$, see equation (4.14).
and $\hat{\Omega}_2$ must follow, and vice versa. In fact, Google Trends has already been widely employed to proxy news magnitude and predict asset volatilities (e.g. Vlastakis and Markellos, 2012, Dimpfl and Jank, 2016). In this paper, the global identification is achieved with the help of GTI, too. The search terms are also the same as Kohonen (2013), which can be found in Table 4.6.1. The trends of the GTIs for these search terms can be found in Figure 4.6.4. Obviously, this plot shows the all the markets in this study receive extreme attention during the Eurozone crisis period.

**Table 4.6.1:** Search Terms for the Google Trends Indices

<table>
<thead>
<tr>
<th>Country</th>
<th>Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>Greece</td>
<td>Greece economy OR Greece debt OR Greece stock market</td>
</tr>
<tr>
<td>Germany</td>
<td>Germany economy OR Germany debt OR Germany stock market</td>
</tr>
<tr>
<td>Italy</td>
<td>Italy economy OR Italy debt OR Italy stock market</td>
</tr>
<tr>
<td>Spain</td>
<td>Spain economy OR Spain debt OR Spain stock market</td>
</tr>
<tr>
<td>Ireland</td>
<td>Ireland economy OR Ireland debt OR Ireland stock market</td>
</tr>
</tbody>
</table>

This table shows the search terms for Google Trends Indices for each country.

In order to apply the Google Trends data to the T-SVAR model, we have to first extract the information on news magnitude from the raw data series, and then use it to pin down the correct permutations of the columns of $\hat{\Omega}_1$ and $\hat{\Omega}_2$.

Considering that the increases of GTI are likely to be caused by news, whereas the decreases of GTI are usually results of no news, I define the news magnitude for a market as the mean of the positive logarithm changes of respective GTI. Moreover, as I consider two different regimes in my model, the news magnitude under each regime is actually measured by the weighted mean of the positive logarithm changes of GTI, where the weights are the probability for the system to be in respective regime at each time. Specifically, the news magnitude for market $i$ in Regime $n$ is given by

$$\text{NewsMagnitude}_{i,n} = \frac{\sum_{t=1}^{T} \tau_{n}^{t} \exp(i \text{Traxx}; \hat{\beta}) \max \{ \Delta \ln \text{GTI}_t, 0 \}}{\sum_{t=1}^{T} \tau_{n}^{t} \exp(i \text{Traxx}; \hat{\beta})}.$$

Since the GTI data are at a weekly frequency, there is a possibility that some news may have
This figure shows the Google Trends Indices for the search terms listed in Table 4.6.1 from January 2nd, 2006 to May 26th, 2015. All series keep their original scales.
long-tail influence in the following week. Therefore I also try to replace \( \Delta \ln \text{GTI}_t \) with the residuals of the AR(1) processes formed by them, and the results turn out to be the same. I report the results in Table 4.6.2.

**Table 4.6.2: News Magnitude**

<table>
<thead>
<tr>
<th>Regime</th>
<th>Country</th>
<th>Log Change Measure</th>
<th>AR(1) Residual Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>News Magnitude</td>
<td>Rank</td>
</tr>
<tr>
<td>1</td>
<td>Greece</td>
<td>0.1361</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Germany</td>
<td>0.0973</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Italy</td>
<td>0.0941</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Spain</td>
<td>0.0971</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Ireland</td>
<td>0.0826</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Greece</td>
<td>0.1123</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Germany</td>
<td>0.0808</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Italy</td>
<td>0.1013</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Spain</td>
<td>0.0984</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Ireland</td>
<td>0.0744</td>
<td>1</td>
</tr>
</tbody>
</table>

This table shows the news magnitude for each country in Regimes 1 and 2 estimated from logarithm changes of Google Trends Indices and the AR(1) residuals thereof. Then they are ranked from the smallest to the largest.

Now that we have the orders of the news magnitude for each market in under the two regimes, we need to find correct permutations of the columns of \( \hat{\Omega}_1 \) and \( \hat{\Omega}_2 \) such that the diagonal elements of \( \hat{\Sigma}_e^{(1)} \) and \( \hat{\Sigma}_e^{(2)} \) are in the same order as the news magnitude orders in Table 4.6.2. Since there are \( 5! = 120 \) possible column permutations either for \( \hat{\Omega}_1 \) or \( \hat{\Omega}_2 \), what we need to do now is to calculate the corresponding \( \hat{\Sigma}_e^{(1)} \) and \( \hat{\Sigma}_e^{(2)} \) of all the 120 possible permutations. After that, we can pick out the permutations which correspond to \( \hat{\Sigma}_e^{(1)} \) and \( \hat{\Sigma}_e^{(2)} \) with correct ordering of diagonal elements. Luckily, in this example, for both Regimes 1 and 2, there are unique correct permutations of \( \hat{\Omega}_1 \) and \( \hat{\Omega}_2 \). The unique global identification of the T-SVAR model is thus achieved.

Once we achieve the global identification, we can calculate the estimations of \( \Upsilon_1 \) and \( \Upsilon_2 \) in equation (4.15), which immediately gives us the instantaneous response of each variable to unit-sized structural shocks. The estimation results are reported in Table 4.6.3. The upper half of the table corresponds to \( \hat{\Upsilon}_1 \) and the lower half to \( \hat{\Upsilon}_2 \). For each half, the element in
i-th row and j-th column indicates the instantaneous response of the i-th market to the unit-sized structural shock originating from Market j. For example, the element in the 1st row and 2nd column in the upper half of the table is -0.319, then it means in Regime 1, when a unit-sized structural shock takes place in the German market, the Greek instantly has a response that has a size of -0.319. The off-diagonal elements actually capture the cross-market spillover effects under each regime. It is interesting to note that for both regimes the response of the Greek market to positive German news and the response of Irish market to Spanish news is negative. This may reflect the roles played by the two countries in the observation periods as the likely places where contagion can originate.

**Table 4.6.3: Instantaneous Responses to Unit-Sized Structural Shocks**

<table>
<thead>
<tr>
<th>Matrix</th>
<th>Responding Variable</th>
<th>Structural Shock Label</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GR</td>
<td>DE</td>
</tr>
<tr>
<td>( \tilde{\Upsilon}_1 )</td>
<td>GR</td>
<td>3.355</td>
</tr>
<tr>
<td></td>
<td>DE</td>
<td>1.268</td>
</tr>
<tr>
<td></td>
<td>IT</td>
<td>1.270</td>
</tr>
<tr>
<td></td>
<td>ES</td>
<td>1.141</td>
</tr>
<tr>
<td></td>
<td>IE</td>
<td>0.978</td>
</tr>
<tr>
<td>( \tilde{\Upsilon}_2 )</td>
<td>GR</td>
<td>3.642</td>
</tr>
<tr>
<td></td>
<td>DE</td>
<td>1.376</td>
</tr>
<tr>
<td></td>
<td>IT</td>
<td>1.378</td>
</tr>
<tr>
<td></td>
<td>ES</td>
<td>1.239</td>
</tr>
<tr>
<td></td>
<td>IE</td>
<td>1.061</td>
</tr>
</tbody>
</table>

This table shows estimations of \( \tilde{\Upsilon}_1 \) and \( \tilde{\Upsilon}_2 \), which stand for the instantaneous responses of each market return to unit-sized structural shocks. The upper half of the table corresponds to \( \tilde{\Upsilon}_1 \) and the lower half to \( \tilde{\Upsilon}_2 \). For each half, the element in i-th row and j-th column indicates the instantaneous response of the i-th market to the unit-sized structural shock originating from Market j. For example, the element in the 1st row and 2nd column in the upper half of the table is -0.319, then it means in Regime 1, when a unit-sized structural shock takes place in the German market, the Greek instantly has a response that has a size of -0.319.

### 4.6.5 The Spatial Model

Now that we have obtained the estimations of the T-SVAR model, we can proceed to the T-SAR model implied by the T-SVAR by applying Propositions 2 and 3. The MD
estimations of $W_1$ and $W_2$, the spatial weight matrix under Regimes 1 and 2 can be found in Table 4.6.4.

**Table 4.6.4: Estimations of Spatial Weight Matrices**

<table>
<thead>
<tr>
<th>Matrix</th>
<th>Spatial Weight, To</th>
<th>Spatial Weight, From</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{W}_1$</td>
<td>GR</td>
<td>DE</td>
</tr>
<tr>
<td>GR</td>
<td>0</td>
<td>0.144</td>
</tr>
<tr>
<td>DE</td>
<td>0.037</td>
<td>0</td>
</tr>
<tr>
<td>IT</td>
<td>0.054</td>
<td>0.238</td>
</tr>
<tr>
<td>ES</td>
<td>0.040</td>
<td>0.092</td>
</tr>
<tr>
<td>IE</td>
<td>0.059</td>
<td>0.081</td>
</tr>
</tbody>
</table>

| $\hat{W}_2$ | GR | DE | IT | ES | IE |
| GR | 0 | 0.360 | 0.013 | 0.034 | 0.568 |
| DE | 0.278 | 0 | 0.330 | 0.312 | 0.079 |
| IT | 0.482 | -0.041 | 0 | -0.014 | 0.573 |
| ES | 0.435 | 0.281 | 0.050 | 0 | 0.233 |
| IE | 0.241 | 0.707 | 0.216 | -0.164 | 0 |

This table shows estimations of the spatial matrices $W_1$ and $W_2$. The upper half of the table corresponds to $\hat{W}_1$ and the lower half to $\hat{W}_2$. Since these matrices are asymmetric, I differentiate the weights from a country to the weights to a country.

It is much easier to interpret the spatial weights than $\Sigma_1$ and $\Sigma_2$. Notice that the rows in $W_1$ and $W_2$ indicate how strong the influences from other countries can affect a certain country, and the columns indicate how strong the influences from the corresponding country to other countries. Now compare $\hat{W}_1$ and $\hat{W}_2$ in Table 4.6.4 column by column, and it is not difficult to find some interesting facts.

First, in Regime 1 the Greek market has quite small influences on every other market, as none of the weights of Greece to other markets exceeds 0.1. However, in Regime 2, the influences of the Greek market surge to more than 0.2. This finding confirms the common impression that Greece plays a very important role in exporting contagion during the recent crises. At the same time, many would also have the impression that Germany is the main supporter of the countries in crises, and this is confirmed by the second column in $\hat{W}_2$ as well, as the weights of Germany to all other markets except Italy are quite large. Especially, it is Ireland, rather than Greece, receives the largest influence from Germany, and the weight 0.71 is surprisingly high, which hints a very close Germany-Ireland financial
bond. However, the significant role played by Germany seems to only exist in Regime 2, and in Regime 1, German influences are much smaller. In other words, Germany is like the firefighter for the Eurozone who only appears when there is fire. As are shown by the third and fourth column, the influences of the Italian and the Spanish markets generally decay from Regime 1 to Regime 2, which reflects that the crises in both countries are relatively weaker. Under Regime 2, although the closeness from Ireland to Greece and Italy increases, the closeness from Ireland to both Germany and Spain decreases, and in general, the role of the Irish market does not seem to change too much across different regimes. A panoramic view of the interconnections among the five countries revealed by the spatial weigh matrices can be found in Figure 4.6.5. We can easily find in this figure that the lines in the right sub-plot representing Regime 2 are darker than the left sub-plot, which indicates generally closer connections among variables in Regime 2.

Figure 4.6.5: An Overview of the Connectedness among the Five-Country System

This figure is an overview of the five-country system studied in this paper. The lines represent the closeness among the countries revealed by the T-SAR model. Dashed lines indicate negative values. A darker line indicates a closer connection between the countries connected by the line. And the sizes of the balls where country names are printed show the PN-centralities of each country.
I then try to test whether the interconnections among the markets are statistically different in different regimes. The null hypothesis is thus set as

\[ H_0^{\text{Equal}} : \quad W_1 = W_2, \]

versus the alternative hypothesis \( H_1 \) : the T-SAR model without any additional restriction on \( W_1 \) and \( W_2 \). This hypothesis is exactly the example shown in equation (4.24) in Section 4.4.2. Since the reduced-form autoregressive coefficients are assumed to be constant across regimes in this application, we can use the results of specification B in Section 4.4.2. For the test of this null hypothesis, we have \( LR = 3.133 \times 10^8 \). Compared with a \( \chi^2 \)-square distribution with 20 degrees of freedom, the null hypothesis \( H_0^{\text{Equal}} \) is rejected with the \( p \)-value less than \( 1 \times 10^{-3} \).

Next, I present a test for a more informative null hypothesis: the contagion, defined as the difference between the connectedness of the markets under Regimes 1 and 2, is spread through the cross-border interbank market network. If we regard Regime 1 as the regime without contagion, and Regime 2 as the regime with contagion, then this hypothesis can be tested with the T-SAR model. Specifically, in Regime 1, shocks are spread through the network described by \( W_1 \), therefore \( W_1 \) represents natural non-contagion-contingent interconnections among the markets even when there is no contagion. Under Regime 2, the non-contagion-contingent interconnections represented by \( W_1 \) are still there. However, now there is the contagion. Besides the existing interconnections \( W_1 \), shocks can now be spread through an additional new contagion-contingent network, which is denoted as \( \mathcal{V}_{\text{Interbank}} \). As a result of the co-existence of the non-contagion-contingent network \( W_1 \) and the contagion-contingent network \( \mathcal{V}_{\text{Interbank}} \) in Regime 2, interconnections among the markets now become \( W_2 \). Therefore, if we assume the contagion is spread through network \( V \), then we should be able to express \( W_2 \) as a linear combination of \( W_1 \) and \( \mathcal{V}_{\text{Interbank}} \). Formally, we can write the null hypothesis as \( H_0^{\text{Diff}} \) : there exists at least one \( \xi_1^{\text{Diff}} \in \mathbb{R}^{1} \) and one \( \xi_2^{\text{Diff}} \in \mathbb{R}^{1} \) such that

\[ W_2 = \xi_1^{\text{Diff}} W_1 + \xi_2^{\text{Diff}} \mathcal{V}_{\text{Interbank}}. \]

We can see that this null hypothesis is also exactly the example shown in equation (4.25) in Section 4.4.2. As long as \( \mathcal{V}_{\text{Interbank}} \) is given, this null hypothesis can be tested against \( H_1 \) :
there is not any additional restriction on $W_1$ and $W_2$ in the T-SAR model.

In this paper, I construct $V_{\text{Interbank}}$ from the cross-border interbank claim data reported by Bank for International Settlements in the same way as Tonzer (2015). Suppose the element in the $i$-th row and $j$-th column of $V_{\text{Interbank}}$ is $v_{ij}^{\text{Interbank}}$. Then

$$v_{ij}^{\text{Interbank}} = \frac{\sum_{t=1}^{T} \text{ForeignClaim}_{ijt}}{\sum_{t=1}^{T} \sum_{j \neq i} \text{ForeignClaim}_{ijt}},$$

(4.33)

where $\text{ForeignClaim}_{ijt}$ is the interbank positions of country $i$’s banking system towards banks in country $j$ at time $t$. The actual value of $V_{\text{Interbank}}$ is reported in Table 4.6.5. Not unexpectedly, all the countries have large claim proportions in Germany and small claim proportions in Greece.

**Table 4.6.5: Value of $V_{\text{Interbank}}$**

<table>
<thead>
<tr>
<th></th>
<th>GR</th>
<th>DE</th>
<th>IT</th>
<th>ES</th>
<th>IE</th>
</tr>
</thead>
<tbody>
<tr>
<td>GR</td>
<td>0</td>
<td>0.690</td>
<td>0.136</td>
<td>0.060</td>
<td>0.114</td>
</tr>
<tr>
<td>DE</td>
<td>0.028</td>
<td>0</td>
<td>0.350</td>
<td>0.377</td>
<td>0.245</td>
</tr>
<tr>
<td>IT</td>
<td>0.052</td>
<td>0.841</td>
<td>0</td>
<td>0.119</td>
<td>0.033</td>
</tr>
<tr>
<td>ES</td>
<td>0.008</td>
<td>0.506</td>
<td>0.372</td>
<td>0</td>
<td>0.113</td>
</tr>
<tr>
<td>IE</td>
<td>0.070</td>
<td>0.431</td>
<td>0.287</td>
<td>0.211</td>
<td>0</td>
</tr>
</tbody>
</table>

This table reports the value of $V_{\text{Interbank}}$ used in this paper, For the meaning of the elements, see equation (4.33).

Now, with the help of results of Specification B in Section 4.4.2, we obtain the LR statistic for the test of $H_{0}^{\text{Diff}}$: $LR = 2.979 \times 10^6$. Compared with a $\chi^2$-square distribution with 21 degrees of freedom, the null hypothesis $H_{0}^{\text{Equal}}$ is rejected with the $p$-value less than $1 \times 10^{-3}$. Note that even after the adjustment due to the difference in degrees of freedom between $H_{0}^{\text{Equal}}$ and $H_{0}^{\text{Diff}}$, the LR statistic for $H_{0}^{\text{Diff}}$ is still far larger than $H_{0}^{\text{Equal}}$. This implies that it is even farther from the fact to assume the contagion is purely spread through the cross-border interbank network than to assume there is no contagion at all.

However, it should be noticed that by this test I do not argue that the Eurozone contagion is not spread through the cross-border interbank network. I only argue that $V_{\text{Interbank}}$ cannot
solely explain the channel of the Eurozone contagion. Actual contagion channels may be obtained by compounding $V^{\text{Interbank}}$ with some other networks.

### 4.6.6 Network Analyses

In this part I present the results of the network analyses. All the measures mentioned in this section are out-measures.

Centrality measures are shown in Table 4.6.6. This table can be regarded as a formal summary of the analyses of the spatial weight matrices in the previous section. Those findings are all reconfirmed here. First, the centrality of Greece grows tremendously from Regime 1 to 2, and the centrality of Germany also sees a large increase. These are consistent with the impression about the important roles played by Greece and Germany in the recent Eurozone crisis. The centralities of Italy and Spain drop significantly from Regime 1 to 2, and Ireland maintains highly central in the network in both regimes. All these findings are in full accordance with the analyses of spatial weight matrices in Section 4.6.5. Additionally, in Regime 2, the two most central countries are Greece and Ireland. We can say Greece and Ireland are the two most important origins of contagion among the selected five countries. These findings do not vary no matter whether we use the simple degree centrality measure or the PN centrality measure, which does not only measure how much influence a country has on its neighbors, but also how much influence the country’s neighbors have on neighbors’ neighbors.

**Table 4.6.6: Centrality Measures**

<table>
<thead>
<tr>
<th>Country</th>
<th>Degree Measure</th>
<th>PN Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Regime 1</td>
<td>Regime 2</td>
</tr>
<tr>
<td>GR</td>
<td>0.111</td>
<td>1.071</td>
</tr>
<tr>
<td>DE</td>
<td>0.326</td>
<td>0.975</td>
</tr>
<tr>
<td>IT</td>
<td>1.182</td>
<td>0.473</td>
</tr>
<tr>
<td>ES</td>
<td>0.409</td>
<td>0.126</td>
</tr>
<tr>
<td>IE</td>
<td>0.911</td>
<td>1.084</td>
</tr>
</tbody>
</table>

This table reports the degree and PN centralities for each market in different regimes.
Then I present the results of clustering analyses in Table 4.6.7. First look at the average clustering coefficients of the whole network. The result here is not surprising at all. Since average clustering coefficients measure the overall connectedness of the entire network, it is no wonder that the markets in the five countries become more interconnected in Regime 2, which is associated with crises. However, when we look at the clustering coefficients of each individual country, we will notice more interesting facts that cannot be revealed by centrality analyses. In Regime 1, the country with the highest clustering coefficient is Greece. Recall that a clustering coefficient indicates how the neighbors of a node are interconnected, the high clustering coefficient of Greece actually suggests that it plays a rather peripheral role in Regime 1 as other countries are closely interconnected even without Greece. In contrast, Italy has the smallest clustering coefficient in Regime 1, which may be a result of its strong centrality under the regime. However, when the system enters Regime 2, the clustering coefficients of these countries change drastically. Italy has become the most peripheral country with the highest clustering coefficient; and now it is Germany that has the smallest clustering coefficient and contributes the most in conglutinating other countries together. Moreover, although the centrality of Ireland increases from Regime 1 to 2, its clustering coefficient increases, too. This hints that the importance of Ireland may not necessarily increase from Regime 1 to 2.

<table>
<thead>
<tr>
<th>Country</th>
<th>Regime 1</th>
<th>Regime 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>GR</td>
<td>1.092</td>
<td>0.891</td>
</tr>
<tr>
<td>DE</td>
<td>1.043</td>
<td>0.867</td>
</tr>
<tr>
<td>IT</td>
<td>0.535</td>
<td>1.373</td>
</tr>
<tr>
<td>ES</td>
<td>1.000</td>
<td>1.203</td>
</tr>
<tr>
<td>IE</td>
<td>0.711</td>
<td>0.885</td>
</tr>
<tr>
<td>Network Average</td>
<td>0.876</td>
<td>1.043</td>
</tr>
</tbody>
</table>

This table displays the clustering coefficient for each country as well as the network average clustering coefficient in different regimes.

And finally, the estimations of $\hat{\rho}_1$ and $\hat{\rho}_2$ measure the interconnection intensity among the markets as well. For Regime 1, the estimation $\hat{\rho}_1 = 0.876$, and for Regime $2, \hat{\rho}_2 = 1.044$. For another time, we find the markets in Regime 2 are more closely interconnected than in
A broad class of economic and financial models involves the investigation of connectedness of economic variables or entities. Historically, different strands of literature, such as vector auto-regressions, spatial models, and network models are all related to connectedness in different ways. In recent years, there have been a few pioneering attempts to integrate these different strands of literature in one unified framework. This paper attempts to bridge the structural vector auto-regression, spatial auto-regression, and network models together.

In many cases, we care about not only static interconnections, but also transitions cross different regimes of connectedness in different circumstances. A typical example is the empirical modeling of financial contagions, which are usually defined as the difference between the interconnections among entities under in different financial environments, rather than the interconnections as such. Catering to this need, I try to upgrade the SVAR-spatial-network “tripod” model with a transition mechanism by introducing thresholds into the model, which provide considerable convenience for studying the dynamics of connectedness of a system.

The SVAR-spatial-network “tripod” model in this paper can be broken down to three major parts, namely the SVAR model, the spatial model, and the network analyses. For the SVAR part, first I show how to identify SVAR model without strong prior assumptions on the causal structure by assuming normal mixture distributions of error terms. This establishes the groundwork for the whole model. Second, I introduce dynamics to the SVAR model by upgrading the SVAR to a threshold SVAR (T-SVAR) model. The threshold function relies on some exogenous variables and can take various convenient forms. For the part of spatial modeling, I demonstrate that a threshold spatial auto-regression (T-SAR) is equivalent to the T-SVAR model with certain linear restrictions. This allows us to estimate the spatial weight matrices from the T-SVAR model even without any prior knowledge of the spatial weights. Spatial weights are good measures of the connectedness. I then show how to use the T-SAR model to construct likelihood-ratio tests of a broad class of hypotheses. Finally,
based on the intrinsic similarity between spatial models and networks, I show that we can use some newly devised techniques for network analyses to reveal additional information of the model.

I then present an application of the model to the analysis of contagion during recent crises in the Eurozone. I select the stock markets in five countries, namely Greece, Germany, Italy, Spain, and Ireland as representative markets. In the SVAR part, I establish a crisis indicator based on a European CDS index, which determines the probability for an observations to be in crisis according to different interconnection patterns of the five markets. Also, with the help of the normal mixture assumption of error terms, I show how the markets respond instantaneously to shocks in each regime. In the spatial part, I obtain the estimations of spatial weight matrices in the T-SAR model in different regimes. The spatial weight matrices are good measures of the connectedness among the five markets. Furthermore, I construct two formal statistical tests. Defining the contagion as the difference between the connectedness of the five-country system under the two regimes, the first test rejects the hypothesis that contagions do not exist; and the second test rejects the hypothesis that the contagion channel can be solely explained by a cross-border interbank network in Europe. In the end, I carry out centrality, clustering, and $\rho$-analyses of the networks revealed by the T-SAR model. Network average clustering coefficients and $\rho$-s in different regimes further confirm that the markets are interconnected differently in different regimes; and centrality measures and individual clustering coefficients show the distinctive roles played by the selected five markets in different regimes.

Certainly, the application to study the contagions in the Eurozone is still far from the depletion of the potential of my model, and there is still considerable room for the improvement of the model in the future. For example, the spatial auto-regression model considered in this paper has a very simple form compared with many other spatial-time series models. Viewing from the angle of the network literature, the network considered in this paper is also quite simple among the immense world of multilayer networks. There is much left for future research.
4.A Mathematical Proofs

4.A.1 Proof of Proposition 2

Let $\mathcal{H}_m = \left(1^\prime_K \otimes \mathcal{R}\right) \otimes \text{vec}(G_n^{\prime 1})_{1_{K-1}}$, then we have

$$\mathcal{H} = \text{diag}(\mathcal{H}_{11}, \mathcal{H}_{22}, \ldots, \mathcal{H}_{NN}).$$

Further let $G_n$ be the $((n-1)(K^2 - K) + 1)$-th to the $n(K^2 - K)$-th components of $G$, then

$$\mathcal{H}G = \begin{pmatrix}
\mathcal{H}_{11}G_1 \\
\mathcal{H}_{22}G_2 \\
\vdots \\
\mathcal{H}_{NN}G_N
\end{pmatrix}.$$

By Theorem 3.1 in Scidá (2016), we know that for $n = 1, 2, \ldots, N$, $\mathcal{H}_{nn}G_n = o_{K-1}$ is equivalent to $RA_{n1_K} = o_{K-1}$, then $\mathcal{H}G = o_{N(K-1)}$ is equivalent to

$$RA_{1_NK} = \begin{pmatrix}
RA_{11_K} \\
RA_{21_K} \\
\vdots \\
RA_{N1_K}
\end{pmatrix} = o_{N(K-1)}.$$

4.A.2 Proof of Proposition 3

The optimization of the MD program can be solved by the Lagrangian method. Suppose that $\mathcal{U}$ has $U$ components. Consider the Lagrangian function

$$\mathcal{L} = (\hat{G} - G^*)\Sigma_{\hat{G}}^{-1}(\hat{G} - G^*) - \ell'(\mathcal{H}\hat{G} - \mathcal{U}),$$
wherer $\ell$ is the Lagrangian multiplier. First-order conditions follow:

$$\frac{\partial L}{\partial \hat{G}'} = 2\Sigma_\hat{G}^{-1}(\hat{G} - G^*) - \mathcal{H}'\ell = o_{N(K'-K)},$$

(4.34)

and

$$\frac{\partial L}{\partial \ell} = \mathcal{H} \hat{G} - \mathcal{U} = o_{\mathcal{U}}.$$  

(4.35)

Premultiplying equation (4.34) by $\mathcal{H}\Sigma_{\hat{G}}^*$, we get

$$2H(\hat{G} - G^*) - \mathcal{H}\Sigma_{\hat{G}}^*\mathcal{H}'\ell = \mathcal{U}.$$ 

Using equation (4.35), it follows that

$$\ell = 2(\mathcal{H}\Sigma_{\hat{G}}^*\mathcal{H}')^{-1}\mathcal{H} \hat{G} - 2(\mathcal{H}\Sigma_{\hat{G}}^*\mathcal{H}')^{-1}\mathcal{U}.$$ 

Substituting it back to equation (4.34), we have

$$\hat{G} = G^* - \Sigma_{\hat{G}}^*\mathcal{H}'(\mathcal{H}\Sigma_{\hat{G}}^*\mathcal{H}')^{-1}\mathcal{H}G^* + \Sigma_{\hat{G}}^*\mathcal{H}'(\mathcal{H}\Sigma_{\hat{G}}^*\mathcal{H}')^{-1}\mathcal{U}.$$ 

(4.36)

Following equation (4.36), it is obvious that

$$\hat{G}^{T-SAR} = G^* - \Sigma_{\hat{G}}^*\mathcal{H}'(\mathcal{H}\Sigma_{\hat{G}}^*\mathcal{H}')^{-1}\mathcal{H}G^*,$$

when $\mathcal{H} = \mathcal{H}$, and $\mathcal{U} = o_{N(K-1)}$. 

\[\square\]
Financial markets are like test tubes where traders and assets react day by day. Often, it is easy to see the outcome of a financial “reaction”, but the mechanism of the reaction remains unclear. This problem is particularly difficult because financial market variables usually respond to shocks very promptly. As a result, causalities will be lost in the intervals among observations due to problems such as reverse causality, if an empirical study uses relatively low-frequency data.

This thesis shows how novel methodologies can bring people new knowledge about causalities in financial markets with novel identification strategies. In two different approaches, Chapters 2 and 3 investigate the relationship between traders’ behaviors and
CONCLUSION

recent large price volatilities in major commodity markets. Chapter 4 develops a novel econometric model that portrays dynamics of interconnections among financial entities. As an illustration, this chapter also shows how the model shed new light on the contagion problem during the recent Eurozone crisis.

In the first chapter of this thesis we find that the recent price patterns of commodity futures markets can be partly attributed to the concurrence of trading strategies of positive feedback traders and commodity index traders, and the theoretical model about positive feedback tradings proposed by De Long et al. (1990) fits the historical market data well. We show that the United States Commodity Futures Trading Commission (CFTC) nonreportable long to total open interests ratios can be a decent proxy for the positive feedback tradings as nonreportable traders prefer more long short positions, and such a strategy empirically proves to be unwise because it will be taken advantage of by rational arbitrageurs and significant losses will follow. As a byproduct of rational arbitrageurs’ utilization of the positive feedback strategy, commodity prices will also depart from fundamental levels. Furthermore, expected price effects of positive feedback trading relies on the presence of passive investors who decide demands negatively related with prices to keep a constant risk-bearing capacity of the economy. We discover that commodity index investors rightly traded in this way and accordingly their changes of demands are followed by commodity return changes in the same direction.

While Chapter 1 relies on the assumption of the De Long et al. (1990) model and classical linear regressions, Chapter 2 studies the mutual interactions between traders’ behaviors and the pricing of a most important commodity – crude oil in a model-free approach. This chapter designs a novel identification strategy for structural vector autoregression models, which enables us to capture (some of) the short-run dynamics between variables that we observe at a lower frequency. This identification strategy relies on an underlying Markov-Switching VAR model with two or more volatility states and the assumption that the impulse-response functions have the same shape across different states. If we are interested in the relationship between two variables, the mutual instantaneous impulse-responses between them can be identified with an additional “catalyst” variable, which only interacts instantaneously with one of two variables and not with the other. Likewise, we can identify a system with more variables, with the inclusion of other catalyst
variables. In the application to the crude oil futures market, we use inventory as the catalyst variable. Based on the estimated index trader positions and other weekly data, we find that an index trader inflow shock can cause a positive response of the WTI futures excess returns instantly, and the response remains significant in several following weeks. Commodity index traders themselves also positively respond to shocks of crude oil futures excess returns instantly, though the responses will not last for long and are only significant in the immediate week. These findings are robust even after we add some extra control variables that are likely to affect oil prices and index traders into our model.

A broad class of economic and financial models involves the investigation of connectedness of economic variables or entities. Different strands of literature, including vector auto-regressions, spatial models, and network models are all related to the connectedness in different ways. Some pioneering literature has attempted to integrate these different strands of literature in one unified framework. This unified model can be called the SVAR-spatial-network “tripod” model.

However, sometimes we care about not only static interconnections, but also transitions cross different regimes of connectedness in different circumstances. A typical example is the empirical modeling of financial contagions, which are usually defined as the difference between the interconnections among entities under in different financial environments, rather than the interconnections as such. Catering to such problems, Chapter 5 upgrades the SVAR-spatial-network “tripod” model with a transition mechanism by introducing thresholds in the model, which provide considerable convenience for studying the dynamics of connectedness of a system.

The model in this chapter can be broken down to three major parts, namely the SVAR model, the spatial model, and the network analyses. For the SVAR part, first I show how to identify SVAR model without strong prior assumptions on the causal structure by assuming normal mixture distributions of error terms. This establishes the groundwork for the whole model. Second, I introduce dynamics to the SVAR model by upgrading the SVAR to a threshold SVAR (T-SVAR) model. The threshold function relies on some exogenous variables and can take various convenient forms. For the part of spatial modeling, I demonstrate that a threshold spatial auto-regression (T-SAR) is equivalent to
the T-SVAR model with certain linear restrictions. This allows us to estimate the spatial weight matrices from the T-SVAR model even without any prior knowledge of the spatial weights. Spatial weights are good measures of the connectedness. I then show how to use the T-SAR model to construct likelihood-ratio tests of a broad class of hypotheses. Finally, regarding network analyses, based on the intrinsic similarity between spatial models and networks, I show that we can use some newly devised techniques for network analyses to reveal additional information of the model. I then present an application of the model to the stock markets in selected European countries, which sheds new light on the contagion during recent crises in the Eurozone.

Many problems in finance revolve around causalities that cannot be easily captured by regular methods. But causality is always there. People should not fear problems such as reverse causality so much. As this thesis shows, there are, and will be more new ways to identify causalities from a given dataset.
A financial phenomenon is often easy to see, but not so easy to understand. However, for a policy maker, good policies must be made based on the sufficient understanding of a phenomenon, which requires knowledge on not only what the phenomenon is like, but also how it takes place.

Chapters 2 and 3 provide a good example. The large price volatility in recent years has raised wide public concern. While the price fluctuation concurs with flows of certain classes of traders such as commodity index traders, this thesis shows that the causality between traders’ behaviors and the price pattern can be not as simple as it looks like. Chapter 3 finds that the causality goes bidirectionally between commodity index traders and crude oil
futures prices, and Chapter 2 shows that the picture can be even more complex – indeed, commodity index traders play an indispensable role in the formation of commodity prices, but the causality only exists with the co-presence of positive feedback traders. In fact, in the mechanism revealed in Chapter 2, the more immediate cause of the deviation of commodity prices from the fundamentals is the positive trading. If a policy maker finds large volatilities of commodity prices undesirable, based on the findings in this thesis, he or she can find cleverer ways to regulate index traders than adding simple trading restrictions. For example, he or she can make better conditions for rational speculators to correct the pricing of commodities by asking traders to report more information to the public.

The fourth chapter of this thesis studies the contagion during recent Eurozone crises with a novel econometric model. I pick the stock markets in five countries, namely Greece, Germany, Italy, Spain, and Ireland as representative markets. I establish a crisis indicator based on a European CDS index, which decides the probability for an observation to be in crisis according to different interconnection patterns of the five markets. Also, with the help of the normal mixture assumption of error terms, I make a clear record of the spillover effects among the markets in each regime. Then I obtain the measures of the connectedness among the five markets. Furthermore, I construct two formal statistical tests. Defining the contagion as the difference between the connectedness of the five-country system under the two regimes, the first test rejects the hypothesis that contagions do not exist; and the second test rejects the hypothesis that the contagion channel can be solely explained by a cross-border interbank network in Europe. In the end, centrality and clustering analyses of the networks revealed by the model further show the distinctive roles played by the selected five markets in different regimes. This chapter shows different aspects of interconnections among selected Eurozone stock markets based on causalities, which helps making policies more pointed to the source and channel of the contagion.

The contribution of this thesis is both empirical and methodological. While the empirical findings in this thesis may help policy makers make better policies concerning commodity markets and the contagion problem in particular, the application of the methodologies proposed by this thesis is in no way limited to these two issues. They can be used in studying the mechanism of many similar issues that are related to the interaction between traders and markets or among different markets. Good policies can always be made upon
good understanding of mechanisms. In this sense, the impact of the methodological part of this thesis is even broader.
References


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


Hang Sun, also known as Sūn Háng, was born on 21st June, 1990 in Laixi, China. He earned his bachelor in Statistics from Wuhan University in 2011. In the same year, he also received bachelor in Accounting from Wuhan University. The titles of the theses for the two bachelor degrees are respectively “Extreme Value Theory and the Estimation of Value-at-Risk” and “On Accounting Principles for Small Business.” In 2013 he obtained his Master in International Economics from Wuhan University with the master thesis titled “Causalities in the Crude Oil Market: A Counter-Factual Approach.”

Hang joined the Finance Department of Maastricht University to pursue a doctorate degree under the supervision of Prof. dr. Jaap W. B. Bos and co-supervision of Dr. Paulo Rodrigues. There he supervises master theses and “marble” bachelor theses.

Hang’s research has been presented at a series of conferences including the 2014 European Association of Agricultural Economists Congress 2014, the 2016 China Finance Annual Meeting, the 2017 Annual Symposium of the Society for Nonlinear Dynamics and Econometrics, and the 2017 Barcelona GSE Summer Forum.

Staring from September, Hang will be Tenure-Track Lecturer in Finance at Dongbei University of Finance and Economics.