# Identification and Robustness in Central Banking and Supply Chain

by

Bomin Jiang

Submitted to the Institute for Data, Systems, and Society
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Social and Engineering Systems
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
September 2021



Author . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Institute for Data, Systems, and Society
July 30, 2021

Certified by . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Professor Munther A. Dahleh
William Coolidge Professor of Electrical Engineering and Computer Science
Thesis Supervisor

Certified by . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Professor Roberto Rigobon
Society of Sloan Fellows Professor of Management
Thesis Supervisor

Certified by . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Professor Hazhir Rahmandad
Schussel Family Professor of Management Science
Thesis Committe Member

Accepted by . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Professor Fotini Christia
Chair, Social and Engineering Systems Doctoral Program

**Identification and Robustness in Central Banking and Supply Chain**

by

Bomin Jiang

Submitted to the Institute for Data, Systems, and Society
on July 30, 2021, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Social and Engineering Systems


## Abstract

In this thesis, we study identification and robustness issues in central banking and supply chains. First, we present a methodology to identify multiple linear financial networks when only an aggregate outcome is observed, and use the method to assess financial networks among financial institutions in the United States. We discover that the data is better explained by a mix of distinct networks, each of which corresponds to a different transmission mechanism. Second, we investigate the effect of bounded uncertainty in central banking, and derive robust decision rules for central banking policymaking. When bounded uncertainty is passed through a conditional expectation channel, we find that committing not to use a policy tool is sometimes optimal for the central bank. An asset purchasing model and a forward guidance model are examined in depth to illustrate our point. Third, we study a stylized supply chain model where large aggregate shock hits and prices are not adjusted due to anti-price gouging laws. We show that individual producers, in this case, will not diversify for aggregate shock due to the externality of fixed prices. Multinational corporations, on the other hand, would still diversify their supply chain due to continuation value. Furthermore, a robustness analysis shows that individual producers will not diversify even when they adopt robust decision rules, but multinational corporations will further diversify their supply chain in this case. The first two chapters tackle the real-world challenge of financial systemic risk reflected by the 2008 financial crisis, and the proliferation of monetary policy tools thereafter. The third chapter tries to analyze the supply chain disruption caused by the outbreak of the Covid-19 global pandemic, and give policy suggestions based on our model.


Thesis Supervisor: Professor Munther A. Dahleh
Title: William Coolidge Professor of Electrical Engineering and Computer Science

Thesis Supervisor: Professor Roberto Rigobon
Title: Society of Sloan Fellows Professor of Management

Thesis Committe Member: Professor Hazhir Rahmandad
Title: Schussel Family Professor of Management Science

# Acknowledgments

First and foremost, I would like to express my sincere gratitude to my advisors Munther Dahleh and Roberto Rigobon, for their guidance, feedback, and support. This work could not have been possible without their suggestions, patience, and encouragement. It has been a true privilege to have both of them as my thesis supervisors. I am so grateful for the freedom and trust Munther gave me to pursue my own research directions. Apart from guidance on specific research problems, he also helped me develop a whole set of tools and skills to find meaningful questions, define problems, and pursue solutions. I cannot thank Roberto enough for his supervision of my Ph.D. research. Working with him has been a real honor and joy; I don't recall ever leaving his office without feeling inspired. He showed me the importance and fascination of macroeconomic research.

I am also very fortunate to have Hazhir Rahmandad on my thesis committee. Although he joined the committee relatively late compared to Munther and Roberto, he asked insightful questions and provided helpful comments to the thesis work. I am also grateful to Brian Anderson, who taught me all the fundamentals of research. His supervision of my Honours thesis at The Australian National University has been invaluable throughout my doctoral studies at MIT. In addition, I would like to acknowledge faculties at MIT that I had a chance to discuss research problems with, including Anna Mikusheva, Victor Chernozhukov, Jean-Noël Barrot, Ali Jadbabaie, Gonzalo Cisternas, and Jonathan Parker.

LIDS has been a friendly home, and I have learned a lot from my fellow students that surround me. It was also a pleasure to learn alongside my colleagues in the SES cohort: Amir, Elizabeth, Paolo, Manxi, and Max, especially when we were overwhelmed by the long list of coursework to take. Fellow students at Munzer's research group also played an important role, particularly my officemates: Flora, Maryann, and Manon.

Thank you also to the organizations that funded my work throughout this PhD: the MIT Presidential Fellows program, WorldQuant LLC, and the Class of 1960

Alumni Fund.

During my graduate studies, my friends in the Boston area gave a significant amount of support and enjoyment. We've gone skiing at Mt Sunapee and played ice hockey during the winters, and we have also biked around Martha's Vineyard and played soccer during the summers.

Last but not least, my most profound appreciation goes to my parents for their unconditional love and support. I am so pleased to be accompanied by my wife Jiali Wang over the years who always cheers me up and helps with my research.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

## 1.1 Problem Summary

The 2008 world financial crisis started in what was supposed to be a relatively small and segmented market. At that time, the outstanding value of Sub-Prime debt was a tad below a trillion dollars, while the market value of credits in both the formal and shadow US financial sectors was higher than 23 trillion dollars. So, a default on 4 percent of an isolated market (the high-risk mortgages) was expected to have an equal minuscule effect. Not surprisingly, the Obama administration only requested 800+ billion dollars for the rescue package. It is evident, today, that the systemic consequences of the sub-prime crisis were unexpectedly more substantial. The shock propagated to other financial sectors and countries. The total world losses reached several trillion. This phenomenon has spurred research on systemic risk, particularly on the estimation of the underlying financial networks governing the propagation of shocks — or as it is sometimes known as financial contagion.

After the above financial crisis, central banks worldwide embarked on a massive easing of monetary policy. They started relying on the monetary policy tool they used for decades: open market operations, lowering the interest rate in the short run. Very soon, interest rates got close to zero, and in many countries, it got to precisely zero. Under the assumption that negative rates were unfeasible, [1] central banks

[1] This assumption was proven to be wrong many years later when Switzerland, Denmark, and

started to expand their monetary policy tools and deploy alternative interventions. For example, one of those tools, the forward guidance, is a central bank's commitment to keeping interest rates around zero for an extended period. If the market trust this commitment, then this provides monetary liquidity and expands the economy (Campbell et al., 2012). However, the forward guidance tool did not incentivize the economy and push inflation consistently above the 2% target as expected. The Phillips curve, as part of the widely accepted New Keynesian Model of central banking, has been reported to be weakly identified and has become flatter and flatter over the years (Mavroeidis et al., 2014). As such, we believe that the problem of model uncertainty and robustness need to be carefully investigated in central banking models.

Optimal control theory in engineering has an extensive influence on economic models. However, optimal control assumes that the system dynamics are accurately known. When there is model uncertainty, robust control is preferred. Robust control optimizes the controller performance for a set of models instead of a single nominal model. In this thesis, I borrow from robust control theory from engineering to investigate central banking model uncertainty issues. In the robust control theory, we no longer assume perfect knowledge of the model. Instead, we focus on a baseline model of the system and then incorporate model uncertainty. After that, the optimal policy for the worst-case scenario is obtained via minimax optimization. Some counter-intuitive policy implications arise from our analysis using robust control.

Considering robustness in economic modeling is not entirely new. In the same year of the financial crisis, Lars P. Hansen, a laureate of the Nobel Memorial Prize in Economic Sciences, published his book 'Robustness', summarizing some research of robust control techniques applied to economic models. Since then, people have been using those techniques to revisit economic models. Generally speaking, a baseline statistical model is considered, and a multiplier penalty is given according to the entropy deviation (i.e., Kullback–Leibler divergence) is attached to the optimization objective function. However, due to the lack of robustness tools with expectation channels, not much has been done on central banking's robustness issues.

---

other European countries set interest rates to negative numbers.

Precisely, we consider the following three different but related identification and robustness issues in central banking and supply chain:

The first problem is the identification of financial networks. In this thesis, I develop a methodology to estimate hidden linear networks when only an aggregate outcome is observed. The aggregate observable variable is a linear combination of the different networks, and it is assumed that each network corresponds to the transmission mechanism of different shocks. The identification problem is challenging because the number of the parameter is large. An EM algorithm is implemented to estimate the networks, and an analytic solution of a Wishart MLE is obtained to enable the estimation of networks one by one. Finally, we implement the methodology to estimate financial networks among US financial institutions. Credit Default Swap rates are the observable variable, and we show that more than one network is needed to understand the dynamic behavior exhibited in the data.

The second problem is the robustness of the central banking policy. After the 2008 financial crisis, more and more unconventional monetary policy tools are used by central banks worldwide. The underlying assumption supporting those moves is that the central bank can do a better job controlling the economy by using more tools simultaneously. However, as shown in this thesis, that is not true when model uncertainty and an expectation channel exist simultaneously. The robust monetary policy is, therefore, to commit never to use certainty policy tools. Our model implication is aligned with many empirical studies; as they found, US forward guidance has a mixed effect on the economy: on the one hand, it signals a weak macroeconomic perspective and slows down consumer spending; on the other hand, it could provide more liquidity to the market and stimulate economic activities. Data show that both effects are significant, and the combined result can be either (Andrade et al., 2019). Intuitively, the central banking system and market are playing a game, and both rely on the expectation of the opponent's strategy to generate its strategy. As such, a new tool introduced by the central bank could change the expectation of strategies and shift the existing equilibrium of the game. With model uncertainty, the change of equilibrium could offset the original intention of introducing the tool, and the combined effect could go either way. A new theoretical model is provided to capture such intuition to make robust monetary policies in the future.

The third problem is the robustness issue of the global supply chain. During the global pandemic, it is evident that the global supply chain is not optimized for robustness. This thesis introduces the idea that a decentralized supply chain could become vulnerable to aggregate shocks in the pursuit of individual efficiency. Furthermore, when the model has an uncertainty, and all agents adopt a min-max robust decision rule, the equilibrium result of the decentralized supply chain deviates even further away from the socially optimal allocation. In other words, in our setup, we reflect the well-known tradeoff between efficiency and robustness. This thesis will present research ideas to identify financial networks and understand the impact of model uncertainty on central banking and supply chain. I will apply statistical tools and optimization methods to identify financial networks and apply tools from game theory and robust control theory to answer essential central banking and supply chain questions. Through this research, I strive to investigate the challenges and implications brought by considering model uncertainty. Furthermore, practical recommendations for central bankers, other policymakers, and regulators are provided.

The thesis uses research tools from macroeconomics, game theory, econometrics, optimization, and robust control theory. The first part of the thesis proposes a way to identify multiple financial networks. The identification process uses Identification through Heteroskedesiticity, a tool from econometrics, and EM-algorithm, a widely adopted optimization method. The second and third part of the thesis applies robust design approaches to existing central banking and supply chain models, combining the robustness analysis from control theory with macroeconomic models.

## 1.2 Summary of Individual Chapters

The remaining chapters of the thesis are organized as follows:

### 1.2.1 Contingent Linear Financial Networks

As part of the response to the financial crises in the developed world, policymakers have significantly increased central banks' regulatory authority. Nowadays, one of the essential roles of a central bank is the regulators for systemic risks. In particular, the estimation of the underlying financial networks has received more and more attention.

Observing and understanding the contracts underlying the links between banks is sometimes challenging, however. There are two mainstream approaches in the literature: either using financial contracts to estimate asymmetric networks or using aggregate data to estimate symmetric networks. When using the first approach, direct lending from one bank to another is straightforward to document, while relationships across other contracts are usually missed out. When using the second approach, directional information is never recovered. Furthermore, because financial connections generally depend on the type of shock that hits the market, an identification method that can identify multiple networks corresponding to different shocks is required. However, to the best of my knowledge, no existing identification methods can achieve that.

The conventional method usually cannot identify multiple networks at once due to two difficulties. One is that the data provide not enough sufficient statistics. We solve this problem by assuming heteroskedasticity regimes in the data and use that to provide constraints. The other difficulty is due to a large of parameters and the nonconvexity of the optimization problem. We solve this problem by using an EM algorithm and developing a procedure that decouples the estimation process of different networks.

Chapter 2 develops a methodology to estimate hidden linear networks when only an aggregate outcome is observed. We implement the methodology to estimate financial networks among US financial institutions. After controlling for aggregate shocks, we find the data is explained more accurately with a mixture of different networks, where each network corresponds to a different transmission mechanism. Credit Default Swap rates are the observable variable and we show that more than one network is needed to understand the data. Furthermore, multiple network identification is implemented using a EM algorithm. We developed a systematic methodology to find the optimal number of networks and discussed how it can help macro-prudential policy making. The estimation of different networks allows the implementation of a macro prudential and robust approach to the systemic risk in the financial system. We show that if the monetary authority estimates a single network it will underestimate the systemic risk of the top 10 banks in the US by 71 percent.

### 1.2.2 Uncertainty and Robustness in Central Banking

While these policies played an essential role in the economic recovery after the financial crisis, it is an important question whether more tools can always help the central bank control the economy. At first glance, having more tools at the Central Bank's disposal is always better. However, there are cases where the Central Bank should commit never to use a given mechanism to better manage the economy. The presents of bounded uncertainty instead of risk is the key to this phenomenon. Unlike risk, which has a distribution, bounded uncertainty we study in this thesis does not have a distribution, and a robust decision rule has to apply.

Furthermore, bounded uncertainty is not uncommon in macroeconomic models. A significant number of macroeconomic models require linearization around an equilibrium before being calibrated; hence, different equilibrium values would result in different models and different implications. However, equilibrium values are generally difficult to measure since the market is constantly evolving, and equilibrium is never indeed achieved in the real world. In fact, the Humphrey—Hawkins report of the Fed uses the word 'long term values' to indicate the equilibrium value that the economy would hypothetically achieve in the very far future if the current environment lasts. As a result, equilibrium values are usually reported as a range in which predictions based on various models fall.

Chapter 3 seeks to exam whether more policy tools are always better in an environment with bounded uncertainty. We find that when bounded uncertainty is passed through a conditional expectation channel, sometimes it is optimal for a central bank to not using a policy tool. Two monetary policy models for asset purchasing and forward guidance are examed in detail, and the cause for this phenomenon is discussed.

### 1.2.3 Robust Global Supply Chain

During the 2019-2020 coronavirus outbreak, we learned a valuable lesson about the global supply chain. Producing all of the parts with the low-cost suppliers is not optimal. Diversification is required to reduce risk. It is also evident that the global supply chain is not diversified for the kind of aggregate risk. This phenomenon is, however, not what the classic asset pricing theory predicts. In the canonical macro and asset pricing models, the demand is such that when the quantities tend to zero, prices rise to infinity. Those models with Cobb-Douglas or CES functions have a pricing system reflecting the scarcity. When an aggregate shock hits, the state should have a higher Arrow-Debru price that incentivizes market participants to diversify production to higher-cost regions. This diversification, on the other hand, maximizes social welfare.

This efficient allocation relies on the assumption of complete markets. In particular, it relies on a floating price. During the Covid-19 outbreak, it was evident that those assumptions do not hold due to anti-price gouging efforts from both the government and the general public. In this case, the companies that diversify their production in advance will not be rewarded with higher profit. Even worse, they will face increasing litigation risk of anti-price gouging enforcement. In this case, diversifying and preparing for aggregate shocks is not financially optimal for individual companies.

Chapter 4 studies a very stylized model of a supply chain, where we study how the decision of a multinational corporation changes in the presence of anti-price gouging and uncertainty. This chapter developed a simple model where large multinationals, even with the fixed price, still diversify their production plant to higher cost regions because of continuation values. Our model is inspired by a class of adaptive market and probability matching literature. It implies that in a market with aggregate shock and anti-price gouging laws, having a small number of large firms is more efficient than having a large number of small firms. We also study a robust supply chain which concentrates on the worst-case, and is not achieved by simply increasing the size of shocks. Our model rationalizes "probability matching" behavior observed in the experimental literature.

# Chapter 2

# Contingent Linear Financial Networks

*This work was performed in collaboration with Roberto Rigobon and Munther Dahleh.*

## 2.1 Introduction

The 2008 world financial crisis incentivized much research in financial systemic risk, which is caused by interconnections among large financial entities. Banks can be interconnected through many different channels. One type of link is related to the exposure banks have to similar microeconomic or industry shocks. For example, two bank's balance sheets can be interconnected because both are lending to the same firm or sector, and it suffers a shock. The second type of channel is related to interbank contracts. Two banks can be interrelated because one bank lends to the other, or they hold each other's liabilities. Therefore, a deterioration in the balance sheet of the borrowing bank affects the quality of the assets of the lending bank. Finally, banks can be exposed to similar macroeconomic shocks such as exchange rate, inflation rate, interest rates, economic activity, real estate, etc. In sum, there are many possible ways in which banks are linked to each other.[1] One implication

[1]See Allen and Gale (2000), Freixas et al. (2000) for earlier contributions, and Acemoglu et al. (2015), Allen et al. (2012), Caballero and Simsek (2013), Cabrales et al. (2017), Elliott et al. (2014), Gai and Kapadia (2010), and Gai et al. (2011) for recent theoretical papers. Empirical papers that

of the different transmission mechanisms is that they could be reflected in different financial networks.

The information required to estimate these financial networks is quite extensive and unlikely to be obtained, especially when these connections are not written in formal contracts.[2] The granularity and detail of the information, as well as the global disclosure requirements, are unfeasible, even to regulators and central banks. Therefore, any estimation of a financial network tends to rely on outcome variables such as public financial statements and market data from banks and financial institutions. Those outcomes capture the total impact confounded through the different networks as opposed to pairwise relationships.

In order to approximate the underlying network without using pairwise data, two main approaches have been taken in the literature. The first one uses information-theoretic principles and aggregate data to fill in the blanks.[3] For example, one can read the total inter-bank debt from Form 10-Ks. After that, the worst-case scenario can be evaluated by inferring the pairwise debt structure using a maximum entropy principle. While used in some stress tests, those methods only give a rough estimate of the debt structures and do not fully utilize the information from time-varying data. The second approach uses the correlation of prices of financial contracts (Onnela et al., 2004) to construct interbank financial networks. However, correlation matrices do not have a structural interpretation[4] and cannot reflect the direction of shock transmission.

The main contribution of the chapter is the estimation of a mixture model of multiple endogenous asymmetric linear networks.[5] The intuition behind the estima-

---

estimate financial networks include Billio et al. (2010), Merton et al. (2013), Adrian and Brunnermeier (2016), and Girardi and Ergün (2013).

[2]For instance, assume a bank lends to Intel, who sells to a South Korean firm, that manufactures a monitor sold to a California firm, who gets a loan from a different bank. In this setting, the two banks are related through the South Korean firm. However, the requirements on reporting and disclosure to be able to uncover such relationship are impossible — i.e., the US regulator can't force the South Korean company to reveal its clients and suppliers.

[3]See Upper (2011); Elsinger et al. (2013) for maximum entropy and Anand et al. (2015) for minimal density

[4]Additionally, the estimates based on correlations can be biased in the presence of heteroskedastic shocks. See Forbes and Rigobon (2002).

[5]The procedure follows Rigobon (2003) and is able to deal with asymmetric responses, contem-

tion comes from the fact that in a linear network, the correlation structure across banks is the result of a linear mixture of the covariance implied by each network - where the weights are the relative variances of each shock. When the variance of those shocks changes through the sample, the weights change, leading to variation in the aggregate covariances. From that variation of the second moment, the underlying structure of the networks can be recovered. Therefore, the identification requires finding periods where the shocks experience different volatilities. To do so, we use statistical identification where the changes in the variance in the observed data determine the "regimes". This procedure is very similar to the one developed by Sentana and Fiorentini (2001). The advantage of this methodology is that it allows estimating the asymmetry of large versus small shocks - where the statistical identification truly captures the differences in the propagation mechanisms. In the case of financial systemic risk both macroeconomic and idiosyncratic shocks might be important. In this chapter, we concentrate on the transmission of idiosyncratic shocks once the macroeconomic shocks have been controlled for.

We apply our methodology to estimate the US financial network among the 10 largest financial institutions in the country. The data collected are their credit default swaps (CDSs). The CDSs have been used widely in the literature studying risk contagion. Most notably, the CDSs of sovereign debt have been examined by various papers to understand international risk propagation. E.g., Kalbaska and Gątkowski (2012) investigates the eurozone contagion via a regression on CDS spread changes. In general, if a risk contagion mechanism exists among certain countries, their CDS spread will co-move. From the trend of CDS spread, the author reaches a conclusion that Sovereign debt risk is mainly limited to EU countries. Similarly, Caporin et al. (2018) also studies European debt crisis by examining the CDS spread. Through a Bayesian quantile regression that incorporates shock heteroskedasticity, they conclude that the increases in the correlation of CDS come from heteroskedasticity instead of structural changes of risk propagation mechanism. In addition, Eichengreen et al. (2012) studies the Subprime Crisis, most notably the market regime shift and risk

poraneous relationships, and parameter instability.

contagion during that time, using CDS data.

This chapter develops a methodology to estimate a mixture of multiple linear financial networks. On the theoretical part, we proved the solution uniqueness of identifying a single network through heteroskedasticity using GMM, and derived an Expectation-Maximization based method to identify multiple linear networks. On the empirical part, we applied our method on credit default swap data of US banks, and find that three networks best describe the data. Furthermore, each one has a different centrality ranking.

Identifying multiple networks offers central bankers the opportunity to design the proper policy response in case a systemic failure exists. The central bank is unlikely to know exactly which network is at play when the shock hits the economy. Therefore, a robust approach is desirable. When there are multiple networks a robust policy making process requires central bankers to stay ready for the worst case. We compare the systemic impact of banks when a robust response is used in the presence of three networks, versus the response when a single network is estimated. We find that the systemic impact is underestimated severely when a single network is used. Indeed the systemic impact is on average 41 percent smaller with one network that with three networks. A robust approach would imply a more supportive monetary policy reaction than what the single network entails. In other words, this chapter provides an improved method of identifying financial networks and monitoring systemic risk, and helps central banks build a more robust policy making process.

This chapter is organized as follows: In Section 2.2, a problem of the financial network is formulated and the problem of identification is discussed. The section starts with our network model, and then provides a proof of solution uniqueness in the single network case. Section 2.3 discusses the case when a single network is estimated. We present the statistical identification method and explains how heteroskedastic regimes are defined. Section 2.4 presents a test for the single network hypothesis, and a EM algorithm to estimate multiple networks. In the 10 banks case, 3 networks is required to explain the data. The results of the estimated networks is then discussed in this section. Finally, Section 2.5 concludes.

## 2.2 Modeling Financial Networks

The estimation of financial networks is not new. A significant proportion of papers, however, concentrate on the estimation of symmetric single networks — non-directional, and non-contingent graphs. In this section, we present first, some preliminary evidence based on correlations. We then introduce a linear financial network model, and show: 1. how the network model can be uniquely identified with variance-covariance data; 2. together with heteroskedastic shocks, how it can explain the volatility in correlation models.

### 2.2.1 The Time-Varying Correlations

It is desired to find out how and why banks are related to each other, and which banks are "crucial" in the risk transmission mechanisms, and which ones are less relevant. One natural measure for those purposes is the correlation coefficients. In modern portfolio theory (MPT), correlation coefficients also play an essential role. MPT quantitatively formalizes the concept of diversification via the statistical notion of covariance, or correlation.

However, most of those applications of linear correlation models make an over-simplified assumption, that the correlation coefficients of any two given financial instruments are time-invariant, or at least time-invariant in the period of analysis. [6] In practice, there are numerous examples where such assumption does not apply. In this chapter, we are concerned with the systemic financial risks which primarily propagate through large banks, so we use the Credit Default Swap (CDS) data of the top 10 US banks as an example. [7]

[6] Although most of the literature estimates the strength of the network transmission by using simple correlations, there are notable exceptions that are worth highlighting. Adrian and Brunnermeier (2016) proposes a new measure of comovement called CoVar — which is defined as the value at risk conditional on the bank being in distress. Girardi and Ergün (2013) extend that measurement to expand the definition of distress. These types of measures are consistent with networks being contingent.

[7] It is important to highlight that CDS might exhibit excessive comovement due to the presence of a government guarantee. (Merton et al., 2013) uses a different approach to measure the credit risk of the banks. They use contingent claims analysis instead of CDS. The CDS are partially guaranteed by government policy — for instance, deposit insurance. Future research should evaluate the robustness

The CDSs of banks have been widely used in studying interbank risk propagation. As pointed out by Eichengreen et al. (2012), the CDS spreads of major banks co-move and reflects market economic prospects. Furthermore, during the Subprime Crisis and the following crisis of Lehman Brothers, the common factor in the factor model become more dominant, i.e., the absorption ratio (as defined in (Kritzman et al., 2011)) is higher. The fact that major financial crises are reflected by CDS regime changes is the exact characteristic we want to study. In our contingent linear financial network model, a significant regime shift is required for unique identification. As shown in (Eichengreen et al., 2012), the common factor accounted for 62% of the variance of major bank CDSs before the 2007 breakout of Subprime Crisis and raised to 77% during the crisis, signaling a major regime shift. Apart from stronger co-movements, we also want to discover whether the change of regime is a sole result of heteroskedasticity of shocks, or a result of both the heteroskedasticity of shocks and the structural change of risk propagation mechanisms. [8]

of the results presented in this chapter when different measures of financial performance are used.

[8]See Appendix for details of the CDS data used.

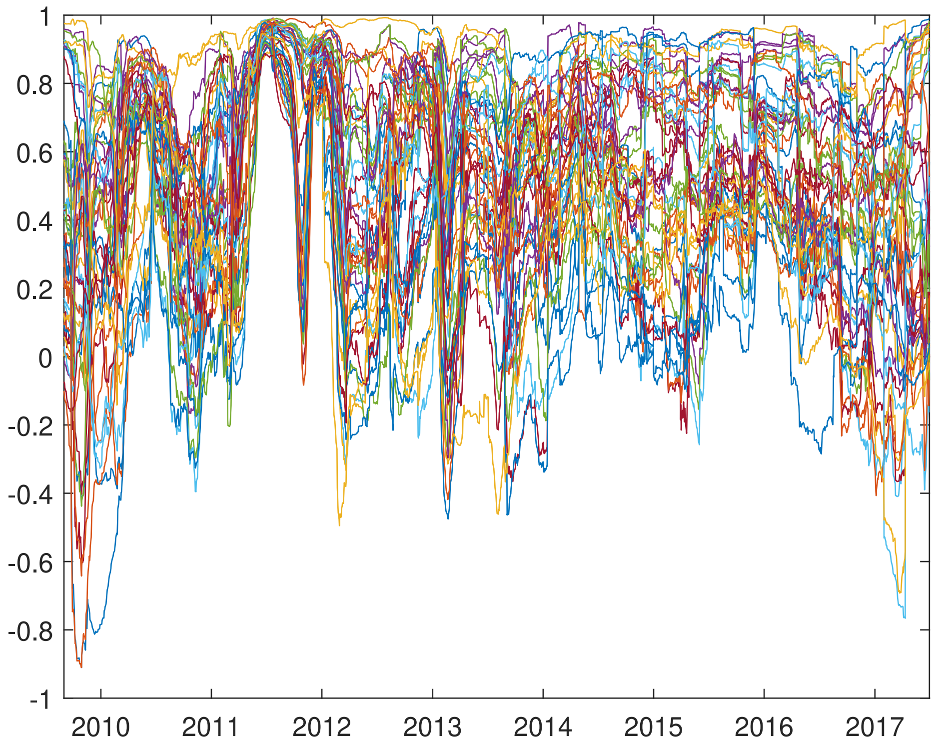



Figure 2-1: The 45 pairwise correlation coefficients of CDS of the top 10 US banks. The stock exchange id numbers of the 10 banks are JPM, BAC, WFC, C, GS, MS, COF, HSBC, AXP, and CSGN.

The change of volatility could be due to either time-varying linear regression coefficients (a.k.a. linear financial networks, as we will define later in the chapter), or heteroskedestistic shocks. Figure 2-1 shows the 45 pairwise correlation coefficients of CDS of the top 10 US banks calculated using a 200 days moving window. It is immediately notable that the correlation cannot be considered time-invariant. At an extreme, the correlation coefficient rises from -0.8 to 0.8 within 200 days.

### 2.2.2 A Linear Network Model

Suppose there are $N$ financial institutions (indexed by $n = 1, 2, \cdots, N$) in a contingent financial network with $M$ possible networks, which are denoted by the directed graphs $\mathcal{G}_m = \{\mathcal{V}, \mathcal{E}_m\}$ for $m \in \{1, 2, \cdots, M\}$. $\mathcal{V}$ is the set of common nodes in all the graphs and each node $v_n \in \mathcal{V}$ corresponds to a financial institution $n$. Furthermore, each edge $e_{ijm} \in \mathcal{E}_m$ denotes the risk spillover through $m$th network between node $i$ and node $j$.

There are two possible — simple — assumptions on how to implement the contingent networks. One in which the shocks are idiosyncratic and hit each bank individually and then they are propagated in the network, or the second one where the shocks are hitting the system as a whole but deferentially each bank. The first assumption is one in which the systemic risk of an individual shock effect is determined by the propagation between one bank and the other, while in the second assumption the relative importance of the aggregate shocks is what makes them systemic.

Our empirical strategy is to control for aggregate shocks and concentrate on the propagation of the idiosyncratic shocks. We believe this is the most intuitive application of our methodology.[9] In particular, we assume that the system is affected by shocks denoted as $\epsilon_t$, a $N$-by-1 vector representing $N$ shocks at time $t$. We assume that the shocks are independent from each other — i.e. $E[\epsilon_{t,n_1} \cdot \epsilon_{t,n_2}] = 0 \quad \forall \quad n_1 \neq n_2$. Assume that at each time instance $t$, each node receives a shock, that then it is transmitted through one of the possible contingent networks, $m$. The total impact is the only statistic that can be observed — which is a combination of direct and indirect effects.

Formally, assume that the impact of shocks propagating through network $m$ is not observable but the total risk measure, defined as the mixture of impact of different networks is observed and given by:

$$X_t = \Theta z + \sum_{m=1}^{M} w_{mh} F_{m,t}$$

where $z$ denotes observed aggregate shocks (such as inflation, S&P500, WTI, etc). $\Theta$ denotes exposure to aggregate shocks. $F_{m,t}$ is the impact of idiosyncratic shocks propagating through network $m$, and $w_{mt}$ is an indicator random variable. $w_{mt} = 1$ if network $m$ dominates at time $t$, and $w_{mt} = 0$ otherwise. Assume that the

[9] Future research should study the implications of aggregate contingent networks

indicator random variable $w_{mh}$ is 1 with probability $p_m$. Note this $p_m$ is a fixed prior distribution in the mixture model that does not depend on $h$. Let $h$ indicate different heteroskedastic regimes, that will be needed to solve the identification problem.

In order to control for aggregate shocks, we first do a linear regression of $z$ against $X$, and model the residuals as follows

$$\bar{X}_t = \sum_{m=1}^{M} w_{mh} F_{m,t}$$

We define $\Gamma_m$, $m \in \{1, 2, \cdots, M\}$ as the weighted-directed adjacency matrix of network $m$. We assume there cannot be self-loops in the network hence all diagonal entries of $\Gamma_m$ are zero. Now the impact of idiosyncratic shocks propagating through network $m$ satisfies

$$F_{m,t} = \alpha_m + \Gamma_m F_{m,t} + \epsilon_t. \tag{2.1}$$

In the above equation, we can see the direct and indirect effects of the shock. The indirect impact of $\epsilon_t$ is through $\Gamma_m$; which represents the network on how financial institutions have exposure through the balance sheet of the other financial institutions.

For simplicity of exposition, the rest of this chapter assumes that our data is demeaned, hence

$$F_{m,t} = \Gamma_m F_{m,t} + \epsilon_t. \tag{2.2}$$

This is a decomposition of any shock to the system as the exogenous part ($\epsilon_t$) and the endogenous network part ($\Gamma_m$). In this model, the time subscript is the same on both sides of the equation, assuming we are at the equilibrium. This is a recursive assumption that was first posed by Christiano et al. (1999).

Note that rearranging (2.2) we obtain

$$F_{m,t} = [I - \Gamma_m]^{-1} \epsilon_t. \tag{2.3}$$

Hence the observable variable can be written as

$$X_t = \sum_{m=1}^{M} w_{mt} \left[I - \Gamma_m\right]^{-1} \epsilon_t, \tag{2.4}$$

This describes the observed variable ($X_t$) as a linear mixture of different unobservable networks ($\Gamma_m$) and different shocks ($\epsilon_t$). The estimation problem is to uncover the unobserved networks from the moments of the observed variables.

### 2.2.3 The identification problem

The identification procedure we use in this chapter is related to the identification through heteroskedasticity developed in Rigobon (2003). The intuition of the identification can be developed in a two by two endogenous system of equations, see Appendix 2.6.1. In this section, we focus on the identification of a multi-bank financial network.

A crucial ingredient is the presence of heteroskedasticity. Therefore, we assume that there are regimes of economic environments, denoted by $h \in \{1, 2, \cdots, H\}$. Let $\mathcal{R}_h$ be the set of time instances $t$ that belongs to heteroskedastic regime $h$. Furthermore, define $n_h = |\mathcal{R}_h|$ as the number of samples in regime $h$. If time $t \in \mathcal{R}_h$, then the shock at that time instance is distributed $\epsilon_t \sim \mathcal{N}(0, \Xi_h)$. Let $\Xi_h$, a diagonal matrix, be the variance of shock $\epsilon$ in regime $h \in \{1, 2, \cdots, H\}$, be unknown constants.

The model parameters we need to identify are

$$\psi = \begin{bmatrix} \{\Gamma_m\}_{m=1}^{M} \\ \{\Xi_h\}_{h=1}^{H} \end{bmatrix},$$

and the number of parameters for each network $m$ is $N(N-1) + N \sum_h w_{mh}$. The first term comes from the networks. There are $N\,(N-1)$ elements in each network (diagonal are ones). The second term comes from the variance of the structural shocks. The shock affects $N$ banks for each regime — and there are $\sum_h w_{mh}$ regimes in which a given network dominates.

In this section, we will start by assuming $w_{mh}$ is known. Later, we will show how it can be estimated. The observed moments (moment constraints) are given by the variance-covariance matrix of $X_i$ in each regime $h$. Assume, without loss of generality, that $E[F_{m,t}] = 0$. Then the following constraint must hold for the covariance of $X_t$

$$E_{t\in\mathcal{R}_h} X_t X_t^\top - \sum_{m=1}^{M} w_{mh} \left( \left(I - \Gamma_m\right)^{-1} \Xi_h \left(I - \Gamma_m\right)^{-\top} \right) = 0,$$

where $E_{t\in\mathcal{R}_h} X_t X_t^\top$ denotes the expected value of the matrix $X_t X_t^\top$ given that $t$ belongs to regime $h$. Because $w_{mh} \in \{0, 1\}$ and $\sum_m w_{mh} = 1$, the above constraints can be written as

$$E_{t\in\mathcal{R}_h} X_t X_t^\top - \left( \left(I - \Gamma_m\right)^{-1} \Xi_h \left(I - \Gamma_m\right)^{-\top} \right) = 0, \forall w_{wh} = 1. \tag{2.5}$$

In the above, each matrix $E_{t\in\mathcal{R}_h} X_t X_t^\top - \sum_{m=1}^{M} \left( \left(I - \Gamma_m\right)^{-1} \Xi_h \left(I - \Gamma_m\right)^{-\top} \right)$ is a $N$-by-$N$ symmetric matrix, therefore each regime $h$ provides $\frac{N(N+1)}{2}$ moment constraints. In total, $\sum_h w_{mh}$ regimes implies $\frac{N(N+1)}{2} \sum_h w_{mh}$ moment constraints.

The condition for an exact or over-identified model is

$$N(N-1) + N \sum_h w_{mh} \leq \frac{N(N+1)}{2} \sum_h w_{mh} \tag{2.6}$$

which implies that the data must have at least

$$\sum_h w_{mh} \geq 2 \tag{2.7}$$

for just-identification and

$$\sum_h w_{mh} \geq 3 \tag{2.8}$$

for over-identification. i.e., we need each network to dominate in at least 3 regimes to achieve over-identification. This identification is the order condition: how many equations are needed for the system to have less unknowns than knowns. In the single network case, this is simple: the single network will dominate all regimes, hence $H = \sum_h w_{mh}$. In the multiple network case, we need to figure out how many different mixture of networks (M) present in the data, and which network dominate which regimes. Those two questions are fundamental questions for the identification of many mixture models, and we will discuss them in more detail later in Section 2.4

### 2.2.4 Solution Uniqueness

When there is a single network, parameters in the above model can be estimated via the Generalized Method of Moments (GMM). We define the score as a vector function

$$g(\Gamma, \Xi_h) = \left\{ X_t X_t^\top - \left( (I-\Gamma)^{-1} \Xi_h \, (I-\Gamma)^{-\top} \right) \right\}_{t \in \mathcal{R}_h, h=1,\cdots,H} \tag{2.9}$$

where $\{\cdot\}_{h=1,\cdots,H}$ denote the $H$-column vector whose entry are evaluated at $h = 1, \cdots H$. Estimation is achieved by solving the optimization problem[10]

$$\begin{aligned} \min_{A, \Xi_h} \quad & E_{t \in \mathcal{R}_h} g(\Gamma, \Xi_h)^\top V^{-1} E_{t \in \mathcal{R}_h} g(\Gamma, \Xi_h) \\ \text{s.t.} \quad & g(\Gamma, \Xi_h) = \left\{ X_t X_t^\top - \left( (I-\Gamma)^{-1} \Xi_h \, (I-\Gamma)^{-\top} \right) \right\}_{t \in \mathcal{R}_h, h=1,\cdots,H} \\ & \Xi_h \text{ are diagonal.} \end{aligned} \tag{2.10}$$

GMM requires that the score function is zero iff the system parameters are correct. We can prove that this requirement is satisfied when there is only one network, i.e. $M = 1$. To prepare for the proof, we define $\Xi = \begin{bmatrix} diag(\Xi_1) & diag(\Xi_2) & \cdots & diag(\Xi_H) \end{bmatrix}$, where $diag(\cdot)$ denotes the diagonal entries of the matrix in the form of a column vector. In addition, define $\Omega_h = \mathbb{E}_{t \in \mathcal{R}_h} X_t X_t^\top$

**Definition 1. Kruskal rank** (Stegeman and Sidiropoulos, 2007) Kruskal rank $k$ of a matrix $A$ is the maximum value of $k$ such that ANY $k$ columns and rows of the matrix $A$ are linearly independent.

**Lemma 1.** *If (i). $M = 1$, (ii). $I - \Gamma$ has full rank and (iii). The Kruskal rank of $\Xi$ is 2 or higher, then $\Omega_h - \left( (I-\Gamma)^{-1} \Xi_h \, (I-\Gamma)^{-\top} \right) = 0$ has a unique solution.*

[10] $V$ is a weighting matrix in the GMM. The GMM is valid as long as $V$ is positive definite, but an optimal $V$ is propotional to the variance-covariance matrix of the score.

Note unlike regular matrix rank, which requires it exists $K$ columns or rows that are linearly independent, the Kruskal rank requires ANY $k$ columns and rows are linearly independent. In the context of our identification problem, this means that the variance of shocks need to present enough heteroskedasticity to obtain a unique solution.

Intuition of the proof:

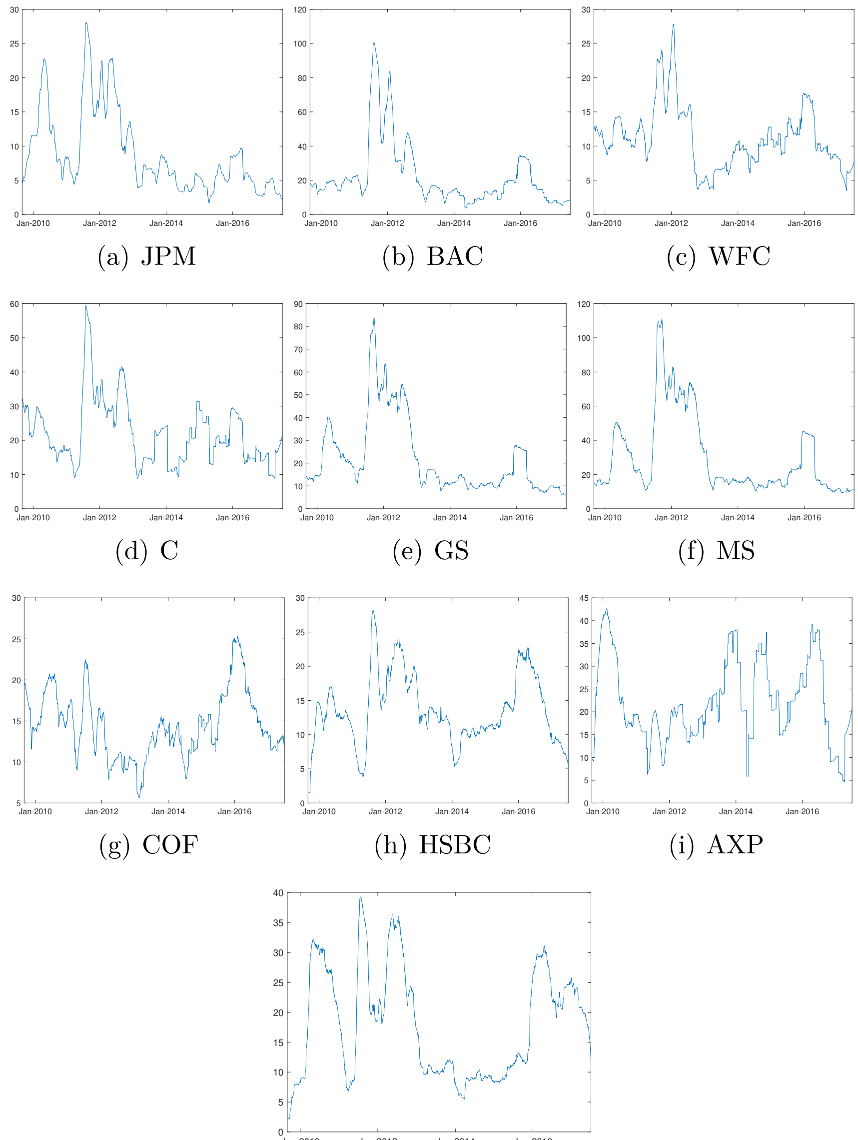
The identification problem is equivalent to a tensor decomposition problem (for details of this equivalence, refer to the Appendix). According to the Kruskal's rank condition, if

$$Krank(a_r) + Krank(a_r) + Krank(\xi_r) \geq 2R + 2$$

where $Krank(\cdot)$ stands for Kruskal rank, then the tensor decomposition problem has a unique solution. In the above equation, given assumption (ii), we have $Krank(a_r) = N$. Furthermore, $R$ is the rank of the tensor in (2.28), which equals $MN$. Inserting the numbers into the Kruskal's rank condition and considering assumption (i) and (iii), we obtain the solution uniqueness.

## 2.3 Results: Single Network Estimation under Endogeneity

Throughout this chapter we use Credit Default Swap (CDS) as a measure of the risk of bank. Figure 2-2 shows the 100-day moving volatility of the 10 banks of interest. The first thing to notice is that they exhibit a significant regime shift of heteroskedasticity over time. Around the time of Nov. 2009, the volatilities of CDS of all 10 banks are high, while the volatilities are significantly lower in mid-2010. In addition, their volatility co-move most of the time, despite regime shifts. The fact that they co-move, allow us to identify a network that is well connected. Apart from similarities in the overall trend, the CDS volatilities exhibit a certain level of bank-specific movements. For example, in early-2016, the CDS of Capital One Financial Corporation has very high volatility that is comparable with that around late 2009. In contrast, the CDS



(a) JPM (b) BAC (c) WFC

(d) C (e) GS (f) MS

(g) COF (h) HSBC (i) AXP

(j) CSGN

Figure 2-2: 100 days moving volatility of the CDS of the 10 banks of interest. CDS data of 10 largest banks in US are collected from Sep. 1, 2009 to June. 20 2017. The stock exchange id tickers of the 10 banks are JPM, BAC, WFC, C, GS, MS, COF, HSBC, AXP, and CSGN.

of Morgen Stanley is at a relatively low historical level. That kind of bank-specific characteristic could allow as explore the different risk propagation mechanisms among different banks.

We control for aggregate shock $z$ by first do an linear regression, and then study the residuals. The aggregate shocks We study include 9 different macroeconomic variables that try to capture different aggregate shocks that typically afflict economies: (1) commodity price shocks (WTI Oil Prices), (2) Inflation (PriceStats daily inflation index), (3) Economic Activity (Non-farm payroll), (4) Stock Market Prices (S&P), (5) Risk (VIX), (6) Housing Prices (Case & Shiller index), (7) liquidity provision (short term interest rates), (8) exchange rates (nominal trade-weighted exchange rate), and (9) the yield curve (the difference between the long and short interest rates). Altogether, the 9 aggregate shocks account for about 50% of the variances of CDS spread.

### 2.3.1 Selection of Regimes

As mentioned in previous sections, our model is not uniquely identifiable if the variance of shock is constant. When the variance of shock is not constant, we can divide our data into different heteroskedastic regimes. Because the network parameters are constant over regimes, additional regimes could offer more additional constraints than additional unknowns. With enough heteroskedastic regimes, our model can be uniquely identified (Rigobon, 2003). In practice, however, dividing data into heteroskedastic regimes is not trivial. An effective regime division method should achieve heteroskedasticity among regimes and maintaining homogeneity within regimes. There are two broad categories of methods to identify heteroskedastic regimes: regimes divided by statistical properties of the data itself, and regimes divided by other exogenous variables (in our case, it makes sense to use macroeconomic factors). An identification process using regimes defined by statistical properties is named statistical identification, and an identification process using regimes defined by macroeconomic factors is named macroeconomic identification. In this chapter, we have already controlled for macroeconomic shocks, so statistical identification is preferred to give us enough heteroskedasticity among regimes for better identification.

Statistical identification requires we first split our data into different regimes according to statistical properties of the data. Because we want to separate the variance of shock, it makes sense to look at the quantile level of CDS data volatility. To maximize the separation of unobserved networks, one could use the quantile level of volatility of CDS to define regimes. For example, if there are two banks A and B in the network, one can define four regimes: bank A's volatility is at top 20% quantile level while B is not; bank B's volatility is at top 20% quantile level while A is not; both banks' volatility is at top 20% quantile level, and neither banks' volatility is at top 20% quantile level. Regimes with insufficient number of samples are not used in the identification process. It is clear that the overall financial network cannot be the same in different regimes.

However, regimes divided by a fixed quantile level is usually very unbalanced, i.e., the low volatility regimes have far more data points than the high volatility regimes. From an economic perspective, this is fine because exceptionally high volatility only occurs during crises. For identification purposes, however, this is not optimal, and many regimes have so few data points to be used in the identification. Alternatively, we could use unsupervised learning techniques, such as K-mean and Gaussian Mixture Model to divide data into groups according to their volatility levels.

Intuitively, a clustering algorithm group data points at different time instances into subsets, and maximize the similarities of the volatility vector in each subset. Due to the inherent limitation of most clustering algorithms, the global optimal grouping is usually very difficult to find, and the algorithm is sometimes trapped at local optimal solutions. If the dimension of the volatility vector is very high, i.e., we are identifying the network for a large number of banks, then it is even harder for the algorithm to find the global optimal solutions. In this case, we can reduce the dimension of the volatility vector by applying PCA prior to clustering.

### 2.3.2 Estimates: Statistical Identification

| | JPM | BAC | WFC | C | GS | MS | COF | HSBC | AXP | CSGN |
|---|---|---|---|---|---|---|---|---|---|---|
| JPM | | -0.0 | 0.4 | -0.06 | 0.13 | -0.02 | 0.01 | -0.03 | -0.03 | 0.02 |
| | | (0.01) | (0.06) | (0.02) | (0.02) | (0.02) | (0.02) | (0.02) | (0.01) | (0.02) |
| BAC | 0.77 | | 0.99 | 0.09 | 0.01 | 0.09 | 0.13 | 0.08 | -0.05 | 0.02 |
| | (0.14) | | (0.08) | (0.04) | (0.07) | (0.06) | (0.06) | (0.09) | (0.04) | (0.05) |
| WFC | 0.43 | 0.08 | | 0.04 | -0.01 | 0.03 | 0.16 | 0.0 | 0.04 | -0.04 |
| | (0.04) | (0.01) | | (0.02) | (0.01) | (0.01) | (0.04) | (0.02) | (0.01) | (0.01) |
| C | 0.03 | 0.2 | -0.11 | | 0.12 | 0.18 | 0.03 | 0.04 | 0.06 | -0.01 |
| | (0.14) | (0.03) | (0.14) | | (0.09) | (0.07) | (0.06) | (0.05) | (0.02) | (0.04) |
| GS | 0.92 | 0.06 | 0.04 | 0.15 | | 0.39 | 0.02 | 0.26 | -0.04 | -0.15 |
| | (0.12) | (0.05) | (0.1) | (0.05) | | (0.07) | (0.05) | (0.1) | (0.03) | (0.06) |
| MS | -0.07 | 0.13 | 0.2 | 0.14 | 0.94 | | 0.02 | -0.08 | 0.08 | 0.14 |
| | (0.13) | (0.04) | (0.15) | (0.06) | (0.07) | | (0.09) | (0.07) | (0.03) | (0.05) |
| COF | -0.06 | -0.08 | 0.59 | 0.14 | 0.09 | -0.14 | | -0.01 | 0.01 | 0.14 |
| | (0.08) | (0.02) | (0.08) | (0.04) | (0.03) | (0.03) | | (0.04) | (0.03) | (0.03) |
| HSBC | 0.04 | 0.03 | 0.06 | -0.08 | 0.21 | 0.02 | 0.0 | | -0.09 | 0.17 |
| | (0.08) | (0.02) | (0.09) | (0.03) | (0.03) | (0.02) | (0.03) | | (0.02) | (0.03) |
| AXP | 0.04 | -0.03 | 0.27 | 0.46 | -0.21 | -0.03 | 0.39 | -0.05 | | -0.02 |
| | (0.17) | (0.04) | (0.14) | (0.05) | (0.08) | (0.06) | (0.07) | (0.1) | | (0.05) |
| CSGN | -0.03 | -0.12 | -0.02 | -0.01 | -0.09 | 0.14 | 0.13 | 0.83 | 0.14 | |
| | (0.16) | (0.04) | (0.11) | (0.05) | (0.07) | (0.06) | (0.08) | (0.07) | (0.03) | |

Table 2.1: Estimates of the network structure. Standard deviations (in brackets) are obtained by bootstrapping (2000 resamples) across regimes. In this case, regimes are decided by CDS quantile. There are $H = 20$ regimes.

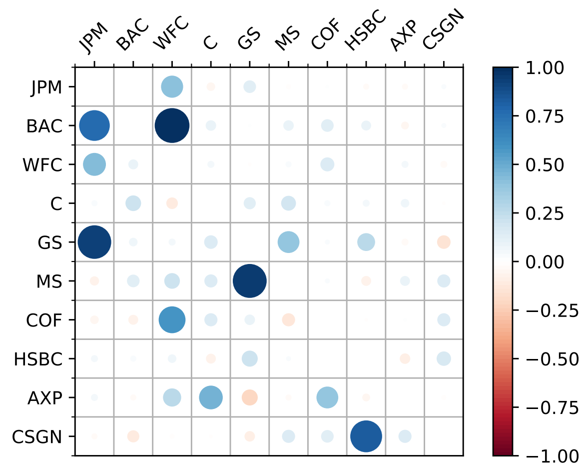



Figure 2-3: Visualization of the single network identified through heteroskedasticity

We now turn to the estimation results for our linear network model. As defined in Section 2, we estimate the linear shock propagation channel among the top 10 US banks. In our estimates, we assume that the linear structural parameters that we estimate is always between -1 and 1. i.e.,

$$\gamma_{i,j,m} \in [-1, 1], \forall i, j, m. \tag{2.11}$$

Mathematically, this constraint will remove any non-unique solutions due to columns permutations. Economically, this means that the bank which receives a shock directly is most affected. This approach of removing permutation solutions will be problematic if a lot of the estimated structural parameters are on the boundary (i.e., the constraint (2.11) is binding for a lot of parameters). However, at least for the CDS spread dataset that we study, only 2 parameters are on the boundary in statistical identification.

Table 2.1 shows the estimates of the 10-by-10 network structure using statistical identification. [11] After controlling for aggregate shocks, Standard deviations (in brackets) are obtained by bootstrapping (2000 resamples) across regimes. In this

[11]We do not use Lasso in this small scale estimation problem, because Lasso may introduce bias that affects our statistical tests in the next section. Furthermore, in our identification using CDS data, a reasonable amount of Lasso penalty will not change the result much. However, in a larger scale estimate (e.g., a banking network with 1000+ nodes), Lasso will be very useful in obtaining sparse networks.

case, as the name "statistical identification" suggested, regimes are decided by CDS quantile levels discussed in Section 3.1. The matrix in Table 2.1 can be regarded as a weighted directed graph. Each column shows where the shock is originated from, and each row shows where the shock propagates to. For example, the structural parameter in the 4th column (Citigroup) and the 1st row (JP Morgan) represents the channel where shock propagates from Citigroup to JP Morgan. In an earlier example in Section 2.1, we used the correlation between Citigroup to JP Morgan as an example to show that the correlation between banks can be very volatile. Here the structural parameter, on the other hand, is reliably estimated with bootstrapping standard deviation of only 0.02.

The estimated matrix of structural parameters is asymmetric in general. However, this does not mean any causal relationship between each pair of banks. In our original model, without (2.11), any column permutation of $\Gamma$ will give a new solution and change the direction of edges of the weighted directional graph in Table 2.1. Now with constraint (2.11), the directions of the edges of the graph are pinned down by the constraint, but not by any inherent causality in the data.

Figure 2-3 gives a visualization of the same network. The values of structural parameters are represented by the color and size of the corresponding circles. Positive structural parameters are displayed in blue and negative structural parameters are displayed in red. In addition, color intensity and the area of the circles are proportional to the absolute values of the structural parameters. The visualization helps identify patterns in the risk transmission mechanism. For example, Wells Fargo is exposed to a number of different shocks, while Bank of America is more resilient to shocks transmitted from other banks.

## 2.4 Multiple Contingent Network Estimation

This section first presents a test for the multiple network assumption. In particular, we propose a testing procedure that compares the consistency of the estimates of structural variables in a single network case versus a multiple network case. In a point-wise test, one can obtain the distribution of differences and carry out the test without any assumptions on distributions. However, a point-wise test cannot provide a summary of the results. If we further assume that structural variables follow independent but not necessarily identical Gaussian distributions, we can test the sum of normalized residuals, which follows a Chi-squared distribution.

After the single network hypothesis is rejected by test, this section further presents a methodology to idenfity multiple networks. In the case of a single network, the solution is established through GMM in the last section. In this section, in the case of multiple networks, the expectation-maximization (EM) algorithm, which is commonly used for estimating mixture models, is adopted. The EM algorithm alternates between an expectation (E) step, which updates the probability of a network dominating a regime, and a maximization (M) step, which estimate each network through heteroskedasticity based on the probability updated in the E step. A Wishart distribution is assumed for sample variance-covairance matrices so that the log-likelihood function used in the EM algorithm is well-defined.

### 2.4.1 Chi-square Test for Network Contingency and the Rejection of the Single Network Hypothesis

In this subsection, we construct an intuitive Chi-square Test that is used to reject the single network hypothesis. Suppose one observes two sequences of data $\{X_t\}_{t\in D_1}$ and $\{X_t\}_{t\in D_2}$, and estimates structural parameters $\Gamma_{D_1}$ and $\Gamma_{D_2}$. We want to know whether the two sets of structural parameters are consistent. Let $\gamma_{i,j,1}$ and $\gamma_{i,j,2}$ denote the $i$th row, $j$th column entry of the network estimated from data set $D_1$ and $D_2$ respectively.

We begin with a number of assumptions

Asymptotic Assumptions: (a) The parameter space $\Psi$ of $\psi = \begin{bmatrix} \Gamma \\ \{\Xi_h\}_{h=1}^H \end{bmatrix}$ is a compact subset of $\mathbb{R}^d$, and the true value $\psi_0$ lies in the interior of the parameter space $\Psi$. (b) The moment function $\psi \to g(\psi)$ defined in (2.9) identifies $\psi_0$: $g(\psi) = 0$ iff $\psi = \psi_0$. (c) The empirical moment function $\psi \to \hat{g}(\psi)$ converges uniformly in probability to the moment function $\psi \to g(\psi)$, namely $\sup_{\psi \in \psi} \|\hat{g}(\psi) - g(\psi)\| \to_p 0$ (d) The empirical Jacobian $\hat{G}(\psi) = \frac{\partial}{\partial \psi}\hat{g}(\psi)$ is continuous and is uniformly consistent for the Jacobian matrix, $G(\psi) = \frac{\partial}{\partial \psi} g(\psi)$, i.e., $\sup_{\psi \in \psi} \|\hat{G}(\psi) - G(\psi)\| \to_p 0$ (e) The matrix $G(\psi_0)^\top G(\psi_0)$ is positive definite. (f) The empirical moment function evaluated at the true parameter value obeys a central limit theorem:

$$\sqrt{n}\hat{g}(\psi_0) \sim N(0, \Omega)$$

asymptotically, where $n$ is the number of samples.

Note those assumptions are inherently the same with the assumptions in the original GMM paper by Hansen (1982). Note that the key to our single network test is the solution uniqueness assumption (b) of GMM summarized above. Failing to establish a result for solution uniqueness will cause incorrect estimates of parameter distribution. Generally speaking, the bootstrap estimated parameter distribution will have a larger variance and therefore fails to reject the hypothesis even if the hypothesis is incorrect. Lemma 1 in this chapter deals with this problem.

Under the null hypothesis that

$$H_0 : \gamma_{i,j,1} \text{ and } \gamma_{i,j,2} \text{ are the same}$$

their difference

$$\gamma_{i,j,1} - \gamma_{i,j,2} \tag{2.12}$$

should follows a distribution with zero mean. If the estimated value of $\hat{\gamma}_{i,j,1} - \hat{\gamma}_{i,j,2}$ lies in the 0.05 left or right quantiles of the bootstrapping distribution, we can reject the null hypothesis and claim that with 90% confidence

$$H_1 : \Gamma_{D_1} \text{ and } \Gamma_{D_2} \text{ are different}$$

The above point-wise test has the advantage of distribution-free. However, without a summarizing statistic, one cannot draw conclusions on the overall network. Suppose further that $\gamma_{i,j,1} - \gamma_{i,j,2}$ follows Gaussian distribution $\mathcal{N}(\overline{\gamma}_{i,j}, \xi_{i,j}^2)$. Under the null hypothesis that

$$H_0 : \Gamma_{D_1} \text{ and } \Gamma_{D_2} \text{ are the same} \tag{2.13}$$

their squared difference $\left(\gamma_{i,j,1} - \gamma_{i,j,2}\right)^2 / \xi_{i,j}^2$ follows a Chi-squared difference with degree of freedom 1.

In addition, under the independence assumption, the sum of squared differences

$$\sum_{i \neq j} \left(\gamma_{i,j,1} - \gamma_{i,j,2}\right)^2 / \xi_{i,j}^2 \sim \mathcal{K}\left(N(N-1)\right)$$

where $\mathcal{K}$ denotes a Chi-squared distribution.

If the estimated value of $\sum_{i \neq j} \left(\gamma_{i,j,1} - \gamma_{i,j,2}\right)^2 / \xi_{i,j}^2$ lies in the 0.1 right quantiles of the bootstrapping distribution, we can reject the null hypothesis and claim that with 90% confidence

$$H_1 : \Gamma_{D_1} \text{ and } \Gamma_{D_2} \text{ are different}$$

Now we are able to compare the network contingency given any two sets of data. In this subsection, we divide our dataset according to quantile levels of each bank's CDS and test the network contingency to those factors. This procedure is analogous to a sensitivity test of our model.

| | Mean diff | Mean std | Num. rejections | Chi-square stat |
|---|---|---|---|---|
| JPM | 0.231852 | 0.134570 | 25.0 | 526.356267 |
| BAC | 0.214357 | 0.124162 | 24.0 | 541.589106 |
| WFC | 0.231365 | 0.142895 | 30.0 | 429.927184 |
| C | 0.240336 | 0.138054 | 27.0 | 517.066271 |
| GS | 0.212381 | 0.128582 | 25.0 | 534.629568 |
| MS | 0.183259 | 0.130501 | 20.0 | 341.691040 |
| COF | 0.169083 | 0.137471 | 22.0 | 327.151715 |
| HSBC | 0.175459 | 0.162129 | 19.0 | 261.359446 |
| AXP | 0.198182 | 0.123283 | 24.0 | 508.955578 |
| CSGN | 0.228512 | 0.154193 | 23.0 | 457.519043 |

Table 2.2: Network changes driven by risk levels of each bank

Among the 10 tests, network structural changes are all very significant. In all cases, the p-value is less than 0.01 and the single network hypothesis is rejected.

### 2.4.2 Estimating mixture models using EM-Algorithm

Although we build the model with a mixture of multiple networks, so far, the identification process has mostly been with a single network. In the case of a single network, the mixture random variable $w := \{w_{mh}\}_{m=1,h=1}^{M\quad,H}$ is trivial, and the identification is through the GMM algorithm. Because the last subsection showed that a single network is not sufficient to describe the data, we move towards the model setup of a multi-network mixture.

Recall the main characteristics of our mixture model as follows: the data $X_t$ in each regime $t \in \mathcal{R}_h$ is generated by

$$X_t = \sum_{m=1}^{M} w_{mh}(1-\Gamma_m)^{-1}\epsilon_t \tag{2.14}$$

where

$$\epsilon_t \sim \mathcal{N}(0, \Xi_h)$$

$\Gamma := \{\Gamma_m\}_{m=1}^M$ and $\Xi := \{\Xi_m\}_{h=1}^H$ are parameters to be identified. In addition, $w_{mh}$ denote the indicator random variable that equals to 1 if network $m$ dominates in regime $h$. Assume $w_{mh} = \begin{cases} 1 & \text{w.p. } p_m \\ 0 & \text{o.w.} \end{cases}$. Because only one network dominates in each regime,

$$\sum_{m=1}^{M} p_m = 1 \tag{2.15}$$

Our objectives is to estimate parameters $\Gamma$ and $\Xi$ by observing $X_t$.

Similarly to the case of single network identification, we construct

$$\Omega_h = \frac{1}{n_h} \sum_{t \in \mathcal{R}_h} X_t X_t^\top$$

the maximum likelihood estimator (MLE) of population variance of $X_t$ in each heteroskedasticity regimes $h$. As it will become clear later in this subsection, using $\Omega_h$ will allow us to separate the identification of $\Gamma$ and $\Xi$. Furthermore, it allows us to estimate each $\Gamma_m$ separately.

The above mixture model is difficult to identify, because $w_{mh}$ is a random variable which is not observed. In this case, the Expectation-Maximization (EM) Algorithm can be used to make the problem tractable. Note that $\Omega := \{\Omega_h\}_{h=1}^H$ given $w$ is Wishart distributed, even though $\Omega$ itself is not. The EM Algorithm takes this advantage by iterating between the E step and M step. In the E step, it computes the discrete distribution of $w_{mh}$ given current parameter estimates, and calculate the function

$$Q(\Gamma, \Xi \mid \Gamma^{(current)}, \Xi^{(current)}) = \mathbb{E}_{w|\Omega,\Gamma^{(current)},\Xi^{(current)}}[\log L(\Gamma, \Xi; \Omega, w)]$$

where $L(\Gamma, \Xi; \Omega, w)$ is the likelihood function assuming $w$ is observable. In the M step, it computes the optimal parameters $\Gamma$ and $\Xi$ to maximize the $Q(\cdot)$ function. It is a general result of EM that improving $Q(\cdot)$ improves the likelihood function of the mixture model $L(\Gamma, \Xi; \Omega)$, see Little and Rubin (2019).

A critical step of implementing the EM algorithm is to calculate the likelihood function of $\Omega$ given $w$.

Define $V_{mh} := \frac{1}{n_h}(1-\Gamma_m)^{-1}\Xi_h(1-\Gamma_m)^{-\top}$. Given $w_{mh} = 1$, i.e. network $m$ dominates in regime $h$, we have

$$\Omega_h \sim \mathit{Wishart}(V_{mh}, n_h) \tag{2.16}$$

The Wishart distributed $\Omega_h$ has a probability density function (PDF)

$$f_{\mathbf{w}}(\Omega_h) = \frac{1}{2^{n_h N/2}\left|V_{mh}\right|^{n_h/2}\mathbb{\Gamma}_N\left(\frac{n_h}{2}\right)}\left|\Omega_h\right|^{(n_h-N-1)/2}e^{-(1/2)\operatorname{tr}(V_{mh}^{-1}\Omega)} \tag{2.17}$$

where $n_h$ is the number of samples in regime $h$, $N$ is the number of banks in the network, $|\cdot|$ is the determinant of a matrix, and $\mathbb{\Gamma}_N(\cdot)$ is the multivariate Gamma function with dimension $N$.

Now we can write the likelihood function $L(\Gamma, \Xi; \Omega, w)$ using the Wishart PDF

$$L(\Gamma, \Xi; \Omega, w) = \prod_m \prod_h \Big(f_{\mathbf{w}}(\Omega_h)\Big)^{\mathbb{I}_{w_{mh}=1}} \tag{2.18}$$

Recall that we can separate the estimation of the network matrices $\Gamma$ with the shock variances $\Xi$ in the single network case. Similarly, we can also do that in EM, but for a different reason: the optimal diagonal $\Xi_h$ that maximize (2.17) can be obtained analytically. i.e.,

**Lemma 2.** *Let $\Omega_h$ be a given $N$-by-$N$ positive semidefinite matrix. Define $V_{mh} := \frac{1}{n_h}(1-\Gamma_m)^{-1}\Xi_h(1-\Gamma_m)^{-\top}$. Also define $\mathbb{PD}$ as the set of positive semidefinite diagonal matrices. Then*

$$\begin{aligned}
\Xi_h^* :&= \arg\max_{\Xi_h \in \mathbb{PD}} \frac{1}{2^{n_h N/2}\left|V_{mh}\right|^{n_h/2}\mathbb{\Gamma}_N\left(\frac{n_h}{2}\right)}\left|\Omega\right|^{(n_h-N-1)/2}e^{-(1/2)\operatorname{tr}(V_{mh}^{-1}\Omega)} \\
&= \mathit{diag}\Big((1-\Gamma_m)\Omega(1-\Gamma_m)^\top\Big)
\end{aligned} \tag{2.19}$$

The proof is given in Appendix 2.6.3.

The above Lemma means that we do not have to identify the join of $\Gamma$ and $\Xi$. We can just identify $\Gamma$, and the optimal $\Xi$ can be obtained analytically. Now that we have separated the identification of the network $\Gamma$ with $\Xi$, we focus on a EM-algorithm that identifies $\Gamma$. At this point, we are looking at a mixture model with parameters $\Gamma$, hidden random variables $w$, and observed random variables $\Omega$. We define the function $Q$ as follows

$$\begin{aligned} Q(\Gamma \mid \Gamma^{(current)}) &= \mathbb{E}_{w|\Omega,\Gamma^{(current)}}[\log L(\Gamma;\Omega,w)] \\ &= \sum_m \sum_h \mathbb{E}_{w|\Omega,\Gamma^{(current)}} \mathbb{I}_{w_{mh}=1} \log f_{\mathbf{W}|\Xi^*}(\Omega) \\ &= \sum_m \sum_h p_{mh} \log f_{\mathbf{W}|\Xi^*}(\Omega) \end{aligned} \tag{2.20}$$

where $f_{\mathbf{W}|\Xi^*}$ is the Wishart PDF given optimal $\Xi$, and $p_{mh} := \mathbb{P}(w_{mh} = 1|\Omega_h;\Gamma)$ is the probability that network $m$ dominates in a specific regime $h$. Inserting (2.19) into (2.17), we have

$$f_{\mathbf{W}|\Xi^*}(\Omega) = \frac{1}{\left|(1-\Gamma_m)^{-1} diag\Big((1-\Gamma_m)\Omega(1-\Gamma_m)^\top\Big)(1-\Gamma_m)^{-\top}\right|^{(n_h/2)}} \cdot constant\ not\ depend\ on\ \Gamma$$

Hence

$$\begin{aligned} &Q(\Gamma \mid \Gamma^{(current)}) \\ &= \sum_m \sum_h p_{mh}^{(current)} \left(-\frac{n_h}{2} \log\left|(1-\Gamma_m)^{-1} diag\Big((1-\Gamma_m)\Omega(1-\Gamma_m)^\top\Big)(1-\Gamma_m)^{-\top}\right|\right) \\ &+ constant \end{aligned} \tag{2.21}$$

where $\mathbf{v}_{mh} = n_h(1-\Gamma_m)\Omega_h(1-\Gamma_m)^\top$.

Another step that plays an important role in EM algorithm is the estimate of $p_{mh}$ and $p_m$. Note that $p_{mh}$ and $p_m$ are two different quantities: $p_{mh}$ is the posterior probability of $w_{mh} = 1$ for a specific $h$. On the other hand, $p_m$ is the prior distribution of $w_{mh}$, which is the same for all regimes. According to Bayes rule

$$\begin{aligned} p_{mh} &:= \mathbb{P}\left(w_{mh}=1|\Omega_h;\Gamma^{(current)}\right) \\ &= \mathbb{P}(w_{mh}=1;\Gamma^{(current)})\cdot\mathbb{P}(\Omega_h|w_{mh}=1;\Gamma^{(current)})\cdot constant \\ &= p_m^{(current)}\cdot f_{\mathbf{V}}^{(current)}(\mathbf{v}_{mh})\cdot constant \end{aligned} \tag{2.22}$$

where $p_m^{(current)}$ can be estimated by $p_m^{(current)} = \frac{1}{H}\sum_{h=1}^H p_{mh}^{(current)}$, $f_{\mathbf{V}}$ is defined in (2.17) and the normalizing constant is decided by $\sum_{m=1}^M p_{mh} = 1$[12].

EM-Algorithm

Take an initial guess of $\Gamma$, then iterate between the following E-step and M-step.

E-step:

Update $p_{mh}$, the probability that network $m$ dominates in regime $h$, given current estimates of $\Gamma$, according to (2.22). After that, we use the updated $p_{mh}$ to construct

$$Q(\Gamma\mid\Gamma^{(current)}) = \sum_m\sum_h p_{mh}^{(current)}\left(\frac{n_h-N-1}{2}\log|\mathbf{v}_{mh}|-\frac{n_h}{2}\log|diag(\mathbf{v}_{mh})|\right)+constant \tag{2.23}$$

where $\mathbf{v}_{mh} = n_h(1-\Gamma_m)\Omega_h(1-\Gamma_m)^\top$.

M-step:

Update estimates of $\Gamma$ by maximizing $Q(\Gamma\mid\Gamma^{(current)})$. It is sufficient to update $\Gamma_m$ separately by

$$\max_{\Gamma_m}\sum_h p_{mh}^{(current)}\left(\frac{n_h-N-1}{2}\log|\mathbf{v}_{mh}|-\frac{n_h}{2}\log|diag(\mathbf{v}_{mh})|\right)$$

for each $m$.

### 2.4.3 The number of networks

To obtain the optimal number of networks, we apply the Bayesian information criterion. The Bayesian information criterion is defined as

$$\mathrm{BIC} = \ln(n)k - 2\ln(L).$$

[12] This process is similar to the case of Gaussian mixture models, see Chapter 9 of Bishop (2006)

where $n$ is the number of samples, $k$ is the number of parameters, and $L$ is the maximum mixture log-likelihood function, see Fraley and Raftery (1998).

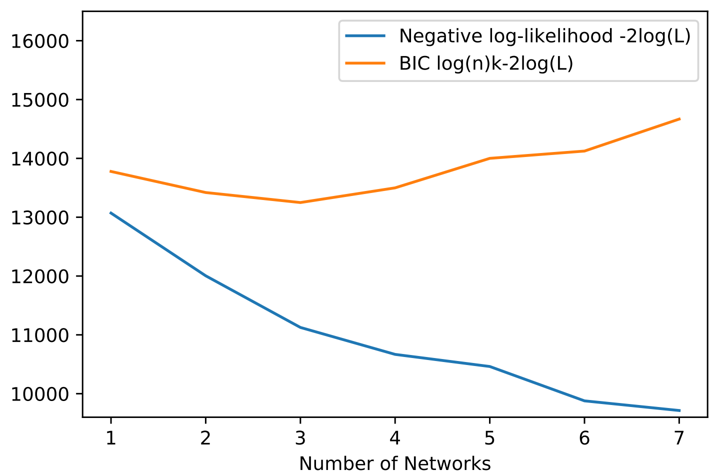



Figure 2-4: Negative log-likelihood and BIC vs. No. of networks

As shown in Figure 2-4, the optimal number of networks selected by Bayesian information criteria is 3.

The estimated 3 networks are given in Figure 2-5. In the 3 networks, network 3 is the most similar one with the single network estimated in Section 2.3.2 in terms of Frobenius Norm.

In summary, we find the data can be explained by 3 networks in the financial network among the top 10 banks in the US with our criteria. We reject the hypothesis of 1 network using an Chi-square Test and then use the Bayesian information criteria to conclude that 3 networks are optimal. With 4 or more networks, the model complexity penalty term in the BIC would standout and reject the models. We are only applying our identification method on financial networks though, other applications of our identification method could give 4 or more networks as the optimal solution.

### 2.4.4 Network Centrality and Policy Implications

There are many ways to estimate systemic risk in a financial network. As an example, we use the Katz centrality because it measures the relative degree of influence of a

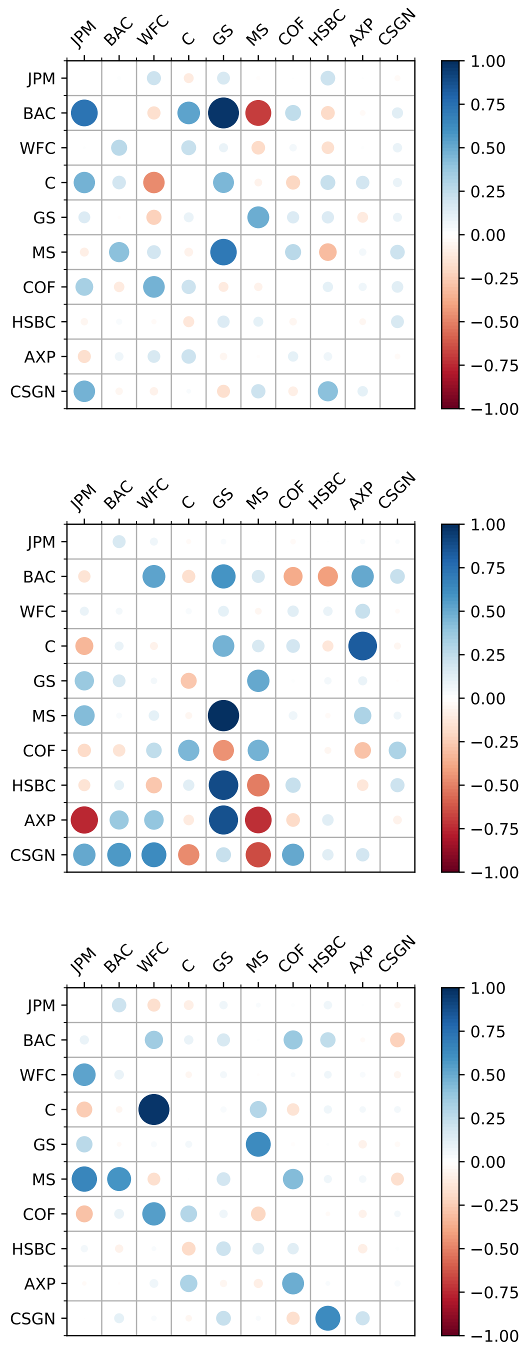



Figure 2-5: Estimated 3 networks using EM Algorithm

bank in the whole financial network, taking into account both the direct and indirect impact — much like our model assumption in this chapter. The Katz centrality (Katz, 1953; Junker and Schreiber, 2008) of a network with adjacency matrix $\Gamma$ is defined as

$$\overrightarrow{C}_{\text{Katz}} = ((I - \alpha\Gamma)^{-1} - I)\overrightarrow{I}, \tag{2.24}$$

where $\alpha$ is a damping factor that satisfies $0 \leq \alpha < 1/|\lambda_{\max}|$.

The Katz centrality is a generalization of the degree centrality. Intuitively, a node in the graph is more important if it more often receives shocks from other nodes. Furthermore, the Katz centrality considers both the direct impact from other nodes as well as the cascade impact many steps away. It assumes that both direct and indirect impacts affect the importance of a node, given that indirect impacts are discounted by a factor of $\alpha$ after each step. Apart from original applications in social networks and biological networks, Katz centrality has also been applied to evaluate systemic risk in financial networks, see (Thurner and Poledna, 2013; Temizsoy et al., 2017). In our estimates, because our matrix $I - \Gamma$ is invertible by itself, there is no need to add the damping factor, i.e., $\alpha = 1$ in our application. The Katz centrality of the top 10 banks in the US in the three estimated networks is shown in Figure 2-5.

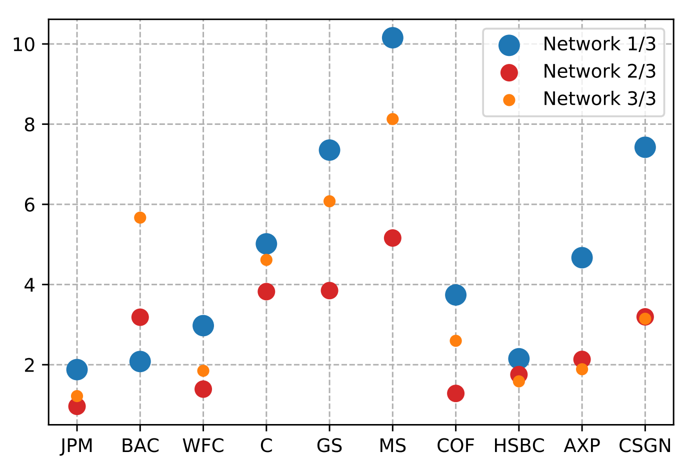



Figure 2-6: Estimated Katz centrality of the 3 estimated networks.

Notice the large variation in the relative importance of the centrality in the different networks. The networks we have estimated reflect an underlying structure that we do not observe. In principle, they should have a structural interpretation, but unfortunately, there is no information to verify this claim. In other words, based on the identification discussion in previous sections, the structural parameters should be recovered. That information is not enough to provide a structural interpretation of the network.[13]

Given that there are three distinct networks, the immediate question is the appropriate monetary policy response. In principle, during a given economic event, the central bank and the regulators do not know which network is the one that is responsible for the transmission of shocks. One way to deal with this uncertainty is to take a robust optimization approach. A robust optimization approach aims to find the optimal policy response in the worst-case scenario. Our estimation strategy helps central banking policymakers design such robust policy responses.

In Figure 2-7 we present the centrality as computed using only one network or using our robust approach with multiple networks. The blue estimates are those from using the heteroskedasticity identification but imposing a single network.[14]. The red estimates are the maximum centrality measures from Figure 2-6 For example, without the multi-network assumption, we would have underestimated every financial institution's systemic impact. Furthermore, the resulting centrality measures, or any other systemic risk measures, will not reflect the fact that there are three networks with potentially different systemic risk rankings. For example, Bank of America's centrality rank assuming one network is less than Citibank, but it is more important than it in the worst case. In the middle of a financial crisis, when central bankers have

[13]If our method is used in conjunction with macroeconomic shocks — such as inflation, exchange rate, stock market movements, etc. — then the network can be interpreted as the transmission of stock market shocks, or the transmission of inflation shocks. Such application requires extending the current model and data set to be able to incorporate those shocks formally. In this chapter, we have just controlled for the macroeconomic shocks leaving only the transmission of idiosyncratic shocks to be at play. More data will be required to deal with macro shocks because it is likely that more than three networks will be found.

[14]In order to make a meaningful comparison with the multiple network case, the network is estimated by maximizing the Wishard likelihood function in Section 2.4.2, instead of using GMM in Section 2.2.4

to decide on supporting banks that are systemically important, a robust approach would be consistent with a macro-prudential approach.

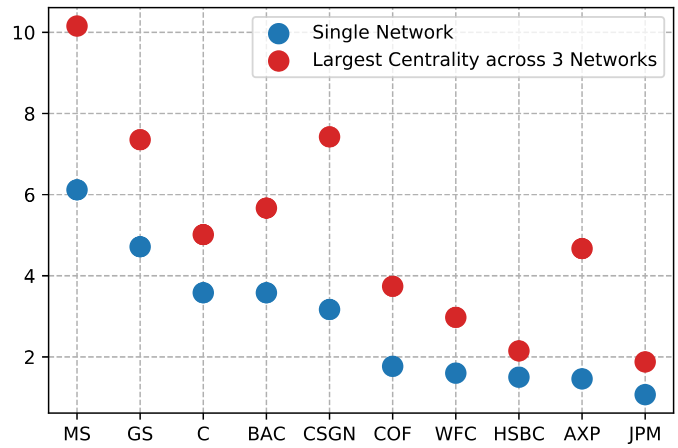



Figure 2-7: Estimated Katz centrality in the case of assuming only 1 network vs. the worst case of three networks.

In summary, our method can identify multiple linear financial networks using outcome data. It offers central bankers an opportunity to examine more granular details of financial networks, thus making more informed decisions. It is worth noting that this chapter's key contribution is an improved method of identifying financial networks and monitoring systemic risk, not the ability to specify a particular bank being systemically important. We believe that this approach will help central banks build a more robust policy-making process.

## 2.5 Conclusions

Understanding the financial system's interconnections has been a first-order concern in developed economies since the 2008 global financial crises. Macroeconomic prudential regulation needs to determine which banks and financial institutions are systemically important to supervise the systemic risks closely. In the network language, it would mean that such financial institutions have a large centrality. Most of this analysis has been done either by concentrating on symmetric responses (computation based on correlations) or by observing a subset of financial contracts. This approach has been quite fruitful. In our view, however, both approaches might be incomplete.

The first approach, which estimates networks using correlation information, tends to obviate the directionality of the effects and lack of structural interpretations. The second approach, which concentrates on the detailed description of the contingent contracts across banks, could represent a solution to this problem. However, it is virtually impossible to observe all possible contracts. Therefore, a market price of CDS is conceivably a more reliable measure of the actual exposure. This ambivalence implies that each procedure has a weakness that we have tried to address in this chapter. Furthermore, as has been shown in the literature, the nature of the transmission mechanisms changes with the shocks hitting the economy – meaning that the network is contingent on the state of the economy. We argue that the estimation of an asymmetric and contingent network requires a different identification method. We develop a methodology based on identification through heteroskedasticity. Applying this estimation method on CDS data of 10 large banks, we construct a financial network model and find that the data generating the model is consistent with three networks.

Our results indicate that the systemically important banks depends on the type of shock that hits the economy – which is ultimately transmitted through a different network. Without our contingent financial network model, it is not possible to identify the importance of each bank in the financial system when a specific type of shock hits the economy. Indeed, we reject that the data is explained by a single network – suggesting that a policy designed based on that network would be inappropriate when a different transmission mechanism governs the dynamics of the system.

From the regulatory point of view, It is crucial to understand the relative rankings on the financial institutions and how such ranking shifts in the sample when systemic risk is present. Our data is short, and therefore we are limited in our ability to observe shocks that have not happened yet. For instance, we have not observed large positive productivity shocks, or relatively high inflation rates, or even high interest rates.

Therefore, our conclusions are conditional on the sample we have seen. Within that sample, though, it is easy to identify much more granular network structures. Finally, the application of our identification method for contingent networks is not limited to policymaking. For example, asset management practitioners could use our method to estimate the contingent network and allocate assets according to the dominant shock in a period of time. Macroeconomists could use our method to evaluate the impact of macroeconomic interventions. In general, how to model, estimate and intervene in shock contingent networks is still an open and important topic for future research.

## 2.6 Appendices

### 2.6.1 A Single Network Endogenous Model

Assume two banks are related according to the following system of equations and shocks

$$x_t = \beta y_t + \epsilon_t \tag{2.25}$$

$$y_t = \alpha x_t + \eta_t \tag{2.26}$$

with reduced form,

$$\begin{aligned} x_t &= \frac{1}{1-\alpha\beta}\left(\beta\eta_t + \epsilon_t\right) \\ y_t &= \frac{1}{1-\alpha\beta}\left(\eta_t + \alpha\epsilon_t\right) \end{aligned}$$

where $\eta_t$ and $\epsilon_t$ are the structural shocks and $\alpha$ and $\beta$ are describing the network.

As it is, this model cannot be estimated from the data. Equations (2.25) and (2.26) describe the behavior of the data entirely with 4 parameters/variables: two shocks $\epsilon$ and $\eta$ and two parameters $\alpha$ and $\beta$. These four constitute the unknowns of the system. The problem of identification arises because there are three equations in four unknowns. The observable variables $x$ and $y$ have zero mean and in the data only three moments can be estimated; all from the variance-covariance matrix. What are the solutions to the problem? In economics, solutions tend to create circumstances in which an additional equation is added to the system of equations. For instance, the exclusion restriction in the instrumental variable approach boils down to assuming that one parameter is zero (the exclusion assumption). Randomized controlled trials assume that all the variation is due to the treatment — again, this is implicitly assuming that there is no feedback effect ($\beta = 0$). This is a very reasonable assumption when the experiment is properly designed. All these solutions are making a parameter assumption (usually that a parameter is equal to zero). The identification through heteroskedasticity has a slightly different flavor. The easiest way to explain how identification through heteroskedasticity works is to show the system of equation. Assume that the parameters are stable and that the data has heteroskedasticity. For simplicity assume that there are two heteroskedastic regimes. In this case, it is possible to estimate one variance-covariance matrix in each regime.

$$\begin{aligned}
\Omega_1 &= \begin{bmatrix} var(x_{t,1}) & covar(x_{t,1}, y_{t,1}) \\ & var(y_{t,1}) \end{bmatrix} = \frac{1}{(1-\alpha\beta)^2} \begin{bmatrix} \xi^2_{\epsilon,1} + \beta^2 \xi^2_{\eta,1} & \alpha\xi^2_{\epsilon,1} + \beta\xi^2_{\eta,1} \\ & \alpha^2\xi^2_{\epsilon,1} + \xi^2_{\eta,1} \end{bmatrix} \\
\Omega_2 &= \begin{bmatrix} var(x_{t,2}) & covar(x_{t,2}, y_{t,2}) \\ & var(y_{t,2}) \end{bmatrix} = \frac{1}{(1-\alpha\beta)^2} \begin{bmatrix} \xi^2_{\epsilon,2} + \beta^2 \xi^2_{\eta,2} & \alpha\xi^2_{\epsilon,2} + \beta\xi^2_{\eta,2} \\ & \alpha^2\xi^2_{\epsilon,2} + \xi^2_{\eta,2} \end{bmatrix}
\end{aligned}$$

There are six unknowns in the system. The two parameters ($\alpha$ and $\beta$), and four variances ($\xi^2_{\epsilon,1}$, $\xi^2_{\epsilon,2}$, $\xi^2_{\eta,1}$, and $\xi^2_{\eta,2}$). As can be seen, there are six equations in six unknowns. This means that the system of equations is just identified.

Notice that even though in each regime the system is under-identified (fewer equations than unknowns) the system as a whole is identified. The key assumptions are two: that the structural shocks are indeed structural (they are uncorrelated) and that the parameters are stable. In the end, the parameter stability allows the heteroskedasticity to add additional equations — which helps solve the identification problem.

The intuition of the two endogenous variables case is as follows. First, the graphical representation of the joint residuals in this model always takes the form of a rotated ellipse. Second, the rotation is summarized by the variance-covariance matrix.

In equations (2.25) and (2.26), the only meaningful moment we can compute to estimate the degree of contagion is the covariance matrix. An important question is then, what does the variance-covariance matrix represent? The errors in these models are distributed as a multinomial and their contours are ellipses. To fix concepts, let us start with a simple endogenous system of equations (2.25) and (2.26). The $95^{th}$ percentile of the errors is distributed as a rotated ellipse. We can solve for two independent normal distributions from the structural equations as follows (with some abuse of notation)

$$\begin{aligned}
\phi_1 = \frac{x_t - \beta y_t}{\xi_\epsilon} &\sim N(0,1) \\
\phi_2 = \frac{y_t - \alpha x_t}{\xi_\eta} &\sim N(0,1)
\end{aligned}$$

Because $\phi_1$ and $\phi_2$ are independent with mean zero and variance one, it is possible to describe the $\zeta$ confidence interval as $\phi_1^2 + \phi_2^2 = \zeta$. This is exactly an ellipse. Substituting

$$\left(\frac{x_t - \beta y_t}{\xi_\epsilon}\right)^2 + \left(\frac{y_t - \alpha x_t}{\xi_\eta}\right)^2 = \zeta \tag{2.27}$$

The two axes of the ellipse cannot be computed in closed-form solution, but they depend on the slope of the curves (structural parameters) as well as the relative variances of the shocks. In Figure $2-8$ a graphical representation is shown. Suppose that the blue curve represents the supply and the red is the demand (when there are no shocks). Then $x_t$ represent quantity and $y_t$ represent price. Furthermore, the points reflect some random realization of structural shocks that leads to a point far from the depicted schedules. The ellipse represents the 90th percentile. In this particular case $\beta$ is assumed to be negative (representing the “demand"), while $\alpha$ is positive. In Figure 2-8, the variance of the demand shocks is larger than the variance of the shocks to the supply, hence, the ellipse is closely aligned with the supply curve. In the limit, if the variance of the demand is infinitely large, the ellipse would coincide exactly with the supply curve.

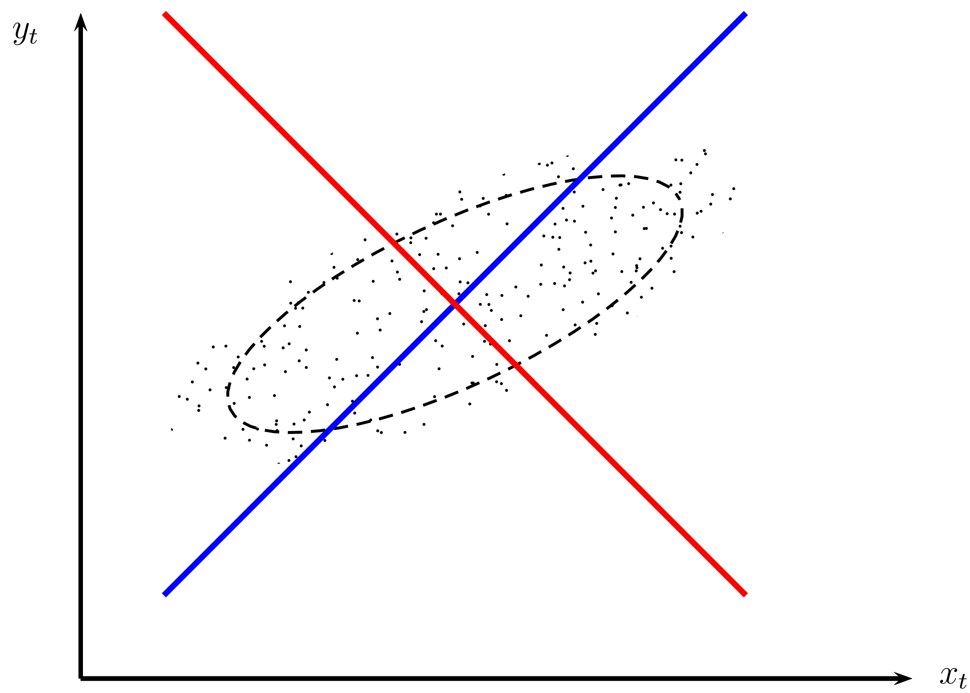



Figure 2-8: Distribution of Errors

The form of the ellipse is also summarized by the variance-covariance matrix computed in the reduced form. Additionally, most of the methodologies we study are based on the variance-covariance matrix. Therefore, all the sources of bias can be tracked to it. Finally, as mentioned previously, in this model the only statistic that can be computed from the data — that allows us to recover the structural parameters — is the variance-covariance matrix.

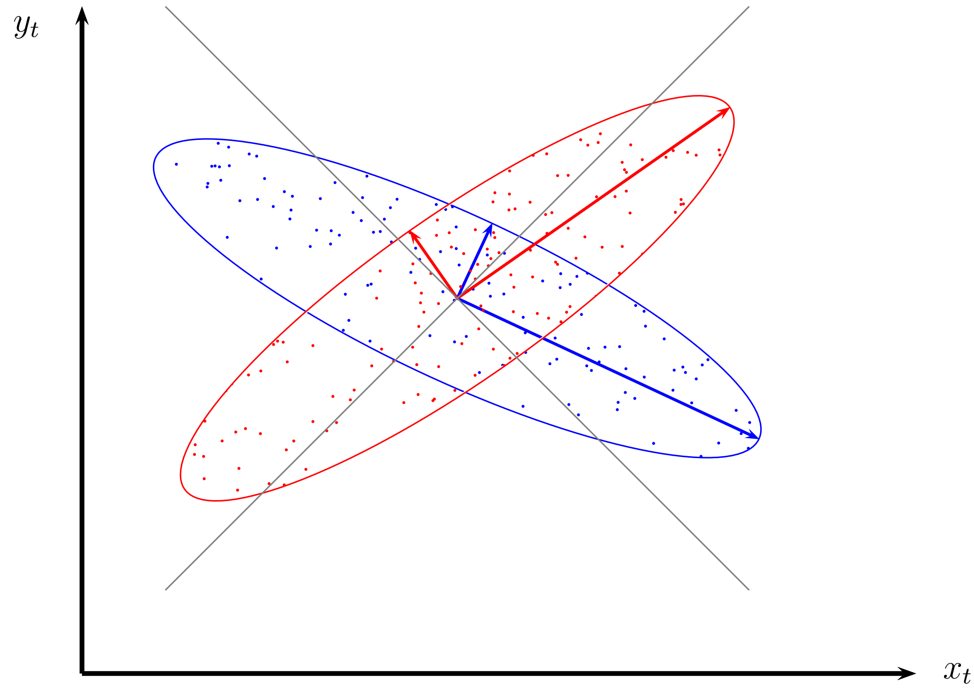



Figure 2-9: Identification Through Heteroskedasticity

The intuition behind the identification through heteroskedasticity comes from the rotation of the residual ellipses. When the variances change, for the same parameters, the ellipses rotate. In Figure $2-9$, we show two cases: One when the shocks to the demand dominate (red), and one when the shocks to the supply dominate (blue). In particular, when the shocks to the demand dominate, then the ellipse approximates the supply curve. In fact, it is identical to the supply curve if the variance of the demand is infinite relative to the supply. Conversely, when the supply shocks are larger, then the long axis of the ellipse tilts toward the demand curve. It is this rotation of the ellipses when the relative variances shift that provides the identification.

It is instructive to re-state the underlying assumptions: structural shocks are uncorrelated (quite uncontroversial) and parameters need to be stable across the regimes (so, this is a good technique to measure spillovers).

### 2.6.2 Equivalent Formulation via Tensor Decomposition

We first show that the identification problem is equivalent to a tensor decomposition problem.

In previous sections, we identify multiple layers by matching the second moments

$$\Omega_h = (1 - \Gamma_1)^{-1} \Xi_h (1 - \Gamma_1)^{-\top}$$

We define a new $N$-by-$M$ matrix

$$A = (1 - \Gamma_1)^{-1}$$

Then we can write the moment matching equation as

$$\Omega_h = A \Xi_h A^\top .$$

Because the matrix $\Xi_h$ is diagonal, we can further write

$$\Omega_h = \sum_{r=1}^{N} \overrightarrow{a}_r \xi_{rh} \overrightarrow{a}_r^\top$$

where $\overrightarrow{a}_r$ is the $r$th column of $A$ and $\xi_{rh}$ is the $r$th diagonal entry of $\Xi_h$. Because vector outer products can be written as tensor products, we can also write

$$\Omega_h = \sum_{r=1}^{N} (\overrightarrow{a}_r \otimes \overrightarrow{a}_r) \, \xi_{rh}$$

where $\otimes$ is the tensor product. Now if we stack all the second moments $\Omega_h$ along a third dimension, we obtain a $N$-by-$N$-by-$H$ tensor $[\Omega_h]$ and it holds that

$$[\Omega_h] = \sum_{r=1}^{N} \overrightarrow{a}_r \otimes \overrightarrow{a}_r \otimes \overrightarrow{\xi}_r \tag{2.28}$$

where $\overrightarrow{\xi}_r = \begin{bmatrix} \xi_{r1} & \xi_{r2} & \cdots & \xi_{rH} \end{bmatrix}^\top$.

We can obtain estimates of $\Gamma_1$ by taking the rank-$N$ tensor decomposition of $[\Omega_h]$. According to the Kruskal's rank condition, if

$$Krank(a_r) + Krank(a_r) + Krank(\xi_r) \geq 2R + 2$$

then the tensor decomposition problem has a unique solution.

### 2.6.3 Proof of Lemma 2

*Proof.* First, constants are irrelevant, so we just have to prove

$$\begin{aligned} \Xi_h^* :&= \arg\max_{\Xi_h \in \mathbb{PD}} \frac{1}{|V_{mh}|^{n_h/2}} |\Omega|^{(n_h - N - 1)/2} e^{-(1/2)\operatorname{tr}(V_{mh}^{-1}\Omega)} \\ &= diag\Big((1-\Gamma_m)\Omega(1-\Gamma_m)^\top\Big) \end{aligned} \tag{2.29}$$

where $V_{mh} := \frac{1}{n_h}(1-\Gamma_m)^{-1}\Xi_h(1-\Gamma_m)^{-\top}$.

Then, the trace term in the exponential can be reduced to

$$\begin{aligned} \operatorname{tr}(V_{mh}^{-1}\Omega) &= \operatorname{tr}\left(\left[\frac{1}{n_h}(1-\Gamma_m)^{-1}\Xi_h(1-\Gamma_m)^{-\top}\right]^{-1}\Omega\right) \\ &= \operatorname{tr}\left(n_h(1-\Gamma_m)^\top \Xi_h^{-1}(1-\Gamma_m)\Omega\right) \\ &= \operatorname{tr}\left(n_h \Xi_h^{-1}(1-\Gamma_m)\Omega(1-\Gamma_m)^\top\right) \\ &= \operatorname{tr}\left(n_h \Xi_h^{-1} diag\Big((1-\Gamma_m)\Omega(1-\Gamma_m)^\top\Big)\right) \end{aligned} \tag{2.30}$$

Define $D$ to be the new optimization variable with

$$D = n_h \Xi_h^{-1} diag\Big((1-\Gamma_m)\Omega(1-\Gamma_m)^\top\Big)$$

Note by definition, $D$ is a diagonal matrix. Now we insert

$$\Xi_h = n_h D^{-1} diag\Big((1-\Gamma_m)\Omega(1-\Gamma_m)^\top\Big)$$

and change the optimization variable to $D$. Now the optimization problem becomes

$$\begin{aligned} D^* &= \arg\max_{D\in\mathbb{PD}} \frac{|\Omega|^{(n_h-N-1)/2} e^{-(1/2)\operatorname{tr}(D)}}{\left|(1-\Gamma_m)^{-1} n_h D^{-1} diag\Big((1-\Gamma_m)\Omega(1-\Gamma_m)^\top\Big)(1-\Gamma_m)^{-\top}\right|^{n_h/2}} \\ &= \arg\max_{D\in\mathbb{PD}} \frac{|\Omega|^{(n_h-N-1)/2} e^{-(1/2)\operatorname{tr}(D)}}{\Big(|(1-\Gamma_m)^{-1}|^2 \cdot n_h \cdot |D^{-1}| \cdot |diag((1-\Gamma_m)\Omega(1-\Gamma_m)^{-\top})|\Big)^{n_h/2}} \end{aligned} \tag{2.31}$$

Now because $\Omega$ is positive definite, $|\Omega| > 0$ and $|diag((1-\Gamma_m)\Omega(1-\Gamma_m)^{-\top})| > 0$. Eliminating positive terms that does not involve $D$ will not change the optimization problem. Hence, the problem reduces to

$$D^* = \arg\max_{D\in\mathbb{PD}} \frac{e^{-(1/2)\operatorname{tr}(D)}}{|D^{-1}|^{n_h/2}}$$

and we just have to prove that the optimal $D^*$ is a diagonal matrix with all entries equal to $n_h$. Let $d_i$ be the diagonal entries of $D$. Note that $|D| = \prod_i d_i$ and $\operatorname{tr}(D) = \sum_i d_i$. Now the optimization problem becomes

$$d_i^* = \arg\max_{d_i>0} (\prod_i d_i)^{n_h/2} e^{-\frac{1}{2}\sum_i d_i} \tag{2.32}$$

Setting the first derivative to 0, we obtain the solution of the optimization problem $d_i^* = n_h$. $\square$

### 2.6.4 Credit Default Swap Data Details and Data Retrieval Process

To better assist understanding our results or reproducing our results, we list carefully the details of data we used in this chapter.

We obtain the par mid spread of the credit default swap of the 10 target banks through the Thomson Reuters Eikon excel tool.

Step 1: CDS Ticker search

| Bank RIC | Primary CDS RIC |
|---|---|
| JPM | JPM5YUSAX=R |
| BAC | BAC5YUSAX=R |
| WFC | WFC5YUSAX=R |
| C | C5YUSAX=R |
| GS | GS5YUSAX=R |
| MS | MS5YUSAX=R |
| COF | COF5YUSAX=R |
| HSBC.K | HSBA5YEUAM=R |
| AXP | AXP5YUSAX=R |
| CSGN.S | CSGN5YEUAM=R |

Table 2.3: RIC tickers for primary CDS products of target banks.

Due to the variety of CDS product exist in the market, we first need to decide which CDS we use. Thomson Reuters has decided a primary CDS for each bank through the ticker searching service. Input the command

=@TR("JPM;BAC;WFC;C;GS;MS;COF;HSBC.K;AXP;CSGN.S", "TR.CDSPrimaryCDSRic","CH=Fd RH=IN",B2)

into the Eikon excel tool, then we have the tickers of primary CDS products of the target companies. The obtained tickers are in Table 2.3.

Most of those produces are 5 years CDS contracts traded in US. For HSBC and CSGN, they are 5 year CDS contracts traded in Euroupe.

Step 2: Retrieve par mid spread data for CDS

After obtaining those tickers, we request the actual par mid spread of those CDS products in the target date range. The command for retrieving spread data is

=@TR("JPM5YUSAX=R;BAC5YUSAX=R;WFC5YUSAX=R;C5YUSAX=R; GS5YUSAX=R;MS5YUSAX=R;COF5YUSAX=R;HSBA5YEUAM=R;AXP5YUSAX=R; CSGN5YEUAM=R","TR.PARMIDSPREAD","Frq=D SDate=20090901 EDate=20170630 CH=Fd;IN RH=date",B2)

# Chapter 3

# Uncertainty and Robustness in Central Banking

*This work was performed in collaboration with Roberto Rigobon and Munther Dahleh.*

## 3.1 Introduction

After the 2008 Global Financial Crisis, Central Banks expanded their monetary policy tools when the conventional policy tool, namely short-term interest rates, approached or reached zero. Heavy asset purchase interventions, all under the umbrella of "quantitative easing", and massive signaling efforts on what became known as "forward guidance", were some of the examples used by central bankers worldwide. [1]

Furthermore, the monetary policy response to the Coronavirus Pandemic in 2020 has led the monetary authorities to replicate some of the actions started a decade before. Central Bankers in developed nations responded by lowering interest rates to zero, and continuing with promises of massive asset purchases. Between March and May of 2020, the FED had already purchased assets worth about 1.5 trillion dollars — doubling the speed of asset purchasing followed after the 2008 crisis.

Although these policies were crucial in the management of monetary policy in the

[1] Equivalent operations include, but are not limited to, "Operation Twist" by the Reserve Bank of India and targeted longer-term refinancing operations (TLTRO) by European Central Bank

aftermath of the financial crisis, and conceivably they will be useful in the management of the pandemic, it is worth asking whether or not the central bank can do a better job controlling the economy by using many tools simultaneously. Or is their ability to control the economy enhanced by committing to using one tool at a time or even no tools at all. At first glance, the immediate reaction to this question seems obvious: having more tools at the disposal of the Central Bank is always better. However, that is not the case. There are conditions in which the ability to manage the economy is improved when the Central Bank is able to commit never to use a given mechanism. These conditions, surprisingly, are not far-fetched. The conditions for having a detrimental impact on the economy when there are too many tools are two: First, the system suffers from uncertainty — not knowing the distribution — rather than risk [2]. Second, the uncertainty changes the market equilibrium[3] and therefore it is incorporated into the system through market expectations. When these conditions exist, it is possible that in the economy, volatility is smaller when the central bank is using fewer tools rather than more.

The first condition involves uncertainty with unknown distribution as opposed to risk with known distribution. Uncertainty with unknown distribution could come from the fact that economic models are not perfectly calibrated. Generally speaking, policy decisions are made based on a model with some mismatch. For example, when productivity increases are described as "between 0.5 and 1 percent" or when the natural rate of unemployment is described as "between 3 and 5 percent" we are describing a range of possible outcomes rather than its mean and standard deviation. Furthermore, the discussion in early 2019 about the steady state balance sheet of the FED is also framed in terms of a range of possible outcome. In general, most macroeconomic models tend to assume the distribution is known, but a large literature[4] has studied the changes implied by deviating from the canonical assumptions.

[2] In this Chapter, we use the term uncertainty to refer to a variable with unknown distribution, and risk to refer to a random variable with known distribution

[3] We will define this equilibrium in a game played by the market participants and the central bank

[4] See (Hansen and Sargent, 2008) for robust control in general, (Vitale, 2018) for robust Kyle, (Miao and Rivera, 2016) for robust contract, (Maccheroni et al., 2013) for robust mean-variance portfolio choice, (Strzalecki, 2011) for the axiomatization of multiplier preferences, and many others.

The second condition considers the interaction between the central bank and the market as a Bayesian game with incomplete information. In such a game, the central bank's action not only will have a direct effect on the market, but also will have an indirect effect by signaling and changing the equilibrium of the game. In the language of central banking, this can be interpreted as changing the optimal target. Any model uncertainty of the central bank, in this scenario, can be passed on and become an uncertain target. Therefore, the volatility of the economy increases by the possibility of using the policy tool. For example, the above signaling effect is already observed in US forward guidance. Andrade et al. (2019) shows that the forward guidance has a mixed impact on the economy: First, it could stimulate economic activities by providing liquidity to the market. Second, it also signals a weak macroeconomic perspective, thus slowing down consumer spending. Andrade et al. (2019) shows that both effects are significant.

The Central Bank can affect the market in several ways. In this chapter, we explore two models to highlight the interaction between bounded uncertainty and market equilibrium. The first model is about interventions in the form of asset purchases, while the second one is about signaling as the outcome of forward guidance.

The idea of forward guidance is inherent in the development of the New Keynesian Model. In the book by (Galí, 2015), two cases are discussed in separate sections: the fully discretionary central banking case, where the central bank's policy decision has no impact on the market expectation of inflation, and the fully committed central banking case, where the market expectation of inflation is computed from the central bank's committed interest rate rule. The fully discretionary case has been a relatively good model in practice because the central bank did not stick to any specific monetary policy rules, and the interest rate is set on board meetings. However, the discussion of the fully committed central banking case in the book depicts how commitments, such as forward guidance, might impact the market. Nevertheless, the actual situation after forward guidance being introduced is much more complicated. On the one hand, the central bank's forward guidance is not legally binding, and a discretionary interest rate setting is allowed. On the other hand, the central bank has to follow its own forward guidance to make it credible. Otherwise, it becomes cheap talk and will not affect market expectation anymore.

Mixing discretionary rule and committed rule in the New Keynesian model brings another complication regarding the forward expectation term[5]. With the discretionary rule in mind, many empirical studies (e.g. (Rudd and Whelan, 2005) and (Galı and Gertler, 1999)) assumes that the central bank's policy decision has no impact on the market expectation of inflation. As a result, the forward expectation $\mathbb{E}_t[\pi_{t+1}]$ equals to $\pi_{t+1} + \epsilon_{t+1}$, where $\epsilon_{t+1}$ is a shock that has zero mean and unobservable until time $t+1$. On the contrary, with committed central banking rule in mind, the theoretical work by Farmer et al. (2007) derive the forward expectation term from the interest rate rule and forward-looking propagation. When a discretionary rule and committed rule appear in the same model, both effects need to be considered.

The ineffectiveness of forward guidance tools has raised many research interests. Apart from our explanation through model uncertainty and robustness, others combine nominal price rigidities with a model of precautionary savings [6] to explain the weak power of forward guidance to expand the economy (McKay et al., 2016). In that paper, agents' precautionary savings come from an incomplete market assumption. The incomplete market assumption means agents cannot allocate their assets optimally according to their risk preference; therefore, a precautionary saving appears to prepare for the uninsured risk, and reduce the effectiveness of forward guidance. However, the example showing the power of forward guidance tool in the first section of (McKay et al., 2016) made some unrealistic assumptions. It did not fully reveal the mechanism of forward guidance: it assumes that the central bank can directly change the real rate, and show that an expected real rate adjustment in the future can affect the output gap by a large margin. In practice, the central bank does not have the tool to affect the real rate directly; it can only affect the nominal interest rate and then propagate the effect through price stickiness or cash-in-advance constraints. As a result, the forward guidance tool may be very weak even in a complete market.

[5]Here we refer to $\mathbb{E}[\pi]$ and $\mathbb{E}[x]$ in New Keynesian Phillips Curve (NKPC) and Euler equation.

[6]This is also used to analyze many other issues. See e.g. (Guerrieri and Lorenzoni, 2017) for credit crisis, (Oh and Reis, 2012) for transfer spending effect.

The effect of the expectation channel along has been studied quite intensively in the literature. For example, Stein and Sunderam (2018) examined the relationship between interest rate policies and the bond market through the expectation channel. In this paper, it is assumed, just as we do in this chapter, that the central bank has private information about its long-run target rate. With this assumption, the central bank would use the interest rate policy with much more precaution and gradualism because the expectation channel will magnify the system's volatility. However, without uncertainty and robustness concerns, the central bank will still prefer always to use a policy tool whenever it is available.

The intuition of the results is simple. When there is uncertainty instead of risk, it implies a much more a dampened action than when the distribution is known. Furthermore, there exist a set of conditions that the optimal policy is inaction — which in our case it implies committing not to use an instrument.

The rest of this chapter is organized as follows: Section 2 shows the impact of bounded uncertainty on central bank decision rule. A Kyle model of markets is adopted to study the behaviour of central banker, market maker and informed trader. Section 3 illustrates the effect of bounded uncertainty on forward guidance tools, where a linearized New Keynesian model with asymmetric information is used. Section 4 discusses several intuitions of the phenomenon presented previously in this chapter, as well as the relevance of bounded uncertainty modeling in central banking.

## 3.2 Central Bank Intervention: An Asymmetric information Perspective

This section considers a market with noise traders and asymmetric information as in the traditional Kyle model. We extend the model in two directions. First, we add a central bank as a player in the game. The central bank has more accurate observation or modeling of the behavior of the noise traders. The central bank's objective is to reduce market inefficiency. Hence, the central bank intervenes in the market by purchasing assets in order to achieve its policy goals.[7] The second extension is the inclusion of an uncertain trader. The uncertain trader acts as a noise trader except that their actions cannot be described by a known distribution and forces others to take a robust optimization approach.

We start the section by analyzing the Kyle model when a central bank intervention occurs. We call this the nominal model — in the sense that the distributions of all the shocks are known. Then we introduce the uncertain trader into the mismatched model, and analyze its impact on the equilibrium and on the robust policy rules of each agent.

### 3.2.1 Nominal Model: Kyle model with Central Bank

This section presents the classical Kyle model and its extension when a central bank is included. This model we identify as the Nominal Model.

Consider a single period security pricing model as follows.

- Nature decides the realization of valuation $v$. $v \sim \mathcal{N}(p_0, \sigma_v^2)$.
- A noise trader first submit to the exchange an order to buy $u$ shares of a risky asset with end of period payoff $v$, where $u \sim \mathcal{N}(0, \sigma_u^2)$
- An informed trader, knowing the realized value of $v$, submit an order to buy $x$ shares of the risky asset. The order size is chosen to maximize the informed trader's expected payoff
$$\max_x E[x(v-p)|v] \tag{3.1}$$
- A market maker, after observing the total order size $x+u$, choose a market price $p(x+u)$ to clear the market. Market makers are fully competitive and risk neutral, therefore the equilibrium price is given by
$$p = E[v|x+u] \tag{3.2}$$

[7] The modelling of central bank intervention using a modified Kyle model is not unique; see Pasquariello et al. (2011) and Pasquariello (2017) for similar constructions.

In this model, the noise traders are individuals and institutions entering market purely for liquidity purposes. For example, a index fund must sell its constituent stocks when its clients cash out their investments. When the economy is hit by a major aggregate shock, noise traders will execute correlated trades with a macro-level impact. Furthermore, the informed traders are individuals and institutions with information advantages over average market participants. Examples of them include large mutual funds that do extensive research in a particular sector or company. Finally, market makers are individuals and institutions making both bids and asks and provide liquidity to the market. Examples of them include large securities companies.

The above model is a signaling game with Perfect Bayesian Equilibrium. The optimal order for the informed agents, and the market clearing prices are:

$$x = \frac{\sigma_u}{\sigma_v}(v - p_0) \tag{3.3}$$

$$p = p_0 + \frac{1}{2}\frac{\sigma_v}{\sigma_u}(u + x) \tag{3.4}$$

We define market inefficiency $\Psi$ as follows

$$\Psi = \mathbb{E}[(v - p)^2 + \lambda(v - p)x] \tag{3.5}$$

The market inefficiency is measured by the sum of price mismatch and informed trader's profit. The first part says the end-of-period price should reveal the valuation of the asset as much as possible. The second part says the price of information should be as small as possible.

Insert the equilibrium strategies into the definition, we have

$$\begin{aligned}\Psi_{\text{Kyle}} = \mathbb{E}[(v-p)^2 + \lambda(v-p)x] &= \mathbb{E}[(v - p_0 - \frac{1}{2}\frac{\sigma_v}{\sigma_u}(x+u))^2] + \lambda\mathbb{E}[(v - p_0 - \frac{1}{2}\frac{\sigma_v}{\sigma_u}(x+u))\frac{\sigma_u}{\sigma_v}(v-p_0)] \\ &= \mathbb{E}[(\frac{1}{2}(v-p_0) - \frac{1}{2}\frac{\sigma_v}{\sigma_u}u)^2] + \lambda\mathbb{E}[(\frac{1}{2}(v-p_0) - \frac{1}{2}\frac{\sigma_v}{\sigma_u}u)\frac{\sigma_u}{\sigma_v}(v-p_0)] \\ &= \frac{1}{2}(\lambda\frac{\sigma_u}{\sigma_v}+1)\sigma_v^2\end{aligned}$$

In summary, the market inefficiency in the case of the standard Kyle model is given by

$$\Psi_{\text{Kyle}} = \frac{1}{2}(\lambda\frac{\sigma_u}{\sigma_v}+1)\sigma_v^2 \tag{3.6}$$

We extend the previous model to introduce a Central Bank. We model the central banker as an agent with information about the noise traders (signal $h$), whose objective is to reduce the market inefficiency by performing an intervention (order $z$).

We assume that the central bank observes the noise trader's submitted order with some noise

$$h = u + \epsilon, \epsilon \sim \mathcal{N}(0, \sigma_g) \tag{3.7}$$

The government agency submit an order to buy the asset with size $z$ to minimize the market inefficiency

$$E[(v-p)^2 + \lambda(v-p)x|h]$$

The only change to the Kyle Model is on equation 3.2. The set of fully competitive, risk-neutral market makers, set prices after observing the total order size $x + u + z$. The equilibrium price is given by

$$p = E[v|x+u+z] \tag{3.8}$$

Conceptually, it is optimal for the government agency to do a Bayesian inference and reduce the effect of noise trader. The optimal intervention is given by

$$z = -\frac{\sigma_u^2}{\sigma_u^2 + \sigma_g^2} h$$

and the joint noisy-central bank orders follow,

$$u - z = u - \frac{\sigma_u^2}{\sigma_u^2 + \sigma_g^2} h = \frac{\sigma_g^2}{\sigma_u^2 + \sigma_g^2} u + \frac{\sigma_u^2}{\sigma_u^2 + \sigma_g^2} \epsilon \sim \mathcal{N}(0, \frac{\sigma_u^2 \sigma_g^2}{\sigma_u^2 + \sigma_g^2})$$

Define

$$\sigma_{cb} = \frac{\sigma_u \sigma_g}{\sqrt{\sigma_u^2 + \sigma_g^2}} \tag{3.9}$$

as the adjusted standard deviation after the central bank intervention. The PBE is given by

$$x_{\text{nom}} = \frac{\sigma_{cb}}{\sigma_v}(v - p_0) \tag{3.10a}$$

$$p_{\text{nom}} = p_0 + \frac{1}{2}\frac{\sigma_v}{\sigma_{cb}}(x + u + z) \tag{3.10b}$$

$$z_{\text{nom}} = -\frac{\sigma_u^2}{\sigma_u^2 + \sigma_g^2} h \tag{3.10c}$$

Here the subscripts "mm" refers to the solutions with model mismatch.

**Lemma 3.** *The set of strategies in* (3.10) *is a perfect Bayesian equilibrium of the insider trading problem with government intervention.*

See proof in the appendix.

The market inefficiency, measured by the price mismatch and informed trader's profit, is given by

$$\Psi_{\text{nom}} = \frac{1}{2}(\lambda \frac{\sigma_{cb}}{\sigma_v} + 1)\sigma_v^2 \tag{3.11}$$

Notice the differences between equations 3.6 and 3.11. The effect of central bank intervention is equivalent to having reduced the variance of the noise traders. $\sigma_{cb} < \sigma_u$. The inefficiency has two parts, one is the profit made by the informed trader, and the other is the mismatch between market clearing price and the value of assets. For the first part, reducing the variance of noise trader will reduce the market capacity of the informed agents. Hence, reducing their profits. For the second part, the price mismatch is independent of the noise trader's variance and, therefore, independent of the actions taken by the central bank. Consequently, the central bank responds by attenuating the impact of the noise traders (equation 3.10c). From the equilibrium point of view, this is equivalent to reducing the volatility of the market. Lastly, Central Bank intervention is always good in this model.

### 3.2.2 Model Mismatch and Robustness with an Uncertain Trader

We extend the Kyle model to include an uncertain trader. The uncertain trader puts an order $\delta$ at the same time as the noise trader. We model the order as belonging to the interval $[-\Delta, \Delta]$ with unknown distribution. After this order is added, the overall order size becomes $u + x + \delta$ without a central bank, and $u + x + z + \delta$ with a central bank.

To understand the implications of the uncertainty in the model, we proceed in steps. First, we introduce the impact of a constant unmodeled $\delta$. Unmodeled implies that the central bank's optimal reaction functions and the informed agents do not consider the existence of $\delta$. That is an extreme simplification but allows us to provide a simple intuition on how the uncertainty enters the market inefficiency.

The following step is to analyze the optimal robust response of the central bank and the informed agents. We show that the uncertainty introduces a non-trivial region of inaction, i.e., it is optimal for an agent to set its action to a constant zero. That means sometimes the market efficiency is improved by either the central bank committing not to intervene. Economies operating under those circumstances achieve lower inefficiencies by not using the policy tools at all.

#### Robustness Analysis with an Uncertainty Trader

First, we would like to highlight the implications of the uncertain trader assuming that the informed agent and the central bank do not modify their optimal responses.

In the model without the central bank, the only change is on the market-clearing price.

The price is given by

$$\tilde{p}_{mm} = p_0 + \frac{1}{2}\frac{\sigma_v}{\sigma_u}(u + x + \delta) \tag{3.12}$$

and the market inefficiency is given by

$$\tilde{\Psi}_{mm} = \frac{1}{2}(\lambda\frac{\sigma_u}{\sigma_v} + 1)\sigma_v^2 + \frac{1}{4}\frac{\sigma_v^2}{\sigma_u^2}\delta^2 \tag{3.13}$$

Here the subscripts "mm" refers to the solutions in the nominal model.

Notice how the uncertainty term $\delta$ enters differently in the expression of the inefficiency (compare with the volatility of the noise traders in (3.6)). Remember we have assumed that the optimal $x$ has not changed from the Kyle model (equation 3.3), and therefore in the computation of the expectation $E[\lambda(v - \tilde{p}_{mm})x]$, $\delta$ drops out. The distribution of $u$, as part of the statistical model, is known to all agents. But the model uncertainty, $\delta$, is not part of the decision of any of the agents. It is treated as an unknown constant. Therefore, its covariance with the optimal actions is zero. Hence, $\delta$ only enters the inefficiency through its impact on prices.

In this model, the informed trade does not have a robust decision rule, so it does not take into account the model uncertainty $\delta$. If it does, the results will differ. As a matter of fact, uncertainty is different from the risk in terms of preference as well. In a model with risk, agents optimize payoffs using expectations. With model uncertainty, agents optimize the worst-case payoff using minimax. We show the minimax optimization in Section 3.2.2.

The inclusion of the central bank implies a market clearing price is set to

$$\tilde{p}_{mm_{cb}} = p_0 + \frac{1}{2}\frac{\sigma_v}{\sigma_{cb}}(u + x + z + \delta)$$

and inefficiency of

$$\tilde{\Psi}_{mm_{cb}} = \frac{1}{2}(\lambda\frac{\sigma_{cb}}{\sigma_v} + 1)\sigma_v^2 + \frac{1}{4}\frac{\sigma_v^2}{\sigma_{cb}^2}\delta^2 \tag{3.14}$$

Notice that we can now compare the inefficiency when the central bank exists, and when the central bank does not exists. This is the comparison between equations (3.13) and (3.14). As opposed to the classical Kyle model, when a central bank intervention always reduces the degree of inefficiency, this is not the case when uncertainty exists.

Note that since $\sigma_u > \sigma_{cb}$ for a relatively small mismatch ($\delta$ close to zero), it is better for the economy that the central bank exists and intervenes. However, when the mismatch is large (large values of $\delta$), the economy is better off not having a central bank. The market inefficiency as a function of the degree of mismatch is shown in Figure 3-1.

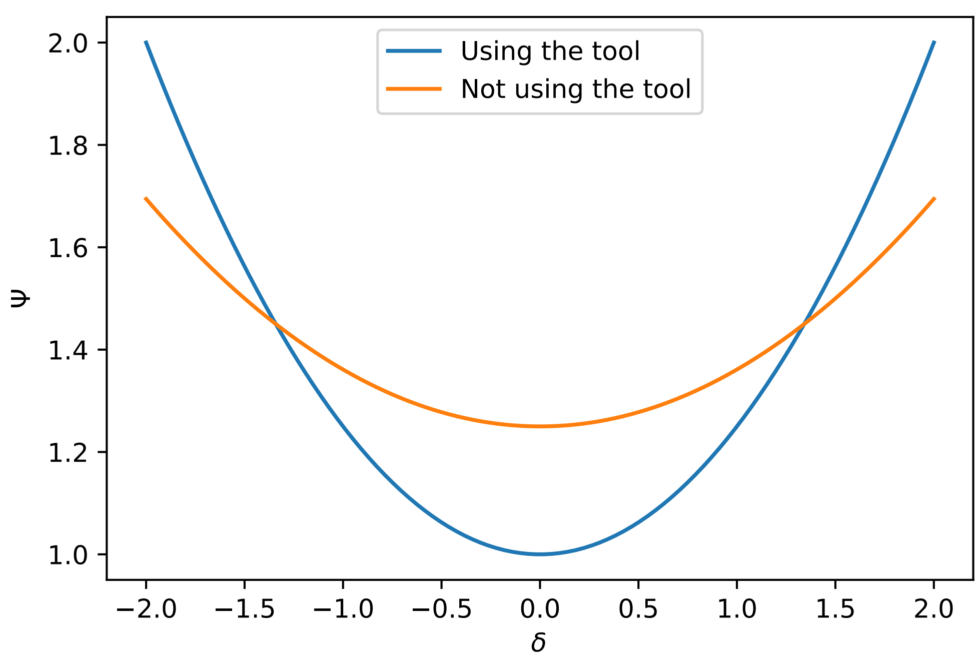



Figure 3-1: Market inefficiency in the Kyle model with model mismatch

What does it mean “not to have a central bank"? It means that a central bank needs to commits to set $z = 0$ at all times. The reason is that the existence of a central bank changes the expectations of the market makers.

As shown in the equations above, the central bank increases the inefficiency when we have a fixed order $\delta$. That is not a surprise. The central bank's strategy is optimized for the nominal model. When the actual model is mismatched with the nominal model, the strategy is no longer optimal and will cause higher inefficiency compared to the nominal model. However, the interesting thing is that the model mismatch deteriorates the inefficiency more significantly when the central bank intervenes compared to when the central bank does not intervene. This phenomenon has several different aspects. First, because the central bank's nominal model has a mismatch with the true model, its decision is not guaranteed to be optimal. Decisions based on a mismatched model could increase, not reduce the market inefficiency. Second, with central bank intervention, agents (including the informed trader and the market maker) expect the market with central bank intervention to be more efficient. When the model is wrong, however, the market is also more efficiently wrong. As a result, the market is better without central bank intervention. Third, this example echoes a series of literature in robust control theory. In many engineering systems, it is observed that optimizing for efficiency reduces robustness. In our central banking model, the expectation channel amplifies the downside robustness issue so that the market is better without central bank intervention.

### Robust Strategies for the Central Bank and Other Participants

All prior derivations assumed that informed agents and central banks kept their policy actions unchanged in the presence of uncertainty. In fact, the uncertainty was treated as a *constant* displacing the distribution of the noise traders. In this section, we explore the implications to the optimal responses of the central bank and the informed agent. We use robust control and assume each agent designs a policy choice considering the worst possible realization of $\delta$ within the existing bounds.

We study the central bank's robust response, the informed agent's robust response, and the market maker's robust response.

Suppose the central bank is aware of the potential model mismatch and wants to adopt a robust decision rule. On the other hand, other agents are not aware of the model mismatch and the fact that the central bank is using a robust decision rule.

Here, choosing a optimal linear decision rule $h \to z$ is equivalent with choosing a optimal variance $\sigma$ for $u+z$ that solves the following minmax problem

$$\min_{\sigma\in[\sigma_{cb},\sigma_u]}\max_{\delta\in[-\Delta,\Delta]}\frac{1}{2}(\lambda\frac{\sigma}{\sigma_v}+1)\sigma_v^2+\frac{1}{4}\frac{\sigma_v^2}{\sigma^2}\delta^2$$

Here the central bank faces a tradeoff: market efficiency versus market stability. Applying a robust control rule can increase market stability, but reduces market efficiency. This trade off is to be considered when the policy tool is introduced.

The robust decision rule of the central bank applies a minimax preference. This results in different decision rules compared to the risk term $u$, which has a known distribution. An immediate difference is that the robust decision rules with model uncertainty generates an inaction region, where the central bank prefer to set its order size to constant zero. As discussed before, when model uncertainty is large, the market inefficiency is less without central bank intervention. Therefore, the solution of the above minimax problem is $z=0$ if $\Delta$ is larger than a fixed threshold.

Now suppose the central bank and the informed agents are both robust. This time, other agents are also aware of the model mismatch and the fact that the central bank uses a robust decision rule. Now a new perfect Bayesian equilibrium has to be established.

First of all, the max-min problem of the informed trader is given by

$$\max_x\min_\delta E[x(v-p)|v]$$

While the minimax optimization problem for the informed trader is straightforward, the robust decision rule of the market maker is less obvious. In the original model, the market maker sets the price as the expected value of assets given the observed combined order $x+u+z$. In the robust control setting, it is natural to assume that it sets the price as a minimax estimator of the conditional expectation of value as follows:

$$p=\arg\min_p\max_{\delta\in[-\Delta,\Delta]}|\mathbb{E}[v|x+u+z+\delta]-p| \tag{3.15}$$

Due to the symmetric nature of the minimax problem, the solution of the market maker's robust decision rule is given by

$$p = \mathbb{E}[p|x + u + z + \delta, \delta = 0]$$

i.e., the market maker acts as if the model uncertainty does not exist.

Furthermore, the central bank solves the following optimization problem

$$\min_{z} \max_{\delta \in [-\Delta, \Delta]} \mathbb{E}[(p_\delta - v)^2 | h] + \lambda \mathbb{E}[(v - p_\delta) x | h]$$

Simulation results show that when the informed trader and market makers are also robust, the optimal decision of the central bank is to use the tools even less.

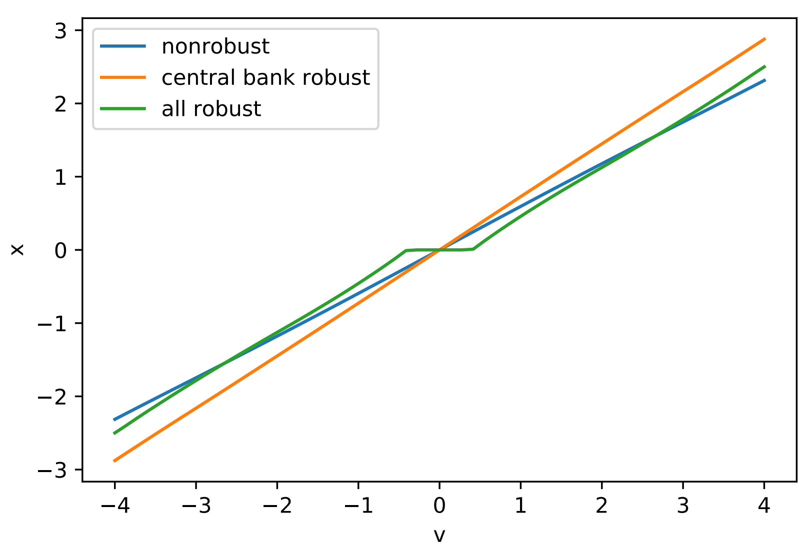



Figure 3-2: The optimal response of the informed trader in three scenarios: non-robust, only the central bank being robust, and everyone being robust. The informed trader observes asset value $v$ and then make a decision on purchase volume $x$.

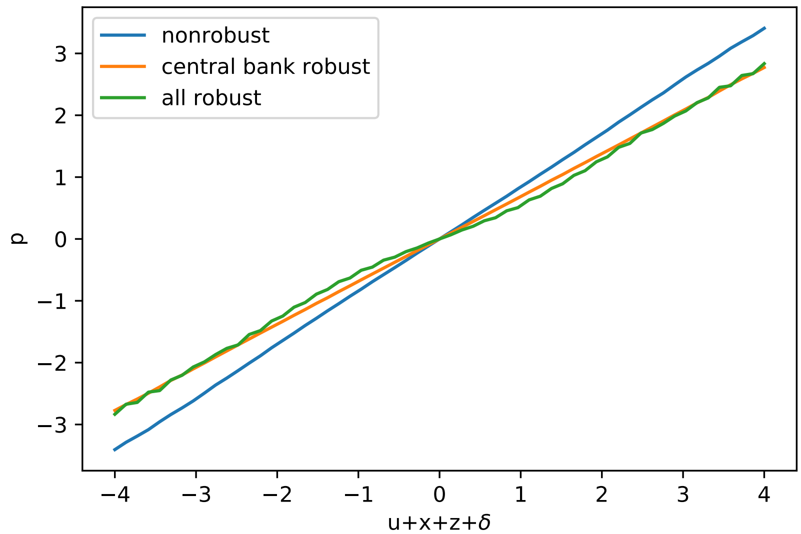



Figure 3-3: The optimal response of the market maker in three scenarios: non-robust, only the central bank being robust, and everyone being robust. The market maker observes combined order volume $u + x + z + \delta$ and then decides market clearing price $p$.

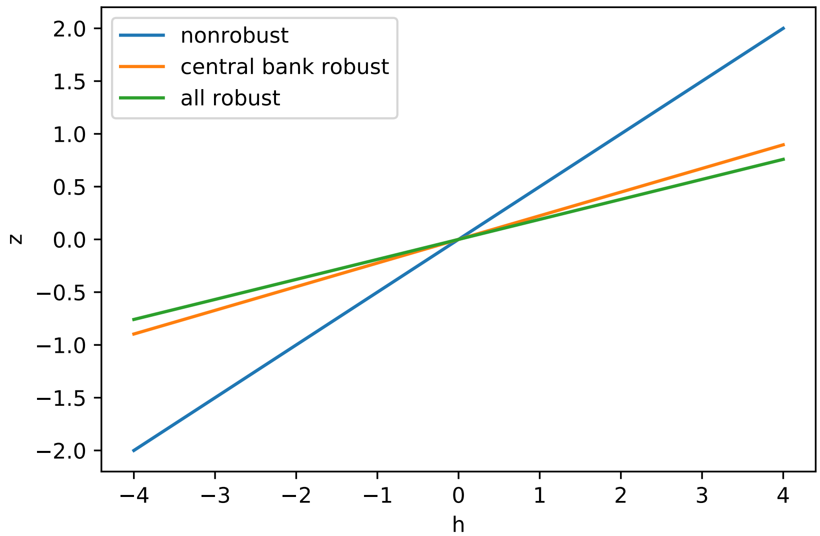



Figure 3-4: The optimal response of the central bank in three scenarios: non-robust, only the central bank being robust, and everyone being robust. The central bank takes an observation $h$ of the noise trade $u$ and then decides policy intervention $z(h)$.

As shown in Figure 3-2, when all agents adopt robust decision rules, the informed trader's optimal response has an inactive region. The inactive region has the following intuition: when the known asset value $v$ is not too far away from zero, unmodelled turbulence (uncertainty term $\delta$) could distort the market so that a trade would lose money in the worst-case scenario. Since the informed trader is robust against that uncertainty, it would prefer no to trade and wait for a better trading opportunity that secures a positive profit. This inactive region of the informed trader, just like the inactive region of the central bank, is caused by the robust response of model uncertainty. However, they are not exactly the same. The central bank is always active when the uncertainty is small, and always inactive when the uncertainty is large. The informed trader, on the other hand, is always inactive for an infinitesimal $v$ as long as the uncertainty exists. The difference in the robust decision is due to the difference in the way that uncertainty term enters their optimization objective function. The uncertainty term enters the objective function of the central bank in the quadratic form, so that direction does not matter; on the other hand, the uncertainty term enters the objective function of the informed trader in a way that direction matters.

In summary, the inactive region of the central bank and the informed trader have different causes: the inactive region of the central bank is caused by the fact that uncertainty deteriorates the market more when the central bank intervenes, and the inactive region of the informed trader is caused by the fact that the uncertainty term enters linearly in the profit of the informed trader.

## 3.3 Uncertainty and Robustness in forward guidance tools

This section models the forward guidance as a signaling game between the central bank and market participants. Following the line of Cisternas (2018), we consider the monetary policy of the central bank a signal to the market. While Cisternas (2018) considers the problem of long-run commitment of central banks, we consider the forward guidance commitment as a short-run commitment instead. Our view is that forward guidance is a signal with dual effect on the economy. First, it encourages borrowing and expands the economy. Second, it also signals a weak macroeconomic perspective, thus discourage borrowing. The signal can be modeled in a signaling game between the central bank (the sender) and the market participants (the receiver). As a result, the signaling game in forward guidance, just like the signaling game in asset purchasing tools, makes central banking more efficient but less robust. A common problem with signaling games is the danger of getting a cheap talk. As pointed out by the original signaling game paper by Holmström (1999), if the sender's action (central bank) and the observation by the receiver are both costless, the signal becomes cheap-talk, meaning it could not affect the payoffs of the game. To overcome this issue, Cisternas (2018) assumes quite strongly that the monetary expansion process is not observed by the market. In his particular setting, the assumption might seem artificial. In our problem of forward guidance, however, this is not an issue: First, Central Banks have a preference to use interest rate based tools, so, the fact that the interest rate is constrained below implies that at least to the Central Bank, there is a cost of setting the nominal rate too low. In other words, there is a cost of using a non-standard monetary policy tool. Second, forward guidance is a statement about the stance of monetary policy in the future, and the Central Bank might have better information about the equilibrium of the economy than the market. If that is the case, even when the market observe the Central Bank actions, it can not fully infer the equilibrium.

### 3.3.1 The benchmark model without forward guidance

We first consider a benchmark case without forward guidance. In this case, the central bank chooses an interest rate rule $i_t$ to minimize the expected square error of inflation $\pi_t$ under economic shock $u_t$. We assume that $u_t \sim \mathcal{N}(0, \sigma_u^2)$ is observed by the central bank. The central bank solves the following optimization problem

$$\begin{aligned} \min_{i_t(\cdot), g_t(\cdot)} \quad & \mathbb{E}[\pi_t^2] \\ \text{subject to} \quad & \pi_t = B(i_t - r_t^n) + u_t \\ & i_t \geq \underline{i} \end{aligned} \tag{3.16}$$

where $r_t^n \in \mathcal{N}(0, \sigma_r^2)$ is also observed by the central bank. Here $\underline{i}$ is the lower bound on feasible interest rate. In practice, this lower bound is 0 as many central banks consider negative interest rates unfeasible. This is a common setting where central banking and the zero bound on interest rates has been studied.[8]

The goal of this section is to study how robustness affects the usefulness of forward guidance. In order to study such case, we need to simplify the model even further. In order to obtain a linear solution, we replace the constraint $i_t \geq \underline{i}$ with a multiplier penalty $\lambda(i_t - r_t^n)^2$ in the cost function. This assumption is important to be able to simplify the learning model under robustness. Now the central bank solves the

[8]See CITATIONS

optimization problem

$$\begin{aligned} \min_{i_t(\cdot)} \quad & \mathbb{E}[\pi_t^2 + \lambda(i_t - r_t^n)^2] \\ \text{subject to} \quad & \pi_t = B(i_t - r_t^n) + u_t \end{aligned} \tag{3.17}$$

The solution of the problem is a linear interest rule

$$i_t = au_t + r_t^n \tag{3.18}$$

where

$$a = -\frac{B}{B^2 + \lambda}.$$

. Furthermore, the expected optimal cost is given by

$$\frac{\lambda\sigma_u^2}{B^2 + \lambda}.$$

We denote (3.17) as the nominal model. As we did in Section 3.2, we study the case when there is a model mismatch, and derive the central bank robust decision.

Assume there is a model mismatch, given by $\delta \in [-\Delta, \Delta]$, on the real interest rate. Now the real interest rate follows a Gaussian distribution $r_t^n \sim \mathcal{N}(\delta, \sigma_r^2)$ instead. The central bank solves a minimax problem as follows to obtain its robust decision rule

$$\begin{aligned} \min_{i_t(\cdot)} \max_{\delta \in [-\Delta, \Delta]} \quad & \mathbb{E}[\pi_t^2 + \lambda(i_t - r_t^n)^2] \\ \text{subject to} \quad & \pi_t = B(i_t - r_t^n) + u_t \end{aligned} \tag{3.19}$$

In our baseline model, because the central bank can directly observe the realizations of the natural real interest rate $r_t^n$, the model mismatch does not change the optimization problem it faces. i.e., the optimization problem (3.19) is the same with (3.17). Later on, when we introduce the forward guidance tool in the next subsection, this is no longer true as a mismatched model will impede the formation of market expectations.

### 3.3.2 The model of forward guidance

We consider a forward guidance model with asymmetric information as follows:

- Suppose $u_t \sim \mathcal{N}(0, \sigma_u^2)$ (iid) and $r_t^n \sim \mathcal{N}(0, \sigma_r^2)$. The distribution is known to every agent in the game.
- At the beginning of the game, nature chooses the realization of $u_t$, $u_{t+1}$ and $r_t^n$.
- The central bank, upon observing the realization of $u_t$, $u_{t+1}$ and $r_t^n$, chooses the current period interest rate $i_t$ as well as the next period forward guidance $f_t$ to minimize the cost function.
- The market participants, upon observing $i_t$ and $f_t$, compute rational expectations $\mathbb{E}_t^m[\cdot]$. Here the expectation $\mathbb{E}_t^m[\cdot]$ is taken over the information set of the market (superscript $m$) at time $t$ (subscript $t$).
- The cost of central bank at time $t$ is given as $\mathbb{E}[\pi_t^2]$, with the following constraints

$$\begin{aligned} \pi_t &= \beta \mathbb{E}_t^m[\pi_{t+1}] + \kappa x_t + u_t \\ x_t &= -\frac{1}{\theta}(i_t - \mathbb{E}_t^m[\pi_{t+1}] - r_t^n) + \mathbb{E}_t^m[x_{t+1}] \\ i_t &\geq \underline{i} \end{aligned} \tag{3.20}$$

where $i_t$ is the nominal interest rate in the current period, $g_t$ is the forward guidance of the nominal interest rate for the period $t+1$, $\beta$ is the discount rate, $\pi_t$ is inflation rate, $x_t$ is output gap, $u_t$ is an inflation shock, $r_t^n$ is the exogenous natural interest rate, and $\underline{i}$ is the lower bound for nominal interest rate.

Because the inequality constraint on $i_t$ makes Bayesian inference impractical, in this section, we assume that a quadratic cost $(i_t - r_t^n)^2$ is used to prevent the central bank from setting the actual nominal interest rate too low in an expending monetary policy regime.

In our model the objective cost function of the central bank is $\mathbb{E}_t^c[\pi_t^2 + \lambda(i_t - r_t^n)^2]$. The first part is a simplified current period inflation target only cost function, and the second part represent the quadratic cost function to prevent the central bank from setting the actual nominal interest rate too high or low. The central bank's overall optimization problem is given by

$$\begin{aligned} \min_{i_t, f_t} \quad & \mathbb{E}_t^c[\pi_t^2 + \lambda(i_t - r_t^n)^2] \\ \pi_t &= \beta \mathbb{E}_t^m[\pi_{t+1}] + \kappa x_t + u_t \\ x_t &= -\frac{1}{\theta}(i_t - \mathbb{E}_t^m[\pi_{t+1}] - r_t^n) + \mathbb{E}_t^m[x_{t+1}] \end{aligned} \tag{3.21}$$

Because the objective function and constraints are both forward looking, all information until and include time $t-1$ are not relevant. At time $t$, central bank observes $u_t$, $u_{t+1}$ and $r_t$, hence $\mathbb{E}_t^c[\cdot] := \mathbb{E}[\cdot | u_t, u_{t+1}, r_t]$. Similarly, at time $t$, market participants observes $i_t$ and $f_t$, hence $\mathbb{E}_t^m[\cdot] := \mathbb{E}[\cdot | i_t, f_t]$.

### 3.3.3 The reduced form system equations

Consider the set of New Keynesian models as follows

$$\begin{bmatrix} x_t \\ \pi_t \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{\theta} \\ \kappa & \beta + \frac{\kappa}{\theta} \end{bmatrix} \mathbb{E}_t^m \begin{bmatrix} x_{t+1} \\ \pi_{t+1} \end{bmatrix} + \begin{bmatrix} -\frac{1}{\theta} \\ -\frac{\kappa}{\theta} \end{bmatrix} (i_t - r_t^n) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_t \tag{3.22}$$

Define $A_O = \begin{bmatrix} 1 & \frac{1}{\theta} \\ \kappa & \beta + \frac{\kappa}{\theta} \end{bmatrix}$ and $B_O = \begin{bmatrix} -\frac{1}{\theta} \\ -\frac{\kappa}{\theta} \end{bmatrix}$. Iterating the system equations, we obtain

$$\begin{aligned} \begin{bmatrix} x_t \\ \pi_t \end{bmatrix} =& B_O(i_t - r_t^n) + \sum_{\tau=1}^{\infty} A_O^\tau B_O \mathbb{E}_t^m (i_{t+\tau} - r_{t+\tau}^n) \\ &+ \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_t + \sum_{\tau=1}^{\infty} A_O^\tau \begin{bmatrix} 0 \\ 1 \end{bmatrix} \mathbb{E}_t^m[u_{t+\tau}] \end{aligned} \tag{3.23}$$

Here we need a technical assumption that as $\tau \to \infty$, both $\mathbb{E}_t^m[u_{t+\tau}]$ and $\mathbb{E}_t^m[i_{t+\tau} - r_{t+\tau}^n]$ go to zero faster than the largest eigenvalue of $A_S$. This is a transversality condition.

Given those assumptions, $\forall \tau \geq 2$, $\mathbb{E}_t^m[i_{t+\tau} - r_{t+\tau}^n] = 0$, and $\forall \tau \geq 2$, $\mathbb{E}_t^m[u_{t+\tau}] = 0$. Therefore, the system equations reduce to

$$\begin{bmatrix} x_t \\ \pi_t \end{bmatrix} = B_O(i_t - r_t^n) + A_O B_O \mathbb{E}_t^m(i_{t+1} - r_{t+1}^n) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_t + A_O \begin{bmatrix} 0 \\ 1 \end{bmatrix} \mathbb{E}_t^m[u_{t+1}] \tag{3.24}$$

Because the above equations represent a static system, and the output gap $x_t$ does not appear in the objective function of the central bank, we only need to concentrate on the inflation part. The equation regarding $\pi$ can be simplified (see Appendix 3.6.2), hence the optimization problem of the central bank is

$$\begin{aligned} &\min_{i_t(\cdot), g_t(\cdot)} \quad \mathbb{E}[\pi_t^2 + \lambda(i_t - r_t^n)^2] \\ &\text{subject to} \quad \pi_t = A_S B_S \mathbb{E}_t^m(i_{t+1} - r_{t+1}^n) + (A_S - 1)\mathbb{E}_t^m[u_{t+1}] \\ &+B_S(i_t - r_t^n) + u_t \end{aligned} \tag{3.25}$$

where $A_S = 1 + \beta + \kappa/\theta$ and $B_S = -\kappa/\theta$

### 3.3.4 Equilibrium

For simplicity, we consider only linear decision rules in the form of

$$\begin{aligned} i_t &= a_1 u_t + a_2 u_{t+1} + r_t^n \\ f_t &= b_1 u_t + b_2 u_{t+1} + \mathbb{E}_t[r_{t+1}^n] \end{aligned} \tag{3.26}$$

We propose a linear solution as above because the optimization objective function is quadratic and the signaling game involves Bayesian inference of Gaussian random variables.

**Proposition 1.** Consider a forward guidance game where i.i.d. shocks $u_t \sim \mathcal{N}(0, \sigma_u^2)$ and $r_t^n \sim \mathcal{N}(0, \sigma_r^2)$ hit a system at each time instance $t$. At time $t$, the central bank observes $u_t$, $u_{t+1}$ and $r_t$, and chooses $i_t$ and $f_t$ according to the linear rules (3.26) to solve optimization problem (3.25). On the other hand, market participants knows the distributions of $u_t$, $u_{t+1}$ and $r_t$, knows the linear rules (3.26), and observes realizations of $i_t$ and $f_t$. Market participants form expectations by calculating $\mathbb{E}_t^m[\cdot] := \mathbb{E}[\cdot|i_t, f_t]$.

Then, it exist at least one set of parameters $\{a_1, a_2, b_1, b_2\}$ such that market participants form conditional expectations and central bank minimizes loss.

Here are some insights about the solutions

- At the optimal solution, $a_2$=0. Because the role of $u_{t+1}$ is to reveal the private information of central bank to market participants, and to make the forward guidance credible. Therefore, it only needs to appear in the second term.

- if $(a_1, b_1, b_2)$ is a tuple of solutions, then $(a_1, \alpha b_1, \alpha b_2)$, $\forall \alpha \in \mathbb{R}$ is also a solution. Because it is the portion of information revealed that changes the cost. Any constant multiplied on it can be reversed by market participants.

*Proof.* Let

$$M = \begin{bmatrix} a_1 & 0 & \frac{\sigma_r}{\sigma_u} \\ 1 & b_2 & 0 \end{bmatrix}$$

It holds that

$$\begin{bmatrix} i_t \\ f_t \end{bmatrix} = \begin{bmatrix} a_1 & 0 & 1 \\ 1 & b_2 & 0 \end{bmatrix} \begin{bmatrix} u_t \\ u_{t+1} \\ r_t^n \end{bmatrix} = M \begin{bmatrix} u_t \\ u_{t+1} \\ \frac{\sigma_u}{\sigma_r} r_t^n \end{bmatrix} \tag{3.27}$$

In the above equation, we rescale $r_t^n$ by $\frac{\sigma_u}{\sigma_r}$ so that all three random variables in $\begin{bmatrix} u_t & u_{t+1} & \frac{\sigma_u}{\sigma_r} r_t^n \end{bmatrix}^\top$ have the same variance. Upon observing this period $i_t$ and $g_t$, the next period market expected nominal interest rate minus real natural interest rate is given by

$$
\begin{aligned}
\mathbb{E}_t^m[i_{t+1} - r_{t+1}^n] &= \begin{bmatrix} 0 & a_1 & 0 \end{bmatrix} M^\top (MM^\top)^{-1} \begin{bmatrix} i_t \\ f_t \end{bmatrix} \\
&= \begin{bmatrix} 0 & a_1 & 0 \end{bmatrix} M^\top (MM^\top)^{-1} M \begin{bmatrix} u_t \\ u_{t+1} \\ \frac{\sigma_u}{\sigma_r} r_t^n \end{bmatrix}
\end{aligned}
\tag{3.28}
$$

We plug in the result of the inference problem of market participants. Now the central bank solves the following optimization problem

$$
\begin{aligned}
\min_{a_1, b_2} \quad & \mathbb{E}[\pi_t^2 + \lambda (i_t - r_t^n)^2] \\
\text{subject to} \quad & \pi_t = \mathcal{E} + B(i_t - r_t^n) + u_t
\end{aligned}
\tag{3.29}
$$

where

$$
\mathcal{E} = \begin{bmatrix} 0 & ABa_1 + (A-1) & 0 \end{bmatrix} M^\top (MM^\top)^{-1} M \begin{bmatrix} u_t \\ u_{t+1} \\ \frac{\sigma_u}{\sigma_r} r_t^n \end{bmatrix}
$$

and

$$
i_t = a_1 u_t + r_t^n
$$

The central bank's optimization problem reduces to

$$
\begin{aligned}
\min_{a_1, b_2} \frac{1}{a_1^2 b_2^2 \sigma_u^2 + (b_2^2+1)\,\sigma_r^2} \Big[ & \sigma_r^2 \sigma_u^2 \Big( a_1^2 \left( b_2^2 \left( \left( A_S^2 + 1 \right) B_S^2 + \lambda \right) + 2 b_2 A_S B_S^2 \right. \\
& \left. + B_S^2 + \lambda \right) + 2 a_1 B_S \left( b_2^2 \left( A_S^2 - A_S + 1 \right) + b_2 \left( 2A_S - 1 \right) + 1 \right) \\
& + b_2^2 \left( A_S^2 - 2A_S + 2 \right) + 2 b_2 \left( A_S - 1 \right) + 1 \Big) + a_1^2 b_2^2 \sigma_u^4 \Big( a_1^2 \left( \left( A_S^2 + 1 \right) B_S^2 + \lambda \right) \\
& + 2 a_1 \left( A_S^2 - A_S + 1 \right) B_S + A_S^2 - 2A_S + 2 \Big) \Big]
\end{aligned}
\tag{3.30}
$$

which can be numerically solved. $\square$

In the following simulations, we use a default parameter setting of $\sigma_u = 1$, $\sigma_r$=1, $\lambda = 0.1$, $\kappa = 0.25$, $\theta = 0.5$ and $\beta = 0.8$.

First, we vary the value of $\lambda$.

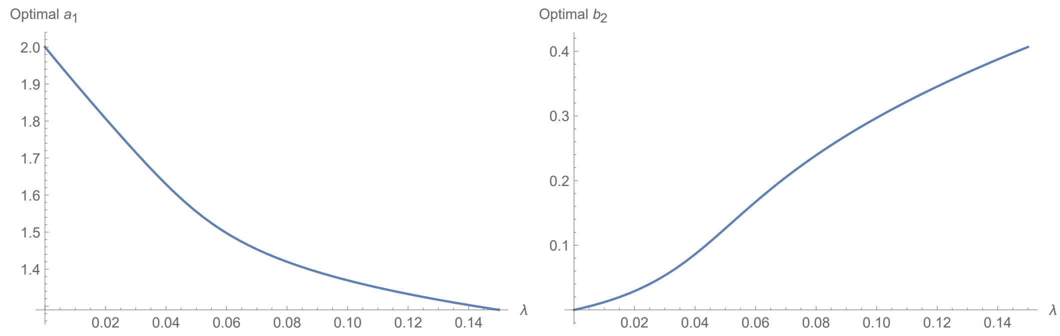



Figure 3-5: The optimal response of the central bank with different values of $\lambda$

As shown in Figure 3-5, as it becomes more and more costly to use interest rate policy (i.e., $\lambda$ increases), the optimal policy for the central bank is to use more forward guidance. Using forward guidance would achieve similar effects with setting the interest rate in terms of inflation targetting. However, those two tools come with different limitations: the interest rate cannot be set too far away from the neutral rate (controlled by $\lambda$ in our model), and the forward guidance only has limited power (see e.g., Del Negro et al. (2012)) because of the way the market expectation is formed. As we will show later, the expectation channel in the forward guidance becomes more problematic with the presence of model uncertainty.

Another factor that affects the optimal policy is the ratio between $\sigma_u$, the volatility of the shocks to the Phillips Curve, and $\sigma_r$, the volatility of the real interest rates. Note if we scale the volatility of the two types of shock together, the optimal policies $i_t$ and $f_t$ will not change. Define $\gamma$ as

$$\gamma = \log \frac{\sigma_r}{\sigma_u} \tag{3.31}$$

and vary the value of $\gamma$.

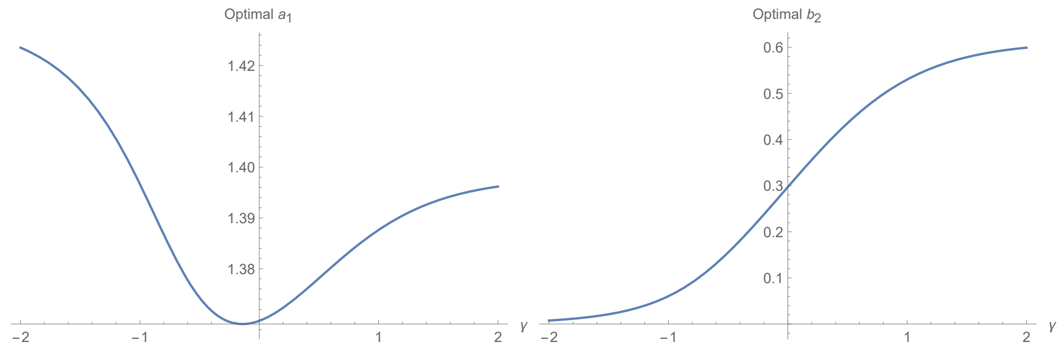



Figure 3-6: The optimal response of the central bank with different values of $\gamma$

The ratio $\gamma$ has two aspects in our model. First, it is the relative scale of shocks to the economy, and second, it is also the signal-to-noise ratio (SNR) in the forward guidance signaling game. As shown in Figure 3-6, when $\gamma$ is large, both tools are used with increasing intensity as $\gamma$ increases, reflecting the first perspective. When $\gamma$ is small, the forward guidance is more effective and the interest rate tool is used less as $\gamma$ increases, reflecting the second perspective.

### 3.3.5 Model uncertainty and Robust Monetary Policy

The New Keynesian model we use in this chapter is linearized around the natural real rate $r_t^n$. However, the natural real rate cannot be directly observed. Models are derived to estimate its values (e.g., the Laubach and Williams (2003) Model), hence model uncertainty on the natural real rate and its potential impact is essential to investigate.

Suppose that the central bank's model on natural interest rate $r_t^n$ is uncertain. More specifically, assume there is a model mismatch, given by $\delta \in [-\Delta, \Delta]$, on the real interest rate. Now the real interest rate follows a Gaussian distribution $r_t^n \sim \mathcal{N}(\delta, \sigma_r^2)$ instead. The central bank solves a minimax problem as follows to obtain its robust decision rule

$$\min_{i_t(\cdot),g_t(\cdot)} \max_{\delta\in[-\Delta,\Delta]} \mathbb{E}[\pi_t^2 + \lambda(i_t - r_t^n)^2]$$

$$\text{subject to} \quad \pi_t = A_S B_S \mathbb{E}_t^m(i_{t+1} - r_{t+1}^n) + (A_S - 1)\mathbb{E}_t^m[u_{t+1}] \tag{3.32}$$

$$+B_S(i_t - r_t^n) + u_t$$

Figure 3-7 shows the optimal cost of the central bank with different levels of model uncertainty under three circumstances: no forward guidance tool, using forward guidance tool with the nominal model, and using forward guidance with robust decision rules. It is evident that when model uncertainty is large, the robust optimal cost is close to the one of not using forward guidance tool.

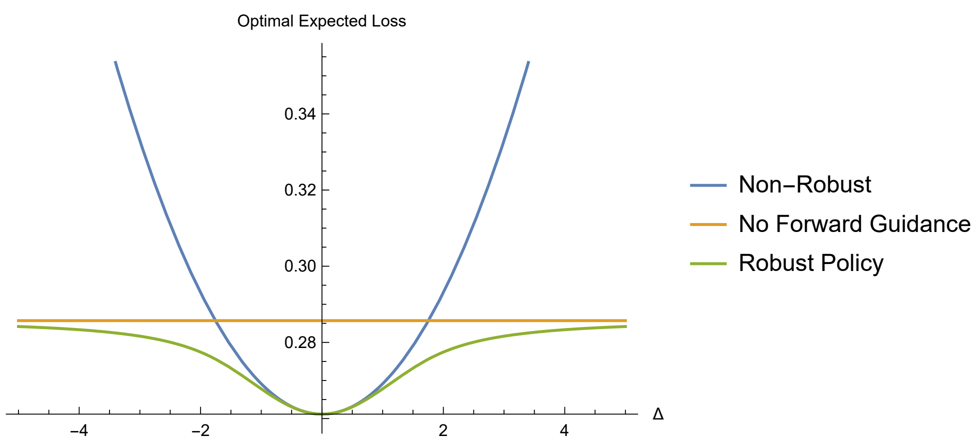



Figure 3-7: The optimal value of cost function when the model of $r_t$ is mismatched

### 3.3.6 Discussion about the Forward Guidance Model

In this section, we present a forward guidance model with model uncertainty. In conclusion, when model uncertainty is small, forward guidance, as an additional policy tool, always helps in stabilizing the economy. On the other hand, when model uncertainty is large, the economy is more stable without the forward guidance tool. This result reassembles many similar aspects with the asset purchasing model in the previous section.

However, this model also shows some unique features. First, the central bank faces a constraint that it cannot set the interest rate to negative as we mentioned in the introduction. This constraint means there is a cost for the central bank to set interest arbitrarily low. In fact, additional policy tools is only beneficial if different monetary policy tools has different costs to the central bank. Second, the central bank has better data in terms of inter-bank transactions than the market, therefore, has a better model of the economy than market participants. We assume that this is reflected in the natural interest rate of $r_t^n$. As a result, even though the market can observe precisely the central bank's action, i.e., the nominal fed fund interest rate $i_t$, it cannot infer the optimal $i_t - r_t^n$ directly, because it is the difference $i_t - r_t^n$, not $i_t$ that carries the information.

## 3.4 When is the bounded uncertainty important?

In our previous analysis of central banking models, we provided two examples where using fewer policy tools is better than using more policy tools in the presence of model uncertainty. However, that is not saying that the bounded uncertainty term always has that effect. In the two central banking models that we show, the bounded uncertainty term both go through an expectation channel. More specifically, in the asset pricing model, market makers take the expectation of asset value conditional on total market size, and the total market size includes a bounded uncertainty term. In the forward guidance model, market participants take the expectation of future inflation conditional on the forward guidance, and the forward guidance includes a bounded uncertainty term. In summary, bounded uncertainty in the model deserves scrutiny when it goes through an expectation channel.

Furthermore, the type of analysis in this chapter is most useful when a quantity of interest is best characterized by a falling in a range instead of following a distribution. For example, according to Humphrey—Hawkins Full Employment Act of the US, the Federal Reserve Board of Governors must submit a Monetary Policy Report to Congress twice a year outlining its monetary policy. Historical Humphrey—Hawkins testimonies and detailed Monetary Policy Report as early as July 1996 are available

| | Federal Reserve Governors and Reserve Bank Presidents | | Administration |
|---|---|---|---|
| | Range | Central Tendency | |
| **1996** | | | |
| *Percent change, fourth quarter to fourth quarter*[1] | | | |
| **Nominal GDP** | 4¾ to 5¾ | 5 to 5½ | 5.0 |
| **Real GDP** | 2½ to 3 | 2½ to 2¾ | 2.6 |
| **Consumer price index**[2] | 3 to 3¼ | 3 to 3¼ | 3.2 |
| *Average level in the fourth quarter, percent* | | | |
| **Civilian unemployment rate** | 5¼ to 5¾ | About 5½ | 5.6 |
| **1997** | | | |
| *Percent change, fourth quarter to fourth quarter*[1] | | | |
| **Nominal GDP** | 4 to 5½ | 4¼ to 5 | 5.1 |
| **Real GDP** | 1½ to 2½ | 1¾ to 2¼ | 2.3 |
| **Consumer price index**[2] | 2½ to 3¼ | 2¾ to 3 | 2.8 |
| *Average level in the fourth quarter, percent* | | | |
| **Civilian unemployment rate** | 5½ to 6 | 5½ to 5¾ | 5.7 |

Table 3.1: Economic Projections for 1996 and 1997, taken from the Monetary Policy Report of July 1996.

Percent

| Variable | Median | | | | | Central tendency | | | | | Range | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2020 | 2021 | 2022 | 2023 | Longer run | 2020 | 2021 | 2022 | 2023 | Longer run | 2020 | 2021 | 2022 | 2023 | Longer run |
| Change in real GDP.... | -2.4 | 4.2 | 3.2 | 2.4 | 1.8 | -2.5–-2.2 | 3.7–5.0 | 3.0–3.5 | 2.2–2.7 | 1.7–2.0 | -3.3–-1.0 | 0.5–5.5 | 2.5–4.0 | 2.0–3.5 | 1.6–2.2 |
| September projection | -3.7 | 4.0 | 3.0 | 2.5 | 1.9 | -4.0–-3.0 | 3.6–4.7 | 2.5–3.3 | 2.4–3.0 | 1.7–2.0 | -5.5–1.0 | 0.0–5.5 | 2.0–4.5 | 2.0–4.0 | 1.6–2.2 |
| Unemployment rate.... | 6.7 | 5.0 | 4.2 | 3.7 | 4.1 | 6.7–6.8 | 4.7–5.4 | 3.8–4.6 | 3.5–4.3 | 3.9–4.3 | 6.6–6.9 | 4.0–6.8 | 3.5–5.8 | 3.3–5.0 | 3.5–4.5 |
| September projection | 7.6 | 5.5 | 4.6 | 4.0 | 4.1 | 7.0–8.0 | 5.0–6.2 | 4.0–5.0 | 3.5–4.4 | 3.9–4.3 | 6.5–8.0 | 4.0–8.0 | 3.5–7.5 | 3.5–6.0 | 3.5–4.7 |
| PCE inflation .......... | 1.2 | 1.8 | 1.9 | 2.0 | 2.0 | 1.2 | 1.7–1.9 | 1.8–2.0 | 1.9–2.1 | 2.0 | 1.1–1.4 | 1.2–2.3 | 1.5–2.2 | 1.7–2.2 | 2.0 |
| September projection | 1.2 | 1.7 | 1.8 | 2.0 | 2.0 | 1.1–1.3 | 1.6–1.9 | 1.7–1.9 | 1.9–2.0 | 2.0 | 1.0–1.5 | 1.3–2.4 | 1.5–2.2 | 1.7–2.1 | 2.0 |
| Core PCE inflation[4].... | 1.4 | 1.8 | 1.9 | 2.0 | | 1.4 | 1.7–1.8 | 1.8–2.0 | 1.9–2.1 | | 1.3–1.5 | 1.5–2.3 | 1.6–2.2 | 1.7–2.2 | |
| September projection | 1.5 | 1.7 | 1.8 | 2.0 | | 1.3–1.5 | 1.6–1.8 | 1.7–1.9 | 1.9–2.0 | | 1.2–1.6 | 1.5–2.4 | 1.6–2.2 | 1.7–2.1 | |
| Memo: Projected appropriate policy path | | | | | | | | | | | | | | | |
| Federal funds rate ..... | 0.1 | 0.1 | 0.1 | 0.1 | 2.5 | 0.1 | 0.1 | 0.1 | 0.1–0.4 | 2.3–2.5 | 0.1 | 0.1 | 0.1–0.4 | 0.1–1.1 | 2.0–3.0 |
| September projection | 0.1 | 0.1 | 0.1 | 0.1 | 2.5 | 0.1 | 0.1 | 0.1 | 0.1–0.4 | 2.3–2.5 | 0.1 | 0.1 | 0.1–0.6 | 0.1–1.4 | 2.0–3.0 |

Table 3.2: Economic projections of Federal Reserve Board members and Federal Reserve Bank presidents, under their individual assumptions of projected appropriate monetary policy, December 2020, reported in the Monetary Policy Report of February 2021.

at Board of Governors of the Federal Reserve System (2021). In each Monetary Policy Reports, projections to key economic factors such as Change in real GDP, Unemployment rate, PCE inflation, and Core PCE inflation are given in ranges. See Table 3.1 for the reported ranges in the Monetary Policy Report of July 1996, and Table 3.2 for the reported ranges in the Monetary Policy Report of February 2021. Another example of such uncertainty is model uncertainty. Chen et al. (2019) points out that all economic model have a potential model mismatch, and discusses the impact of such mismatch and possible measures of such mismatch[9].

## 3.5 Conclusions

In response to the 2008 Global Financial Crisis, Central Banks worldwide embarked on an extensive easing of monetary policy. They greatly expanded their monetary policy tool arsenal due to the zero limit of interest rate. Furthermore, in response to the COVID 19 Global Pandemic, monetary authorities further expanded their monetary policy tools. All of these actions rely on the assumption that more policy tools are always better. This chapter seeks to examine whether this assumption holds true where this is bounded uncertainty as oppose to risk. Unlike risk, which has a distribution, bounded uncertainty we study in this chapter does not have a distribution, and a robust decision rule has to apply. We find that when bounded uncertainty is passed through a conditional expectation channel, sometimes it is optimal for a central bank to not using a policy tool. In this chapter, two monetary policy models for asset purchasing and forward guidance are examed in detail.

The first model, the asset purchasing model, is derived from the Kyle model, and bounded uncertainty is added to its expectation channel. The central bank and market participants react to the bounded uncertainty by applying robust decision rules. Under robust decision rules, sometimes it is optimal for the central bank to not using its policy tool. The second model, the forward guidance model, is derived

[9] For example, in the Section II of that paper, model mismatch is modeled as bounded uncertainty in the form of $b(t/n)$, such that $\sup_{u\in[0,1]}|b(u)|\leq 1$

from the linearized New Keynesian model, and bounded uncertainty is added to the natural real interest rate and passed through its expectation channel. In reality, just like our modeling, the natural real interest rate is generally estimated to fall within a range without a known distribution. Under robust decision rules, when the range of uncertainty grows, it is optimal for the central bank to use less and less the forward guidance tool.

Our results indicate that when bounded uncertainty interacts with conditional expectation, it is optimal for the central bank to use less to no additional tools. This result raises the potential problem of expanding monetary policy tools in an environment with uncertainty, and points out that robust and macro-prudential policymaking sometimes contradicts more policy tools.

Finally, our finding should also apply to other applications where bounded uncertainty and robust decision rule are concerned. An example of those includes optimal currency areas, where the optimal movement of capital and labor can fall in a range of continuum multi-equilibria without any distribution. In this application, uncertainty might imply that it is optimal to eliminate a currency. We seek to investigate the interplay of bounded uncertainty and expectation channel in those applications in the future.

## 3.6 Appendices

### 3.6.1 Proof of Lemma 3

#### IC condition for the government

We first check the IC condition for the government. We evaluate each term separately. The second term provides no incentives as

$$
\begin{aligned}
\mathbb{E}[(v-p)x|h;z] &= \mathbb{E}[(\frac{1}{2}(v-p_0) - \frac{1}{2}\frac{\sigma_v}{\sigma_{cb}}(u+z))\frac{\sigma_{cb}}{\sigma_v}(v-p_0)] \\
&= \frac{1}{2}\frac{\sigma_{cb}}{\sigma_v}\sigma_v^2
\end{aligned}
$$

For the first term

$$
\begin{aligned}
\mathbb{E}[(p-v)^2|h;z] &= \mathbb{E}[(p_0 + \frac{1}{2}\frac{\sigma_v}{\frac{\sigma_u\sigma_g}{\sqrt{\sigma_u^2+\sigma_g^2}}}(x+u+z) - v)^2|h] \\
&= \mathbb{E}[(p_0 + \frac{1}{2}\frac{\sigma_v}{\frac{\sigma_u\sigma_g}{\sqrt{\sigma_u^2+\sigma_g^2}}}(\frac{\frac{\sigma_u\sigma_g}{\sqrt{\sigma_u^2+\sigma_g^2}}}{\sigma_v}(v-p_0)+u+z) - v)^2|h] \\
&= \frac{1}{4}\mathbb{E}[(p_0-v)^2] + \frac{1}{4}\frac{\sigma_v^2}{\sigma_{cb}^2}\mathbb{E}[u^2|h] + \frac{1}{4}\frac{\sigma_v^2}{\sigma_{cb}^2}z^2 \\
&+ \frac{1}{2}\mathbb{E}[p_0-v]\frac{\sigma_v}{\sigma_{cb}}\mathbb{E}[u|h] + \frac{1}{2}\mathbb{E}[p_0-v]\frac{\sigma_v}{\sigma_{cb}}z + \frac{1}{2}\frac{\sigma_v^2}{\sigma_{cb}^2}z\mathbb{E}[u|h] \\
&= \frac{1}{4}\sigma_v^2 + \frac{1}{4}\frac{\sigma_v^2}{\sigma_{cb}^2}\mathbb{E}[u^2|h] + \frac{1}{4}\frac{\sigma_v^2}{\sigma_{cb}^2}z^2 + \frac{1}{2}\frac{\sigma_v^2}{\sigma_{cb}^2}z\mathbb{E}[u|h]
\end{aligned}
$$

Note that

$$
\mathbb{E}[u|h] = \frac{\sigma_{cb}^2}{\sigma_g^2}h
$$

The optimal strategy for the central bank satisfies the first order condition

$$
\frac{1}{2}\frac{\sigma_v^2}{\sigma_{cb}^2}z + \frac{1}{2}\frac{\sigma_v^2}{\sigma_{cb}^2}\frac{\sigma_{cb}^2}{\sigma_g^2}h = 0
$$

Hence the optimal strategy is given by

$$
z = -\frac{\sigma_{cb}^2}{\sigma_g^2}h
$$

**IC condition for the market maker**

Now we check the IC condition for market maker. Because the set of market makers are fully competitive, we have

$$
\begin{aligned}
p &= \mathbb{E}[v|x+u+z] \\
&= \mathbb{E}[\frac{\sigma_v}{\sigma_{cb}}x + p_0|x+u+z]
\end{aligned}
$$

Note that

$$u + z = u - \frac{\sigma_u^2}{\sigma_u^2 + \sigma_g^2} h = \frac{\sigma_g^2}{\sigma_u^2 + \sigma_g^2} u + \frac{\sigma_u^2}{\sigma_u^2 + \sigma_g^2} \epsilon \sim \mathcal{N}(0, \sigma_{cb}^2)$$

and

$$x = \frac{\frac{\sigma_u \sigma_g}{\sqrt{\sigma_u^2 + \sigma_g^2}}}{\sigma_v} (v - p_0) \sim \mathcal{N}(0, \sigma_{cb}^2)$$

Hence

$$\mathbb{E}[x | x + u + z] = \sigma_{cb}^2 / (\sigma_{cb}^2 + \sigma_{cb}^2)(x + u + z) = \frac{1}{2}(x + u + z)$$

and

$$p = \mathbb{E}[\frac{\sigma_v}{\sigma_{cb}} x + p_0 | x + u + z] = \frac{1}{2} \frac{\sigma_v}{\sigma_{cb}} (x + u + z) + p_0$$

**IC condition for the informed trader**

The informed trader maximize its total payoff

$$\max_x \mathbb{E}[x(v - p) | x]$$

The total payoff

$$\begin{aligned} \mathbb{E}[x(v - p) | x] &= \mathbb{E}[x(v - p_0 - \frac{1}{2} \frac{\sigma_v}{\sigma_{cb}} (x + u + z)) | x] \\ &= -\frac{1}{2} \frac{\sigma_v}{\sigma_{cb}} x^2 + x(v - \frac{1}{2} \frac{\sigma_v}{\sigma_{cb}} \mathbb{E}[u + z] - p_0) \end{aligned}$$

The first order condition is given by

$$-\frac{\sigma_v}{\sigma_{cb}} x + (v - \frac{1}{2} \frac{\sigma_v}{\sigma_{cb}} \mathbb{E}[u + z] - p_0) = 0$$

Hence the optimal strategy for the informed agent is given by

$$x = \frac{\sigma_{cb}}{\sigma_v} (v - p_0)$$

This ends the proof.

### 3.6.2 Simplifying Equation (3.24)

Inserting the definitions of $A_O$ and $B_O$ into equation (3.24) we obtain

$$\begin{bmatrix} x_t \\ \pi_t \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sigma} \\ -\frac{\kappa}{\sigma} \end{bmatrix} (i_t - r_t^n) + \begin{bmatrix} 1 & \frac{1}{\sigma} \\ \kappa & \beta + \frac{\kappa}{\sigma} \end{bmatrix} \begin{bmatrix} -\frac{1}{\sigma} \\ -\frac{\kappa}{\sigma} \end{bmatrix} \mathbb{E}_t^m (i_{t+1} - r_{t+1}^n) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_t + \begin{bmatrix} 1 & \frac{1}{\sigma} \\ \kappa & \beta + \frac{\kappa}{\sigma} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \mathbb{E}_t^m [u_{t+1}]$$

Evaluating the coefficients of all the terms in the equation regarding $\pi$, we obtain

$$\pi = -\frac{\kappa}{\sigma}(i_t - r_t^n) - \frac{\kappa}{\sigma}\left(1 + \beta + \frac{\kappa}{\sigma}\right) \mathbb{E}_t^m (i_{t+1} - r_{t+1}^n) + u_t + \left(\beta + \frac{\kappa}{\sigma}\right) \mathbb{E}_t^m [u_{t+1}]$$

To simplify the expression, we define $A_S = 1 + \beta + \kappa/\sigma$ and $B_S = -\kappa/\sigma$, and the equation regarding $\pi$ can be written as

$$\pi_t = A_S B_S \mathbb{E}_t^m (i_{t+1} - r_{t+1}^n) + (A_S - 1)\mathbb{E}_t^m [u_{t+1}] + B_S (i_t - r_t^n) + u_t$$

# Chapter 4

# From Just in Time, to Just in Case, to Just in *Worst*-Case

*This work was performed in collaboration with Daniel Rigobon and Roberto Rigobon.*

## 4.1 Introduction

"Just-in-Time" (JIT) manufacturing was introduced in Japan during the late 1940s and early 1950s by Toyota, with the purpose of reducing inventories, reducing setup times, and saving costs in other aspects of the supply chain. The reasons why it started in Japan are not completely clear. Some have argued that it was a combination of limited natural resources, a lack of physical space to hold the inventories, and financial constraints that the Japanese industry faced at the end of the war.[1] The cost reduction and efficiency gains of JIT became well known worldwide. Indeed, it became the standard of excellence in a short period of time and was adopted by many corporations. In fact, the globalization of the manufacturing of goods and services that started in the 1980s was, for the most part, inspired by JIT premises. Interestingly, even from the very beginning, Toyota suggested that the main risk of

[1] See the Toyota Production System, where one of the two pillars for production is just-in-time: a type of production where "only the necessary products, at the necessary time, in the necessary quantity are manufactured, and in addition, the stock on hand is held to a minimum." Also see Plenert (2007).

this strategy was its excessive reliance on suppliers — which could be less resilient and flexible than Toyota itself. Hence, a successful JIT implementation required a large emphasis on supplier development. Toyota argued that the JIT's biggest weakness was this vulnerability.

In response to those risks and seeking resilience, companies have explored other alternatives. These strategies, dubbed "Just-in-Case", usually recommend actions such as larger inventories, diversification of the production network, and harmonization of parts. In the end, the advice is simple: to develop flexibility and redundancy in the supply chain. Regardless of all these efforts, the world's supply chains proved to be unreliable during the Covid-19 pandemic. Either international trade frictions, quarantine restrictions, large shifts in demand (such as protective equipment), or even panic demand purchases of some products (such as toilet paper and disinfectant wipes in the US) highlighted the fragility of supply chains in the world. As a result, many countries experienced supply disruptions for various products during the pandemic. We believe this collapse is due to a design flaw: while the just-in-case approach might be appropriate for idiosyncratic shocks, it seems to have failed in the presence of an aggregate shock.

Many have argued that the solution is to have greater resilience and robustness.[2] However, what exactly does it mean to have a robust supply chain? How does it differ from assuming a more severe shock? In this chapter, we argue that a robust supply chain is one that addresses *uncertainty* instead of *risk*. The presence of uncertainty requires a decision-maker to solve a *minimax* optimization problem, in which they optimize the worst case of a set of outcomes – see for instance Gilboa and Schmeidler (1989) for an axiomatic treatment. In this sense, robustness means more than just assuming larger shocks; it means considering the worst possible outcome of a set of models; it means shifting to "Just-in-*Worst*-Case".

In this chapter, we present a very simple model. It introduces the idea that, in the pursuit of efficiency, a decentralized supply chain could become vulnerable to aggregate shocks. In other words, our setup reflects the well-known trade-off between

[2]See The FT Editorial Board (2020), Evans (2020) and Long (2020)

efficiency and robustness. In particular, a design with greater resilience to shocks will sacrifice efficiency during normal times. We are interested in understanding the consequences of natural disasters, environmental shocks, and pandemics on the global supply chain. These shocks tend to be substantial and infrequent, but also widespread — affecting many countries and regions at the same time. Just-in-Case's standard supply chain analysis studies firms facing large, frequent idiosyncratic shocks with known distribution. We study a different problem, one in which the shocks are infrequent, aggregate, and with unknown distribution. The simple model analyzes the survival of a multinational firm that purchases from small global suppliers. The suppliers decide where to locate, and locations are subject to aggregate shocks — which for simplicity are correlated with location.[3] We compare two types of global supply chain arrangements and two different types of shocks. First, the small suppliers individually decide their location, and the multinational purchases the surviving suppliers' products. Second, we study the case when the multinational can choose all its suppliers' locations, thereby internalizing the location decision (i.e., the suppliers are subsidiaries of the multinational firm). From the perspective of the shocks, we compare the situation of risk versus uncertainty through two settings of random aggregate shocks. In the first case, we assume the distribution is given, while in the second case, only the distribution's support is known.

Our setup replicates many well known results: just-in-time, a multinational internalizing an externality, etc. The purpose of reproducing those results is to compare them to the strategy implied by robust control. The robust strategy is in the spirit of a behavioral result known in psychology as *probability matching*. A rational individual facing a choice between two options should always choose the option with the higher payoff probability. For example, consider an individual with two choices: option A delivers one dollar 70 % of the time (zero otherwise), while option B pays one dollar only 30 % of the time (and zero otherwise). A rational agent should choose option A in all instances — whether they are playing once or many times. However,

[3] Our aggregate shock is one that affect all suppliers in the world, but not all of them negatively. In other words, some suppliers could be benefited from the shock while others could be hurt. It is a form of aggregate shock in the sense that when one location is affected the other location is not.

experimental research has shown that this is not the case. When participants are faced with a series of these choices, they often choose by replicating the probabilities. In other words, they would pick option A 70 % of the time and option B 30% of the time. Hence, the strategy is called probability matching. Several possible explanations of this phenomenon have been offered. First, one may describe the behavior as irrational. Another explanation appeals to bounded rationality. Intuitively, because it is costly for humans to process information, they will recur to heuristics when making decisions.[4] A heuristic allows for a fast resolution, and in fact replicates many of the (seemingly) irrational behaviors observed in sports and gambling. For example – the fallacy that an event must occur because it has not happened for a while: "a coin has flipped four heads in a row, it is very unlikely that another one will occur", or thinking that a batter will get a hit because they have failed to do so for a while. In this case the heuristic is to match the decision to the representativeness of choices. The last explanation for probability matching comes from evolutionary psychology.[5] Observation of foraging species documents that the allocation of individuals matches the distribution of resources, and that this behavior maximizes the survival probability of the species while also reducing the possibility that competing species invade the resource. [6] In this chapter, we present a model in which a fully rational agent who is ambiguity averse and internalizes the survival probability will replicate the probability matching behavior.

Our basic model has two locations: the Mountain and the Valley.[7] The two locations differ only in the probability that an aggregate shock hits. We will assume the shock is extreme, such that all suppliers in the affected location perish.[8] Without

[4]See Kahneman et al. (1982) and Kahneman and Tversky (1979)

[5]See Todd and Gigerenzer (2012)

[6]See Seth (2007) for simulations replicating probability matching — that in econology is known as the ideal free distribution.

[7]This example is inspired by Lo (2017) discussion on probability matching.

[8]In practice, aggregate shocks may not be so severe or permanent. In keeping with the theme of this chapter, we assume the worst possible outcome. See, for instance, Hallegatte (2015) and Tran et al. (2020) for discussion on the extent and duration of economic shocks due to natural disasters. For example, in November of 2020, it was still very difficult to purchase masks, hand sanitizer, and sanitary wipes in the US. This suggests that even in a developed nation, the supply disruptions can last for a long period of time.

loss of generality, we also assume that the shock is more likely to occur in the Mountain than in the Valley. Formally, if an aggregate shock occurs in our model, then the conditional probability of the shock affecting the Mountain (or the Valley) will be $\theta$ (or $1-\theta$), where $\theta > 1/2$.

Each supplier produces a single unit, or a "part". Both the cost of production and the sell price per part are held constant: in both locations, for all producers, and in any state of the world. In particular, the prices are independent of both the realization of the aggregate shock, and the number of surviving suppliers.[9] However, in canonical macroeconomic and international models, this would not be the case. The demand is usually chosen such that when quantities tend to zero, prices increase and can even tend to infinity. Models based on Cobb-Douglas or CES functions have this feature, and the pricing system reflects scarcity. Nonetheless, if firms are concerned about the consequences of increasing their prices after natural disasters, or there is a law that does not allow prices to increase after such an event, then the price required to achieve the efficient allocation may never be realized. If this were known ex-ante, then it would affect the willingness of suppliers to diversify into risky locations. Our model takes an extreme assumption — prices are fixed — to capture this feature of regulations and institutions. Moreover, below we include some anecdotal evidence of law enforcement and consumers' negative reactions to price gouging during the Covid-19 pandemic.

The number of suppliers grows and depends on the number of surviving suppliers. We treat each suppliers' part as a different intermediate good, and the multinational purchases as many parts as possible to sell the final product internationally. The product is more desirable the more parts it has — but it can be manufactured with a subset of the parts. This setting is equivalent to assume that the quality of the item

[9]There are a few motivations for this assumption, which are discussed more in detail by Section 4.2. First, a natural disaster or other disruptive event may lead to some degree of price stickiness. While fixed prices are an extreme of stickiness, partially-adjusting prices will still replicate the qualitative results of our model. Additionally, we can appeal to consumer anger in response to price gouging after natural disasters. For consumer anger see Rotemberg (2002, 2011)) and for price gouging laws Executive Order 13910 of March 23, 2020, *Preventing Hoarding of Health and Medical Resources To Respond to the Spread of Covid-19.* For examples of price gouging laws in the US, see King and Spalding (2020)

increases with the number of parts it has.[10]

The parameters in the model are such that each supplier's location decision has a dominant strategy, which is not collectively optimal when the probability of global survival is taken into account. This discrepancy comes from the inability of the pricing system to compensate firms properly for moving into the Mountain. In this setting, the multinational wants to ensure its suppliers survive. Putting this differently, the multinational cares about survival, while the suppliers do not. This difference implies that the multinational might be willing to set production facilities in the Mountain to insure itself against an aggregate shock in the Valley. This part of the model captures a simple externality and the need for a diversified supply chain, but not yet a robust one.

This is where the nature of the shock matters. When the shocks have known distributions — what is known as risk or the *nominal* model — the multinational will exhibit behavior that takes into account all the sources of risk. This setting implies a desire for diversification, and one of the implications, for example, is that the multinational's optimal allocation of firms to the Mountain depends on the number of suppliers that have survived. There is both a marginal benefit and cost of diversification, and in general, an internal solution is found (at least under our assumptions). The multinational's policy changes dramatically when the shock has bounded uncertainty — meaning that the distribution is unknown — in the *robust* model. In this setting, an ambiguity-averse multinational will perform a robust control optimization. As we will see, with sufficient uncertainty the optimal allocation of firms is independent of the number of surviving firms. The robust supply chain decision, therefore, looks very different from the nominal model.

This chapter includes several theoretical results worth highlighting. First, we compare the centralized and decentralized solutions to the model. We show a corner solution of the decentralized allocation (all firms locate themselves in the Valley) – exposing the multinational to an aggregate shock to the Valley. This result contrasts

[10]It is very common in supply chain management to assume that if one good or part is missing, the whole product can't be manufactured. By relaxing this assumption, we can eliminate the typical assumptions behind just-in-case theories and concentrate on the robustness aspect.

with the internal solution (a proportion larger than zero of firms in the Mountain) of the centralized allocation. Individual suppliers maximize efficiency (or productivity) while the multinational maximizes survival. This result is known and simple to understand — the multinational internalizes the survival externality.

Second, we study the implications of risk and uncertainty in the probability distribution of the aggregate shock. We depart from the assumption of the first part of our chapter: that the value of $\theta$, the conditional probability that the aggregate shock affects the Mountain, is known. We first study what occurs when $\theta$ is risky; for example we assume it is distributed between $[\bar{\theta} - \Delta, \bar{\theta} + \Delta]$ according to some a priori known distribution. This exercise represents a risky setup — the nominal model. We compare it to the uncertain setup where the multinational only knows that $\theta \in [\bar{\theta} - \Delta, \bar{\theta} + \Delta]$, but the distribution is unknown. This is the case of uncertainty, and the optimal control problem requires a robust approach. Being averse to ambiguity, the individual suppliers and the multinational maximize the expected profit assuming the worst-case value of $\theta$. For example, if $\bar{\theta} > 1/2$ but the support of $\theta$ contains $1/2$, then the optimal robust control will optimize as if $\theta = 1/2$. The diversification resulting from robust control replicates probability matching, and is very different from the diversification obtained in a traditional expected utility maximization problem with large variance or risk aversion.

In traditional supply chain literature of Just-in-Time and Just-in-Case, most of the analysis concentrates on idiosyncratic shocks of suppliers. As a result, building a supply chain with a precisely calculated amount of inventory can maximize efficiency and profit. Since the Covid-19 outbreak and the consequently supply chain disruptions, it is evident that aggregate shock on macroeconomic situation matters. In macroeconomics, the analysis of aggregate shock relies heavily on mathematical models, which are subject to modeling assumptions and model calibrations. Recently, some macroeconomic research starts to consider model uncertainty and robustness [11]. With model uncertainty, robust decision rules seek to maximize the payoff in the

[11] See (Hansen and Sargent, 2008) for robust control in general, and (Strzalecki, 2011) for the axiomatization of multiplier preferences, and many others.

worst-case scenario of a set of potential models.

This modeling choice is also supported by theoretical and empirical literature in ambiguity aversion. Ambiguity aversion refers to the behavioral phenomenon where people prefer variations with known distribution over unknown distribution. This phenomenon can be explained by people maximizing their utility for the worst-case scenario when the distribution is unknown, as in Gilboa and Schmeidler (1989). Since global supply chain disruptions have catastrophic impacts on social welfare, and the probability of such disruptions is not accurately known, uncertainty and robust decision rules are the proper tools for analysis and policy recommendations. In our case, the diversification implied by robust decision rules is very different from the diversification that could be obtained in a standard risk-averse model by either increasing the variance or the risk aversion to infinity.

Finally, we study what happens when prices and costs of production differ across locations. We compare three settings for the global supply chain: (i) the decentralized myopic setting that always chooses a corner solution, except when prices and costs are at the knife-edge when the value of both locations is the same; (ii) the probability matching heuristic where the allocation of firms coincides with the probabilities of survival in each location; (iii) and the optimal allocation by the multinational. We show that the decentralized solution can replicate the centralized solution when the value in the Mountain is equal to the value in the Valley. Our model justifies why governmental subsidies can help the decentralized economy achieve a robust solution.

For each simple model, we draw policy implications motivated by the recent experience with the pandemic. Summarized in the end, all of these policy implications have a simple message: robustness is under-supplied.

This chapter is organized as follows: Section 4.2 introduces some empirical evidence of supply shocks and price gouging restrictions during the Covid-19 pandemic. Section 4.3 presents our simple model of optimal control of a supply chain, with accompanying simulations. In Section 4.4, we introduce risk and uncertainty, and discuss the differences between the optimal solutions of these settings. Next, Section 4.5 studies the impact of different costs in the Mountain and the Valley, and finally Section 4.7 concludes with policy implications and future research.

### 4.1.1 Literature Review

**Probability Matching:** Our model provides a control problem rationalization of probability matching. When a game similar to ours is presented to individuals, experiments show that individuals tend to locate at the Valley roughly with probability $\theta$, and locate at the Mountain roughly with probability $1-\theta$. This experimental result is known as probability matching in the literature.[12] Many different theoretical approaches to behavior are developed to explain this phenomenon that humans prefer probability matching over rational expected utility maximization. Some early work suggests that it is a behavioral limitation due to bounded rationality, but more recent literature attributes that to learning strategies. Vulkan (2000) and Gaissmaier and Schooler (2008) argue that people would consistently try to learn patterns of the outcome series in a repeated game even when they are informed that the series is completely i.i.d. As a result, if the outcome series is truly i.i.d. as many of the earlier models assume, then probability matching seems irrational. On the other hand, if there is, in fact, a pattern in the series, probability matching claims a higher expected reward in the long run by gradually learning the patterns. Other literature suggests that probability matching is related to the growth pattern of a group of individuals. For example, Brennan and Lo (2011) concludes that if two choices result in similar growth rates, then deterministic decision rule prevails. On the other hand, if two choices result in drastically different growth rates, then probability matching gives an evolutionary advantage over the deterministic decision rules.

In our model, the intuition of the mechanism is simple: what is individually optimal is not collectively optimal. The reason is that the survival of firms in our model has a global externality from the inability of firms to coordinate — the increase in the number of firms that could produce — that given our set up the decentralized market does not take into account. When the multinational solves the model, how-

[12] See, for example, Fiorina (1971), Morse and Runquist (1960), Vulkan (2000), and Brennan and Lo (2011)

ever, it internalizes this effect, and it forces firms to do something that looks locally irrational. As said before, locating firms in the Mountain provides insurance to the multinational when an aggregate shock to the Valley takes place. That insurance is extremely valuable when uncertainty is present.

**Supply Chain:** The literature on the supply chain is extensive and impossible to summarize in a few paragraphs. There are, however, aspects that have been discussed in the literature that are relevant to the model we present here.

Aggregate shocks like Covid-19 has a significant implication of global supply chain risk management. Earlier empirical research describes this as supply chain flexibility, see Vickery et al. (1999). That paper defines flexibility as the ability to adapt to aggregate shocks. It shows through correlation analysis that supply chain flexibility is critical to the long-run survival of an organization. On the other hand, flexibility may affect the immediate competitiveness of the firm in the short run.

The vulnerability of the global supply chain to identical suppliers has already raised some concerns in the industry. Wagner and Bode (2006) studied questionnaires from company executives in Germany and concludes that a firms dependence on single type customers and suppliers is the largest contributor to a firms exposure to supply chain risk.

In recent years, the question of whether to integrate suppliers or not has been receiving more and more attention to supply management. The existing empirical literature has been studying this issue by looking at the elasticity of substitution of produces, see Antràs and Chor (2013) and Alfaro et al. (2019). When the demand for the final product is elastic, and inputs are not substitutable, firms choose not to integrate upstream suppliers. On the other hand, when the demand for the final product is inelastic, and inputs are substitutable, firms choose to integrate upstream suppliers. This finding shows that firms' supply chain decisions are optimal for the deterministic case, but not necessarily when an aggregate shock hits.

Apart from the works highlighted above, two groups of literature align with the spirit of this chapter.

First, there is ample literature discussing the organization of the supply chain. Following Antràs (2020), global supply chains can be viewed through different lenses. First, the value-added approach where firms allocate production internationally, and each stage of production contributes to the final product. In general, this literature concentrates on countries and industries as the unit of analysis. Second, the firm-level perspective — started by Melitz (2003) — offers an alternative to the aggregate view of the first approach. In this literature, the firms are the unit of analysis, and they are the ones that decide whether or not to participate in global supply chains. Both of these approaches assume there is no informational problem. This is relaxed by the relational view of supply chains. In this view, firms and suppliers face contracting problems — moral hazard or incomplete contracts — and therefore enter in relation to solve the informational problem. The main question it addresses is the organizational structure of the firm. The boundary of the firm in the global supply chain started with the seminal contribution of Antras (2003). The author discusses how incomplete contracts determine whether a firm should be integrated internationally versus enter into arms-length negotiations.[13] Finally, Yeaple (2003) studies the vertical and horizontal integration of multinationals.

Second, the literature on supply resilience highlights that resilience can be obtained by organizational robustness or organizational flexibility. See, for instance, Ambulkar et al. (2015), Töyli et al. (2013), Zhao and You (2019), Saenz et al. (2015), Durach and Machuca (2018), Helpman et al. (2004) and the references therein. Most of this literature, however, has two features. One is very related to our model — the literature advises that a robust supply chain can be achieved by working closely with the suppliers. In the spirit of our model, that is equivalent to when the multinational decides the global allocation problem. The second aspect is that most of these papers think about the robustness of a supply chain in response to shocks to the firms — i.e., the robustness to idiosyncratic shocks.

[13] See Antràs and De Gortari (2020), Antràs and Chor (2018) and Antras (2015).

## 4.2 Covid-19 and Supply Chains

### 4.2.1 Heterogeneous Supply Chain Disruptions

As has been argued by many, Covid 19 is a hybrid crisis. It has both supply and demand components, so understanding the magnitude and relevance of a single supply chain disruption is difficult. Furthermore, the demand/supply shock combinations are country-specific.

From the anecdotal point of view, many products suffered shortages during the Covid pandemic — hand sanitizer, toilet paper, meat products, beer, etc. The demand for these products, especially hand sanitizer, did not decline; therefore, it is clear that supply disruptions were present in many of them. The disruptions, interestingly, were not specific to China. In fact, in Figure 4-1 we present the value of trade merchandise for the world (top panel), the USA, and China (left, and right bottom panels). This data comes from the WTO.

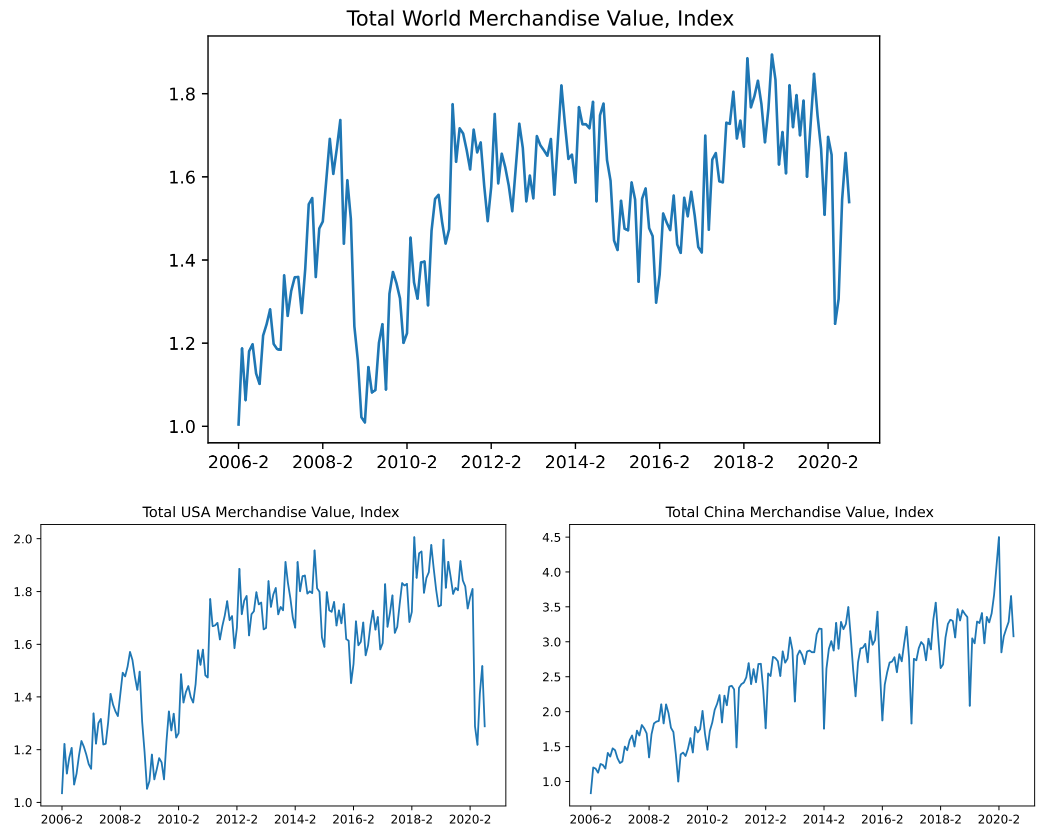



Figure 4-1: Merchandise Value for World, USA and China

For each region, the indices have been normalized to one in January of 2006. As can be seen, the 2008 financial crisis had a huge impact on all of them: world trade declined by almost a half (from 1.73 to 1.01), in the USA the decline was from 1.57 to 1.05, and in China, the value of trade dropped from 2.10 to 0.99.

On the other hand, the impact of Covid was heterogeneous. World trade experienced a small decline; comparing trade volumes in Q3 2020 to Q3 2019, the drop was only 6.2 percent. In the case of China, trade actually increased by 3.7 percent, while it declined by 23 percent for the USA.

In summary, the trade data shows the distinct effects of the pandemic on the supply chains of the US and China. It is exactly this heterogeneity that we seek to capture through our model's use of the Mountain and Valley.

#### 4.2.2 Prices and Supply Shocks

It is difficult to measure the impact of the supply disruption by only the value of trade, but a study of prices proves a much simpler exercise. Figures 4-2 and 4-3 present the inflation rate of different sectors and items. All the data comes from the Bureau of Labor Statistics (BLS), the CPI database. We selected all the seasonal adjusted monthly series aggregated at the US level. There are 317 categories; for illustrative purposes, the figures only present the most aggregate ones (about 97).

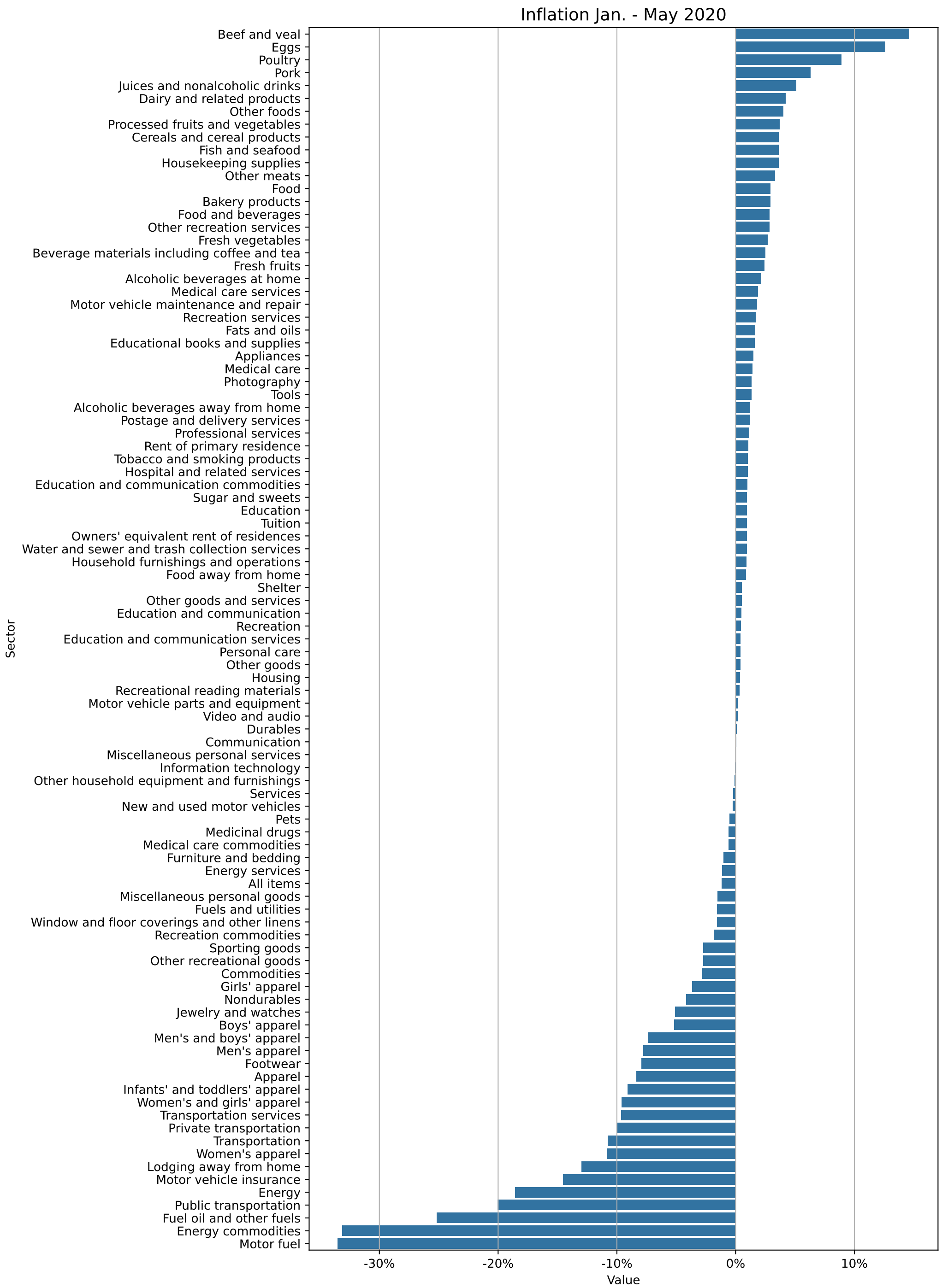



Figure 4-2: US Inflation by sector between January-May of 2020

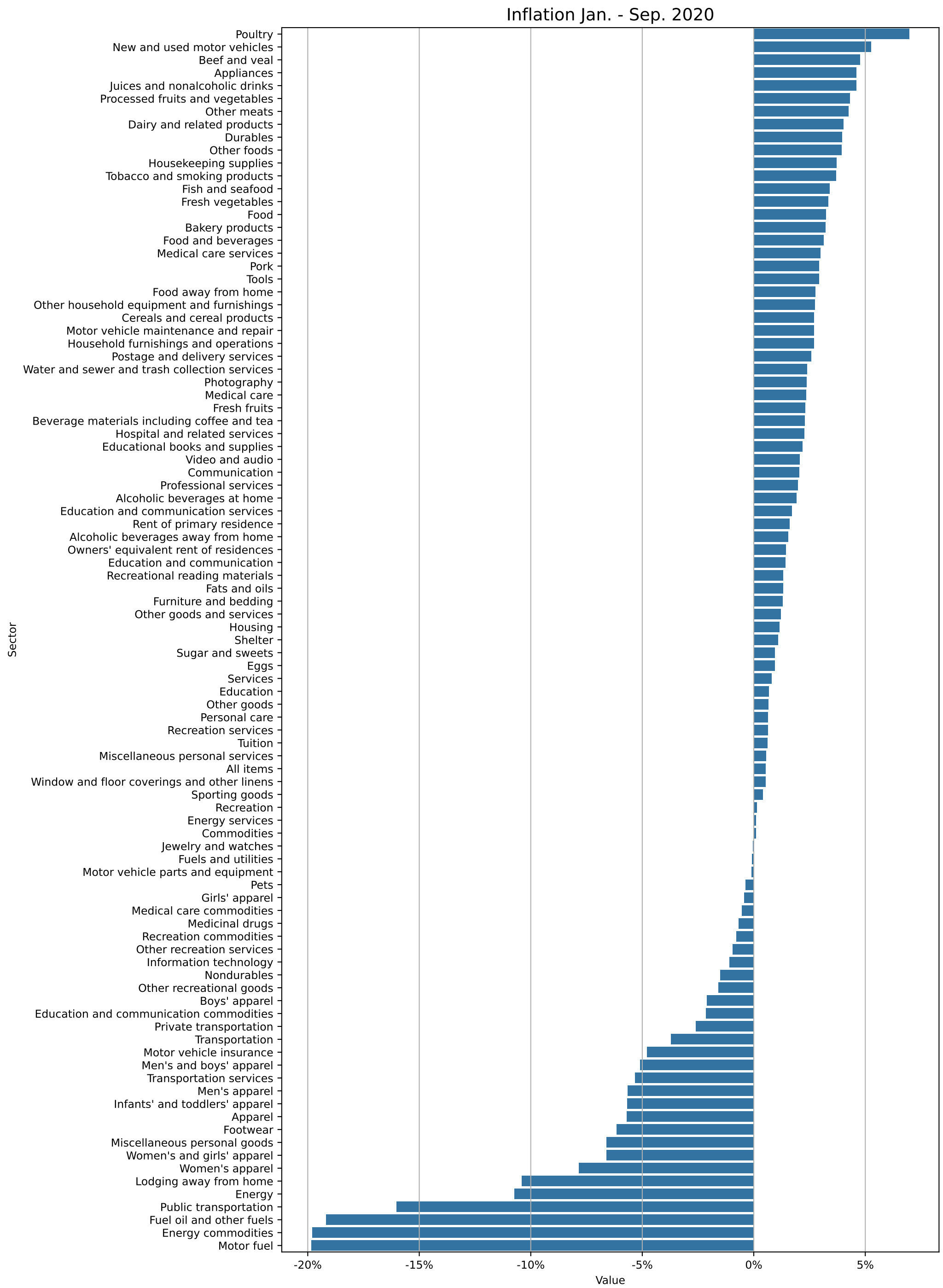



Figure 4-3: US Inflation by sector between January-September of 2020

The figures are sorted by the size of the price deflation or inflation. As can be seen, there is tremendous heterogeneity, although the median annualized inflation rate is north of 10 percent. Several aspects are worth highlighting.

First, if the demand shock of Covid were to dominate any supply disruption, then we should observe most sectors experiencing deflation. Empirically this was not the case – of the 313 BLS series for which inflation could be computed between January and May[14], only 119 experienced deflation. Consequently, more than 60 percent of the indices computed by the BLS experienced inflation at a time when the economy was undergoing strict lockdowns. We believe this inflation reflects the supply chain disruptions occurring during the pandemic.

Second, the sectors where prices declined are concentrated in Energy, Transportation Apparel, Tourism, and Jewelry.[15] The sectors that experienced price increases are related to Food, Household products, Beverages, and Medical Supplies.[16]. As we might have expected, more essential products experienced greater inflation during the pandemic.

Third, extending the analysis to September does not change the qualitative results — except that the supply shock dominates more than the demand shock — which should have been expected given that the US economy opened up. From January to September, only 99 of the 313 series ID's computed by the BLS experienced any form of deflation. Furthermore, the repressed sectors were still Energy, Transportation, Apparel, Tourism, and Recreation. Interestingly, in these sectors there is no report of supply disruptions – either in the US or globally. Deflation in these sectors is exclu-

[14]Several sectors or items have prices collected at different frequencies.

[15]In fact, the sectors with deflation higher than 2 percent are Motor fuel, Energy commodities, Fuel oil and other fuels, Public transportation, Energy, Motor vehicle insurance, Lodging away from home, Women's apparel, Transportation, Private transportation, Transportation services, Women's and girls' apparel, Infants' and toddlers' apparel, Apparel, Footwear, Men's apparel, Men's and boys' apparel, Boys' apparel, Jewelry and watches, Nondurables, Girls' apparel, Commodities, Other recreational goods, and Sporting goods.

[16]The list of the items which experienced more than 2 percent inflation, in order from the lowest inflation (2%) to the highest (15%), are: Medical care services, Alcoholic beverages at home, Fresh fruits, Beverage materials including coffee and tea, Fresh vegetables, Other recreation services, Food and beverages, Bakery products, Food, Other meats, Housekeeping supplies, Fish and seafood, Cereals and cereal products, Processed fruits and vegetables, Other foods, Dairy and related products, Juices and nonalcoholic drinks, Pork, Poultry, Eggs, and Beef and veal

sively driven by the demand shock. The sectors experiencing inflation in September and May are similar, and are concentrated in Food, Beverages, Household Products, and Medical Supplies. In all these sectors, we have reports of supply disruptions — at least anecdotally.

In summary, analysis of prices in the US by sector makes clear that supply chains failed to deliver basic products. Items such as food, beverages, household products, and medical supplies experienced large inflation. In fact, in these types of essential products many countries established much-needed anti-price gouging laws during the pandemic. Therefore, the inflation that we observe is not as high as the one that would have existed without the restrictions. This is the topic of the next subsection.

### 4.2.3 Price Gouging

One of the important assumptions of our model is that prices do not adjust fully to the aggregate shocks. In general, prices may deviate from equilibrium due to stickiness, as has been studied in the relevant macro literature. See, for instance, Caballero and Engel (2006) and references therein for theoretical results, and Anderson et al. (2015) for some empirical evidence. In the case of large aggregate shocks, we also motivate the fixed-price assumption by the existence of price gouging laws in many countries during the pandemic — either because of fairness considerations or consumer anger. In this subsection, we discuss the origin and relevant evidence of this latter motivation.

When facing large aggregate shocks, the price of essential goods cannot float freely as assumed by the classic general equilibrium model. Take Covid-19 as an example; during the pandemic the demand for personal protective equipment, foods, and other essential supplies rose dramatically, which raised the concern of price gouging with both regulators and the general public.

On the side of regulators, Executive Order 13910 of March 23, 2020, *"Preventing Hoarding of Health and Medical Resources To Respond to the Spread of Covid-19"* was issued by the US to deal with the threat of price gouging. Individual States of the US, evidenced in the laws of King and Spalding (2020), were very active in controlling companies pricing for products related to the pandemic.


Figure 4-4: Law enforcements' efforts to control price gouging.

As shown in Figure 4-4, many State Attorney General Offices created procedures to deal with price gouging complaints. The purpose of these laws is quite "benevolent"; regulators seek to prevent hoarding and ensure that the prices of essential goods do not increase beyond that which is considered "reasonable or fair". However, in practice, it is not possible to distinguish which part of the price increase is reasonable (e.g., a price increase which generates profits to compensate for the cost of diversification in normal times) and which part is not reasonable (e.g., price increase due to hoarding).

During Covid-19, the companies that had previously diversified their production, and were hence able to keep producing essential products during the pandemic, were not rewarded with higher profit. They were unable to set prices freely, and were instead penalized by the increasing litigation risk of anti price gouging enforcement – failing to benefit from prices adjusting to a new equilibrium. As a result, diversifying and preparing for aggregate shocks may not a financially optimal decision for companies in classic market equilibrium models. Theoretically, the problem described above is an inefficient allocation due to market incompleteness. Our model, with individual producers and the fixed-price assumption, reflects this issue.

The general public also paid great attention to price gouging. For example, a

Google Images search on October 5th, 2020 for "price gouging" presents the following results.

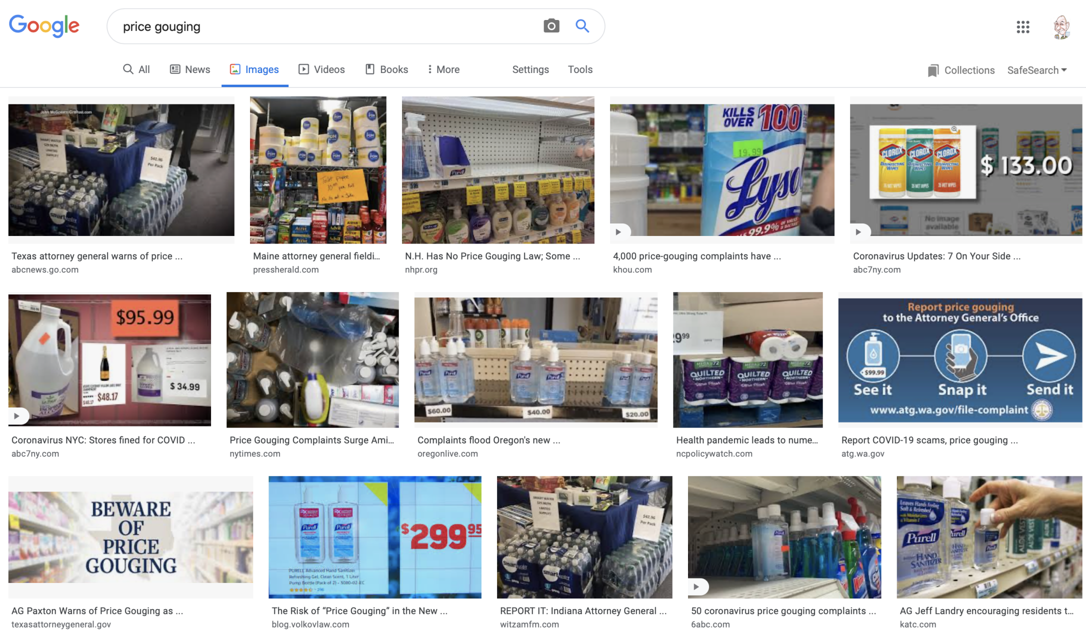



Figure 4-5: Google Search results for Price Gouging.

Some of the images are shown in detail in figure 4-6: consumers compare the price of isopropyl alcohol with a a bottle of champagne, a convenience store clarifies the price of toilet paper is not a joke, and purchase of water is restricted.

When the consumers and law enforcement are so concerned with price increases after a natural disaster, is reasonable to expect that firms are unable to fully benefit from the scarcity, and may even prefer not to supply than to face the public relations nightmare that would require justifying their selling price.

The examples we show here are only in the US, but European countries experienced a similar search for firms violating price gouging laws. See Cary et al. (2020) and UK Competition Authority (2020) for a discussion of the recent law enforcement efforts regarding complains of price abuses in many developed nations.

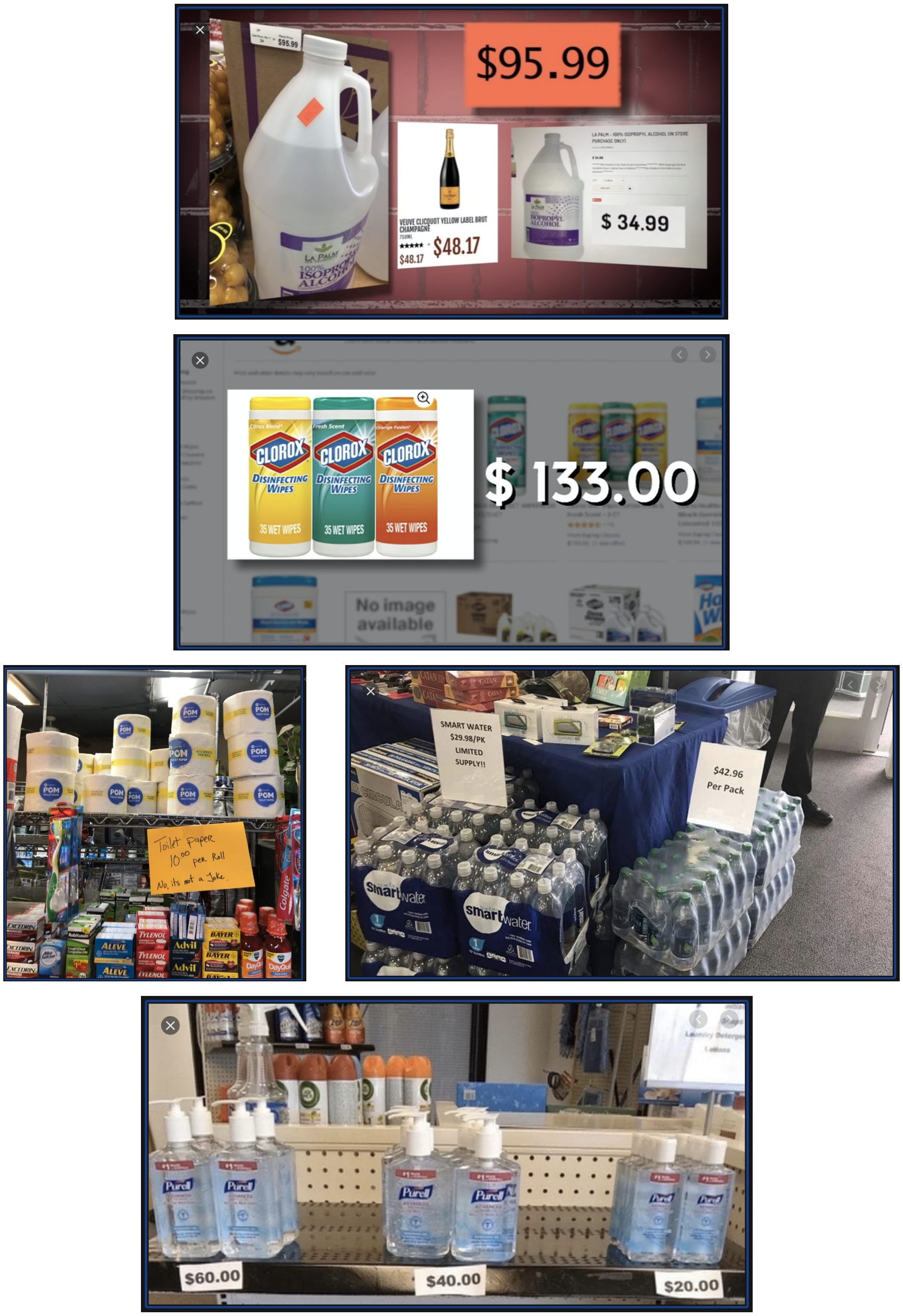



Figure 4-6: Price gouging examples in stores and online.

On the side of the general public, price gouging received immediate attention at the start of the pandemic. Even though the images constitute anecdotal evidence, we can provide evidence on the intensity with which people searched for “price gouging”.

Figure 4-7 shows the Google search frequencies of topics "Coronavirus" and "Price Gouging" in different regions (US, UK, and worldwide) as well as different languages (English and Spanish). It is evident that the public's awareness of "price gouging" rises almost simultaneously with awareness of "Coronavirus" itself.

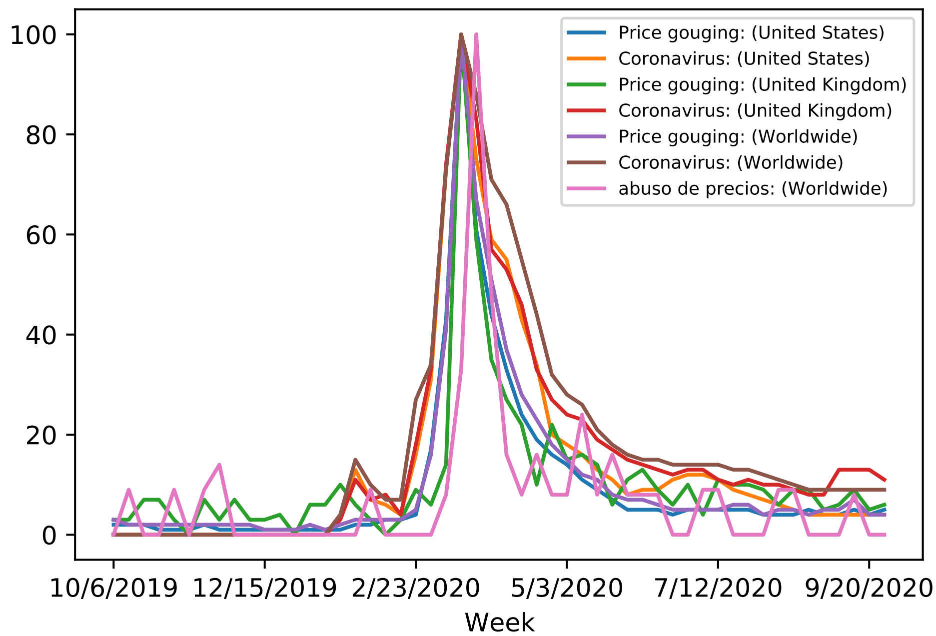



Figure 4-7: Google search frequencies of topics "Coronavirus" and "Price Gouging". Numbers are normalized by 100 at maximum values. *Data Source: https://trends.google.com/trends/*

Supply chain allocation is a durable decision. Nevertheless, we identify price inflexibility as a source of fragile supply chain allocation for the following reasons: First, supply chain allocation is durable and unlikely to adapt during disasters on a large scale, which means the supply disruption is generally as long as the disaster itself. Second, when an aggregate shock hits, the anti-price gouging enforcement actions are as long as the supply-side shock. For example, Anderson and Apfel (2020) summarizes many anti-price gouging enforcement actions worldwide from February to July 2020. Third, even if the period of price distortion during a disaster is relatively short compared to normal times, it eliminates companies' essential motivation to diversify production to less competitive locations and prepare for those shocks. In summary, we believe price inflexibility is an important externality when studying aggregate shocks like the global pandemic and many natural disaster events.

## 4.3 Model

In this section, we present a firm-location-problem model that highlights the vulnerability of the global supply chain to aggregate shocks. It is a simple survival model in which individual firms fail to take into consideration the impact they have on the aggregate — a standard externality argument — and whose decisions change quite substantially once uncertainty is taken into account.

We assume two different forms of organizing the world supply chain. In the first, a multinational asks already established firms (factories) to independently offer the parts required to produce a final product. In this case, the factories decide where to locate themselves. We identify this structure with a global supply chain of *Independent Suppliers* or the *Decentralized* economy. The second organization is one in which the multinational allocates its production facilities — which are the subsidiaries of the multinational. We identify this structure as *Multinational Subsidiaries* or as the *Centralized* economy.

As said before, a second ingredient in our model is the difference between risk and uncertainty. Optimization under risk produces a policy function that is very different from that derived under uncertainty. Our model is a single firm, partial equilibrium model, which concentrates on the existence and response of the supply chain to aggregate shocks. The Covid-19 pandemic was an obvious aggregate shock. However, natural and environmental disasters becoming more prevalent implies that we need a different approach to the understanding of resilience and robustness of the supply chain. A distinct feature of these shocks is their aggregate nature, but also how uncertain they are; we might know that sea level will be rising, but the extent of the damage has tremendous uncertainty, and the distribution itself is likely to be unknown.

We use both ingredients — the internalization of survival (the externality) and the robust approach (uncertainty) — to rationalize supply chains whose level of diversification are an order of magnitude larger than what we observe in practice.

### 4.3.1 Baseline Model

A product sold by a multinational is comprised of many "parts", each produced by a factory, and each factory can be located in two different regions. Time occurs in discrete steps, and the discount rate is $\beta$.

Assume there are $N_t$ firms at the start of period $t$. Each factory/supplier has a location decision: for simplicity, we will identify the locations as the Mountain and the Valley. Factories choose one of these two locations at time $t$ where they set up production. Each factory only produces one unit of the part, which has a constant cost $c$. Suppliers sell the part to the multinational, who produces the final good. The cost is paid before production takes place.

Production is uncertain. In each period, one of two locations might suffer an aggregate shock with arrival probability $\gamma$. Conditional on such a shock, and before production occurs, all firms in the Mountain or Valley perish with probability $\theta$ or $1-\theta$, respectively. We assume that $\theta >> 1/2$. In other words, the Mountain is significantly riskier than the Valley. With probability $1-\gamma$ there is no aggregate shock. Production takes place only by the surviving firms, and the multinational produces the final product with the parts it has access to[17].

At the end of each period, the number of subsidiaries can grow. The growth rate is given by

$$N_{t+1} = A \cdot (N_t^s)^{1-\mu} \tag{4.1}$$

where $N_t^s$ denotes the number of firms that have survived the aggregate shock. Notice that this growth process has a fixed point at $N^* = A^{1/\mu}$. In our model, it makes sense to have a decreasing return to scale due to limited resources. This is a distinction

[17]The reason behind this assumption is that in the canonical model of complementary inputs (e.g., Kremer (1993)) an idiosyncratic shock has macroeconomic consequences. In our model, we want idiosyncratic shocks to be "harmless" and concentrate on the role of aggregate shocks.

with the origional model in Lo (2017). The timing is denoted in Figure 4-8.

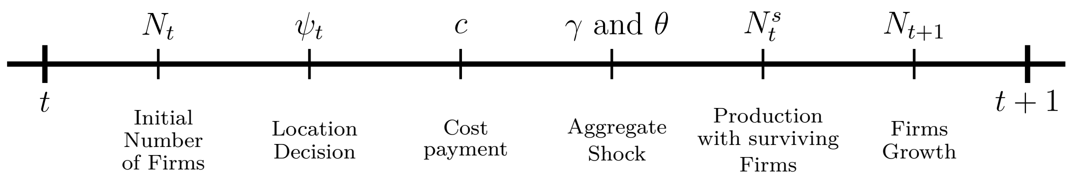



Figure 4-8: Model timing.

The multinational aggregates parts from all the suppliers and produces a final good. The complexity of the final good depends on the number of parts included. This model has an extremely simple demand side; we assume that the final product's revenue when sold is linear in the number of parts it includes. Furthermore, we add a *Survival Constraint* to this model by assuming that the firm needs at least 1 part to be able to produce the final product. In other words, $N_t^s \geq 1$ for the firm to be able to continue operating.

$$\Pi_t = \begin{cases} pN_t^s & \text{if } N_t^s \geq 1 \\ 0 & \text{o.w.} \end{cases}$$

The price per part, $p$, is constant and independent of production and the state of the world. This is equivalent to assuming that all firms are price taker, but it also is capturing the fact that prices rarely move freely after natural disasters — the anti-price gouging laws. This is obviously an extreme assumption, but one that simplifies the exposition. In many countries, there are price gouging regulations that limit the extent to which the pricing system helps ameliorate the supply chain problem. Therefore, the pricing system cannot finance the supply chain reforms required to reestablish production, and instead other actions (such as time or government subsidies) are required to recover the supply chain. During the Covid-19 pandemic, it has been clear that fairness arguments have dominated the discussion. For example, see Executive Order 13910 of March 23, 2020, *Preventing Hoarding of Health and Medical Resources To Respond to the Spread of Covid-19*, or King and Spalding (2020) for a list of price gouging laws in the US. A more detailed discussion of price gouging during Covid-19 was given earlier in Subsection 4.2.3.

Two important aspects of the price assumption are worth highlighting. First, we are assuming a very extreme form of anti-price gouging — prices are completely fixed. In our model, the anti-price gouging behavior is quite important for our results. If prices cannot adjust after a natural disaster, there is no way of compensating the suppliers that locate themselves in the mountain. On the other hand, if firms face a standard demand satisfying the inada conditions, then all suppliers can be compensated when locating themselves in the mountain. In fact, it is optimal to make the expected value of investing in the mountain equal to investing in the valley. The assumption of anti-price gouging laws, however, is not unreasonable. First, they are observed in practice. Second, as argued in Dworczak et al. (2020), anti-price gouging laws can be socially optimal in the presence of income inequality and other inefficiencies.

The second aspect worth highlighting is that the price gouging is as long as the investment horizon. This is also an extreme assumption that allows us to characterize the solution. In our case, the investment horizon is one period (after which the suppliers can relocate without cost). Future research should include the possibility that there is stickiness in the location decision to study the implications of price freezes shorter than the investment horizon.

The fact that inflation is found in some of the sectors could be misconstrued as a rejection of the price gauging assumption. That is not necessarily the case. If prices are not allowed to increase to the market clearing price there is scarcity in the market and inflation at the same time. For most of the products highlighted here, inflation, and rationing and excess demand indeed existed.

Denote $\psi_t$ the proportion of firms that are located in the Valley. The evolution of firms is given by

$$N_{t+1} = \begin{cases} A \cdot (N_t)^{1-\mu} & \text{w/p} \quad (1-\gamma) \\ A \cdot (\psi_t N_t)^{1-\mu} & \text{w/p} \quad \gamma\theta \\ A \cdot ((1-\psi_t) N_t)^{1-\mu} & \text{w/p} \quad \gamma(1-\theta) \end{cases}$$

where the top realization occurs when there is no aggregate shock, and the second (third, respectively) one is when the aggregate shock hits the Mountain (Valley, respectively). The growth of the firms has two components: multiplicative and exponential. As can be seen, the growth of suppliers depends on the total number of surviving suppliers in the world. We assume that the new suppliers are distributed according to the existing number of surviving firms in each location, but that the growth rate depends on the total number of existing firms. This assumption in the basic model is innocuous, but it is essential if the model is extended to introduce adjustment costs — or switching costs. We leave this interesting application for future research.

### Independent Producers

In the independent producers setting, the suppliers decide their location individually, and then the multinational contracts with the firms. We assume that all the revenue from the multinational is transferred to the suppliers — i.e. the multinational has zero profits. The total revenue is equally shared among the surviving suppliers.

Suppliers are small and they do not take into account the impact their decision has on the decision of the location of others ($\psi_t$). As we mentioned before, there is no cost of switching between locations. Therefore the suppliers are solving a static problem — the continuation value is exactly the same for all firms. Firms are maximizing the expected value of Mountain versus Valley and given our assumptions Valley dominates for all firms. Then, the value at time $t$ of locating in the Valley or the Mountain is given by

$$V_t^v = ((1-\gamma)+\gamma\theta)p \qquad - c + \frac{1}{1+\beta}((1-\gamma)+\gamma\theta)V_{t+1} \tag{4.2}$$

$$V_t^m = ((1-\gamma)+\gamma(1-\theta))p - c + \frac{1}{1+\beta}((1-\gamma)+\gamma(1-\theta))V_{t+1}. \tag{4.3}$$

The continuation value for each supplier, conditional on having survived the aggregate shock, is independent of the location. This is a feature of the zero cost of relocation. Therefore, the value of locating in the Valley is always larger than the value of locating in the Mountain. Formally,

$$V_t^v - V_t^m = \gamma(2\theta - 1)\left(p + \frac{1}{1+\beta}V_{t+1}\right) > 0 \tag{4.4}$$

for $\theta > 1/2$, and hence $\psi_t = 1$.

### Multinational Subsidiaries

Assume now that the multinational has all the decision power and it allocates the production units. Two aspects now matter for the multinational firm that were not relevant for the independent suppliers: the multinational takes into account the distribution of firms, and it takes into account the expected value of continuation in all states of the world.

The problem of the multinational firm can be written as follows

$$V(N_t) = \max_{\psi_t}\left\{\left[\begin{array}{ccll} (1-\gamma) & \cdot & (pN_t & +\frac{1}{1+\beta}V\left(A\cdot(N_t)^{1-\mu}\right) \quad ) \\ +\gamma\theta & \cdot & (p\psi_t N_t & +\frac{1}{1+\beta}V\left(A\cdot(\psi_t N_t)^{1-\mu}\right) \quad ) \\ +\gamma(1-\theta) & \cdot & (p(1-\psi_t)N_t & +\frac{1}{1+\beta}V\left(A\cdot((1-\psi_t)N_t)^{1-\mu}\right) \quad ) \end{array}\right] - cN_t\right\} \tag{4.5}$$

where

$$\lim_{N\to 1^-} V(N) = 0 \tag{4.6}$$

is the value matching constraint.

As before, the top line represents the value when the aggregate shock does not occur, and the second (third, respectively) when the aggregate shock hits the Mountain (Valley, respectively). Although the cost is the same $cN_t$, the revenue depends on the number of surviving firms. Recall that the cost of production is paid irrespectively of the aggregate shock.

The first-order condition (FOC) with respect to $\psi_t$, after simplifying, is

$$(2\theta-1)p = A(1-\mu)\frac{N_t^{-\mu}}{1+\beta}\left\{(1-\theta)(1-\psi_t)^{-\mu}V'\left(A((1-\psi_t)N_t)^{1-\mu}\right) - \theta\psi_t^{-\mu}V'\left(A(\psi_t N_t)^{1-\mu}\right)\right\}. \tag{4.7}$$

We simulate the discrete time version of the model to characterize the solution. The parameters used in the simulation are: $\beta = 0.02$, $\gamma = 0.2$, $\theta = 0.6$, $\mu = 0.5$, $A = 5$, $p_m = p_v = 1$ and $c_m = c_v = 0.5$.

In terms of the number of suppliers in equilibrium, the choice of $\mu = 0.5$ implies a fixed point of $N_t = 25$, in the absence of aggregate shocks. We initialize all simulations in this fixed point.

In Figure 4-9, we present the proportion of firms in the Valley as a function of the total number of suppliers (horizontal axis). The orange line represents the decentralized allocation — the individual rationality solution. The blue line indicates the probability matching solution for survival. Finally, The green line indicates the optimal solution of the multinational.

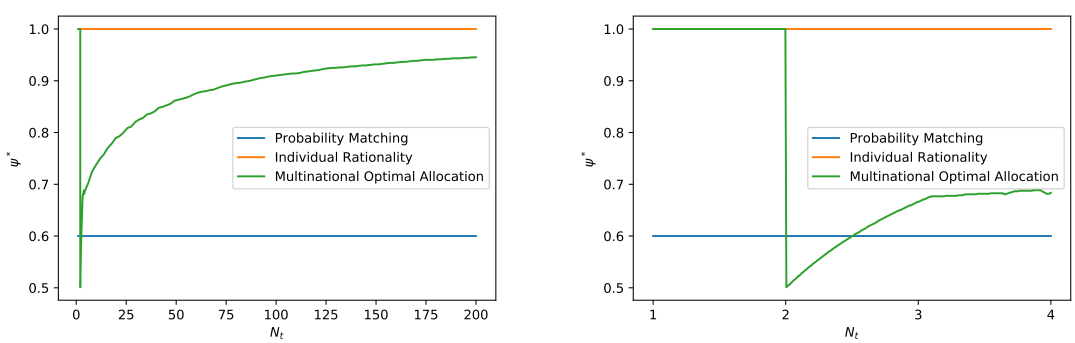



Figure 4-9: Optimal $\psi^*$ as a function of $N_t$.

The multinational has a tradeoff between instantaneous profits (what the individual suppliers maximize) and the probability of survival. The right panel of Figure 4-9 is a closer view of the left panel, but concentrating on relatively small $N$.

As shown in Figure 4-9, the multinational's optimal $\psi$ is a function of the number of production units $N_t$, and has three distinct phases. In the first phase, when $N_t \in [1, 2)$, the multinational's optimal choice is a corner solution, which coincides with individuals' optimum. This occurs because when $N_t < 2$, losing one unit will discontinue the multinational's operation. Hence, there is no way to ensure survival and to realize the benefit of continuation value. In the second phase, when $N_t \in [2, 3)$, the multinational will allocate exactly one production unit to the Mountain to take advantage of the continuation value. As a result, the optimal allocation is given by $\psi^* = 1 - 1/N_t$. In the third phase, when $N_t \geq 3$, the optimal $\psi$ is a concave increasing function of $N_t$. It is increasing because, with guaranteed survival, it is optimal to allocate a greater percentage of production units to the Valley to maximize profit. It is concave because the function $\psi^*(N_t)$ asymptotically approaches a constant less than 1.

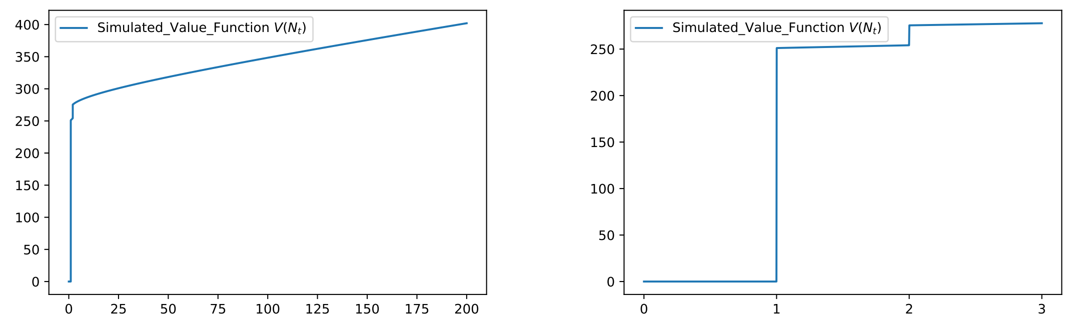



Figure 4-10: Value function $V(N_t)$ with optimal $\psi$.

The value function of the multinational, with the optimal $\psi$, is presented in Figure 4-10. The concavity of the value function comes from the concavity of the growth function of the suppliers, and also from the value matching constraint. If the suppliers grow at a constant rate, the value function would be linear with respect to the number of suppliers — and therefore, even the solution of the multinational would be at a corner.

The value function has two discontinuities points at $N_t = 1$ and $N_t = 2$. The discontinuity at $N_t = 1$ is trivial due to the constraint that $V(N_t) = 0, \forall N_t < 1$. On the other hand, the discontinuity at $N_t = 2$ has an important implication about continuation value. When the number of production units is less than 2, the multinational firm cannot ensure survival. When the number of production units is greater than or equal to 2, the multinational can allocate one production unit on the Mountain to ensure survival. This ensured survival creates a continuation value for the multinational, which is responsible for the jump of value function at $N_t = 2$.

### 4.3.2 Probability Matching

As was shown in Figure 4-9, the optimal proportion of firms in the Valley varies with the number of suppliers. It is interesting to compare the effectiveness of the optimal firm allocation with respect to a naive strategy — a constant *psi* independent of the number of firms. For instance, assuming the probability matching strategy is adopted, the value function of the multinational with constant $\psi = 0.6$ is presented in Figure 4-11.

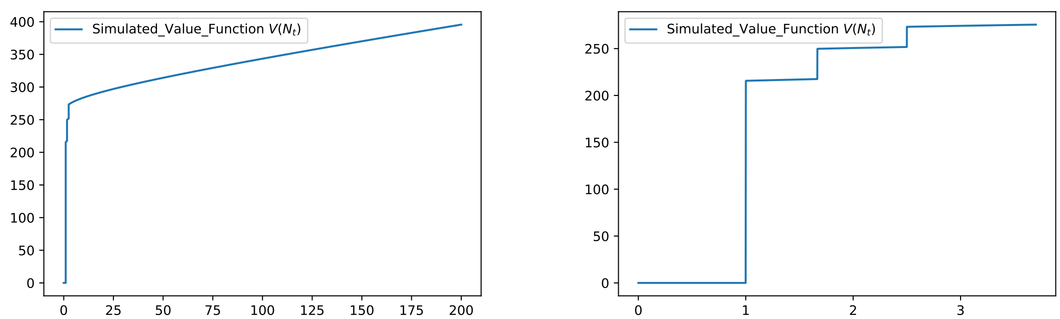



Figure 4-11: Value function $V(N_t)$ with constant $\psi = 0.6$.

Apart from the trivial discontinuity at $N_t = 1$, the value function has two other discontinuous points at $N_t = 5/3$ and $N_t = 2.5$. Below 5/3, the constant $\psi = 0.6$ does not place a single production unit in either location, and the multinational fails when any aggregate shock occurs. As a result, the discontinuity at 5/3 represents the continuation value of firms in the Valley. On the other hand, the discontinuity at $N_t = 2.5$ represents the continuation value of firms in the Mountain. When the number of available production units is less than 2.5, the multinational places fewer than one unit in the Mountain, and will not survive an aggregate shock to the Valley. However, when the number of production units is greater than or equal to 2.5, the multinational can allocate at least one production unit on the Mountain, thereby ensuring survival. The jump of the value function at $N_t = 2.5$ reflects this guaranteed survival.

The difference between the value function using the optimal strategy, and the value function following the probability matching strategy is small. We compare the optimal-strategy value function with the probability-matching value function (when $\psi$ is constant and equal to 0.60). Figure 4-12 shows the percentage increase in the value function when the firm switches between probability matching to optimal.

The x-axis is the number of firms on a logarithmic scale, and for comparison purposes, we concentrated on $N_t > 3$. The discrete jumps in the value function for smaller $N_t$ swamp any possible comparison outside that region. On the y-axis is the percentage difference between the two value functions.

The relationship, as expected, is increasing. The reason is that the optimal $\psi$ increases with the number of surviving firms; therefore, the loss incurred by fixing it at 0.60 is also increasing. Having said this, notice that the magnitudes are small: between one and two percent.

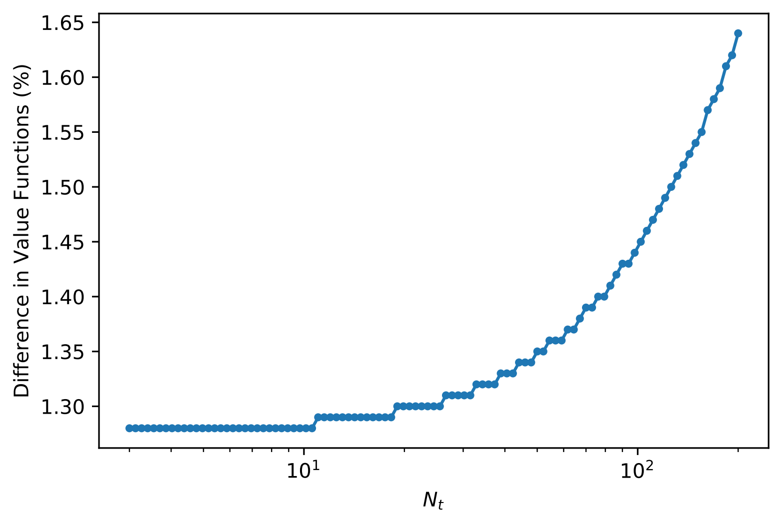



Figure 4-12: Difference between Value function $V(N_t)$ with optimal $\psi$ and Value function with constant $\psi = 0.6$.

One interesting question to ask is how the probability of survival is affected by different possible allocation strategies by the multinational. Recall that Figure 4-9 indicates that the optimal proportion of firms in the Valley is a function of the total number of suppliers that exist. However, we here study a simple, naive allocation strategy in the spirit of probability matching models. For instance, assume the multinational chooses a fixed proportion of suppliers in the Valley regardless of the total number of suppliers that exist.

In Figure 4-13, we present the probability of survival for various fixed values of $\psi$ over different horizons. We define the probability of survival as one minus the probability that the number of suppliers is smaller than 1 for any time step within the horizon, and initialize simulations with the same parameters as above.

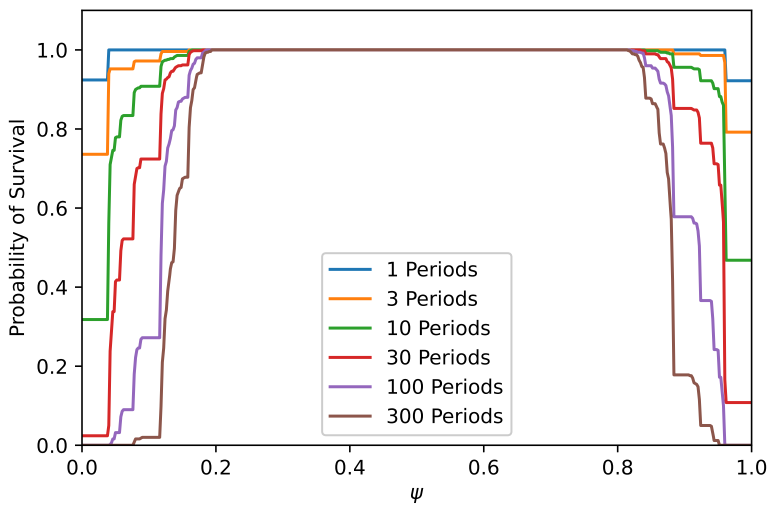



Figure 4-13: Probability of survival

For large time horizons, the probability is either one or zero; interestingly the breakpoints include the probability matching proportion ($\psi = 0.6$).

### 4.3.3 Implications

This model presents a simple contrast between three possible strategies: the decentralized allocation in which firms do not take into account the survival probability of the multinational, the centralized allocation in which the multinational internalizes the decision, and the behavioral response that would use simple probability matching heuristics concentrating on the maximization of the probability of survival.

The behavioral finance literature points out many cases in which individuals will tend to chose the third strategy. In our model, indeed, such a strategy will guarantee the survival of the multinational. However, it is inefficient. A dynamic allocation increases profits, for example, and also guarantees the survival of the firm in equilibrium.

Many questions arise from this framework that we explore further in this chapter and some that are left for future research.

First, how can the probability matching behavior be rationalized in this setting? As shown in figure 4-9 the optimal allocation in the Valley is an increasing function of the number of surviving firms - after $N_t > 2$. So, the optimal solution is an internal solution, and it is dependent on the number of firms. As we discussed before, some jumps happen at small numbers, which are the result of the constraint at which the multinational shuts down. Probability matching implies a constant proportion of firms regardless of the number of surviving firms, which contradicts this feature of the multinational's optimal allocation. In Section 4.4 we will introduce uncertainty and show how the robust optimal response is indeed very close in its spirit to the probability matching.

Second, given the externality, is there something that a government could do such that the decentralized economy could reproduce the centralized allocation? The answer to this question is usually yes, and it either requires taxes or subsidies. We study this point in Section 4.5. We show that in the context of differences in prices, there exists a governmental policy in which the decentralized economy achieves the centralized outcome — or at least — the firms are indifferent between locating in the Mountain versus the Valley, so the centralized allocation is available. The price or cost differences can be interpreted as an *ex-ante* tax or subsidy to the allocation of a firm in a particular destination.

Although trivial, it is worth highlighting that if prices are allowed to adjust, the decentralized allocation will replicate the centralized one. In our model, prices are not allowed to change, and therefore it is impossible to compensate the firms in the Mountain when a shock to the Valley has taken place. If the prices were to adjust, then once an aggregate shock takes place, the revenues of the surviving suppliers would increase. Moreover, because there are fewer firms in the Mountain, the price increase when a shock to the Valley takes place would increase the price of parts more than when the shock occurred in the Mountain. Because survival is very important, any usual demand function — CES for example — implies that the expected value of firms in the Mountain and the Valley are equalized. Under those circumstances, the decentralized economy reproduces the centralized one. The assumption of prices NOT adjusting is crucial. We do believe it is a reasonable assumption when aggregate shocks occur, justified through the prevention of price gouging.

Third, a simplifying assumption in our model is that the growth of firms is related to the total number of surviving firms regardless of where the firms were located. Also, we are assuming that there is no cost of reallocation. These are simplifying, but unreasonable assumptions. Further research should look into the implications when the growth of firms is specific to the location, and there are adjustment costs. We leave this extension to future research.

Finally, our supply chain structure is extremely simple. In reality supply chains look like complex networks.[18] Future research should look at the implications of robustness in a more complex structure.

## 4.4 The *Nominal* and the *Robust* Models

The model in the previous section only deals with risk. In this section, we explore the implication of adding uncertainty into the model. In particular, we will assume the probability of the aggregate shock in the Valley ($\theta$) is uncertain. As has been said before, this section shows that the optimal robust strategy is exactly in the spirit of probability matching: a constant proportion regardless of the number of firms that exist.

We study two cases: in the first, we consider a risky $\theta$. Here we assume that $\theta$ is uniformly distributed in $[\bar{\theta} - \Delta, \bar{\theta} + \Delta]$. Due to the linearity of the value function with respect to $\theta$, we show that the optimal choices for both the individual producers

[18]See Yeaple (2003)

and the multinationals are the same as the baseline case. Though the result is trivial in this case, it allows us to set up the comparison with our second case, where we assume model uncertainty: $\theta$ can be any value between $[\bar{\theta} - \Delta, \bar{\theta} + \Delta]$ and agents deploy a robust decision rule by solving a minimax problem.

### 4.4.1 Risk: the *Nominal* Model

Let us assume that $\theta \in [\bar{\theta} - \Delta, \bar{\theta} + \Delta]$ where $\Delta$ is small enough to guarantee that the support of $\theta$ is contained in $[0, 1]$. We assume that the distribution is uniform and known by all agents. This setting is identified in our discussion as the *nominal* model. We will continue to assume that the Mountain is riskier, therefore, $\bar{\theta} > 1/2$. Because individual suppliers are risk-neutral and the aggregate shock enters linearly, the decentralized equilibrium is identical: all the suppliers choose to locate in the Valley.

Similarly for the multinational, because $\theta$ enters linearly to the value function, the profit maximization problem and its solution is unchanged. More specifically

$$
\begin{aligned}
V(N_t) &= \max_{\psi_t(N_t)} \int_{\bar{\theta}-\Delta}^{\bar{\theta}+\Delta} \left\{ \begin{array}{llll} (1-\gamma) & \cdot & (pN_t & + \frac{1}{1+\beta} V\left(A \cdot (N_t)^{1-\mu}\right) \quad ) \\ +\gamma(\theta) & \cdot & (p\psi_t N_t & + \frac{1}{1+\beta} V\left(A \cdot (\psi_t N_t)^{1-\mu}\right) \quad ) \; - cN_t \\ +\gamma(1-\theta) & \cdot & (p(1-\psi_t)N_t & + \frac{1}{1+\beta} V\left(A \cdot ((1-\psi_t) N_t)^{1-\mu}\right) \quad ) \end{array} \right\} d\theta \\
&= \max_{\psi_t(N_t)} \left\{ \begin{array}{llll} (1-\gamma) & \cdot & (pN_t & + \frac{1}{1+\beta} V\left(A \cdot (N_t)^{1-\mu}\right) \quad ) \\ +\gamma(\bar{\theta}) & \cdot & (p\psi_t N_t & + \frac{1}{1+\beta} V\left(A \cdot (\psi_t N_t)^{1-\mu}\right) \quad ) \; - cN_t \\ +\gamma(1-\bar{\theta}) & \cdot & (p(1-\psi_t)N_t & + \frac{1}{1+\beta} V\left(A \cdot ((1-\psi_t) N_t)^{1-\mu}\right) \quad ) \end{array} \right\}
\end{aligned}
$$

Therefore, under the assumption of risk the solutions of the decentralized and decentralized economy are identical to the baseline model. Of course this is a feature of the assumptions we have chosen to make and where the parameter risk was introduced. We have made these choices for simplicity.

### 4.4.2 Uncertainty: The *Robust* Model

The second case we study is the case of uncertainty in the sense of robust control. Assume that all agents know that $\theta \in [\bar{\theta} - \Delta, \bar{\theta} + \Delta]$, but they do not know the distribution.

Optimal control implies that the optimization maximizes the worst possible case. For individual producers, the values in the Valley and the Mountain are

$$V_t^v = \min_{\delta \in [-\Delta, \Delta]} ((1-\gamma) + \gamma(\bar{\theta} + \delta))p \qquad - c + \frac{1}{1+\beta}((1-\gamma) + \gamma(\bar{\theta} + \delta))V_{t+1} \quad (4.8)$$

$$V_t^m = \min_{\delta \in [-\Delta, \Delta]} ((1-\gamma) + \gamma(1 - \bar{\theta} - \delta))p - c + \frac{1}{1+\beta}((1-\gamma) + \gamma(1 - \bar{\theta} - \delta))V_{t+1} \quad (4.9)$$

Because nature will choose a value of $\Delta$ that minimize producers' value, the worst case scenario for the Valley is when $\delta = -\Delta$, and the worst case scenario for the Mountain is when $\delta = \Delta$. The difference in value between the Valley and Mountain is given by

$$V_t^v - V_t^m = \gamma(2\bar{\theta} - 1)\left(p + \frac{1}{1+\beta}V_{t+1}\right) > 0 \qquad (4.10)$$

for $\bar{\theta} > 1/2$. Note this is identical to the baseline case (Equation 4.4 in Section 4.3. The worst case for each individual firm still implies that the worst case in the Valley is better than the worst case in the Mountain.

The problem of the multinational firm can be written as follows

$$V(N_t) = \max_{\psi_t(N_t)} \min_{\delta \in [-\Delta, \Delta]} \left\{ \begin{bmatrix} (1-\gamma) & \cdot & (pN_t & + \frac{1}{1+\beta} V\left(A \cdot (N_t)^{1-\mu}\right) & ) \\ +\gamma(\bar{\theta} + \delta) & \cdot & (p\psi_t N_t & + \frac{1}{1+\beta} V\left(A \cdot (\psi_t N_t)^{1-\mu}\right) & ) \\ +\gamma(1 - \bar{\theta} - \delta) & \cdot & (p(1-\psi_t)N_t & + \frac{1}{1+\beta} V\left(A \cdot ((1-\psi_t)N_t)^{1-\mu}\right) & ) \end{bmatrix} - cN_t \right\}$$

subject to

$$\lim_{N \to 1^-} V(N) = 0.$$

The above value function has the following characteristics: First, given any fixed

$\delta$, the value function is upper semi-continuous and concave in $\psi$, for all $N_t \geq 2$. For $N_t < 2$ there is a unique corner solution. Second, given any fixed $\psi$, the value function is linear (therefore continuous and convex) in $\delta$. Finally, $\psi \in [0, 1]$ and $\delta \in [-\Delta, \Delta]$ are chosen from compact sets.

As a result, according to Sion's Minimax Theorem Sion et al. (1958), the maximization and minimization are interchangable, and the minimax problem has at least one solution.

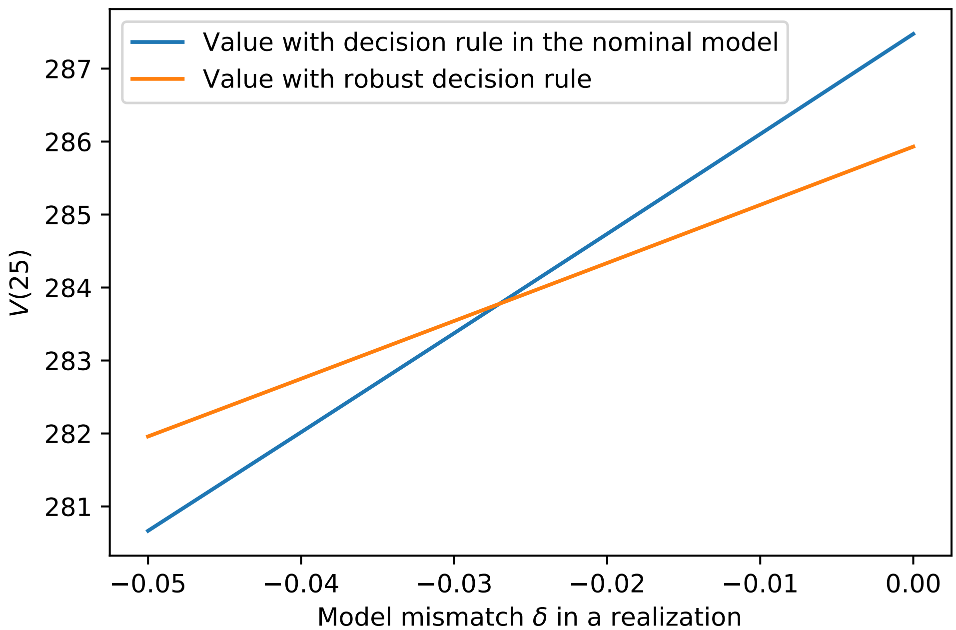
The optimal $\delta$ (meaning the choice that produces the worst possible case for the multinational) is given by

$$\delta^* = \begin{cases} -\Delta & \text{if } \bar{\theta} - \Delta > 1/2 \\ -\bar{\theta} + 1/2 & \text{if } \bar{\theta} - \Delta <= 1/2 \end{cases} \tag{4.11}$$

This then implies that the multinational's optimal response is the $\psi$ from Figure 4-9, but where the shock probability is given by $\theta = \bar{\theta} + \delta^*$.

### 4.4.3 Efficiency vs. Robustness

It may seem too conservative to always considering the worst-case scenario, especially for multinationals trying to maximize profit. Even though the worst-case scenario is known to be $\delta = -\Delta$ for small enough $\Delta$, this scenario may not be considered by the agents in the system who seek efficiency. Naturally, an efficiency versus robustness tradeoff emerges.




(a) Δ=0.05

Figure 4-14: Efficiency versus Robustness. $\Delta = 0.05$ Efficiency is represented by the value in nominal model, and robustness is represented by the value in worse-case model. The number of firms $N_t$ is 25.

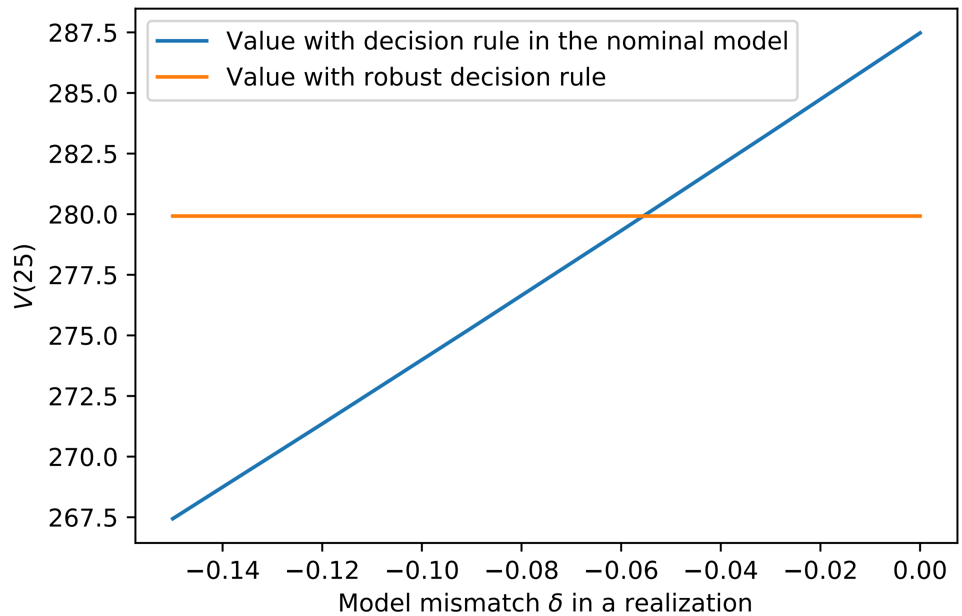
Figure 4-14 compares the case of the nominal and worst-case models. The blue line represents the value function for the nominal model, and the orange the value function for the robust model. The simulations are constructed assuming that the uncertainty parameter $\Delta = 0.05$, and that $\bar{\theta} = 0.6$. So, the range for $\theta$ is between $[.55, .65]$. The nominal model optimizes as if $\delta = 0$, and we plot what the realized value function for $\delta < 0$ instead being zero. In other words, when $\delta = 0$ the blue line reaches the maximum because the real value of theta is exactly the one used by the multinational to optimize. on the other extreme (left) when the value of $\delta = -0.05$ the multinational makes choices thinking that the relative shock parameter is $\bar{\theta}$ when it actually is $\bar{\theta} - 0.05$. Therefore, the nominal value function is subject to potential losses.

The orange line is the robust model. Given the assumption of the bounded range, we know that the multinational assumes that the $\theta = \bar{\theta} - 0.05$. In this case, notice that the orange line is flatter, and the worst-case is better than when the nominal model chooses. In fact, the robust model is optimal when $\delta = -0.05$.

Figure 4-15 depicts the case when $\Delta = 0.15$. In this case, the size of the uncertainty is large enough that the case of $\theta = 0.5$ is in the support. According to equation 4.11 the robust approach assumes that $\theta = 0.5$.

Notice that the blue line — the nominal model — behaves similarly as in the previous case, except that a larger range implies bigger potential losses. Again, in the extreme left, the multinational assumes that $\theta = 0.6$ when it actually is 0.45. In contrast, the orange line assumes a $\theta = 0.5$ and produces a totally flat value function.

The intuition behind this result is simple. The robust allocation of suppliers is to set half in the mountain and half in the valley. This implies that in the presence of an aggregate shock — independently where it occurs — half of the suppliers disappear. The flows and costs are identical. Therefore, the expected value is independent of the true "$\theta$".




(b) Δ=0.15

Figure 4-15: Efficiency versus Robustness. $\Delta = 0.15$. Efficiency is represented by the value in nominal model, and robustness is represented by the value in worse-case model. The number of firms $N_t$ is 25.

These two figures highlight the implications of applying robustness to a decision problem. Robustness is needed when the agents do not know the distribution of the shock they are facing, and therefore need to prepare for the worst. Robustness, then, serves to find a policy that reduces the differences over all possible states of nature. In the limit, the most robust action is one in which the outcomes are identical in all states of nature (as shown in Figure 4-15)

### 4.4.4 Discussion

The results in our model need two ingredients: robustness, and a centralized decision maker that internalizes survival probabilities. Table 4.1 summarizes the relationships between the modeling choices and the characteristics of the policy function.

Table 4.1: Relationship between modeling choices and characteristics of the policy function.

| | Risk | Uncertainty |
|---|---|---|
| Decentralized | Corner Solution **Valley** | Corner Solution **Valley** |
| Centralized | Internal Solution $\psi(N_t)$ | Probability Matching $\psi'(N_t) = 0$ |

When the economy is populated with decentralized decision makers, the optimal policy choices are independent of the nature of the shock. In other words, individual suppliers will choose a corner solution (e.g. locating exclusively in the Valley) irrespectively of whether they are facing risk or uncertainty. A centralized decision maker, on the other hand, tends to prefer internal solutions. When they face risk, the optimal allocation of firms is an increasing function of the number of firms. It is when both uncertainty and centralization are present that the solution is an internal and fixed ratio; very much in the spirit of probability matching.

This result has important implications for the supply chain. In our model, it does not matter how bad the Mountain is relative to the Valley, there is a level of uncertainty for which the multinational allocates half the firms in the Mountain. A robust supply chain is one in which the survival probability is maximized, and where there will be production even in the worst of circumstances. Of course, this is a result that depends on the underlying assumptions of the model, but the intuition should be easy to extend to more realistic circumstances: if the supply chains in the world would have been prepared to supply goods in the worst possible circumstance, then the Covid-19 shock should have produced zero stock-outs. A supply chain that deals with risk but optimizes using the "expected" value is found to be ill-prepared to handle an aggregate shock of the magnitude implied by Covid-19.

Finally, robustness is often compared to the solution of a model with large risk aversion. In some applications, that is indeed the case. In our model, robustness changes the nature of the solution — in fact, the optimal allocation is independent of the number of firms, and the allocation in Mountain and Valley is symmetric. It is common in economics to equate robust approaches to greater risk aversion, which itself leads to higher incentives for diversification. Again, that is not the case in our model. The result we highlight, that implicitly rationalizes probability matching, cannot be achieved with infinite risk aversion.

## 4.5 Price and Cost Differences

Up to now, we have not allowed prices to change depending on the state of the world. This is clearly a simplification that has allowed us to characterize the solution of the model and study the price gouging case. This section, in practice, relaxes the price gouging assumption.[19] The conclusion so far is that even when prices do not adjust — making the decision to locate the suppliers in the Mountain a less profitable decision — the values of continuation and robustness are enough for the multinational to allocate firms in the Mountain. In this section we study the implications of allowing prices and costs in the two locations to be different.

### 4.5.1 Model

We assume that there are heterogeneous costs $c_v \neq c_m$ and prices $p_v \neq p_m$ in the two locations. We use this model to address many different questions. First, can the government define an intervention (either reducing the cost of the Mountain, or increasing its price) for which the decentralized economy achieves the social optimum - even in the case of uncertainty? Second, what is the profit margin of the Mountain at which the robust control strategy ceases to diversify the supply chain? In other words, when is robustness undesirable?

[19]see section 4.2.3 for the justification.

The value functions for the individual suppliers are given by

$$V_t^v = ((1-\gamma) + \gamma\theta)p_v \qquad\quad - c_v \; + \frac{1}{1+\beta}((1-\gamma)+\gamma\theta)V_{t+1} \tag{4.12}$$

$$V_t^m = ((1-\gamma) + \gamma(1-\theta))p_m - c_m + \frac{1}{1+\beta}((1-\gamma)+\gamma(1-\theta))V_{t+1} \tag{4.13}$$

The continuation values still are identical in each of the two locations because there is no cost of relocation of suppliers. The difference between the two locations is given by

$$V_t^v - V_t^m = (((1-\gamma)+\gamma\theta)p_v - c_v) - (((1-\gamma)+\gamma(1-\theta))p_m - c_m) + \gamma(2\theta-1)\left(\frac{1}{1+\beta}V_{t+1}\right) \tag{4.14}$$

There is an expected markup in the Mountain larger than the markup at the Valley at which the firms are indifferent in their location. Intuitively, it is not enough for the markups of the Mountain and Valley to be the same. We need to compensate individuals going to the Mountain for their lower probability of survival. Therefore, given $p_v$, $c_v$ and $c_m$, there exists a cutoff $p_m^*$ at which $V_t^v = V_t^m = V_{t+1}$. Substituting in Equations 4.12 and 4.13, the transition occurs when

$$\frac{(1-\gamma+\gamma\theta)p_v - c_v}{1-\frac{1}{1+\beta}(1-\gamma+\gamma\theta)} = \frac{(1-\gamma+\gamma(1-\theta))p_m^* - c_m}{1-\frac{1}{1+\beta}(1-\gamma+\gamma(1-\theta))}. \tag{4.15}$$

Given the parameters of our simulation, $p_m^* \approx 1.24$. Then, for any $p_m < p_m^*$ all individual firms locate in the Valley, and for $p_v > p_m^*$ all locate in the Mountain. The intuition of Equation 4.15 is simple; it states that the expected markups adjusted by the survival probabilities need to be equated in the two locations.

In this case, the individual allocation implies multiple equilibria due to the indifference between the two locations. Below that markup the dominant strategy is to locate in the Valley, and above it the optimal decision is to locate in the Mountain. We compare this solution to the one chosen by the multinational in the exact same setting.

The problem of the multinational firm can be written as

$$V(N_t) = \max_{\psi_t} \left\{ \left[ \begin{array}{llll} (1-\gamma) & ((p_v\psi_t + p_m(1-\psi_t))N_t & +\frac{1}{1+\beta}V\left(A\cdot(N_t)^{1-\mu}\right) & ) \\ +\gamma\theta & (p_v\psi_t N_t & +\frac{1}{1+\beta}V\left(A\cdot(\psi_t N_t)^{1-\mu}\right) & ) \\ +\gamma(1-\theta) & (p_m(1-\psi_t)N_t & +\frac{1}{1+\beta}V\left(A\cdot((1-\psi_t)N_t)^{1-\mu}\right) & ) \end{array} \right] \right.$$
$$\left. - (c_v\psi_t + c_m(1-\psi_t))N_t \right\} \tag{4.16}$$

subject to the same boundary condition we have imposed before:

$$\lim_{N\to 1^-} V(N) = 0.$$

The instantaneous profits are linear in prices and costs, so we have decided to keep costs constant and only change the Mountain's price ($p_m$) in our simulations. Our first result studies how the multinational's optimal policy $\psi^*(N)$ changes with the Mountain's price $p_m$.

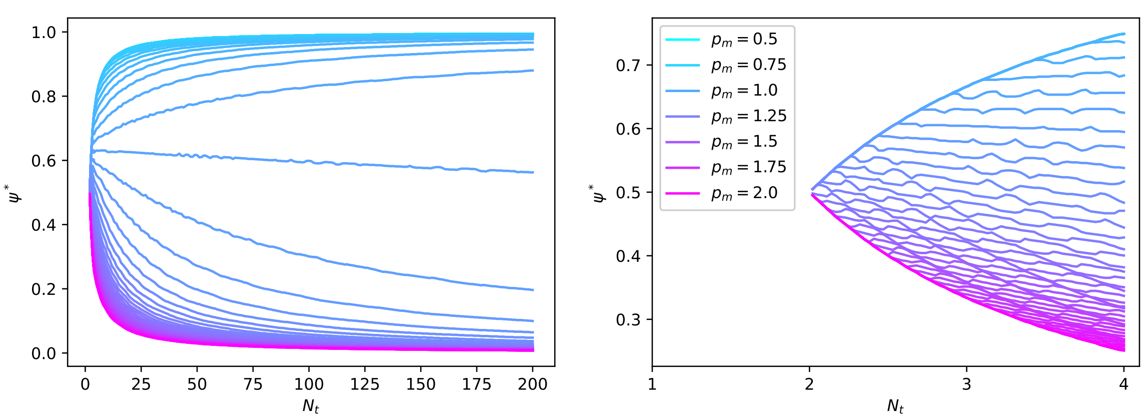



Figure 4-16: Optimal policy as cost of Mountain changes

In Figure 4-16 we present two panels. The left panel shows the optimal allocation for different number of suppliers, and the panel on the right just zooms into the case when there are few suppliers available. Each colored line indicates a different price level, and prices in the Mountain were varied in the range $[0.5, 2.0]$.

As in the baseline model, when there are two firms the multinational allocates one in each location to ensure survival. We decided to start the figure at that point because nothing particularly new occurs for $N_t = 1$. As $N_t$ increases, the multinational allocates firms depending on the prices and the number of available suppliers.

When the prices are low, the optimal allocation is biased toward the Valley, and when the prices are high, the allocation is biased toward the Mountain. Interestingly, there is a price at which the allocation is virtually flat.

Figure 4-17 shows the optimal allocation in the Valley for a given $N$, but different values of $p_m$. The plot has been drawn for $N_t = 10$.

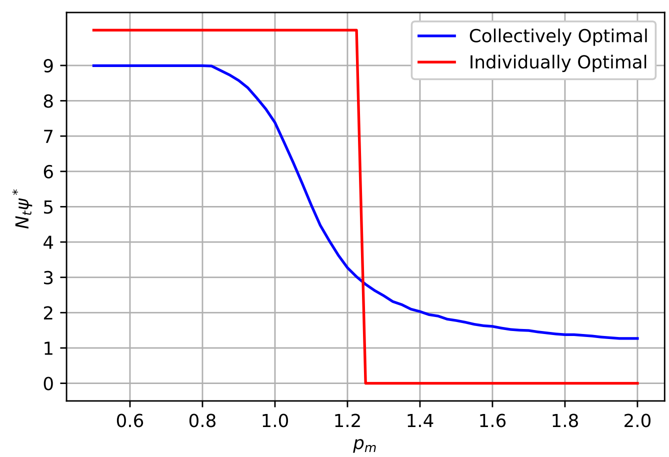



Figure 4-17: Individual vs. Multinational Optimal with Heterogeneous prices.

The multinational transitions smoothly between extreme values of $\psi^*$ as the price of the Mountain varies, but the decentralized equilibrium instantaneously shifts at a critical value of $p_m^* \approx 1.24$. This phase transition from $\psi^* = 0$ to $\psi^* = 1$ occurs when the values of the Mountain and Valley given by Equation 4.12 are equal.

There are two aspects worth highlighting from this simulation. First, the price at which the multinational is indifferent between Mountain and Valley — the point at which it allocates half the firms in the Valley — occurs between 1 and 1.2 (when $N_t\psi = 5$). Notice that the price at this indifferent point is much lower than the price at which the individuals are indifferent. The reason is that the value of continuation is marginally improved when $\psi$ is lowered from 1, and therefore the marginal contribution for the multinational is larger than for the individual firm.

Second, there is a kink at the top left for the multinational. This is the place where the optimal allocation in the Mountain would have implied less than one firm. However, because of the value matching condition, the multinational allocates a maximum of 9 firms to the Valley. A similar kink occurs on the right side of the graph, but the prices required to reach it are large - swamping the details of the graph presented.

## 4.6 Discussion and Policy Implications

Our model is quite minimalist and has the objective of highlighting the contrast between the possible alternatives of approaching the global supply chain problem. The most important result is that robustness in the presence of certain forms of uncertainty replicates a behavior that tends to be considered sub-optimal or irrational – firms follow a probability matching strategy. However, what does robustness looks in practice, and what can policy makers do to achieve such an arrangement?

In this chapter we argue that robustness should be a stronger consideration for the design of supply chains, but that this has not been the case. Therefore, it is not surprising to us that there are many supply chain disruptions — during Covid-19 and even in 2021 with the scarcity of CPU's. It is our opinion that Just-in-Time and Just-in-Case are not enough to remedy those shocks. Moreover, JIT and JIC will not provide assurances for shocks that are yet to come – for instance those related to environmental disasters and social unrest. Therefore, instead of trying to find an example that is either extremely particular or not terribly important, we have decided to look at other industries that are designed for robustness. Those are the financial system and the postal services.

### 4.6.1 Robustness in Practice

The financial system is becoming more and more robust through time. This is driven by the fact that it is a sector where supply disruptions are extremely costly, and aggregate shocks tend to happen repeatedly. The financial industry continues to advance and become more resilient – this is a slow process, but the institutions of the modern banking system are ages ahead of those that existed at the beginning of the 20$^{\text{th}}$ century. The banking system is safer today than ever before and that has allowed for unprecedented growth.[20]

The banking system has deposit insurance, cash and capital requirements, stress tests, specific rules to deal with failure, strict licenses of operation, and a lender or last resort. Most of these institutional improvements have been proposed after financial crises. Interestingly to us, supply disruptions have not tended to improve the institutional environment of the global supply chain.

First, the stress test implemented after the 2008 crisis is clearly a tool to evaluate the bank's performance in a set of extreme circumstances. It is by definition an statistic to measure the impact on the balance sheet in the worst case. In reality, the worst case for each bank can be a different combination of shocks. There are some cases in a supply chain — such as an internal production line — that stress tests are performed. Very rarely is this evaluated internationally, certainly not to the extent of achieving robustness to global shocks.

The second set of tools that are associated with robustness are those that deal with readiness to a shock. In this case, capital requirements and cash reserves are designed to protect the balance sheet to contingent loses. Failure of the stress tests or the provision of capital requirements has consequences for financial institutions. These sometimes include limiting the ability of the risky bank to continue operating, and even the removal of its license to operate.

The third set of policies deal with the "clients" behavior. In the example of the

[20]In fact, if we were to take the current regulatory environment and the size and activities of the banking system of 100 years ago, the probability of a financial crisis would be virtually zero. The increased resilience and trust allows for the financial sector to continue expanding — which creates new opportunities of disruption and subsequently new opportunities to improve it.

banking system, the clients are the depositors. The objective of a federal deposit insurance is to stop bank runs — or at least to reduce their severity and frequency. Again, this is a policy that is entirely designed to deal with an extreme shock.

Finally, the existence of a lender of last resort works in conjunction with the previous regulatory tools. It makes the deposit insurance credible, and it creates the residual claimant of the stress tests and the capital requirements. Therefore, there exists someone in the economy that is really concerned with the rescue of the banks because it is extremely costly for them. One important aspect of the single residual claimant of a banking crises is the ability to handle the informational frictions that are likely to appear. This institutional arrangement offers social insurance to the banking system, which leads to the standard moral hazard and adverse selection problems. It is important in this design to take into account those issues. During the Covid-19 crisis we indeed observed a “lender of last resort” in the form of fiscal and monetary policy support to firms and citizens. However, these were reactionary policies, and if used repeatedly without proper design may create bad incentives.

When we think about the global supply chain, (i) there is a lack of stress tests, (ii) most suppliers are independent and therefore there is little ability to enforce proper firm health (equivalent to a capital requirement) (iii) there is no institution that compensates clients for failures of the suppliers (analogous to the FDIC), and (iv) only in the case when all the suppliers belong to a single firm can we find a residual claimant (the lender of last resort). We hope this makes clear why supply chain failures happen so regularly, and why ex-post these shocks look easy to deal with – such as occurred with isopropyl alcohol and toilet paper.

We are obviously not advocating for a central bank equivalent of global supply chains. We are, however, highlighting the fact that firms can design for and assess robustness ex-ante. Stress testing is simple to implement and can even be a part of the supply contract. Furthermore, capital requirements or financial health is usually evaluated in some international relationships (such as joint ventures) but a process of certification that is transparent and standardized would be more effective than dealing with these issues on a case-by-case basis.

There is a final aspect of robust supply chains that is hard to see in the financial system — even though it exists. Probability matching implies resilience though redundancies. A robust supply chain will have excess capacity that seems very costly in normal times, because that excess capacity exists to be used if the extreme shock occurs. It is exactly the opposite of Just-in-Time and on average contradicts Just-in-Case which presumably implies that only half the time it has excess capacity. The financial system has this excess capacity built into the tools we describe, but this is more easily seen in the postal service. The postal service is a very robust supply chain. Notice the elements: one residual claimant, continuous stress testing, and excess capacity in normal times. In fact, the postal service has so much excess capacity that people in general argue that it is too inefficient. We are not arguing that the postal service is efficient. We are highlighting that such a statement cannot be made in the absence of robust thinking. If the postal service has been designed with robustness in mind, it *needs* to look inefficient during normal times. This problem is even harder to resolve in the global supply chain. The reason is that excess capacity requires not only physical capital but also human capital. Multinationals need to develop production capabilities in countries where, in normal times, it seems like a bad decision to do so or a luxury the firm can avoid. Of course, these assessments lead to underinvestment in production capabilities; when the shock hits, the regret of not having invested settles in.

### 4.6.2 Uncertainty and Price Gouging in Reality

The two most important elements of our model are the existence of uncertainty and the failure of prices to fully adjust. Uncertainty occurs when the distribution of a model parameter is unknown, but it can fall only within a range. In macroeconomics, uncertainty is pervasive, and rightfully so. Many shocks of different natures affect the economy, and are transmitted through complex and unobservable networks that shifting through time. Not surprisingly then, central bankers experience such uncertainty and even "talk in ranges" (See Section 3.4 of Chapter 3).

In fact, as argued by Sargent and Hansen, uncertainty is the natural outcome of an estimation process where multiple models are acceptable. Uncertainty in our economic and policy models should be the norm, although unfortunately it is not. We believe that this failure may lead to bad policy choices – thereby increasing the risk our economies are subjected to and the likelihood that living standards deteriorate.

The second important element in our model is price gouging. For example, if firms knew *ex-ante* that prices would be allowed to increase to their fullest extent, then firms would individually choose to diversify until locating in the Mountain or the Valley has the same expected returns. There were several events during Covid-19 that show that prices did not adjust to this extent. First, many countries have price gouging laws that limit firms' ability to increase prices. So, after a supply shock prices would increase – but not fully. Second, in many countries, lawsuits began accusing firms of abusive behavior – most of which were related to price increases. Third, even when prices for some goods increased, we saw stock-outs and rationing of products — clearly signaling that the demand was higher than the supply and prices could not have been at equilibrium. As we have said before, we know that price gouging as modeled here is an extreme assumption. The supply disruptions and rationing that occurred in developed nations for products such as food, beverages, personal care, etc., indicates that prices never increased enough for it to be profitable for the supply chain to diversify and prepare for large aggregate shocks. to the point at which it is profitable for the supply chain to prepare for large aggregate shocks. Our results in this chapter will qualitatively hold as long as prices do not adjust fully.

Furthermore, we see countries provide price incentives to curtail Just-in-Time practices. As shown in Figure 4-16, there is a price of Mountain-produced goods for which the multinational's optimal policy implies diversification similar to the robust policy. In fact, a similar heterogeneity of prices can make individual suppliers replicate the robust policy as well. Therefore, a subsidy given to firms in riskier locations can help achieve the robust allocation.

This conclusion is important because governments and the private sector might experience risk differently. Typically, governments are the residual claimant in case of natural disasters. Therefore, governments are more likely to prefer a "robust approach" than the private sector would. For instance, if the cost of a natural disaster is very asymmetric, the government is more likely to pay attention to the worst-case than the private sector. If this is indeed the case, the government can align private incentives by providing a small subsidy to the riskier location.

## 4.7 Conclusions

Covid-19 was an aggregate shock that highlighted the weaknesses on the supply chain. Many products suffered disruptions: from personal protective equipment, to toilet paper, and beer. It is clear that the supply chains of the world were not prepared for this event. Many are imploring that "future" supply chains become more resilient or robust — but what exactly is a robust supply chain? and how exactly do firms' decisions change when taking a robust approach?

This chapter studies a very stylized model of a supply chain. A multinational producing a product benefits from many suppliers providing parts, but those firms might not choose the best allocation of resources when aggregate shocks are present. Our model discusses how different arrangements of the supply chain emerge in different settings. In particular, we concentrate on two factors: (i) the internalization of the survival probability — in the spirit of the usual externalities; and (ii) the nature of the shock with either risk or uncertainty.

In this chapter, a robust supply chain is that which optimally deals with uncertainty. It implies designing for the worst-case of a set of aggregate shocks. Robustness yields a strategy that seems to maximize survival probability, and therefore our model rationalizes or explains the well known "probability matching" behavior observed in experimental literature. Probability matching is the result of a group decision process in which the number of people in the group who choose a given strategy is proportional to the probability that that strategy will yield survival. Probability matching is inefficient for individual decision-makers, as it is optimal for them to maximize their own probability of survival. However, probability matching maximizes collective growth. To achieve, or at least to get close to probability matching, we need coordination between small decentralized suppliers — the market does not work properly during a crisis because the price system fails to signal scarcity correctly. A multinational can partially remedy this by considering continuation value, and fiscal support can help replicate the probability matching and robust allocation.

This chapter leaves many research questions that could be addressed in the future. First, the relaxation of the identical price fixing period and investment horizons. Implicitly, the question is to evaluate quantitatively the importance of the anti-price gouging or partial price adjustment effects. Second, our model needs a market inefficiency to exist in order to break the first welfare theorem. This chapter concentrates on such an inefficiency in the pricing system. However, it is conceivable that many other market imperfections could produce similar results.[21]. The role of robustness in those environments — with imperfect information, coordination failures, or externalities — might be a promising area of research.[22]

Finally, as we highlighted in this chapter, robustness is not equivalent to assuming that shocks are larger or more severe; it is a fundamentally different strategy that addresses uncertainty by minimize the losses of the worst case outcome. In doing so, it produces a different form of diversification. Increasing the variance of shocks is equivalent to arguing that the supply chain moves from Just-in-Time to Just-in-Case. Uncertainty, on the other hand, implies that the supply chain moves to Just-in-Worst-Case.

[21]We thank David Baqaee for providing us with these future avenues of research.

[22]Other forms of inefficiencies can affect how costly the supply disruptions can be. In our model is a very simple mechanism, but see Baqaee and Farhi (2020) for a thorough study of misallocation in general equilibrium of different market imperfections

# Chapter 5

# Conclusion and Future Work

A large amount of modern economic analysis consists primarily of building abstract mathematical models for decision-making. However, mathematical models are essential tools that rely heavily on model assumptions and calibrations. As a result, it is an important question how to make decisions based on a collection of potentially true models. This thesis studies this robustness issue in a variety of macroeconomic and supply chain problems. Standing on the shoulder of existing literature, we develop tools and theories to build a more robust macroeconomic decision-making process.

In the first chapter, we study the problem of identifying linear financial networks using aggregate data. Since a financial network may be contingent on the type of shock that hits it, an identification tool for multiple networks is needed. We showed the solution uniqueness of identifying a single network using heteroskedasticity and created an Expectation-Maximization based method to identify multiple linear networks in the theoretical section. We tested our strategy on credit default swap data from US banks and discovered that three networks best represent the data. Furthermore, each one of the networks ranks banks differently in terms of centrality. A macroprudential policy would prepare for the worst case of the three networks.

In the second chapter, we examine whether more policy instruments are always better in a bounded uncertainty setting. Bounded uncertainty is the more appropriate modeling tool for our central banking problem because many quantities in central banking are more commonly communicated in terms of falling in a range rather than following a distribution. When bounded uncertainty is transmitted through a conditional expectation channel, we find that not utilizing a policy tool is sometimes the better option for a central bank. Two monetary policy models for asset purchasing and forward guidance are used to illustrate the phenomenon.

In the third chapter, we look at a stylized supply chain model to see how a global corporation's decision alters in the face of anti-price gouging and uncertainty. This chapter created a basic model in which huge multinationals diversify their manufacturing plants to higher-cost countries due to continuation values despite having a fixed price. Our approach is based on a class of literature on adaptive markets and probability matching. It means that having a small number of large enterprises is more efficient than having a large number of small firms in a market with aggregate shock and anti-price gouging rules. We also look at a robust supply chain that focuses on the worst-case scenario. Our model explains the behavior of "probability matching" reported in the experimental literature.

Sophisticated mathematical models improve our economy's efficiency and sometimes bring with it the problem of model robustness. More and more models are verified, selected, and calibrated with the help of big data, but large economic shocks usually bring fundamental changes not reflected in the data collected in normal times. As a result, the robustness challenge is even more prominent these days in economics, as well as other social science domains. There is a lot of exciting future research in this area, many of which I do not have enough knowledge to comment on. In the following paragraphs, I will focus on a few that are closest to my research expertise.

We discussed a forward guidance model in an environment with bounded uncertainty in the second chapter. The model is based on a linearized New Keneysian Model of central banking. While the linearized model is enough for our purpose and is straightforward to solve, we cannot use it to study the effect of model uncertainty in the inflation dynamics models. In the New Keynesian model, while the Euler equation part is widely accepted and adopted in many different analyses, there are many different ways to model the inflation dynamics part. The canonical one that we embraced assumes sticky prices and no capital accumulation. While those assumptions allow us to derive a linearized New Keynesian Phillips curve in just a few steps, it is a great simplification of reality. It will be interesting to discuss the robust decision rules of the central bank when the New Keynesian Phillips curve has a large model uncertainty. This future research will also have a critical real-world impact, since many empirical papers have found that the slope of the Phillips curve has a considerable variation when fitted with data collected in recent years.

We discussed a stylized model of a supply chain facing aggregate shocks in the third chapter. A critical assumption in this chapter is the fixed price due to price rigidity and anti-price gouging. Although this assumption is enough to demonstrate our point and makes the model straightforward to solve, we cannot analyze or measure the amount of price distortion and the induced allocation distortion without a model with a partially adjusted price system. Some established price rigidity models can be used to improve our analysis. While the model may become much harder to solve, it might suggest a way to estimate or measure the fragility of supply chains in different industries. This would then help build a warning system for industry-wide catastrophic supply chain failures.

Furthermore, many global games model gives a continuum of multiple equilibria. Although more granular modeling of the dynamics of equilibrium formation can often pin down the unique equilibrium, the information required may not always be available to policymakers. As a result, policymakers have to make decisions based on a range of potential equalibria results. In this case, the equilibrium value cannot be modeled as a random variable with some distribution. Instead, policymakers should model it as a bounded uncertainty term and apply a robust decision-making rule.

# Bibliography


Acemoglu, D., A. Ozdaglar, and A. Tahbaz-Salehi (2015). Systemic risk and stability in financial networks. *The american economic review 105*(2), 564–608.

Adrian, T. and M. K. Brunnermeier (2016). Covar. *American Economic Review 106*(7), 1705–1741. Predicting and measuring a financial institution's contribution to systemic risk that internalizes externalities and avoids procyclicality.

Alfaro, L., D. Chor, P. Antras, and P. Conconi (2019). Internalizing global value chains: A firm-level analysis. *Journal of Political Economy 127*(2), 508–559.

Allen, F., A. Babus, and E. Carletti (2012). Asset commonality, debt maturity and systemic risk. *Journal of Financial Economics 104*(3), 519–534.

Allen, F. and D. Gale (2000). Financial contagion. *Journal of political economy 108*(1), 1–33.

Ambulkar, S., J. Blackhurst, and S. Grawe (2015). Firm's resilience to supply chain disruptions: Scale development and empirical examination. *Journal of operations management 33*, 111–122.

Anand, K., B. Craig, and G. Von Peter (2015). Filling in the blanks: Network structure and interbank contagion. *Quantitative Finance 15*(4), 625–636.

Anderson, D. and P. Apfel (2020, July). Covid-19 global impact: A world tour of competition law enforcement (4 "tracker maps"). *www.concurrences.com*.

Anderson, E., N. Jaimovich, and D. Simester (2015). Price stickiness: Empirical evidence of the menu cost channel. *Review of Economics and Statistics 97*(4), 813–826.

Andrade, P., G. Gaballo, E. Mengus, and B. Mojon (2019). Forward guidance and heterogeneous beliefs. *American Economic Journal: Macroeconomics 11*(3), 1–29.

Antras, P. (2003). Firms, contracts, and trade structure. *The Quarterly Journal of Economics 118*(4), 1375–1418.

Antras, P. (2015). *Global production: Firms, contracts, and trade structure.* Princeton University Press.

Antràs, P. (2020). *Conceptual aspects of global value chains.* The World Bank.

Antràs, P. and D. Chor (2013). Organizing the global value chain. *Econometrica 81*(6), 2127–2204.

Antràs, P. and D. Chor (2018). *On the measurement of upstreamness and downstreamness in global value chains*. Routledge.

Antràs, P. and A. De Gortari (2020). On the geography of global value chains. *Econometrica 88*(4), 1553–1598.

Baqaee, D. R. and E. Farhi (2020). Productivity and misallocation in general equilibrium. *The Quarterly Journal of Economics 135*(1), 105–163.

Billio, M., M. Getmansky, A. W. Lo, and L. Pelizzon (2010). Econometric measures of systemic risk in the finance and insurance sectors. Technical report, National Bureau of Economic Research.

Bishop, C. M. (2006). *Pattern recognition and machine learning*. springer.

Board of Governors of the Federal Reserve System (2021, Feb). Monetary policy report.

Brennan, T. J. and A. W. Lo (2011). The origin of behavior. *The Quarterly Journal of Finance 1*(01), 55–108.

Caballero, R. J. and E. Engel (2006). Price stickiness in ss models: basic properties. *unpublished, MIT, October*.

Caballero, R. J. and A. Simsek (2013). Fire sales in a model of complexity. *The Journal of Finance 68*(6), 2549–2587.

Cabrales, A., P. Gottardi, and F. Vega-Redondo (2017). Risk sharing and contagion in networks. *The Review of Financial Studies 30*(9), 3086–3127.

Campbell, J. R., C. L. Evans, J. D. Fisher, A. Justiniano, C. W. Calomiris, and M. Woodford (2012). Macroeconomic effects of federal reserve forward guidance. *Brookings Papers on Economic Activity 2012*(1), 1–80.

Caporin, M., L. Pelizzon, F. Ravazzolo, and R. Rigobon (2018). Measuring sovereign contagion in europe. *Journal of Financial Stability 34*, 150–181.

Cary, G. S., M. J. F. M. Dolmans, B. Hoffman, T. Graf, L. Brannon, R. Pepper, H. Mostyn, A. R. B. Lazda, S. Haynes, K. Georgieva, and J. Przerwa (2020, April). Exploitative abuses, price gouging & covid-19: The cases pursued by eu and national competition authorities. *www.concurrences.com*.

Chen, H., W. W. Dou, and L. Kogan (2019). Measuring dark matter in asset pricing models. Technical report, National Bureau of Economic Research.

Christiano, L. J., M. Eichenbaum, and C. L. Evans (1999). Monetary policy shocks: What have we learned and to what end? *Handbook of macroeconomics 1*, 65–148.

Cisternas, G. (2018). Two-sided learning and the ratchet principle. *The Review of Economic Studies 85*(1), 307–351.

Del Negro, M., M. P. Giannoni, and C. Patterson (2012). The forward guidance puzzle. *FRB of New York Staff Report* (574).

Durach, C. F. and J. A. Machuca (2018). A matter of perspective–the role of interpersonal relationships in supply chain risk management. *International Journal of Operations & Production Management 38*(10), 1866–1887.

Dworczak, P., S. D. Kominers, and M. Akbarpour (2020). Redistribution through markets.

Eichengreen, B., A. Mody, M. Nedeljkovic, and L. Sarno (2012). How the subprime crisis went global: evidence from bank credit default swap spreads. *Journal of International Money and Finance 31*(5), 1299–1318.

Elliott, M., B. Golub, and M. O. Jackson (2014). Financial networks and contagion. *American Economic Review 104*(10), 3115–53.

Elsinger, H., A. Lehar, and M. Summer (2013). Network models and systemic risk assessment. *Handbook on Systemic Risk 1*, 287–305.

Evans, D. (2020, Feb). Coronavirus shows that supply chains are outdated and unfit for modern manufacturing. *www.forbes.com*.

Farmer, R. E., D. F. Waggoner, and T. Zha (2007). Understanding the new-keynesian model when monetary policy switches regimes. Technical report, National Bureau of Economic Research.

Fiorina, M. P. (1971). A note on probability matching and rational choice. *Behavioral Science 16*(2), 158–166.

Forbes, K. J. and R. Rigobon (2002). No contagion, only interdependence: measuring stock market comovements. *The journal of Finance 57*(5), 2223–2261.

Fraley, C. and A. E. Raftery (1998). How many clusters? which clustering method? answers via model-based cluster analysis. *The computer journal 41*(8), 578–588.

Freixas, X., B. M. Parigi, and J.-C. Rochet (2000). Systemic risk, interbank relations, and liquidity provision by the central bank. *Journal of money, credit and banking 32*(3), 611–638.

Gai, P., A. Haldane, and S. Kapadia (2011). Complexity, concentration and contagion. *Journal of Monetary Economics 58*(5), 453–470.

Gai, P. and S. Kapadia (2010). Contagion in financial networks. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 466*(2120), 2401–2423.

Gaissmaier, W. and L. J. Schooler (2008). The smart potential behind probability matching. *Cognition 109*(3), 416–422.

Galí, J. (2015). *Monetary policy, inflation, and the business cycle: an introduction to the new Keynesian framework and its applications.* Princeton University Press.

Galı, J. and M. Gertler (1999). Inflation dynamics: A structural econometric analysis. *Journal of monetary Economics 44*(2), 195–222.

Gilboa, I. and D. Schmeidler (1989). Maxmin expected utility with non-unique prior. *Journal of mathematical economics 18*(2), 141–153.

Girardi, G. and A. T. Ergün (2013). Systemic risk measurement: Multivariate garch estimation of covar. *Journal of Banking & Finance 37*(8), 3169–3180.

Guerrieri, V. and G. Lorenzoni (2017). Credit crises, precautionary savings, and the liquidity trap. *The Quarterly Journal of Economics 132*(3), 1427–1467.

Hallegatte, S. (2015). *The indirect cost of natural disasters and an economic definition of macroeconomic resilience.* The World Bank.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society 50*(4), 1029–1054.

Hansen, L. P. and T. J. Sargent (2008). *Robustness.* Princeton university press.

Helpman, E., M. J. Melitz, and S. R. Yeaple (2004). Export versus fdi with heterogeneous firms. *American economic review 94*(1), 300–316.

Holmström, B. (1999). Managerial incentive problems: A dynamic perspective. *The review of Economic studies 66*(1), 169–182.

Junker, B. H. and F. Schreiber (2008). *Analysis of biological networks*, Volume 2. Wiley Online Library.

Kahneman, D., S. P. Slovic, P. Slovic, and A. Tversky (1982). *Judgment under uncertainty: Heuristics and biases.* Cambridge university press.

Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica 47*, 263–291.

Kalbaska, A. and M. Gątkowski (2012). Eurozone sovereign contagion: Evidence from the cds market (2005–2010). *Journal of Economic Behavior & Organization 83*(3), 657–673.

Katz, L. (1953). A new status index derived from sociometric analysis. *Psychometrika 18*(1), 39–43.

King and Spalding (2020). Covid-19 survey of federal and state price gouging laws. *www.kslaw.com.*

Kremer, M. (1993). The o-ring theory of economic development. *The Quarterly Journal of Economics 108*(3), 551–575.

Kritzman, M., Y. Li, S. Page, and R. Rigobon (2011). Principal components as a measure of systemic risk. *The Journal of Portfolio Management 37*(4), 112–126.

Laubach, T. and J. C. Williams (2003). Measuring the natural rate of interest. *Review of Economics and Statistics 85*(4), 1063–1070.

Little, R. J. and D. B. Rubin (2019). *Statistical analysis with missing data*, Volume 793. John Wiley & Sons.

Lo, A. W. (2017). *Adaptive markets*. Princeton University Press.

Long, S. (2020, May). Coronavirus pandemic exposes fatal flaws of the 'just-in-time' economy. *www.abc.net.au*.

Maccheroni, F., M. Marinacci, and D. Ruffino (2013). Alpha as ambiguity: Robust mean-variance portfolio analysis. *Econometrica 81*(3), 1075–1113.

Mavroeidis, S., M. Plagborg-Møller, and J. H. Stock (2014). Empirical evidence on inflation expectations in the new keynesian phillips curve. *Journal of Economic Literature 52*(1), 124–88.

McKay, A., E. Nakamura, and J. Steinsson (2016). The power of forward guidance revisited. *American Economic Review 106*(10), 3133–58.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica 71*(6), 1695–1725.

Merton, R. C., M. Billio, M. Getmansky, D. Gray, A. W. Lo, and L. Pelizzon (2013). On a new approach for analyzing and managing macrofinancial risks (corrected). *Financial Analysts Journal 69*(2), 22–33.

Miao, J. and A. Rivera (2016). Robust contracts in continuous time. *Econometrica 84*(4), 1405–1440.

Morse, E. B. and W. N. Runquist (1960). Probability-matching with an unscheduled random sequence. *The American journal of psychology 73*(4), 603–607.

Oh, H. and R. Reis (2012). Targeted transfers and the fiscal response to the great recession. *Journal of Monetary Economics 59*, S50–S64.

Onnela, J.-P., K. Kaski, and J. Kertész (2004). Clustering and information in correlation based financial networks. *The European Physical Journal B 38*(2), 353–362.

Pasquariello, P. (2017). Government intervention and arbitrage. *The Review of Financial Studies 31*(9), 3344–3408.

Pasquariello, P., J. Roush, and C. Vega (2011). Government intervention and strategic trading in the us treasury market. *Journal of Financial and Quantitative Analysis*, 1–91.

Plenert, G. (2007). *Reinventing lean: introducing lean management into the supply chain.* Oxford, UK: Butterworth-Heinemann.

Rigobon, R. (2003). Identification through heteroskedasticity. *Review of Economics and Statistics 85*(4), 777–792.

Rotemberg, J. J. (2002). Customer anger at price increases, time variation in the frequency of price changes and monetary policy. Technical report, National Bureau of Economic Research.

Rotemberg, J. J. (2011). Fair pricing. *Journal of the European Economic Association 9*(5), 952–981.

Rudd, J. and K. Whelan (2005). New tests of the new-keynesian phillips curve. *Journal of Monetary Economics 52*(6), 1167–1181.

Saenz, M. J., X. Koufteros, C. F. Durach, A. Wieland, and J. A. Machuca (2015). Antecedents and dimensions of supply chain robustness: a systematic literature review. *International Journal of Physical Distribution & Logistics Management 45*(1), 118137.

Sentana, E. and G. Fiorentini (2001). Identification, estimation and testing of conditionally heteroskedastic factor models. *Journal of econometrics 102*(2), 143–164.

Seth, A. K. (2007). The ecology of action selection: insights from artificial life. *Philosophical Transactions of the Royal Society.*

Sion, M. et al. (1958). On general minimax theorems. *Pacific Journal of mathematics 8*(1), 171–176.

Stegeman, A. and N. D. Sidiropoulos (2007). On kruskals uniqueness condition for the candecomp/parafac decomposition. *Linear Algebra and its applications 420*(2-3), 540–552.

Stein, J. C. and A. Sunderam (2018). The fed, the bond market, and gradualism in monetary policy. *The Journal of Finance 73*(3), 1015–1060.

Strzalecki, T. (2011). Axiomatic foundations of multiplier preferences. *Econometrica 79*(1), 47–73.

Temizsoy, A., G. Iori, and G. Montes-Rojas (2017). Network centrality and funding rates in the e-mid interbank market. *Journal of Financial Stability 33*, 346–365.

The FT Editorial Board (2020, April). Companies should shift from just in time to just in case. *Financial Times.*

Thurner, S. and S. Poledna (2013). Debtrank-transparency: Controlling systemic risk in financial networks. *Scientific reports 3*, 1888.

Todd, P. M. and G. Gigerenzer (2012). Ecological rationality: Intelligence in the world.

Töyli, J., H. Lorentz, L. Ojala, A. Wieland, and C. M. Wallenburg (2013). The influence of relational competencies on supply chain resilience: a relational view. *International Journal of Physical Distribution & Logistics Management 43*(4), 300–320.

Tran, B. R., D. J. Wilson, et al. (2020). The local economic impact of natural disasters. Technical report.

UK Competition Authority (2020, April). The uk competition authority publishes update on its covid-19 taskforce, including a number of excessive pricing complaints and investigations. *www.concurrences.com*.

Upper, C. (2011). Simulation methods to assess the danger of contagion in interbank markets. *Journal of Financial Stability 7*(3), 111–125.

Vickery, S. n., R. Calantone, and C. Dröge (1999). Supply chain flexibility: an empirical study. *Journal of supply chain management 35*(2), 16–24.

Vitale, P. (2018). Robust trading for ambiguity-averse insiders. *Journal of Banking & Finance 90*, 113–130.

Vulkan, N. (2000). An economists perspective on probability matching. *Journal of economic surveys 14*(1), 101–118.

Wagner, S. M. and C. Bode (2006). An empirical investigation into supply chain vulnerability. *Journal of purchasing and supply management 12*(6), 301–312.

Yeaple, S. R. (2003). The complex integration strategies of multinationals and cross country dependencies in the structure of foreign direct investment. *Journal of international economics 60*(2), 293–314.

Zhao, S. and F. You (2019). Resilient supply chain design and operations with decision-dependent uncertainty using a data-driven robust optimization approach. *AIChE Journal 65*(3), 1006–1021.