PhD THESIS DECLARATION

I, the undersigned

FAMILY NAME	Todorova
NAME	Zornitsa
Student ID no.	1824317

Thesis title:

Essays on Networks in Financial Economics

PhD in	Finance
Cycle	30th
Student's Advisor	Prof. Carlo Favero
Calendar year of thesis defence	2019

DECLARE

under my responsibility:

1) that, according to Italian Republic Presidential Decree no. 445, 28th December 2000, mendacious declarations, falsifying records and the use of false records are punishable under the Italian penal code and related special laws. Should any of the above prove true, all benefits included in this declaration and those of the temporary "embargo" are automatically forfeited from the beginning;

2) that the University has the obligation, according to art. 6, par. 11, Ministerial Decree no. 224, 30th April 1999, to keep a copy of the thesis on deposit at the "Biblioteche Nazionali Centrali" (Italian National Libraries) in Rome and Florence, where consultation will be permitted, unless there is a temporary "embargo" protecting the rights of external bodies and the industrial/commercial exploitation of the thesis;

3) that the Bocconi Library will file the thesis in its "Archivio Istituzionale ad Accesso Aperto" (Institutional Registry) which permits online consultation of the complete text (except in cases of temporary "embargo");

4) that, in order to file the thesis at the Bocconi Library, the University requires that the thesis be submitted online by the student in unalterable format to Società NORMADEC (acting on behalf of the University), and that NORMADEC will indicate in each footnote the following information:

- PhD thesis Essays on Networks in Financial Economics;
- by Todorova, Zornitsa ;
- defended at Università Commerciale "Luigi Bocconi" – Milano in the year 2019;
- the thesis is protected by the regulations governing copyright (Italian law no. 633, 22nd April 1941 and subsequent modifications). The exception is the right of Università Commerciale "Luigi Bocconi" to reproduce the same, quoting the source, for research and teaching purposes;

5) that the copy of the thesis submitted online to Normadec is identical to the copies handed in/sent to the members of the Thesis Board and to any other paper or digital copy deposited at the University offices, and, as a consequence, the University is absolved from any responsibility regarding errors, inaccuracy or omissions in the contents of the thesis;

6) that the contents and organization of the thesis is an original work carried out by the undersigned and does not in any way compromise the rights of third parties (Italian law, no. 633, 22nd April 1941 and subsequent integrations and modifications), including those regarding security of personal details; therefore the University is in any case absolved from any responsibility whatsoever, civil, administrative or penal, and shall be exempt from any requests or claims from third parties;

7) that the thesis is not subject to "embargo", i.e. that it is not the result of work included in the regulations governing industrial property; it was not written as part of a project financed by public or private bodies with restrictions on the diffusion of the results; is not subject to patent or protection registrations.

Date 29.03.2019

Abstract

This thesis consists of three chapters on networks in financial economics.

Chapter 1 investigates whether monetary policy shocks propagate through production networks in the Euro-area. Using the stock market reaction as a laboratory and novel tools from spatial econometrics, this is the first paper to show that between 40% and 50% of the overall European stock market reaction is due to higher-order network effects.

Chapter 2 shows that financial linkages between countries can be used to better predict sovereign CDS spreads. Modeling each sovereign's CDS spread a function of the CDS spreads of its "network neighbors" improves forecasting accuracy by 15 % to 20 %.

Chapter 3 analyzes a correlation network of *residual* stock returns of CRSP and COMPUSTAT firms. It develops a systemic measure of network centrality using principal components analysis, which has predictive power in out-of-sample tests related to the recent financial crisis.

Acknowledgements

I would like to thank my supervisor Prof. Carlo Favero, as well as Prof. Stefano Rossi and Prof. Allan Timmermann for their invaluable academic support. I am indebted to my family, who have given me the luxury of a choice in life and have supported me every step of the PhD way. And, Peter: you are the best! I am lucky to have you.

Chapter 1

Network Effects of Monetary Policy: Evidence from the European Stock Market

1. Introduction

Understanding how monetary policy shocks propagate through the economy is a particularly relevant question in the context of the European Union (EU) and the Eurozone, which involve a common monetary policy for many independent countries and the coordination of economic and fiscal policy. Recent research has shown that the input-output structure of the economy is an important propagation mechanism through which idiosyncratic shocks translate into aggregate fluctuations (Gabaix, (2011) and Acemoglu, Carvalho, Ozdaglar and Tahbaz-Salehi, (2012)). Likewise, the production network can be an important conduit for the propagation of common macroeconomic shocks, such as monetary policy shocks. For example, an expansionary monetary policy may directly increase the demand for cars. Car producers may increase their demand for engines, which on its turn may increase demand for parts and components and metal plates. Another aspect of this propagation mechanism is that the network spillover effects might trigger feedback effects to the car manufacturers, because engine or metal producers demand more vehicles to transport their products. The Eurozone provides a unique setting, because it adds another layer of complexity to the analysis: due to the close economic linkages between European countries, industries are not *only* embedded in production networks in their home countries, but they are also integrated in a Eurozone-wide production network. Consider the case in which the German automotive industry also uses French engines produced with metal plates from Finland. This example clearly demonstrates that spillovers need not be confined to the geographic borders of a country.

This paper studies higher-order demand effects of monetary policy shocks using the Eurozone production network. To test for the importance of network effects the stock market is used as a laboratory, because central banks have a direct and immediate effect on financial markets (Ehrmann and Fratzscher (2004); Bernanke and Kuttner (2005)). Changes in monetary policy are transmitted through the stock market via changes in interest rates, which then influence consumption and investment decisions. Using the property of stock markets to incorporate quickly news about monetary policy, it is possible to calculate the total effect of monetary policy shocks on asset prices in a short event window, eliminating the effect of any other confounding shocks.

Quantifying network effects following monetary policy shocks is important from a policy perspective, because it tells us whether production linkages should enter the models policy makers use. Many canonical macroeconomic models ignore linkages for tractability reasons. However, production linkages across sectors and countries introduce new channels of shocks transmission. If network effects are sizeable, then the structure of the production network becomes crucial for the determination of ECB's monetary policy and its effectiveness. **Figure 1** gives a graphical motivation by depicting the Eurozone intersectoral input-output network in 2014. The graph contains 418 vertices each of which corresponds to a country-industry observation. For every output transaction above 1 % of the total output sales of a sector, a directed link is drawn between that sector and the output customer. Shorter links depict economically stronger connections and color codes represent different countries. Several observations follow immediately from the graph. *First*, the network structure exhibits clusters, populated by industries the majority of which belongs to the same country. This indicates that home country connections are strong and dense. *Second*, clusters are "colorful", which means that a non-negligible amount of cross-border connections is economically large. *Third,* clusters are linked to each other through many long links and form *one* connected component, which is the Eurozone-wide production network. This evidence taken as a whole, points that in order to obtain the full picture of the transmission mechanism, production linkages both within and across countries have to be considered.

Studying the effects of monetary policy is a difficult task, because most of the movements in the policy rates are driven by the systematic component of monetary policy, which reacts endogenously to the state of the economy. Hence, in order to trace the causal impact of monetary policy and its network effects, it is necessary (*i*) to isolate unexpected exogeneous shifts to the monetary policy rates, and (*ii*) to generate responses of financial variables using an econometric model that effectively captures the network structure of the economy. This paper identifies monetary policy shocks using high-frequency movements in interest rate-linked securities around ECB announcement dates. It applies the methodology by Gürkaynak, Sack and Swanson (2005) and Miranda-Agrippino (2016) on euro area data to construct a novel series of monetary policy shocks, which surprise market participants, are unanticipated and account for potential information asymmetries between the market and the ECB. The innovation that this paper proposes is to identify network effects of monetary policy along two dimensions: *target*, which captures surprises about the short-term rate, and *path,* which reflects market expectations about the future stance of monetary policy. Distinguishing between the *target* and *path* shocks is particularly important, given the transparent communication policy of the ECB, which organizes press conferences since 1999 and publishes staff forecasts shortly after they are produced. Moreover, since July 2013 the ECB started actively communicating information about the future stance of monetary policy in the form of *forward guidance.* The *path* factor responds closely to important *forward guidance* announcements, which makes it a useful tool to evaluate more holistically monetary policy announcement decisions. The two informationally-robust factors are used as exogeneous instruments in a spatial autoregressive framework (SAR) to decompose the total effect of the monetary policy shocks into a *direct effect* and a *network effect*.

Spatial econometrics typically identifies spillovers across units or "neighbors", connected to each other through geographic proximity, family and friendship linkages, infrastructure etc. The existence, direction and magnitude of the effect is governed by the spatial weight matrix. In this paper, the spatial weight matrix identifies production linkages between 418 country-industry observations in the Eurozone. Data is organized as annual time series of input-output (IO) tables and collected from the World Input-Output Database (WIOD) (Timmer, Dietzenbacher, Los, Stehrer, & de Vries, 2015). To the best of my knowledge, this is the first paper to apply the data from the World Input-Output Project to study the mechanisms of monetary policy transmission in the European production network. Finally, daily stock returns for individual firms are merged with industry-level IO tables.

The main result of the paper is that network effects account for between 40% to 50% of the total stock market reaction following a monetary policy shock. Importantly, 45 % of the network effect is driven by cross-country industry linkages, which demonstrates that the *Eurozone-wide* production network plays a crucial role for the transmission of monetary policy. In terms of economic magnitudes, a 100 bps higher than expected policy rate (*target*) reduces industry returns by 3.50 percentage points. On the other hand, the effect and importance of the *path* factor is more complex. During the period before the European Sovereign Debt Crisis and the removal of the zero-lower bound (ZLB), the main factor driving asset prices is *target*. The effect of *path* during this period is small and positive. These responses are in line with the interpretation that the ECB communicates good news about the future economic outlook and tightens monetary policy to counteract its impact on the macroeconomy. However, following the removal of the ZLB, when policy changes in the short-term rate are no longer effective, monetary policy announcements are exclusively reflected in the *path* factor. Moreover, the network effect of *path* after the ZLB accounts for as much as 70 % of the total stock market response. This is an important insight, which suggests that long-term monetary policy shocks are transmitted much more powerfully through the network than short-term surprises.

Theoretical work by Acemoglu et al. (2015) predicts that demand shocks, such as monetary policy shocks, travel *only* upstream in production networks *i.e* from customers to their suppliers. The result is a consequence of Cobb-Douglas production functions and consumer preferences, which are assumptions that might not hold in the real economy. This paper provides empirical evidence that monetary policy shocks travel *both* upstream and downstream. However, the upstream transmission is much more powerful than the downstream one. The intuition is that since downstream industries sell most of their output directly to consumers, their response to monetary policy shocks is mainly driven by *direct effects*. On the other hand, the output of upstream industries is mainly used as an intermediate input in the production of downstream industries. In a sense, upstream industries are located further away from the shock's origin and for them most of the impact is coming from *network effects* transmitted through the production linkages.

The presence of a common monetary policy shock introduces comovement of returns on the market. In the standard capital asset pricing (CAPM) framework, the beta coefficient reflects exposure to aggregate market risk. This means that if we were to run the CAPM on days of ECB's announcements, the beta should give the contribution of the "market" for the transmission of monetary policy shocks. Another way to think about the CAPM model is as a special case of the SAR model, where the weights matrix W is determined by market capitalization. The paper provides evidence that the SAR model strongly outperforms the CAPM, which is only able to explain 1.17% of the variation in the data and attributes roughly 7% to network effects. The distinct feature between the two models is the treatment of risk: whereas the CAPM considers the "network factor" as *market-specific*, the network model uses *industry-specific* weights. These results re-emphasize the important role that sector-specific production connections play in determining the transmission of monetary policy shocks.

Next, the paper explores heterogeneous industry effects based on specific characteristics of the network. The starting point of the analysis is the empirical prediction of the SAR model, which shows that higher connectivity indicates higher *relative* importance of network effects. The intuition is that central suppliers are more exposed to shocks passing through the intermediates production network, because they are connected to many customers, which transmit the shock upstream to them. Specifically, the predicted proportion of network effects for a highly connected industry is 34 % higher than for the average industry. Interestingly, although the relative importance of network effects increases with centrality, the *absolute* magnitude of the total effect decreases. To investigate this result, the paper uses data on *global* production networks and documents an important empirical relationship: industries that are highly connected in the Eurozone and also very likely to be important global players. This evidence suggests that part of the reason why central industries react less to the monetary policy shock is that they are less exposed to the shock, because a considerable fraction of their demand is located outside of the Eurozone. The result cannot be explained by lower riskiness of central industries and is robust to alternative definitions of centrality.

The results of the SAR model have important policy implications. First, the model predicts which industries are most "influential" for the transmission of shocks upstream. These are industries, which depend largely on Eurozone demand and are most exposed to the shock, but at the same time are connected enough to transmit the shock to their suppliers upstream. Network effects are three times higher for suppliers connected directly to at least one influential customer than for suppliers that are not connected to an influential industry. Second, drawing on recent asset pricing work by Herskovic (2018), it is possible to show that decreasing network concentration increases aggregate spillover effects. Theoretically, very high levels of concentration signify that a very small subset of industries plays a disproportionately important role in the economy. The important insight here is that in the context of the Eurozone, the concentration factor is closely related to economic convergence. The more low-income countries catch up with high-income countries, the tighter they get embedded in the production network and the stronger the spillovers from monetary policy become.

The results of the paper survive a comprehensive set of robustness checks. Network effects of monetary policy are not subsumed by other mechanisms previously studied in the literature such as the interest rate channel, credit channel and the expectations channel. The result is robust to the inclusion of a battery of industry characteristics such as: exchange rate sensitivity, product demand elasticity, capital intensity, inventory investment, production structure, mean debt-to-capital ratio, firm size composition, industrial output and industry, country and year fixed effects.

Using the identification-through-heteroskedasticity estimator by Rigobon (2003), the paper demonstrates that the results are not biased by the presence of omitted variables or simultaneous movements between policy rates and asset prices. Furthermore, simulation results show that the monetary policy factors are precisely estimated and not contaminated by large measurement errors.

Results are robust to different specifications of the spatial weights matrix. Imposing zero-diagonal connectivity matrix, it is possible to show that network effects are not driven by linkages to one's own industry. Given the importance of global connections, one concern could be that the relative importance of home connections is overestimated. However, normalizing by total supplier sales globally does not change neither the policy shocks estimates nor the proportion of network effects. Finally, the large magnitude of network effects following monetary policy shocks might appear surprising giving that most industries are not directly connected to each other. To address the concern that the model attributes the effects mechanically, the paper shows that a random network with the same sparsity as the Eurozone input-output network generates network effects, which are five times smaller than the baseline estimate. This evidence stresses that the large network effects are driven by higher-order network connections, which cannot be captured by sparsity alone.

Finally, in order to prove a clear structural interpretation, the paper develops a simple production model with intermediate inputs in the spirit of Acemoglu et al. (2012) and Weber and Ozdagli (2016). The theoretial framework shows that the

SAR model emerges as an equilibrium outcome and arrives precisely at the input-output production network as the appropriate weighting matrix.

The paper is closest to Weber and Ozdagli (2017), who use spatial models to study how monetary policy shocks propagate through the US input-output production network. Building on their work, this paper makes the following three contributions to the literature. *First*, the paper introduces an additional dimension of monetary policy in the network analysis. The *path* factor captures expectations about the future monetary policy and allows us to study network effects of long-term monetary policy shocks. *Second*, it applies the model by Weber and Ozdagli (2017) in a multi-country multi-sector framework (the Eurozone) and identifies production linkages using a new dataset from WIOT. This is the first paper to provide evidence on the transmission of monetary policy shocks through production networks in the Eurozone. An important finding of the paper is that spillover effects are not driven by characteristics of the countries, but by the complex structure of European network. *Third*, it relates the heterogeneous spillover responses to the degree of connectivity and importance of a sector in the production network. Moreover, the paper makes the important observation that industries that are central in the Eurozone network tend to be important global players as well. Therefore, it opens new avenues for future research, investigating the importance of global supply chains for the transmission of economic shocks.

2. Literature Review

A fast-growing literature in macroeconomics argues that significant aggregate fluctuations may originate from microeconomic shocks to firms or sectors. The traditional macroeconomic argument has been that at high levels of disaggregation in the economy, individual shocks average out and their idiosyncratic effect will be negligible (Lucas, (1977)). However, one problem with this argument is that it ignores the fact that firms and industries do not function in isolation but are imbedded in intricate production networks. Consequently, the "diversification argument" discards the possibility that network interconnections have an impact on aggregate volatility or asset prices. Recent work by Gabaix (2012) and Acemoglu et al. (2012) shows that when firm size or sectoral interconnections are characterized by a fat-tailed distribution, the law of large numbers fails, and aggregate output does not concentrate around a constant value. Acemoglu, Akcigit and Kerr (2015) and Barrot and Sauvagnat (2016) extend their analysis by showing how federal spending, trade, technology and knowledge shocks are propagated through the production network. Other examples of recent contributions in this field are: Acemoglu, Ozdglar and Tahbez-Salehi (2017), Carvalho (2014), Carvalho and Gabaix (2013), Atalay (2017). On the asset pricing side, Kelly, Lustig and Van Nieuwerburgh (2013) study the relationship between firm-size distribution and stock returns volatility, whereas Herskovic (2018) shows that the production network is a source of systematic risk reflected in asset prices. However, a relatively understudied area in the literature is the amplification mechanism of *common macroeconomic* shocks. This paper fills this gap by investigating how monetary policy shocks propagate through production networks.

Second, the paper relates to a nascent literature on spatial econometrics in economics. Traditionally, spatial autoregressive models been extensively applied to research questions in regional science, economic geography and international trade[1]. SAR models in finance are not widespread, but have began to gather momentum in the recent years. Fernandez (2011) derives a spatial version of the CAPM model and uses the results to perform value-at-risk simulations. Blasques

---

[1] For early reviews relate to the work of Cliff and Ord (1981); Upton and Fingleton (1985). More recent contributions related to theory and estimation methods include, amond others, Anselin and Bera (1998); Anselin, Bera, Florax, and Yoon (1996); Prucha and Kelejian ((1998), (1999), (2004), (2006), (2010); LeSage and Pace (2006), (2009); Arbia ((2006), (2012)); Lee (2004); Bivand, Pebesma and Gómez-Rubio (2013). On the empirical side, SAR models have been used to study regional unemployment differentials (Elhorst, 2003); property tax rates (Elhorst & Allers, 2005), inflation (Elhorst, Heijnen, Samarina, & Jacobs, 2017); air quality improvements (Kim, Phipps, & Anselin, 2003); European regional per capita GDP (Le Gallo & Ertur, 2003); property taxes (Bordignon, Cerniglia, & Revelli, 2003); foreign direct invesment (Bloningen, Davies, Waddell, & Naughton, 2007).

et al. (2016) and Eder and Keiler (2015) study systemic risk and spillovers in the CDS market.

Third, in using market-based monetary surprises the paper connects to a large literature pioneered by Cook and Hahn (1989) and Kuttner (2001). Important contributions include, but are not limited to, Gürkaynak, Sack and Swanson (2005); Bernanke and Kuttner (2005); Ehrmann and Fratzscher (2007), Brandt, Buncic and Turunen (2010); Barakchian and Crowe (2013); Jardet and Monks (2014); Gertler and Karadi (2015); Gilchrist, López-Salido and Zakrajšek (2015). The general message of these studies is that unexpected changes in the monetary policy rate are associated with strong and instantaneous asset price reaction. This paper focuses on a variant of Gürkaynak, Sack and Swanson (2005) applied to forward OIS rates and computes two factors (target and path), which effectively describe the impact of ECB's monetary policy. Using recent advances by Miranda-Agrippino (2016), a cleaner version of the two factors is constructed, which accounts for autocorrelation and informational asymmetries.

Fourth, this work builds on the research studying the transmission mechanisms of monetary policy. Beyond the standard interest rate channel, previous papers have also investigated the credit channel (Bernanke & Gertler, 1995) the exchange rate channel (Taylor, 2001), asset price channel, bank lending channel (Kashyap & Stein, 2000) and risk-taking channel (Borio & Zhu, 2012) among others. This paper focuses on a new distinct channel: higher-order demand effects due to linkages in intermediates production. It is important to emphasize that these channels are not mutually exclusive and can work in parallel to each other.

3. Empirical Framework

3.1 Economic Underpinnings

Theoretical work by Acemoglu et al. (2012), Calvalho (2014), and Weber and Ozdagli (2017), predicts that demand-side shocks, such as monetary policy shocks, propagate only upstream. This means that upstream suppliers are affected more strongly by demand shocks than downstream customers. This pattern results from adjustments in the production levels, and, consequently input demands. For example, a contractionary monetary policy shock decreases demand for goods by the end-customer. Firms optimally respond by decreasing their purchases of intermediate inputs. The producers of intermediate inputs decrease their production to match the decreased demand for their goods, which translates into decreased demand for goods produced further upstream. Due to non-linearities in the production process, the shock could travel back to industries initially hit by the shock. Hence, recessionary monetary policy shocks not only *directly* affect the demand for goods of industries selling to consumers but can also result in *indirect* higher-order demand effects through its impact on intermediates consumption. However, this theoretical result is a consequence of Cobb-Douglas production functions and preferences and might not hold in reality. Therefore, whether monetary policy shocks propagate upstream and/or downstream is an open *empirical* question in the literature.

In production models with intermediate inputs the output of one sector is used as an input to another sector. The presence of production linkages introduces simultaneity in the responses of industries to common monetary policy shocks. Intuitively, once dependence relations between N number of industries are allowed, there are N^2 relations that could arise. It is easy to see that if $N > T$, as is the case here, this results in a system with many more relation-specific parameters than observations. The solution to this over-parametrization problem, proposed originally by Ord (1975), is to impose a parsimonious structure on the dependence relation. This gives rise to a data-generating process known as a *spatial autoregressive process*.

3.2 Spatial Autoregressions

The spatial autoregressive model (SAR) is given by the following empirical specification:

$$ret_{ijt} = \beta_0 + \rho W_{ij} ret_{ijt} + \beta_1 Target_t + \beta_2 Path_t + \lambda_i + \delta_j + \mu_{year} + \gamma controls_{ijt} + \varepsilon_{ijt} \qquad (1)$$

where ret_{ijt} is the event return in the interval $[t-1, t+1]$ around announcement (event) date t of industry i in country j, β_0 is an intercept term, $Target_t$ and $Path_t$ are the two monetary policy surprise factors, λ_i is an industry fixed effect and δ_j is a country fixed effect[2]. To control for period-specific unobservable market-wide shocks year-dummies μ_{year} are included. Finally, $controls_{ijt}$ is a vector of country-industry specific characteristics.

ρ is the spatial parameter and the main object of interest in this model. It indicates the relevance of network connectedness for the propagation of monetary policy shocks and is interpreted as a measure for network spillover effects. Testing for the presence of spillovers is tantamount to the following hypothesis:

$$H_0: \rho = 0$$
$$H_1: \rho \neq 0$$

If the null hypothesis is rejected, values of $\rho > 0$ indicate that the shocks are propagated through the network; if $\rho < 0$ the shock is absorbed by the system[3].

W is a row-normalized spatial weights matrix, whose entries give production linkages between industries[4]. The model assumes that W is exogenously given and stable over time. Rows correspond to supplier-industries and columns correspond to customer industries. Then, the ij^{th} -row of W gives the fraction of output of industry i in country j used in the production in all country-industries ij. Output could be sold to both an industry's home country or to other countries in the Eurozone market. Entries on the main diagonal of W give sales to own industry (e.g. tire producers selling to car manufacturers). For example, if entry $w_{11,21} > 0$,

[2] Since the panel is short and N is large, a full battery of fixed effects coefficients cannot be estimated consistently in a MLE framework due to the incidental parameter problem (Anselin, (2001)). For this reason, all specifications include only industry fixed effects (38 dummies) and country fixed effects (11 dummies)

[3] If the spatial parameter is statistically indistinguishable from zero, then the SAR model collapses to the standard Ordinary Least Squares (OLS) model.

[4] In the traditional application of SAR models, the "spatial" structure refers to geographic distance. Here, instead of modelling explicitly geographic borders between countries, *economic* distance is considered.

this means that industry 1 in country 1 is a supplier to industry 2 in country 1. In the spatial econometrics jargon, the two industries are said to be neighbors. Note that W gives upstream propagation (from customers to suppliers), whereas its transpose W^T gives downstream effects (from suppliers to customers). The empirical counterpart of W is the WIOT input-output matrix, which is described in Section 3.6.

The term $W_{ij,ij} ret_{ijt}$ is constructed as a linear combination of neighboring values to each observation. It is important to note that these entries are observation-specific and capture the heterogeneity of production linkages. In contrast, the shock variables *target* and *path* are global variables, whose entries are common to the entire market.

Stacking in vector form and solving for *ret*, the result is a reduced-form equation:

$$ret_t = (\mathbb{I}_N - \rho W)^{-1}(\alpha \boldsymbol{\iota} + \beta_1 Target_t + \beta_2 Path_t + \varepsilon_t) \qquad (2)$$

$$\varepsilon_t \sim N(0, \sigma^2 \mathbb{I}_N)^5$$

$Wret$ is called a spatial lag and $(\mathbb{I}_N - \rho W)^{-1}$ is called a spatial multiplier. Borrowing from the time-series literature, W is the spatial analog if the lag operator **L**. Whereas $\boldsymbol{L}.ret$ measures the potential spillover from time time *t-1* to *t*, $\boldsymbol{W}ret$ specifies spillovers from one country-industry to another. *Appendix A* contains a stylized example, which explains how to construct the spatial lag.

3.3 Direct and Indirect Network Connections

For a row-normalized matrix W and for values $|\rho| < 1$, the spatial multiplier exists, the model is uniquely specified[6] and the infinite series converges:

[5] The term $N(0, \sigma^2 \mathbb{I}_N)$ denotes a zero mean disturbance process with constant variance σ^2 and zero covariance between the observations. This results in a diagonal variance-covariance matrix $\sigma^2 \mathbb{I}_N$ with $\mathbb{I}_N$ representing an $N \times N$-dimensional identity matrix. The results are robust to relaxing this assumption.

[6] Generally, it is possible to show that the matrix $\mathbb{I}_n - \rho W$ is nonsingular for all values of the parameter ρ in the space $(-\frac{1}{\lambda_{max}}, \frac{1}{\lambda_{max}})$, where λ_{max} is the largest eigenvalue of the matrix W (see

$$V(W) = (\mathbb{I}_N - \rho W)^{-1} = \sum_{q=0}^{\infty} \rho^q W^q = \mathbb{I}_N + \rho W + \rho^2 W^2 + \cdots \qquad (3)$$

The q^{th} power of the matrix W collects the total number of links both direct and indirect in the entire network starting from node i and ending in node j. These powers correspond to observations themselves (zero-order), immediate neighbors (first-order), neighbors of neighbors (second-order) etc. One challenge in understanding the mechanism of propagation is to capture the importance of indirect links.

<u>**Example 1**</u>: Consider the following simplified example with three industries. Connections are given by the matrix A:

$$A = \begin{pmatrix} 0 & 5 & 0 \\ 0 & 1 & 4 \\ 0 & 0 & 0 \end{pmatrix}$$

The matrix reads like this: industry 1 sells 5 units of output to industry 2; industry 2 sells 1 unit to itself and 4 units to industry 3; industry 3 does not sell to other industries. First-order neighbors are: $\boxed{1} \rightarrow \boxed{2}$; $\boxed{2} \rightarrow \boxed{2}$ and $\boxed{3}$. The *second-order* neighbors are neighbors to the *first-order* neighbors. Hence, the *second-order* neighbors to industry 1 are industries 2 and 3: : $\boxed{1} \rightarrow \boxed{2} \rightarrow \boxed{2}$ and $\boxed{3}$.

$$W^1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0.20 & 0.80 \\ 0 & 0 & 1 \end{pmatrix}, W^2 = \begin{pmatrix} 0 & 0.20 & 0.80 \\ 0 & 0.04 & 0.16 \\ 0 & 0 & 0 \end{pmatrix}$$

Intuitively, the condition $|\rho| < 0$ captures the idea that distant connections are less important than direct connections. In this sense the spatial parameter serves as a discount factor, which assigns decreasing influence to higher-order neighbors, where the decay declines geometrically as the order increases. If $\rho = 0.5$, then $\rho^2 = 0.25$, $\rho^3 = 0.125$ etc. In the context of a production network, the magnitude of ρ

Kelejian and Prucha (1998), (1999)). For a row-normalized matrix the condition is trivially satisfied since $\lambda_{max} = 1$ and so, for values $|\rho| < 1$ the model is well-defined.

reflects the degree to which shocks are transmitted locally or to the entire network. Small values for ρ weight the local structure more, while large values take into consideration the position of an industry in the production network as a whole.

3.4 Parameters Interpretation

Conditional on orthogonality of regressors, parameter estimates in a linear regression have straightforward interpretation as partial derivatives of the dependent variable with respect to the independent variables. This arises because of assumed linearity and independence of the observations in the model. When the model contains spatial lags of the dependent variable, interpretation of the parameters becomes more complicated, because the model incorporates a richer information set. In a spatial context, a change in any given explanatory variable will have an impact on the stock returns of the industry itself (direct effect) and potentially an impact of the stock returns of other related industries (indirect effect). For this reason, spatial autoregressive models require special attention to the interpretation of the parameters (LeSage & Pace (2009)).

To see more clearly the complication of parameter interpretation, define

$$S(W) = ((\mathbb{I}_N - \rho W)^{-1} - \mathbb{I}_N)\, \mathbb{I}_N \beta \tag{4}$$

Collect covariates in X and estimates in β and rewrite the model as:

$$ret_t = \beta X_t + S(W)X_t + V(W)\varepsilon_t \tag{5}$$

Consider again the previous example with two industries and two countries at a particular point in time t. The data-generating process can be expanded to the following:

$$\begin{pmatrix} ret_{1t} \\ ret_{2t} \\ ret_{3t} \end{pmatrix} = \mathbb{I}_N \beta \times \begin{pmatrix} X_t \\ X_t \\ X_t \end{pmatrix} + \begin{pmatrix} S(W)_{11} & S(W)_{12} & S(W)_{13} \\ S(W)_{21} & S(W)_{22} & S(W)_{23} \\ S(W)_{31} & S(W)_{32} & S(W)_{33} \end{pmatrix} \times \begin{pmatrix} X_t \\ X_t \\ X_t \end{pmatrix} + V(W)\varepsilon_t$$

with $S(W)_{ij}$ indicates the ij^{th} element of the matrix $S(W)$. Focusing on industry 1, the following obtains:

$$ret_{1t} = \underbrace{\beta X_t}_{Direct\ Effect} + \underbrace{S(W)_{11}X_t}_{Feedback\ Effect} + \underbrace{S(W)_{12}X_t + S(W)_{13}X_t}_{Indirect\ Effect} + V(W)_1\varepsilon_t$$

with $V(W)_1$ referring to the first row of the matrix $V(W)$. The return of industry 1 depends the *direct effect* of the shock as if network connections were severed (β) and two network effects: a *feedback effect*[7] due to the shock hitting industry 1 and coming back to it through the network ($S(W)_{11}$) and an *indirect effect* ($S(W)_{12} + S(W)_{13}$) due to spillovers from intermediates production. For example, $S(W)_{12}$ denotes the response of industry 1's returns to the change in the return of industry 2 due to the monetary policy shock.

Following LeSage and Pace (2006), it is possible to define four scalars, which summarize the *total, direct, feedback* and *indirect* effects:

i. *Average Direct Impact*: β. This is the effect as if all network connections were shut down.
ii. *Average Indirect Effect*: $\frac{1}{N}\iota_N'[S(W) - tr(S(W))]$. The sum across the i^{th} row of $S(W)$ minus the i^{th} entry on the diagonal gives the impact on an individual industry resulting from changing the monetary policy shock by the same amount across all N industries.
iii. *Average Feedback Effect*: $\frac{1}{N}\iota_N' tr(S(W))$. The average of the diagonal elements of $S(W)$ gives the effect of an industry's response travelling back to itself through the network.
iv. *Average Total Effect*: the sum of average direct, feedback and indirect effects.

Therefore, the response of an industry's stock return to the monetary policy shock is determined by the input-output matrix W through its effect on demand for intermediate outputs, the spatial autoregressive parameter ρ, which denotes the strength of the network spillover effects, and the parameter β.

[7] It has to be noted that first-order feedback effects arise because industries sell output to themselves (e.g. 1→1); higher-order feedback effects arise because of impacts passing through "neighboring" industries and coming back to the industry itself: e.g. 1→2→1 (second-order) and 1→2→3→1 (third-order).

<u>Example 1:Upstream vs. Downstream</u> Theory predicts that the relative importance of indirect effects increases upstream. To understand this mechanism better, it is useful to revisit Example 1. In this example, industry 1 is an *upstream* industry because it supplies directly to industry 2 and indirectly to industry 3, but neither of these industries buys from industry 2. Industry 2 can be classified as *midstream* because it supplies to itself and industry 3, but not to industry 1. Finally, industry 3 is *downstream* because it sells only to itself and receives inputs from all the other industries. Suppose that 100 bps higher than expected policy rate reduces returns by 1 % ($\beta = -1$) and that $\rho = 0.5$. Then, the total effect is given by:

$$\begin{pmatrix} upstream \\ midstream \\ downstream \end{pmatrix} = \underbrace{\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}}_{Direct\, Effect} + \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & -0.11 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{Feedback\, Effect} + \underbrace{\begin{pmatrix} 0 & -0.55 & -0.22 \\ 0 & 0 & -0.44 \\ 0 & 0 & 0 \end{pmatrix}}_{Indirect\, Effect}$$

Since the downstream industry does not sell to any industry, direct effects account for 100 % of its the total effect. For midstream industries, indirect effects account for 28 % and a small impact is attributed to feedback effects (7 %). For comparison, the proportion of indirect effects for upstream industries is considerably higher and equals 44 %.

3.5 Monetary Policy Shocks

This paper defines monetary policy shocks as exogenous changes in the policy rate, which (*i*) surprise market participants, (*ii*) are unanticipated and (*iii*) do not reflect the central bank's systematic response to its own assessment of the economic outlook. To this end, a novel dataset of euro area monetary policy shocks is constructed. This dataset contains 198 policy announcements from 2001 to 2017[8]. Operationally, the identification proceeds in three steps.

[8] Interest rates announcements for the years 1999-2000 are not included due to the high volatility of the OIS rates during this period. This number includes scheduled announcements and excludes one unscheduled event on 8th October 2008, when the ECB announced a 50bps cut in its main refinancing rate. This was a joint announcement by several central banks and hence, ECB's surprise component cannot be precisely purged out from the data.

(i) Monetary surprises. Following the methodology by Gürkaynak, Sack and Swanson (2005), monetary surprises are constructed using changes in a set of forward Overnight Index Swaps (OIS) [9] rates in a one-day window bracketing ECB's announcements. The argument in favor of using these rates is that they reflect *expectations* about the future monetary policy stance in the Euro Area[10] and given that surprises are computed within a sufficiently narrow time interval, they can be thought of as a measure of monetary policy shocks. The advantage of using many forward rates is that the entire term structure of expectations can be uncovered. Furthermore, OIS rates are characterized by low levels of credit risk, which means that rates are not likely to be affected by risk premia, even at maturities longer than 1 month.

Here, seven forward rates are considered: $[f_1^0, f_1^1, f_3^3, f_3^6, f_3^9, f_6^{12}, f_6^{18}]$ where f_m^j gives the m month j months ahead OIS rate. The surprise componenet for each forward rate is computed as:

$$\Delta f_{m,t}^j = \left(f_{m,t+1}^j - f_{m,t-1}^j\right) \tag{6}$$

where t is the time when the ECB issues the announcement, $f_{m,t+1}^j$ is the rate 1 day after t and $f_{m,t-1}^j$ is the rate 1 day before t. The surprise series is used to extract factors with principle componenets analysis (PCA).

The Path Factor. The standard assumption in the existing literature has been that the effect of monetary policy announcements on asset prices is adequately captured by the first principal component, which is typically related to changes in the central bank's short-term target rate. However, nowadays ECB's

---

[9] An OIS contract is an interest rate swap derived from exchanging the daily overnight interest rate for a fixed interest rate. In the case of euro OIS contracts, the floating rate is the euro overnight index average (EONIA) rate. The EONIA rate is calculated as a weighted average of overnight interest rates on unsecured operations for a panel of banks. For more details on the euro OIS markets, the reader is referred to the review by Remolona & Wooldridge (2003).

[10] For an overview of studies that document the close relationship between the EONIA rate and ECB's main refinancing rate consider (ECB, 2008)

announcements simultaneously convey information about current policy rates and the central bank's assessment of the future stance of monetary policy and the economic outlook. These communications contain valuable information because they affect consumption and investment decisions, which in their turn are dependent on expectations about future central bank's decisions (Blattner et al., 2008). To corroborate this intuition, note that actual interest rate changes were announced in only 33 out of the total 198 cases. Treating monetary policy as a 0 bps surprise change in the current rate in the other 165 cases would be largely missing the whole story.

The novelty that this paper proposes is to investigate network effects along two dimensions of monetary policy: a short-term and long-term dimension. To do so, two factors are extracted, which are orthogonal by construction and explain a maximal fraction of the variance (approximately 82%)[11]. In order to provide a structural interpretation of the two factors a factor rotation is applied as in Gürkaynak, Sack and Swanson (2005) such that the two factors remain orthogonal and that the second factor does not have an impact on the one-month OIS rate. The first factor is interpreted as the surprise rate of the ECB announcement and is called the *target factor*. The second factor can be regarded as all other information in the central bank's announcements that moves market expectations about the future path of monetary policy without changing the current rate and is denoted as the *path factor*. In this respect, the *path factor* bears a close resemblance to *forward guidance* as in Swanson (2017) and to the information shock as in Jarociński and Karadi (2018). To further faciliate interpretation and comparison of the two factors, the *target* factor is rescaled to move *one-to-one* with the one-

[11] The two factors closely resemble the "level" and "slope" factors from the traditional Nelson-Siegel representation. However, the difference with respect to that literature is that the sample here focuses exclusively on the tight time interval around ECB's announcements.

As a robustness I estimate the optimal number of factors using the reduced rank test by Cragg & Donald (1997) and confirm that two factors are necessary to describe the data. Refer to Jardet & Monks (2014) and Ellen, Jansen, & Midthjell (2017) for recent applications. The appendix of GSS2005 provides details on the estimation procedure for the test statistic.

month OIS rate and the *path* factor is rescaled so that both factors have the same effect on the 3-month 9-month ahead OIS rate, which is approximately 57 bps. The exact rotation procedure is described in detail in *Appendix B: Monetary Policy Shocks*.

Table 1 reports the rotated factor loadings from the estimation. The effect of the *target* factor is strongest for shortest maturities and decreases for longer maturities. The downward slope is consistent with theory: surprises in the key policy rate affect more the short end of the yield curve and die out for longer maturities. The loadings on the *path* factor on the other hand exhibit a concave shape, with the biggest loading on the f_3^9 rate.

(ii) Past monetary surprises and (iii) economic outlook. Identification of the path factor relies on the assumption that the economic forecasts of market participants and the central bank coincide. However, in reality this need not be the case due to information frictions. Previous work by Miranda-Agrippino (2016) and Ramey (2016) shows that monetary policy surprises can reflect anticipatory effects because the market is not able to fully account for the systematic aspect of monetary policy when the shock occurs. To account for this, the surprise factors *target* and *path* are regressed onto their lags to control for sluggish absorption of information and onto forecasts and forecasts revisions[12] about real GDP growth, inflation and unemployment rate to control for the ECB's private information:

$$\begin{aligned} surprise\ factor_t &= \alpha_0 + \sum_{i=1}^{p} \alpha_i surprise\ factor_{t-i} + \sum_{j=-1}^{3} \theta_j F_t^{CB} x_{q+j} \\ &+ \sum_{j=-1}^{2} \vartheta_j \left[F_t^{CB} x_{q+j} - F_{t-1}^{CB} x_{q+j} \right] + \varepsilon_t \quad (7) \end{aligned}$$

[12] Each quarter, The ECB publishes a Survey of Professional Forecasters (SPF). This is a survey of expectations for the rates of inflation, real GDP growth and unemployment in the euro area for several horizons, together with a quantitative assessment of the uncertainty surrounding them.

$F_t^{CB} x_{q+j}$ denotes the survey forecast for quarter $q + j$ made at time t and q is the current quarter. $[F_t^{CB} x_{q+j} - F_{t-1}^{CB} x_{q+j}]$ is the forecast revision between two announcements dates and x_q contains output growth, inflation and unemployment. The residuals from this regressions are used as clean measures of the monetary policy shocks.

Figure 2 plots *target* and *path* on days of monetary policy announcements. Both factors revolve around their mean value of 0 with occasional spikes, particularly in the period 2008-2012. The *target* factor is less volatile with a standard deviation of 4.61%, compared to 17.46% for the *path* factor. A natural interpretation of this evidence is that changes in the policy rate have a clear meaning to market participants, whereas announcements about the future policy and outlook are typically associated with more uncertainty about how they are interpreted.

3.6 Data

Production linkages are identified using data from the World Input-Output Database (WIOD), which provides annual time-series of world input-output tables (WIOT) from 1995 onwards (Timmer, Dietzenbacher, Los, Stehrer, & de Vries, 2015). Although the database has been extensively used in the literature on international trade[13], to the best of my knowledge, this is the first application to a problem in finance.

Countries included in the sample are Austria, Belgium, Germany, France, Spain, Greece, Finland, Ireland, Italy, the Netherlands and Portugal. In each country, the same set of 38 industries is observed. **Table 2** gives an illustrative example with 2 countries and 2 industries. The columns of the WIOT provide information about the supply of inputs used in the production process of a given

[13] Important contributions include but are not limited to: (Adao, Costinor, & Donaldson, 2017); (Fajgelbaum & Khandelwal, 2016); (Timmer, Erumban, Los, Stehrer, & de Vries, 2014); (Costinor & Rodriegez-Clare, 2014); (Ottaviano, Pessoa, Sampson, & Van Reenen, 2014); (Johnson, 2014); (Koopman, Wang, & Wei, 2014); (Los, Timmer, & de Vries, 2015);

industry. The rows provide information about the final usage of outputs. Products can be used either by other industries as intermediate inputs or by consumers and the government (final consumption) or by firms (stocks and gross fixed capital formation).

As a next step, firm-level returns from the entire universe of Eurozone companies are matched to country-industry observations using a concordance table designed specifically for this project. Then, for each country-industry observation an equally-weighted return index is constructed. When time-series data is poor or not available, returns are proxied by an *aggregate* industry index available from Datastream[14]. Finally, data is merged with monetary policy announcement dates. The result of the matching procedure is $38 \times 11 = 418$ country-industry pairs, which I observe over 198 announcement dates from January 2001 to June 2017. *Appendix C: Data and Network Construction* contains more details.

3.7 Estimation and Identification Assumptions

The spatial lag introduces cross-sectional correlation in the error terms and renders OLS estimates inconsistent. Hence, the model is estimated using Maximum Likelihood (MLE) and standard errors for direct, indirect, feedback and total effects are produced using 5,000 simulated parameter values. The identification of the MLE estimates relies on three important assumptions.

First, the production network is determined by technology and not by self-selection (Carvalho and Voigtländer, (2015)). This argument is supported by **Figure 3**, which shows that the distribution of output shares for all sectors is stable over time[15]. Furthermore, during the estimation period 2001-2017 there were no big technological breaks (e.g. the invention of the container, computerization etc.), which could confound the results. All baseline regressions

[14] In principle, such an aggregate return index can be used for all industries. However, these indices are based on large-cap companies and, hence, it would introduce a measurement error in the data.

[15] The figure plots standard deviations, instead of mean shares, because due to the row-normalization the mean share for each sector is by construction $\frac{1}{418} = 0.0024$.

use a W matrix from the year 2000 (W^{2000}), which is entirely predetermined with respect to the sample and addresses the concern that an omitted variable affects simultaneously the dependent variable and the formation of links. Furthermore, W^{2000} is a valid instrument for the production network during 2001-2017 with an adjusted-R^2 of above 80%.

Second, $Wret_t$ is *weakly exogenous*. This means that a significant portion of "neighboring" industries influence another industry in the aggregate, even if the bilateral influence of each individual industry is low (Lee (2002)). This "smallness" condition ensures that all industries receive small enough weights, so that there is no dominant industry structure. Since $Wret_t$ is constructed using many country-industry pairs, the assumption is easily satisfied.

Third, the ECB policy rate does not react to a particular country-industry. Given that each industry is small enough with respect to the market, it is unlikely that the ECB targets a particular sector and so, it is possible to assume that simultaneity does not bias the results.

4. Empirical Results

4.1 Summary Statistics and Spatial Dependence

To get a preliminary idea of the importance of production links for the transmission of policy shocks, it is useful to plot the spatial lag $Wret_t$ against contemporaneous industry returns, a graph also known as Moran's Scatterplot (1950) in the spatial econometrics literature (**Figure 4**). Points appear to be roughly distributed about a line, which suggests that observations with similar values are located "close" to each other. The intuition here is that if the matrix W contained no relevant information, observations would be randomly dispersed.

Table 3 discusses in detail the number and strength of network connections. The average supplier in the sample is connected to 410.81 customers (**Panel A**), which means that there are no isolated industries in the sample. The number of linkages to very small customers (less than 1% of sales) accounts for 393.92 out the total number of non-zero links. The number of links to small customers (1-5 % of sales) is 12.98, to medium-sized ones (5-20 % of sales) is 3.18 and to large customers

(≥ 20 % of sales) is 0.73 respectively. In terms of strength of the connection, **Figure 5** shows that together very small, small and medium-sized customers account for as much as 74.49 % of the total sales, whereas large customers buy 25.51 %. These numbers show that, on average, industries have many connections with relatively small weights, which justifies the "smallness" assumption.

Industries are embedded in two types of networks: first, within the home country ("home") and, second, outside of the home country but within the euro zone ("foreign") **Panel B** shows the decomposition of links between "home" and "foreign". The average supplier has 16.89 links worth above 1 % of sales, out of which 2.58 are to foreign industries. Mean foreign sales account for roughly 15 % of the total volume. These numbers might appear small, but the distribution is heavily skewed. For example, only 30% of the industries account for as much as 90% of all the foreign links above 1 %. Industries such as mining, manufacturing, transportation, computers & electronics have a large number of foreign connections, whereas industries such as agriculture, utilities and public administration tend to have few links. Moreover, **Figure 6** plots the number of home links against the number of foreign links for each supplier and finds that there is a negative relationship between the two measures. This means that those industries, which are well-connected at home are not necessarily those that are also highly connected abroad i.e. some industries are *local* and others are *international*.

Panel C of **Table 3** shows that the average pairwise correlation between industry returns is 0.12 and 0.16 if only the largest counterparty is considered. These numbers reflect the degree to which monetary policy shocks will be propagated in the absence of any production network due to the raw correlation between returns. In order to be able to make any claims about the transmission of shocks through the production network, the value of ρ, given a row-normalized W, has to be significantly larger than 0.16.

5.2 Baseline Results

Table 4 presents the baseline results of this analysis. As a first step, column (1) reports OLS estimates of regressing monetary policy surprise factors on industry returns. The estimates for *target* and *path* are statistically significant and economically large. However, in such a standard setting it is not possible to distinguish between direct and network effects. To do so, the spatial lag is added to the model. Columns (2)-(6) present maximum likelihood estimates for the SAR model[16].

The point estimates β_1 and β_2 represent the *direct* response of industry returns as if network connections are severed. In the baseline specification in column (2), the model is estimated using equally-weighted returns and a constant weight matrix from the year 2000. The spatial autoregressive parameter is positive ($\rho = 0.44$), strongly statistically significant and considerably higher than the mean pairwise correlation coefficient of 0.16. This means that the effect of monetary policy shocks is propagated through the network. Industry returns react to the monetary policy surprise factors ($\beta_1 \neq 0, \beta_2 \neq 0$). Economically, the negative sign of β_1 means that tighter than expected monetary policy leads to a decrease in the stock market. For example, a policy rate that is 25 bps tighter than expected reduces industry returns by approximately 50 bps (direct effect). The coefficient of the *path factor* is smaller and positive. These responses are in line with the interpretation that the central bank communicates good news about the economy and consistent with its reaction function, tightens monetary policy to counteract its effect on the economy. The fact that the *path factor* has a significant impact on the stock market is important, because it demonstrates the effectiveness of ECB's communication.

Panels B and C offer a decomposition of the network effect and also tabulate *total effects*. Recall that $network\ effect = indirect\ effect + feedback\ effect$.

---

[16] Results are estimated using Matlab. I acknowledge the use of some routines from Paul Elhorst and LeSage and Pace's Econometrics Toolbox

Based on t-statistics calculated from a set of 1000 parameter values, all effects are statistically distinguishable from zero. For both factors, network effects account for as much as 44% of the total effect. Importantly, *feedback effects* represent only 5% of the overall effect, which demonstrates that the large network effects are not solely driven by business within one's own industry. These results are comparable in direction and magnitude to Weber and Ozdagli (2016), who find that in the US network effects account for between 50% and 85% of the overall stock market reaction. One reason for the lower magnitude of spillovers in the Eurozone could be due to the fact that the European stock market is less liquid than the American one.

The production network channel of monetary policy is not subsumed by other transmission mechanisms previously studied in the literature. To address the broad interest rate channel the model controls for exchange rate sensitivity, product demand elasticity, capital intensity (gross investment/value added), inventory investment and production structure (value added/output). To proxy for the credit channel, mean debt-to-capital ratio, firm size composition (i.e. Herfindahl index) and industrial output are used. Finally, the volatility index (VSTOXX) is included as a measure of the market expectations channel. Results remain strongly robust. This empirical evidence shows that the production network channel is a distinct channel of monetary policy transmission, which can however work in parallel with other channels.

5.4 The Role of the *Path* Factor

The baseline regression shows that the broad role of the *path* factor is to adjust the response of industry returns by precisely factoring in the effect of expectations about the future economic outlook. Prior to the European Sovereign Debt Crisis monetary policy was mainly active through changes in the policy rate, which explains why most of the effect is picked up by the *target* factor. Having exhausted its conventional set of policy tools following the crisis period, the ECB resorted to several unconventional monetary policy measures[17]. For example, on 4th July 2013 the ECB started communicating information about the future course of monetary policy in the form of forward guidance. During the press conference the president commented: *"Looking ahead, our monetary policy stance will remain accommodative for as long as necessary. The Governing Council expects the key ECB interest rates to remain at present or lower levels for an extended period of time."* On this day, the *path* factor shows a large negative jump and the stock market appreciated by 1 %, suggesting that forward guidance led to an easing of monetary policy. The *target* factor on the other hand did not record any policy change.

Furthermore, in June 2014 the ECB decided to decrease one of its three main policy rates, deposit facility rate, to -10 bps. A further rate cut to -20 bps followed in September. The interbank EONIA rate and OIS rates turned negative, as well as bond yields of short- to medium-term maturities. Following the removal of the zero lower bound (ZLB), values of the *target* factor remain very close to zero (mean value of 0.12 bps), whereas the *path* factor moves between -14 bps and 22 bps. To address this the model is estimated separately on two subsamples: *before ZLB* (before June 2014) and *after ZLB* (after and including June 2014). Excluding events after the removal of the ZLB does not change the baseline results and decompositions (column (4) **Table 4**). Turning to the post ZLB period however, the response of asset prices to the *target* factor becomes insignificant, whereas the response to the *path* factor flips sign and increases substantially in magnitude. A 100 bps higher than expected future policy rate reduces industry returns by 2.13 percentage points. This is consistent with the explanation that when the ZLB is reached, changes to the policy rate have no real bite on asset prices and short rates typically stick to the bound. At the same time, revisions in expectations about the

[17] Given that the two factors are constructed using shifts in forward OIS rates, they cannot capture unconventional monetary policy that has no direct impact on the EONIA rate. One such example is the Securities Markets Programme (SMP), which targeted sovereign yields.

future stance of monetary policy are reflected in long rates and hence, the *path* factor.

Importantly*,* network effects during the *after ZLB* period are significantly higher than the baseline result and account for 70 % of the overall effect. The reason for this is that *path,* through its effect on expectations about the future outlook, has an impact on long-term consumption and investment decisions. Typically, production plans are organized on schedules longer than one or two months ahead, which means that industries are less responsive to adjust their input orders in the short run than in the long run. This evidence suggests that monetary policy shocks related the future economic outlook transmit much more powerfully through the network than shocks related to changes in the short-term rate.

4.2 Relationship to the CAPM Model

Recall that the average supplier is connected to 410 industries *i.e* essentially the entire market. Hence, by definition the presence of a common monetary policy shock introduces comovement in industry returns. The standard capital asset pricing framework (CAPM) relates the expected industry return to its sensitivity to systematic risk:

$$ret_{ijt} - r_{ft} = \beta_0 + \beta_M(ret_{Mt} - r_{ft}) + \beta_1 Target_t + \beta_2 Path_t + \varepsilon_{ijt} \tag{8}$$

where ret_{ijt} is the expected industry return, ret_{Mt} is the expected return on the market and r_f is the risk-free rate. When asset prices are driven by *target* and *path* shocks, β_M measures propagation effects due to industries exposure to the market factor. In empirical work the market factor is proxied by a value-weighted index. In this sense, the CAPM model can be thought of as a special case of a SAR model, where the weights matrix W is determined by market capitalization:

$$ret_{ijt} - r_{ft} = \beta_0 + \beta_M \left(\boldsymbol{W^M}(\boldsymbol{ret_{ijt}} - \boldsymbol{r_{ft}})\right)_{\boldsymbol{t}} + \beta_1 Target_t + \beta_2 Path_t + \varepsilon_{ijt} \tag{9}$$

However, the distinct feature between the two models is the treatment of risk. The CAPM considers the “network factor” as a global variable *i.e.* the rows of W are identical for all industries. On the other hand, the network model uses *industry-specific* weights[18]:

$$ret_{ijt} - r_{ft} = \beta_0 + \rho\left(\boldsymbol{W_{ij}}(\boldsymbol{ret_{ijt}} - r_{ft})\right)_{\boldsymbol{ijt}} + \beta_1 Target_t + \beta_2 Path_t + \varepsilon_{ijt} \tag{10}$$

The estimates in column (5) of **Table 4** show that the SAR model strongly outperforms the CAPM: the CAPM explains only 8.64% of the variation in the data and attributes roughly 7% to network effects. These findings have two important implications. First, they re-emphasize the important role that sector-specific production connections play in determining the transmission of monetary policy shocks. Second, they indicate that the relevant network variable has to be constructed conditional on the type of shock. When asset prices are driven by demand shocks $W^{UPSTREAM}$ should be used; when asset prices are driven by supply shocks $W^{DOWNSTREAM}$ should be used. Since the CAPM always uses the same W^M regardless of the type of shock, the model cannot capture appropriately the magnitude of network effects. To expand on this point, the next section discusses upstream and downstream effects.

5.3 Upstream vs. Downstream Propagation of the Shock

Theory predicts that monetary policy shocks travel only from customers to their suppliers. However, empirically it is also possible that the downstream channel is active. To assess this, a downstream version of the SAR model is estimated, where returns are weighted by W^T (column (6) **Table 4**). Results indicate that downstream effects exist, but their importance is considerably lower than upstream effects.

To investigate the upstream transmission further, industries are split into downstream, midstream and upstream based on the percentage of goods they sell

[18] Baseline regressions use raw returns, because of the assumption that the risk-free is not likely to move a lot at daily frequency. As a robustness risk-free rate adjusted returns are used, where the risk-free rate is proxied by the daily yield on the 10-year German government bond (Bund-rate). The two specifications are equivalent.

directly and indirectly to the end-customer. The model is re-estimated for each group of industries. The precise sorting procedure is contained in *Appendix D*. The findings in **Table 5** show that the *network* impact intensifies the further upstream an industry is located. The idea is that all industries are subject to both direct and indirect impacts, but their relative importance is different depending on the position of an industry in the supply chain. Since downstream industries sell most of their output directly to consumers, they are more affected by changes in final customer demand than upstream industries. On the other hand, the output of upstream industries is mostly used as an intermediate input in the production of downstream industries. Intuitively, upstream industries are located further away from the shock's origin and for them most of the impact is coming from indirect effects transmitted through the network. For comparison, direct effects of the *target* factor constitute 68 % of the total effect for downstream industries as opposed to only 44% for upstream industries. The results for the *path* factor are qualitatively similar. Taken as a whole, this evidence shows that monetary policy shocks transmit mainly upstream.

5.5 Network Effect across Countries

What makes the case of the Euro-zone special and interesting is the fact that there is *one* common monetary policy, which applies to many countries connected in a euro-zone-wide production network. Therefore, it is important to understand how much of the indirect effect is due to foreign connections. For the baseline model (column (2)) in **Table 4**, the matrix $S(W)$ is split into two non-overlapping matrices: home connections are given by $S(W^H)$(i.e. entries *off* the main 38 × 38 block-diagonal are set to 0) and foreign connections are given by $S(W^F)$)(i.e. entries *on* the main 38 × 38 block-diagonal are set to 0). Then, the *Average Home Effect* = $\frac{1}{N}\iota_N'[S(W^H) - tr(S(W^H))]$ and the *Average Foreign Effect* $= \frac{1}{N}\iota_N' S(W^F)$. **Table 6** shows that on average foreign effects account 25 % of the total network effect.

One caveat when estimating these effects is the inherent difference between home and foreign connections. Industries have a small number of large links to home industries and a large number of small links to foreign industries. The baseline model weighs in the relative importance of these groups and the spatial autoregressive parameter reflects the *average* relationship. This is not a problem when the only interest is producing point estimates. However, in order to be precise about the decomposition, the row-normalized W matrix is split into W^H (only home connections) and W^F(only foreign connections) and the following second-order SAR model is estimated:

$$ret_{ijt} = \beta_0 + \rho_1 {W_{ij}}^H ret_{ijt} + \rho_2 {W_{ij}}^F ret_{ijt} + \beta_1 Target_t + \beta_2 Path_t + \lambda_i + \delta_j + \mu_{year} + \gamma controls_{ijt} + \varepsilon_{ijt} \quad (11)$$

The estimated parameter values are $\rho_1 = 0.39$ and $\rho_2 = 0.86$[19]. Note that compared to the baseline case, the magnitude of the indirect effect remains the same in absolute sense, but the importance of foreign effects increases to 45 %. These results support the graphical evidence in **Figures 1** and **2** and demonstrate that alongside with home connections, cross-country production linkages between Eurozone member states are an important determinant for the transmission of the ECB's monetary policy.

5.6 Higher-order Effects

The results in the previous sections reveal that the magnitude of spillover effects is considerably higher than what would be predicted by just looking at the number and strength of *direct* connections between industries. This is particularly relevant for foreign connections, which on average account for only 15 % of total flows, but as much as 45 % of the total indirect effect. To gain further insights into the propagation of the shocks, it is interesting to partition the network effect by order of the neighbors. Recall that according to the infinite series decomposition in Section 5, network orders are given by the powers of *W*: observations themselves (*0-order*), immediate neighbors (*1st-order*), neighbors of immediate neighbors (*2nd-*

[19] The magnitudes of the two parameters are not directly comparable. Following theoretical work by Elhorst et al. (2012), *Appendix E* contains details about the stationarity region of the second-order SAR model.

order) and so on. **Figure 7** shows spatially-partitioned network effects for both *home* and *foreign* connections associated with orders 1 to 5[20].

Feedback effects die out very quickly and after the second order are essentially zero. This is reassuring evidence because it demonstrates that spillovers are not solely driven by connections to one's own industry. *Home* and *Foreign* network effects, on the other hand, exhibit notably slower decay. Importantly, effects for both factors falling on orders greater than one are large and statistically significant. These findings indicate the sparsity of the W matrix alone cannot explain the results, because it disregards the crucial role that higher-order connections play in the transmission process. For comparison, considering only first-order connections would understate the magnitude of network effects by approximately 41%. The key feature of the SAR model, which makes it adept to study network effects of monetary policy, is its ability to capture precisely these types of non-linearities in the network structure.

5. Heterogeneous Network Effects

So far, the model has been estimated assuming that the sensitivity to monetary policy shocks and to network effects is homogeneous across industries. However, industries' exposure is likely to differ for various reasons. Previous research has shown that heterogeneous sensitivities can be explained by industry characteristics such as output durability, financial structure, investment intensity, exchange-rate sensitivity, borrowing capacity, firm size ((Dedola & Lippi, 2005); (Ehrmann & Fratzscher, 2004)) and profitability among others (Weber & Ozdagli, 2017). Revisiting **Figure 1**, it is interesting to observe that connections in the center are much denser than those in the periphery. This suggests that industries differ in their importance as input suppliers to others and

[20] The choice of the first 5 orders is adequate, because the associated effects account for 98.6 % of overall network effect. Note that if the spatially partitioned effects are cumulated over all powers of W until empirical convergence of the infinite series, these would exactly equal the magnitudes reported in the last row of Table 4.

serves as a motivation to study the interaction between network centrality and exposure to monetary policy shocks.

5.1 Network Centrality

Using US BEA data Acemoglu et al. (2012) show that industries differ in their importance as input suppliers to other sectors. This empirical relationship also holds in the case of the Eurozone economy (**Figure 8**): industries such as manufacturing of electrical equipment are *central* and sell to many customer industries, whereas others such as agriculture are not strongly connected. Given the upstream propagation of shocks from customers to suppliers, differences in connectedness predict that more central suppliers are more exposed to network effects of monetary policy than less connected industries. The idea is consistent with previous empirical evidence by Ahern (2013) and also Yang and Zhang (2016), who find that central suppliers are characterized by higher aggregate risk because they are more exposed to shocks transmitted through the intersectoral trade network.

When the network interaction between industries is specified as a spatial autoregressive model, the Bonacich centrality measure (1987) emerges naturally in the reduced-form equation. To see clearly the connection to the SAR model, recall the relationship given by equation (20):

$$ret_t = \beta_0(\mathbb{I}_{\boldsymbol{N}} - \boldsymbol{\rho W})^{-\mathbf{1}} + (\mathbb{I}_{\boldsymbol{N}} - \boldsymbol{\rho W})^{-\mathbf{1}}(\beta_1 Target_t + \beta_2 Path_t) + (\mathbb{I}_{\boldsymbol{N}} - \boldsymbol{\rho W})^{-\mathbf{1}}(\varepsilon_t)$$

The term $b(\rho, W) = \alpha(\mathbb{I}_n - \rho W)^{-1}\boldsymbol{\iota}$ gives the Bonacich centrality score and captures the effect of the position of an industry in the network on its returns[21]. Therefore, once an estimate for ρ is obtained and depending on the heterogeneity of network links (captured by W), the *distribution* of individual Bonacich centralities can be produced for all the nodes in the network. To arrive at a centrality score for an industry in year *T*, I apply the Bonacich centrality measure to the input-output matrix in year *T-1*. Using lagged centrality, I can address potential endogeneity problems with respect to the network formation. A detailed description of the

[21] "$\boldsymbol{\iota}$" denotes a column vector of ones.

Bonacich measure, including the connection to equations (19) and (20) is contained in the *Appendix E.*

As a first step, I divide industries into terciles based on their centrality score and look at the composition of industries (Level 1 of aggregation) and countries in the highest tercile (**Table 7**). The most connected sectors are Manufacturing, followed by Transport and Warehousing and Wholesale and Retail Trade. Within Manufacturing, the most important supplier industries are manufacturers of general inputs such as metal plates, electronic components and machines. 22.37 % and 19.27 % of the most central industries belong to Germany and France. Interestingly though, countries in the "periphery" of the Eurozone, such as Spain and Italy, hold almost an equal share. This suggests that centrality is closely related to size rather than economic stability.

Next, I interact centrality with the network factor and the two policy shock factors and estimate the following specification:

$$ret_{ijt} = \beta_0 + \rho W_{ij} ret_{ijt} + \theta W_{ij} ret_{ijt} \times Centr_{ijt} + \beta_1 Target_t + \beta_2 Path_t + \beta_3 Centr_{ijt} + \beta_4 Target_t \times Centr_{it} + \beta_5 Path_t \times Centr_{ijt} + \delta Controls_{ijt} + \varepsilon_{ijt} \quad (12)$$

In order to show that the effect of centrality is not subsumed by other variables previously studied in the literature, I add controls for industry characteristics (exchange rate sensitivity, product demand elasticity, capital intensity, inventory investment, production structure, mean debt-to-capital ratio, firm size composition and industrial output) and add industry, country and year fixed effects. All variables are standardized. The estimate θ gives the incremental effect of a 1 standard deviation increase in centrality on network spillovers. For example, the magnitude of spillovers for industries that are 1 standard deviation away from the mean equals $\rho + \theta$. As predicted by theory, columns (1) and (2) of **Table 8** show that the importance of network effects increases with centrality ($\theta > 0$): one standard deviation increase in centrality increases the point estimate of the spatial autoregressive parameter by 0.03. Interestingly, central industries respond less to the *target* factor ($\beta_4 > 0$) whereas there is no effect for the *path* factor ($\beta_5 = 0$). Centrality per se is not associated with stock returns in the event interval around policy announcements ($\beta_3 = 0$). **Table 9** shows that the difference between the predicted proportion of network spillovers for the average industry and a highly connected industry is sizeable and roughly equal to 15%. Importantly, although the relative importance of spillover effects increases with centrality, the absolute magnitude decreases.

Since central industries are likely to be bigger, more profitable and have more resources to counteract negative shocks, results could be driven by their lower riskiness. To test this, I divide industries into terciles based on their previous year centrality and look at excess monthly returns during the entire sample period from 2001 to 2017 with respect to the CAPM model and the Fama-French 3-Factor Model. **Table 10** shows that for both models, the *H-L* (high minus low centrality) portfolio earns a large, positive and statistically significant *alpha*. This finding is consistent with Ahern (2013) and rules out a risk-based explanation for the lower responsiveness of central industries to monetary policy shocks.

So far, the analysis has focused on the structure of the Eurozone network. However, industries have also production relations outside of the Eurozone. In fact, **Figure 9** shows that Eurozone industries are not clustered together, but that they are strongly embedded in a *global* production network. This means that strongly connected industries in the Eurozone tend to be globally important players as well. Therefore, in order to gain a deeper understanding of how monetary policy shocks are transmitted via the input-output network, global production links must also be taken into consideration.

5.2 Global Production Networks

Using information about global production linkages between industries for *all* 44 countries in the WIOT database, it is possible to compute how much of an industry's output is consumed *locally*, i.e. in the Eurozone. **Table 11** shows that the percentage of total output *both* consumed directly by end-customers or used as in intermediate input decreases monotonically with centrality. For example, only 21% of the output of industries in the highest tercile of centrality is consumed by end-customers and 39% is used as intermediate input in the Eurozone. For comparison, the respective numbers for industries in the lowest tercile are 39%

and 54%. This evidence suggests that part of the reason why highly connected European industries react less to monetary policy shocks is because they are less exposed to them, since a considerable fraction of their customers are located outside of the Eurozone. To further corroborate this intuition, in column (3) of **Table 8** I interact the share of local end-customer demand with the network lag and the two monetary policy factors. I find that the importance of spillovers is sizably higher for local industries ($\theta = 0.048$) and that they are more responsive to the target policy shocks ($\beta_4 > 0$).

The result is in line with Potjagaiolo (2017), who studies the effect of euro area monetary policy across non-euro area Europe and finds that spillovers are larger for countries that trade more with the eurozone due to increased exposure to foreign demand effects. The same argument holds when industries *within* the eurozone are considered and their trade relations in the global network.

6. Robustness and Discussion

6.1 Threats to Identification

The validity of the analysis relies on the correct identification of the spatial autoregressive parameter $\boldsymbol{\rho}$. A traditional concern in the empirical assessment of network effects is the reflection problem due to (Manski, 1993). The simultaneity in the behavior of interacting industries introduces perfect collinearity between the expected mean outcome and mean background of industries. In this context, it becomes crucial to be able to distinguish between industry i's impact on j and j's impact on i. Bramoullè, Djebbari, & Fortin (2009) and Cohen-Cole et al. (2014) observe that intransitivities in complex networks can be used for identification. For example, i is connected to j, j is connected to k, but i is not connected to k. In this case, $W_{ik} = 0$, while $W_{ik}^2 \geq W_{ij}W_{jk} > 0$. Effectively, $W_{ik}^2$is an instrument for the impact of i on k: it is correlated with i's behavior, but it has no direct effect on k. This means that network effects are identified if there are two industries in the economy that differ in the average connectivity of their direct customers. Empirically, this condition translates into linear independence of the matrices I_N, W, W^2. A complex directed network such as the one used in this paper has a sufficiently rich structure of connections, and identification would practically never fail. Nonetheless, I vectorize each matrix and verify that the matrix formed by concatenating the three vectors has rank three.

Despite the short event window, the two main problems in estimating the correct response of asset prices to monetary policy are the existence of omitted variables and the simultaneous movements in interest rates and asset prices. Addressing the former, manual inspection of event dates indicates that the only major coincidence is the publishing of Initial Jobless Claims in the US. These are reports produced by the US Department of Labor, which measure the number of people that have filed for unemployment benefits for the first time. Median forecast values from Bloomberg's Market Analysis Survey (starting from late 2008 onwards) are used to derive the surprise component of these claims $\Delta\, IJC = actual - forecast$. Monetary policy surprise factors are projected on $\Delta\, IJC$ and the residuals are used as right-hand-side variables in column (1) of **Table 12**. Controlling for Initial Jobless Claims reduces slightly the magnitude of the effects, but the estimates of the *target* and *path* factors remain statistically significant.

To address the simultaneity problem, the paper uses the identification-through-heteroskedasticity estimator proposed by Rigobon (2003). The assumption is that the variance of the monetary policy shocks is higher on days of ECB's announcements, because a bigger portion of the news spread in the market is about monetary policy. Rigobon (2003) shows that this shift in the variance is sufficient to identify the response of asset prices. A convenient feature about this method is that it can be easily applied in an instrumental variable setting Rigobon and Sack (2004). It only requires identifying two subsamples, *policy* and *non-policy* dates, over which the "importance" of policy shocks increases. For the set of non-policy dates the paper takes the dates immediately preceding those included in the baseline event window. This keeps the two samples the same size and minimizes the impact of changes in the variance of the shocks over time. *Appendix H* provides more details about the assumption and the implementation. Column (2) of **Table 12** provides evidence that the baseline finding of the paper remains robust. Contractionary monetary policy decreases industry returns, favorable economic outlook increases returns and the effect is propagated. Importantly, Panels B and C show that higher-order network effects account for 69 % of the total effect.

Since the monetary policy surprise factors are estimated using market rates[22], it is important to check whether results are not driven by white noise in the data. The model is re-run drawing 1000 random samples of event dates without replacement and the reaction of industry returns to the *target* and *path* factors is found to be insignificant (column (3) **Table 12**). The network parameter is positive and significant, which is due to the transmission of shocks, other than monetary policy ones, the presence of which cannot be controlled for.

Network spillovers could be driven by the presence of correlated effects: e.g. industries react similarly because face similar demand cyclicality or output durability (Petersen & Strongin (1996); Braun & Larrain (2005); D'Acunto, Hoang & Weber (2016)). To control for these, industry returns on regressed on industry dummies and the demeaned data in used in the SAR model in column (4) of **Table 12**. To control for differences in market capitalization, column (5) uses value-weighted returns[23]. Both adjustment procedures has very little effect on point estimates and decompositions.

6.2 Spatial Weights Matrix

A major caveat when working with spatial models is the choice of the spatial weights matrix. First, the use of a row-stochastic W implies that all intermediates trade is confined within the borders of the Eurozone. Given the importance of global connections discussed in the previous sections, it is possible that the relative

[22] Following Rudebusch (1998) and later by Kuttner (2001), an alternative way to measure the short-term surprise component of monetary policy is to use changes in the 3-month EURIBOR futures rate. These instruments are actively traded on the EUREX exchange in Frankfurt and have a long time-series starting from 1999. A 100 bps higher than expected policy rate decreases industry returns by approximately 0.80 % and 53 % of the total effect is due to network effects.

[23] Value-weighted returns are calculated using last year's market capitalization. In untabulated tests, I use previous month's market capitalization as weights and obtain very similar results.

importance of home connections is overestimated. Column (1) of **Table 13** normalizes the strength of the connections by total supplier sales using data for all 44 countries in the sample[24]. Estimates of the monetary policy shock factors and the proportion of network effects to total effects remain qualitatively unchanged, which provides evidence that the normalization scheme for the matrix W does not drive the results.

A second possible concern is that the empirical input-output matrix has non-zero elements on the diagonal. Practically, this means that a construction company uses building materials and fixtures to build homes[25]. Although the baseline estimation shows that feedback effects constitute for only 5% of the overall effect, within industry demand effects could bias the magnitude of indirect network effects. In column (2), zero-diagonal entries are imposed keeping intermediate input shares adding up to unity. The spatial autoregressive parameter remains positive and strongly significant and network effects remain close to 50%.

Another doubt is that the large estimate for ρ is mechanical, because industry returns are regressed on a weighted average of industry returns. To address this, a simulated input-output matrix is used as an input to the SAR model. The empirical input-output matrix is sparse and its distribution is heavily right-skewed, which means that a small number of sectors are important suppliers to the rest of the economy (consider **Figure 1** and **2** in this paper and Acemoglu et al. (2012); Gabaix (2012)). Hence, input-output matrix, which inherits those two features, is simulated. Random numbers are drawn from a Generalized Pareto

---

[24] Note that the magnitude of ρ is not directly comparable to the value in the Baseline Model because the parameter space is different. In principle $\rho \in \left(-\frac{1}{\lambda_{max}}, +\frac{1}{\lambda_{max}}\right)$ where λ_{max} is the maximum eigenvalue of W. Since W is not row-stochastic, $\lambda_{max} < 1$ and $\rho \in (-1.85, +1.85)$. All previous arguments hold: $(I - \rho W)^{-1}$ exists and since the maximum eigenvalue of ρW is less than 1, the infinite series $I + \rho W + \rho^2 W^2 + \cdots$, converges. Intuitively, this means that the process has fading memory.

[25] On the subsector level, the ICB code for 'Heavy Construction' is 2357 and for 'Building Materials & Fixtures' is 2353. Both belong to the sector 2350 'Construction & Materials'.

Distribution with shape parameter $\xi = 4.21$, scale parameter $\sigma = 0.0121$ and location parameter $\theta = 0.0001$. To derive these parameters, the squared distance between the empirical distribution function and the estimated distribution is minimized using data from the year 2000. Column (3) shows that the bias in estimating the spatial parameter ρ is likely to be very small. The model obtains a $\rho = 0.095$, which is more than five times smaller than the baseline estimate of 0.44. The decomposition of the overall effect also provides reassuring evidence: indirect effects constitute only 9% for the *target* factor and 11% for the *path* factor.

Finally, the assumption of a constant weight matrix is relaxed and specification (4) employs a time-varying spatial weights matrix[26]. To address endogeneity concerns, W in year k is instrumented with W using data from year *k-1*. Results for β_1, β_2 and ρ remain close to the baseline results.

6.3 Heterogeneous ρ

One possible shortcoming of the SAR model with homogeneous ρ is that it only captures average behavior across spatial units, even if country-industry fixed effects are allowed (Elhorst, (2003)). Given the short time dimension of the panel, allowing the spatial autoregressive parameter to vary fully across spatial units will be computationally taxing and would result in model overfitting. Still, to get an approximate feeling about the bias introduced by imposing parameter homogeneity, it is possible to estimate the SAR model separately for each industry and for each country:

$$ret_{ikt} = \beta_0 + \rho_k W_{ik} ret_{ikt} + \beta_1 Target_t + \beta_2 Path_t + \varepsilon_{ikt} \qquad (13)$$

where ret_{ikt} is the return of country-industry pair i belonging to group k (industry or country) at event date t. **Figure 11** plots the parameter values for the 38 industries (Panel A) and for the 11 countries (Panel B). It is immediate to observe that there is very little dispersion. In fact, for 32 industries and for 10 countries

[26] Note that row-normalization in the case of a dynamic W could bias the impact of changes in the network density. For example, an increase in the number of customers a supplier has would decrease the relative impact of an individual customer. Therefore, following Kelejian & Prucha (1998) I normalize by the maximum of row and column sums.

the 95-percent confidence bounds of the parameter contain the mean parameter value estimated in the baseline regressions. This evidence is reassuring. It demonstrates in that the bias from imposing a homogenous spatial autoregressive parameter is not likely to be large. Moreover, it re-emphasizes that the large network effects found in this paper are not due to particular countries, but due to the complex production network in the euro area.

6.4 Centrality Measures

Network centrality can be calculated in many different ways, which capture different characteristics of the network. To ensure robustness of the empirical results, I consider three additional measures: *outdegree* centrality, *eigenvector* centrality and *closeness* centrality[27]. Outdegree centrality is a simple measure that counts the number of outbound links. Eigenvector centrality is a recursive measure of centrality equal to the eigenvector associated with the largest eigenvalue of the connectivity matrix W. Closeness is defined as the sum of the length of the shortest paths between an industry and all other industries in the graph. **Table 14** shows that results are highly robust to alternative specifications of centrality.

7. Theoretical Interpretation

This section develops a simple static production model with intermediate inputs in which monetary policy has a heterogeneous effect on industries' stock prices. The model provides a structural interpretation of the transmission of monetary policy shocks to the real economy through production networks and motivates the empirical strategy of the paper. The theoretical setup follows closely Carvalho (2014), Acemoglu, Ozdaglar and Tahbaz-Salehi (2017) and Weber and Ozdagli (2017).

7.1 Production Model with Intermediate inputs

Consider an economy consisting of K countries in each of which there are L industries. There are $N = K \times L$ country-industry pairs. Each industry is

[27] For a thorough discussion of different centrality measures (Borgatti, 2003); (Billio, Getmansky, Co, & Pellizon, 2012); (Wassermann & Faust, 1994); (Scott, 2017)

specialized in a different product. These products serve a dual role in the economy: they can either be valued by households as final consumption or can be used as an intermediate input in the production of other goods. Industry i's objective is to maximize profits π_i by choosing a primary factor- in this case homogeneous labor l- and intermediate inputs x_{ji} from industries $j = 1,..,N$ i.e. including itself. Industries take prices $\{p_i\}_{i=1}^{N}$ and wages ω as given. Industries also have a pre-determined fixed cost f_i, which can be regarded as payment of rent or nominal debt. I assume that the production process at each of these industries is given by a Cobb-Douglas production technology. Thus,

$$\pi_i = max\, p_i y_i - \sum_{j=1}^{N} p_j x_{ji} - \omega l_i - f_i \tag{14}$$

$$y_i = l_i^{\lambda} \left(\prod_{j=1}^{N} x_{ji}^{w_{ji}}\right)^{\alpha} \tag{15}$$

where y_i is industry output, λ and α are factor shares and w_{ji} is the share of inputs from industry j used in the production of industry i s.t. $\sum_{j=1}^{N} w_{ji} = 1$.

Defining industry revenue as $R_i = p_i y_i$, the first order conditions are given by:

$$\alpha w_{ij} R_i = p_j x_{ji} \tag{16}$$

$$\lambda R_i = \omega l_i \tag{17}$$

Substitute in the first-order conditions to obtain:

$$\pi_i = (1 - \lambda - \alpha) R_i - f_i \tag{18}$$

Consumers maximize utility:

$$max \sum_{i=1}^{N} \log(c_i) \tag{19}$$

subject to:

$$\sum_{i=1}^{N} p_i c_i = \omega \sum_{i=1}^{N} l_i + \sum_{i=1}^{N} \pi_i + \sum_{i=1}^{N} f_i \tag{20}$$

Equation (7) assumes that fixed costs are a transfer from industries to consumers. Consumers get wages, profits and fixed costs and spend all the resulting income on N products. Using equations (4) and (5), it is possible to express the first-order condition as:

$$c_i = \frac{\omega \sum_{i=1}^{N} l_i + \sum_{i=1}^{N}(\pi_i + f_i)}{N p_i} = \frac{(1-\alpha)\sum_{i=1}^{N} R_i}{N p_i} \tag{21}$$

Given that goods can be either consumed or used as intermediates, the goods clearing condition becomes:

$$y_i = c_i + \sum_{j=1}^{N=1} x_{ij} = \frac{(1-\alpha)\sum_{i=1}^{N} R_i}{N p_i} + \frac{\alpha \sum_{j=1}^{N} w_{ij} p_j y_j}{p_i} \tag{22}$$

Multiplying by p_i, equation (9) simplifies to:

$$R_i = (1-\alpha)\frac{\sum_{i=1}^{N} R_i}{N} + \alpha \sum_{j=1}^{N} w_{ij} R_j \tag{23}$$

Equation (19) is very important from a conceptual point of view because it says that an industry i's revenue is affected by shocks to households' demand, captured by the first term, and by shocks to the revenues of its customers, captured by the second term. The transmission of shocks through the production network depends on the size of the industries that buy intermediates from industry i, R_j, and their importance as customers of i, $\alpha \times w_{ij}$. This means that monetary policy shocks are transmitted *upstream*: from customers to their suppliers.

Define $W \equiv [w_{ij}]$ as the matrix of input shares and $R \equiv (R_1, \dots, R_N)'$ as a vector of revenues. Then,

$$(I - \alpha W)R = (1-\alpha)\begin{pmatrix} \frac{(\sum_{i=1}^{N} R_i)}{N} \\ \vdots \\ \frac{(\sum_{i=1}^{N} R_i)}{N} \end{pmatrix}_{N\times 1} \tag{24}$$

I assume that intermediates are purchased with trade credit, whereas consumer goods are purchased with cash. Money supply affects prices through the following cash-in-advance constraint:

$$\sum_{i=1}^{N} p_i c_i = (1-\alpha)\sum_{i=1}^{N} R_i = M \tag{25}$$

with M being money supply. Combining equations (11) and (12):

$$(I-\alpha W)R = (1-\alpha)\begin{pmatrix} M/N \\ \vdots \\ M/N \end{pmatrix}_{N\times 1} = m \tag{26}$$

Define $\pi \equiv (\pi_1, \dots, \pi_N)'$ and $f \equiv (f_1, \dots, f_N)'$ as vectors of profits and fixed costs respectively. Then, substituting equation (13) in equation (5), the following obtains:

$$\pi = (I-\alpha W)^{-1}(1-\lambda-\alpha)m - f \tag{27}$$

The last expression can be log-linearized[28] to get:

$$\bar{\pi}\hat{\pi} = (I-\alpha W)^{-1}(1-\lambda-\alpha)\bar{m}\widehat{M} \tag{28}$$

Define $\beta \equiv (\beta_1, \dots, \beta_N)'$ with

$$\beta_i = \frac{(1-\lambda-\alpha)\bar{m}}{\bar{\pi}_i} \tag{29}$$

and so:

$$\hat{\pi} = (I-\alpha W)^{-1}\beta\widehat{M} \tag{30}$$

Finally, it is possible to rewrite the expression for the deviation of net income as:

$$\hat{\pi} = \alpha \times W \times \hat{\pi} + \beta \times \widehat{M} \tag{31}$$

Hence, an industry's net income, or stock price in that matter, reacts to changes in money supply, $\widehat{M}$, and the reaction of its customers to the monetary

---

[28] Consider a variable x be with steady state $\bar{x}$ and log-deviation $\hat{x}$ from the steady state $\hat{x} = \log(\mathrm{x}) - \log(\bar{x}) = \log\left(\frac{\hat{x}}{\bar{x}}\right) \approx \frac{\hat{x}-\bar{x}}{\bar{x}}$ such that $x = \bar{x}\exp(\hat{x}) \approx \bar{x}(1+\hat{x})$

shock, $W \times \hat{\pi}$. It is important to observe that equation (27) has exactly the form of a *spatial autoregression*.

7.2 Policy Implications

One of the advantages of the SAR approach is that it provides a convenient framework via which policy makers can understand better the impact of monetary policy decisions using the structure of the network. A key question for policy evaluation here is which industries are most important for the transmission of shocks through the production network. Drawing from the empirical evidence in the previous chapters, the answer is intuitive: these are industries, which depend largely on Eurozone demand and hence, are most exposed to the shock, but at the same time are connected enough to transmit the shock to their suppliers upstream. These industries are called *"influential"*. **Table 15** shows that suppliers connected directly[29] (first-order) to at least one influential customer react nearly three times more than those that are not connected to an influential industry. Suppliers connected indirectly to an influential industry through a customer (second-order) react nearly two times less. A small impact is observed for third-order neighbors, after which the effect largely dies out.

The finding that the aggregate level of network spillovers depends on the topology of the network raises the point whether the past evolution of the network structure can be used to predict the direction and magnitude of spillover effects from monetary policy shocks. To investigate this, the paper borrows from a recent result in the asset pricing literature due to Herskovic (2018). The author shows that two global characteristics of the network matter for asset prices: network concentration and network sparsity. Network concentration measures the concentration of industries' centrality: high levels of concentration signify an economy where a few sectors play a disproportionately important role as input suppliers. Network sparsity on the other hand measures specialization: high sparsity implies that sectors supply to a few large customers.

[29] Only connections above the 90th percentile of i-th row of W, W^2 and W^3 are considered.

Each year *concentration* and *sparsity* factors are constructed using data from the previous year available. The two factors are standardized and interacted with the network lag in the following specification:

$$ret_{ijt} = \beta_0 + \rho W_{ijt} ret_{ijt} + \theta_1 Wret_{ijt} \times Concentration_t + \theta_2 Wret_{ijt} \times Sparseness_t + \beta_1 Target_t + \beta_2 Path_t + \beta_3 Concentration_t + \beta_4 Sparseness_t + \varepsilon_{ijt} \quad (32)$$

The result is that one standard deviation increase in network *concentration* reduces the magnitude of spillover effects by 0.047 (**Table 16**). Network *sparsity* does not produce a significant impact, which is consistent with the main argument of the paper: sparsity cannot rationalize the large network effects of monetary policy, because it ignores higher-order linkages between industries. Intuitively, high concentration means that spillovers due to production linkages will be small for the majority of the sectors, because they are weakly connected. Another way to think about the concentration asset pricing factor in the Eurozone is in terms of economic convergence. The more lower-income countries catch up, the more they get integrated in the Eurozone production network and the stronger their ties become with the rest of the member states. In the network jargon high economic convergence translates into lower network concentration. The insight from this result is that production networks are likely to play an increasingly important role for the transmission of monetary policy in the future.

8. Conclusion

This paper explores the role of production networks for the propagation of monetary shocks in the euro area. The property of equity markets to react quickly to monetary policy news is used as a tool to identify network effects. This paper provides first empirical evidence that between 40% and 50% of the total stock market reaction following an ECB's announcement event is due to spillover effects. The transmission mechanism arises due to intermediate input linkages between sectors, which introduces higher-order demand effects.

The paper emphasizes that these large network effects cannot be solely explained by transmission within individual countries but are rather the result of the complex structure of the euro are production network. Results indicate that not including linkages across countries would lead to underestimate the magnitude of spillover effects by 50 %.

Consistent with theoretical predictions, monetary policy shocks travel mostly upstream from customers to their suppliers. Using the rich data structure from the World Input-Output Tables, the paper documents heterogeneous network effects based on network connectivity and exposure to the global production network. Combing the results on the upstream propagation of monetary policy shocks and network centrality improves our understanding of the transmission mechanism of monetary policy and allows us to draw meaningful policy conclusions. Moreover, a key difference between the Eurozone and the US for example, is the possibility of countries entering and exiting the currency union. Given data on production linkages and estimates for aggregate network spillovers and the concentration factor, the SAR model provides a powerful tool to analyze different counterfactual scenarios. A task for future research is to carefully design policy experiments to evaluate the pass-through of monetary policy through production networks.

Going forward, an interesting extension to this work is to investigate network effects due to the interaction between the specific fiscal policy of the member states and the common euro area monetary policy. Another direction for future research is to explore the role of global production networks for asset prices. The fact that nowadays firms and industries are simultaneously embedded in several layers of networks e.g. home country, Eurozone and global networks calls for a deeper analysis of the impact of globalization for the transmission of economic shocks.

References

Acemoglu, D., Akcigit, U., & Kerr, W. (2015). Networks and the macroeconomy: an empirical exploration. *NBER Macro Annual.*

Acemoglu, D., Carvalho, M., Ozdaglar, A., & Tahbaz-Salehi, A. (2012). The Network Origins of Aggregate Fluctuations. *Econometrica, 80*(5), 1977-2016.

Acemoglu, D., Ozdaglar, A., & Tahbez-Salehi, A. (2017). Microeconomic origins of macroeconomic tail risks. *The American Economic Review, 107*(1), 54-108.

Adao, R., Costinor, A., & Donaldson, D. (2017). Non-parametric counterfactual predictions in neoclassical models on international trade. *American Ecnomic Review, 107*(3), 633-689.

Ahern, K. (2013). Network Centrality and the Cross-Section of Stock Returns. *Working Paper*.

Anselin, L. (2001). A comparison to theoretical econometrics. In B. Baltagi (Ed.), *Spatial Econometrics* (pp. 310-330). Blackwell.

Anselin, L., & Bera, A. (1998). Spatial dependence in linear regression models with an introduction to spatial econometrics. *Statistics Textbooks and Monographs, 155*, 237-290.

Anselin, L., Bera, A., Florax, R., & Yoon, M. (1996). Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics, 26*(1), 77-104.

Arbia, G. (2006). *Spatial econometrics: statistical foundations and applications to regional convergence.* Springer Science and Business Media.

Arbia, G. (2012). *Spatial data configuration in statistical analysis of regional economic and related problems.* Springer Science & Business Media.

Atalay, E. (2017). How important are sectoral shocks? *American Economic Journal*, 254-280.

Barakchian, S., & Crowe, C. (2013). Monetary policy matters: Evidence from new shocks data. *Journal of Monetary Economics, 60*(8), 950-966.

Barrot, J.-N., & Sauvagnat, J. (2016). Input specificity and the propagation of idiosyncratic shocks in production networks. *The Quarterly Journal of Economics, 131*(3), 1543-1592.

Bernanke, B., & Gertler, M. (1995). Inside the Black Box: The Credit Channel of Monetary Policy Transmission. *Journal of Economic Perspectives, 9*(4), 27-48.

Bernanke, B., & Kuttner, K. (2005). What explains the stock market's reaction to Federal Reserve Policy? *The Journal of Finance, 60*(3), 1221-1257.

Billio, M., Getmansky, M., Co, W., & Pellizon, L. (2012). Econometric measures of connectedness and systemic risk in finance and insurance sectors. *Journal of Financial Economics, 104*, 535-559.

Bivand, R., Pebesma, E., & Gómez-Rubio, V. (2013). *Gómez-Rubio* . Springer.

Blasques, F., Koopman, S., Lucas, A., & Schaumburg, J. (2016). Spillover dynamics for systemic risk measurement using spatial financial time series models. *Journal of Econometrics, 195*(2), 211-223.

Blattner, R., Caterano, M., Ehrmann, M., Strauch, R., & Turunen, J. (2008). The predictability of monetary policy.

Bloningen, B., Davies, R., Waddell, G., & Naughton, T. (2007). FDI in space: Spatial autoregressive relationships in foreign direct investment. *European Economic Review, 51*(5), 1303-1325.

Bonacich, P. (1972). Factoring and weighting approaches to status cores and clique identification. *Journal of Mathematical Sociology, 2*(1), 113-120.

Bonacich, P. (1987). Power and Centrality: A Family of Measures. *American Journal of Sociology, 92*(5), 1170-1182.

Bordignon, M., Cerniglia, F., & Revelli, F. (2003). In search of yardstick competition: a spatial analysis of Italian municipality property tax setting. *Journal of Urban Economics, 54*(2), 199-217.

Borgatti, S. (2003). The key player problem. In R. Breiger, K. Carley, & P. Pattison, *Dynamic Social Media Network Modeling and Analysis* (pp. 241-252). New York: National Academy of Sciences Press .

Borio, C., & Zhu, H. (2012). “Capital regulation, risk-taking and monetary policy: a missing link in the transmission mechanism? *Journal of Financial Stability, 8*(4), 236-251.

Bramoullè, Y., Djebbari, H., & Fortin, B. (2009). Identification of peer effects through social networks. *Journal of Econometrics, 150*, 41-55.

Brandt, C., Buncic, D., & Turunen, J. (2010). The Impact of ECB's Monetary Policy Decisions on the Yield Curve. *Journal of the European Economics Association, 8*(6), 1266-1298.

Braun, M., & Larrain, B. (2005). Finance and the Business Cycle: International, Inter-Industry Evidence. *The Journal of Finance, 60*(3), 1097-1128.

Carvalho, V. (2014). From micro to macro via production networks. *The Journal of Economic Perspectives, 28*(4), 23-47.

Carvalho, V., & Gabaix, X. (2013). The great diversification and its undoing. *The American Economic Review, 103*(5), 1697-1727.

Carvalho, V., & Voigtländer, N. (2015). Input Diffusion and the Evolution of Production Networks. *Working Paper*.

Cliff, A., & Ord, K. (1981). *Spatial process: models and applications.* London: Pion.

Cohen-Cole, E., Kirilenko, A., & Patacchini, E. (2014). Trading Networks and Liquidity Provision. *Journal of Financial Economics, 113*, 235-251.

Cook, T., & Hahn, T. (1989). The effect of changes in the federal funds rate target on market interest rates in the 1970s. *Journal of Monetary Economics, 24*(3), 331-351.

Costinor, A., & Rodriegez-Clare, A. (2014). Trade Theory with Numbers: Quantifying the Consequences of Globalization. *Hnadbook of International Economics*, 197-261.

Cragg, J., & Donald, G. (1997). Inferring the rank of a matrix. *Journal of Econometrics, 76*, 223-250.

D'Acunto, F., Hoang, D., & Weber, M. (2016). The effect of unconventional fiscal policy on consumption expenditure. *Working Paper*.

Dedola, L., & Lippi, F. (2005). The monetary transmission mechanism: Evidence from the industries of five OECD countries. *European Economic Review, 49*(6), 1543-1469.

ECB. (2008). The Analysis of the Euro Money Market from a Monetary Policy Perspective. *ECB Monthly Bulletin*, 65-79.

Eder, A., & Keiler, S. (2015). CDS Spreads and Contagion Amongst Systemically Important Financial Institutions–A Spatial Econometric Approach. *International Journal of Finance & Economics, 20*(4), 291-309.

Ehrmann, M., & Fratzscher, M. (2004). Taking Stock: Monetary Policy Transmission to Equity Markets. *Journal of Money, Credit and Banking, 36*(4), 719-737.

Ehrmann, M., & Fratzscher, M. (2007). Communication and Decision-Making by Central Bank Committees: Different Strategies, Same Effectiveness? *Journal of Money, Credit and Banking, 39*(2-3), 509-541.

Elhorst, P. (2003). Specification and Estimation of Spatial Panel Data Models. *International Regional Science Review, 26*(3), 244-268.

Elhorst, P. (2003). The mystery of regional unemployment differentials: Theoretical and empirical explanations. *Journal of economic surveys*.

Elhorst, P. (2012). On model specification and parameter space definitions in higher order spatial econometric models. *Regional Science and Urban Economics, 42*, 211-220.

Elhorst, P., & Allers, M. (2005). ax mimicking and yardstick competition among local governments in the Netherlands. *International tax and public finance, 12*(4), 495-513.

Elhorst, P., Heijnen, P., Samarina, A., & Jacobs, J. (2017). Transitions at Different Moments in Time: A Spatial Probit Approach. *Journal of Applied Econometrics, 32*(2), 422–439.

Ellen, S., Jansen, E., & Midthjell, N. (2017). The effect of ECB's conventional and unconventional monetary policy on Norwegean asset prices. *ECB Working Paper Series*.

Fajgelbaum, P., & Khandelwal, A. (2016). Measuring the unequal gains from trade. *Quarterly Journal of Economics, 131*(3), 1113-1180.

Fernandez, V. (2011). Spatial linkages in international financial markets. *Quantative Finance, 11*(2), 237-245.

Gabaix, X. (2012). The granular origins of aggregate fluctuations. *Econometrica, 79*(3), 733-772.

Gertler, M., & Karadi, P. (2015). Monetary Policy Surprises, Credit Costs, and Economic Activity. *American Economic Journal: Macroeconomics, 7*(1), 44-76.

Gilchrist, S., López-Salido, D., & Zakrajšek, E. (2015). Monetary Policy and Real Borrowing Costs at the Zero Lower Bound. *American Economic Journal: Macroeconomics, 7*(1), 77-109.

Gürkaynak, R., Sack, B., & Swanson, E. (2005). Do Actions Speak Louder than Words? The Response of Asset Prices to Monetary Policy Actions and Statements. *International Journal of Central Banking, 1*(1), 55-93.

Herskovic, B. (2018). Networks in Production: Asset Pricing Implications. *Forthcoming Journal of Finance*.

Herskovic, B., Kelly, B., Lustig, H., & Van Nieuwerburgh, S. (2016). The common factor in idiosyncratic volatility: Quantitative asset pricing implications. *Journal of Financial Economics, 119*(2), 249-283.

Jardet, C., & Monks, A. (2014). Euro Area Monetary Policy Shocks: Impact on Financial Asset Prices during the Crisis. *Working Paper*.

Jarociński, M., & Karadi, P. (2018). Deconstructing monetary policy surprises: the role of information shocks. *ECB Working Paper Series*.

Johnson, R. (2014). Five Facts about Value-Added Exports and Implications for Macroeconomics and Trade Research. *Journal of Economic Perspectives, 28*(2), 119-142.

Kashyap, A. K., & Stein, J. C. (2000). What do a million observations on banks say about the transmission of monetary polily? *American Economic Review, 90*(3), 407-428.

Kelejian, H., & Prucha, I. (1998). A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *The Journal of Real Estate Finance and Economics, 17*(1), 99-121.

Kelejian, H., & Prucha, I. (1999). A generalized moments estimator for the autoregressive parameter in a spatial model. *International Economic Review, 40*(2), 509-533.

Kelejian, H., & Prucha, I. (2004). Estimation of simultaneous systems of spatially interrelated cross sectional equations. *Journal of Econometrics, 118*, 27-50.

Kelejian, H., & Prucha, I. (2010). Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of Econometrics, 157*, 53-67.

Kelejian, H., Tavlas, G., & Hondronyiannis, G. (2006). A Spatial Modeling Approach to Contagion Among Emerging Economies. *Open Economies Review, 17*(4/5), pp. 423-442.

Kelly, B., Lustig, H., & Van Nieuwerburgh, S. (2013). Firm volatility in granular networks. *Working Paper*.

Kim, C., Phipps, T., & Anselin, L. (2003). Measuring the Benefits of Air Quality Improvement: A Spatial Hedonic Appriach. *Journal of Environmental Economics and Management, 45*, pp. 24-39.

Koopman, R., Wang, Z., & Wei, S. (2014). "Tracing Value-Added and Double Counting in Gross Exports" American Economic Review. *American Economic Review, 104*(2), 459-494.

Kuttner, K. (2001). Monetary policy surprises and interest rates: Evidence from the Fed funds futures market. *Journal of Monetary Economics, 47*(3), 523-544.

Le Gallo, J., & Ertur, C. (2003). 422–439. *Papers in regional science, 82*(2), 175-201.

Lee, L.-F. (2002). Consistency and efficiency of least squares estimation for mixed regressive, spatial autoregressive models. *Econometric Theory, 18*, 252-277.

Lee, L.-F. (2004). Asymptotic Distributions of Quasi-Maximum Likelihood Estimators for Spatial Econometric Models. *Econometrica, 72*, pp. 1899-1926.

LeSage, J., & Pace, R. (2006). Interpreting Spatial Econometric Models. *Regional Science Association International North American Meetings.*

LeSage, J., & Page, R. (2009). *Introduction to Spatial Econometrics.* CRC Press Taylor & Francis.

Los, B., Timmer, M., & de Vries, G. (2015). How Global are Global Value Chains? A New Approach to Measure International Fragmentation. *Journal of Regional Science, 55*(1), 66-92.

Lucas, R. (1977). Understanding business cycles. *Carnegie-Rochester conference series on public policy.* Elsevier.

Manski, C. (1993). Identification of Endogenous Social Effects. *Reiew of Economic Studies, 60*, 531-542.

Miranda-Aggripino, S. (2016). Unsurprising Shocks: Information, Premia, and the Monetary Transmission. *Staff Working Paper, Bank of England.*

Moran, P. (1950). Notes on continuous stochastic phenomena. *Biometrica, 37*, 17-23.

Ord, J. (1975). Estimation Methods for Models with Spatial Interaction. *Journal of the American Statistical Association*, 120-126.

Ottaviano, G., Pessoa, J., Sampson, T., & Van Reenen, J. (2014). The Costs and Benefits of Leaving the EU. *LSE Mimeo.*

Petersen, B., & Strongin, S. (1996). Why Are Some Industries More Cyclical Than Others? *Journal of Business and Economic Statistics, 14*(2), 189-198.

Potjagailo, G. (2017). Spillover effects from Euro area monetary policy across Europe: A factor-augmented VAR approach. *Journal of International Money and Finance, 72*, 127-147.

Ramey, V. (2016). Macroeconomic Shocks and their Propagation. *Working Paper 21978, National Bureau of Economic Research.*

Remolona, E., & Wooldridge, P. (2003). The Euro Interest Rate Swap Market. *BIS Quarterly Review*, 47-56.

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

Rigobon, R., & B., S. (2004). The impact of monetary policy on asset prices. *Journal of Monetary Economics, 51*(8), 1553-1575.

Rigobon, R., & Sack, B. (2008). Noisy macroeconomic announcements, monetary policy and asset prices. *NBER Working Paper*.

Rudebusch, G. (1998). Do Measures of Monetary Policy in a VAR Make Sense? *International Economic Review, 39*(4), 907-931.

Saito, Y., Nirei, M., Carvalho, V., & Tahbaz-Salehi, A. (2015). Supply Chain Disruptions: Evidence from the Great East Japan Earthquake. *Unpublished manuscript: University of Cambridge*.

Scott, J. (2017). *Social Network Analysis* (4 ed.). London: Sage Publications.

Su, Y. (2016). The reflection channel of shock transmission in the production network. *Unpublished manuscript: University of Chicago Booth School of Business*.

Swanson, E. (2017). Measuring the Effects of Federal Reserve Forward Guidance. *Working Paper 21816*.

Taylor, J. (2001). The role of the exchange rate in monetary-policy rules. *American Economic Review, 91*(2), 263-267.

Timmer, M. P., Los, B., Stehrer, R., & de Vries, G. J. (2016). An Anatomy of the Global Trade Slowdown based on the WIOD 2016 Release. *GGDC research memorandum number 162, University of Groningen*.

Timmer, M., Dietzenbacher, E., Los, B., Stehrer, R., & de Vries, G. (2015). An Illustrated User Guide to the World Input-Output Database: the Case of Global Automotive Production. *Review of International Economics, 23*(3), 575-605.

Timmer, M., Erumban, A., Los, B., Stehrer, R., & de Vries, G. (2014). Slicing Up the Global Value Chain. *Journal of Economic Perspectives, 28*(2), 99-118.

Timmer, M., Los, B., Stehrer, R., & de Vries, G. (2013). Fragmentation, Income and Jobs. An Analysis of European Competitiveness. *Economic Policy, 28*, 613-661.

Upton, G., & Fingleton, B. (1985). *Spatial Data Analysis by Example Vol 1: Point Pattern and Quantitative Data.* New York: Wiley.

Wassermann, S., & Faust, K. (1994). *Social Network Analysis: Methods and Applications.* Cambridge UK: Cambridge University Press.

Weber, M., & Ozdagli, A. (2017). Monetary Policy through Production Networks: Evidence from the Stock Market. *Working Paper*.

Yang, J., & Zhang, R. (2016). Network Centrality of Customers and Suppliers. *Working Paper*.

List of Figures

Figure 1: Production Network in 2014 by Country

The figure draws the input-output production network as of 2014. Each vertex corresponds to a sector defined by Table 15: Concordance Table. Directed links represent the share of input from industry j used in the production of industry i in country k. Only links, where the transaction is above 1 % of the total output of a sector , are drawn. Shorter links represent stronger connections. The following color code applies: BLUE-Austria, BLACK-Belgium, RED – Germanay, YELLOW – Spain, GREEN-Finland, BROWN – France, OCHRA – Greece, PINK – Ireland, ORANGE – Italy, WHITE – The Netherlands, PURPLE – Portugal. Kamada-Kawaii algorithm is used. The figure is produced using the Pajek software package.

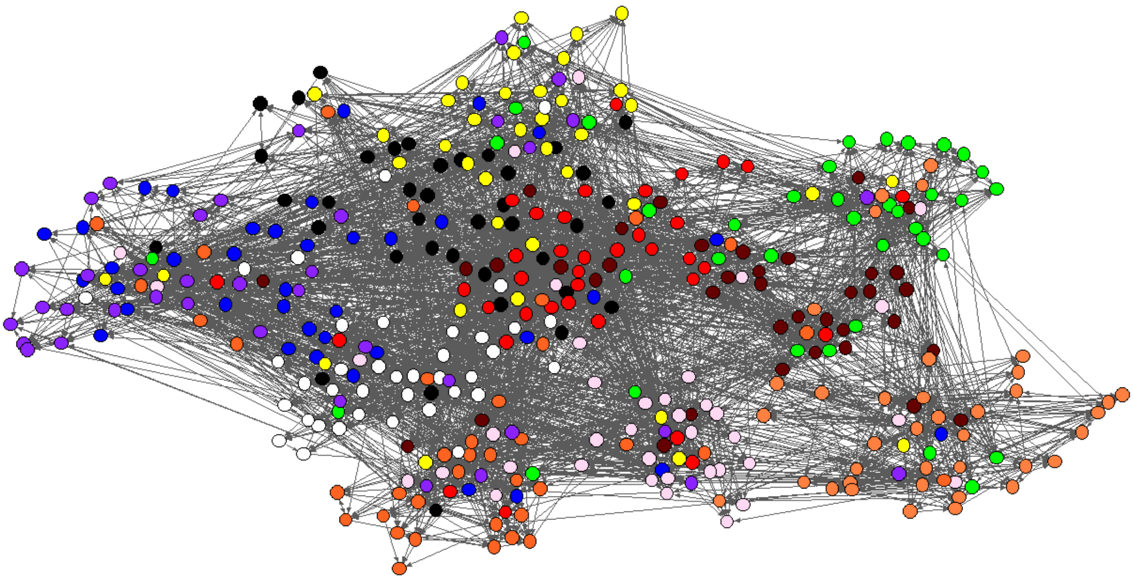


Figure 2: GSS Factors

The figure shows the two rotated GSS2005 factors: *target* and *path*

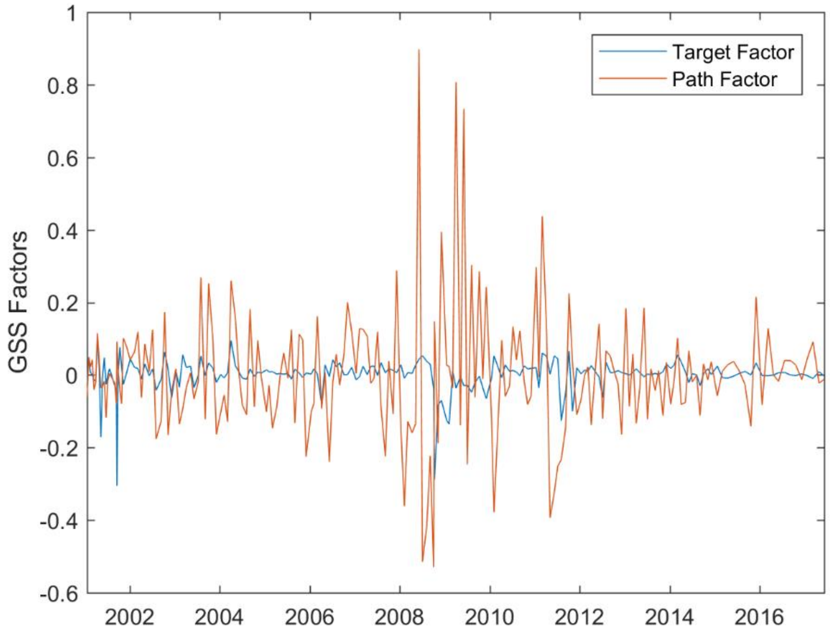


Figure 3: Empirical Density of Intermediate Output Shares (standard deviations)

The figure plots the empirical density of the standard deviation of intermediate output shares. The estimate is based on a normal kernel smoothing function.

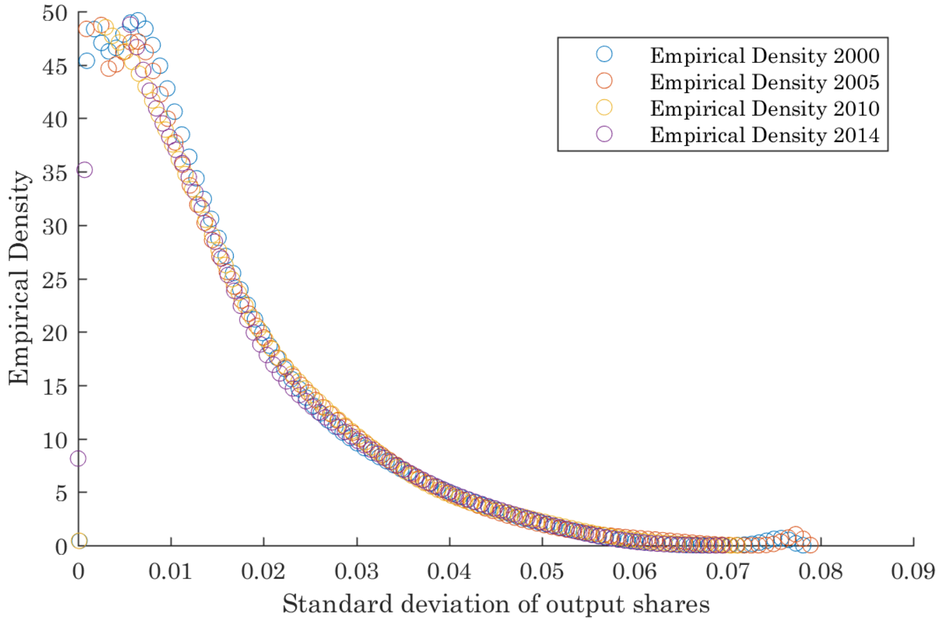


Figure 4: Moran's Scatterplot

The figure shows Moran's Scatterplot. For visualization purposes, only a 10-% sample of the population data is used.

Figure 5: Supplier Connections by Customer Size

The figure plots the contribution of customers of different sizes to average supplier sales. Very small customers are those that buy less than 1 % of total supplier sales; small customers between 1-5 %; medium customers between 5-20 % and large customers ≥ 20 %.

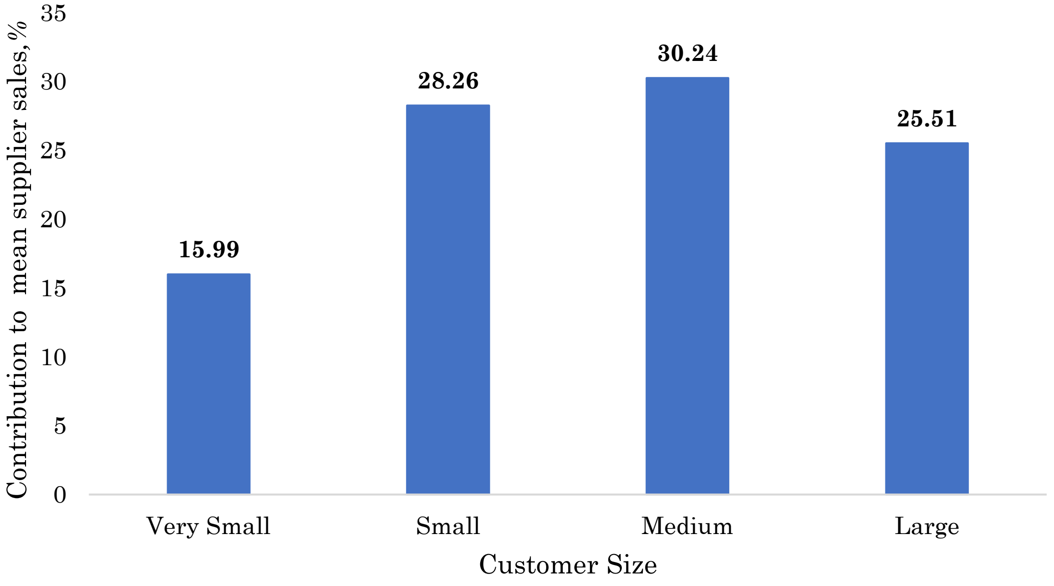


Figure 6: Home vs. Foreign Links

The figure plots links to “Home” and “Foreign” industries. Only connections above 1 % of total ouput are shown.

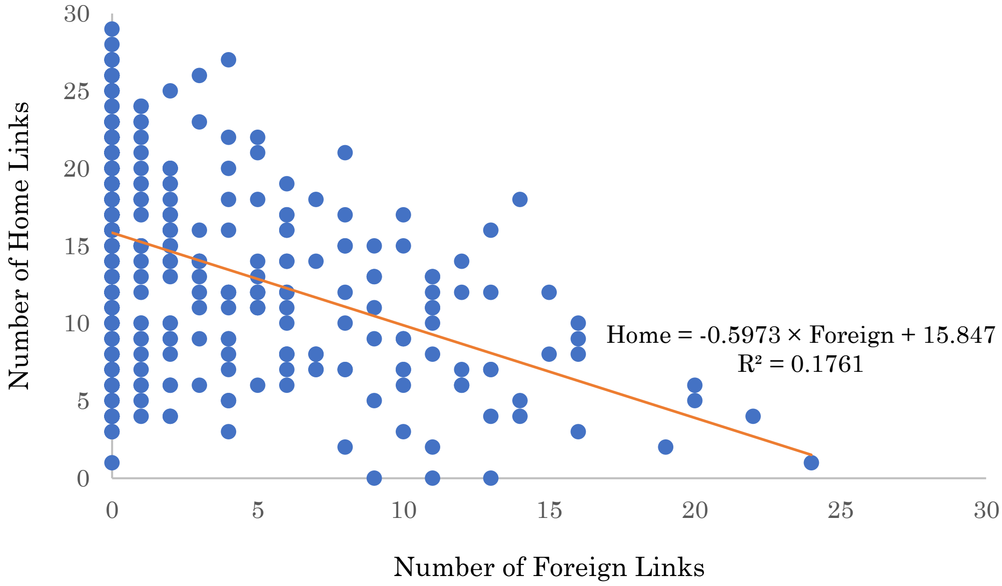
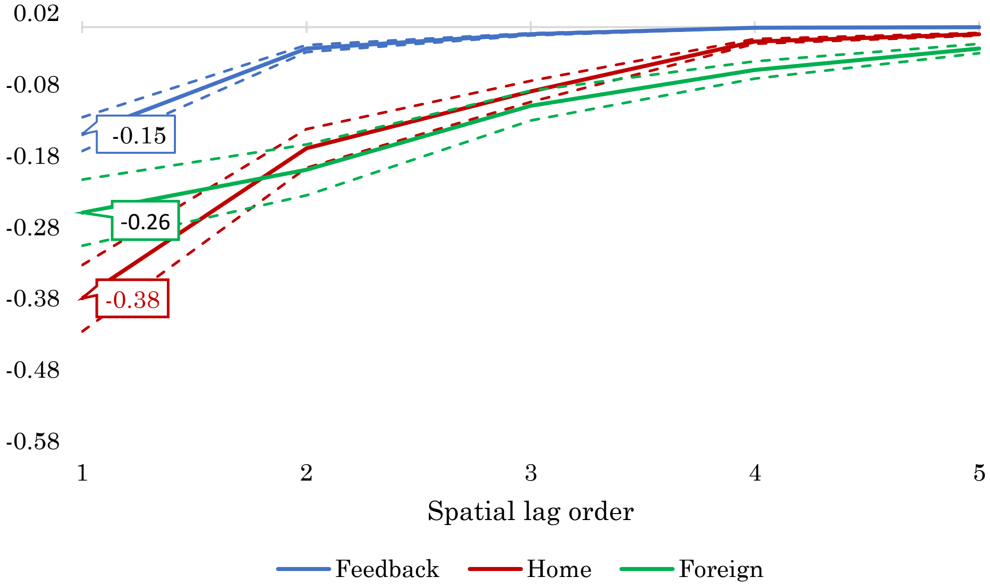


Figure 7: Higher-order Effects

The figure shows effects associated with spatial lag orders 1 to 5 for the target (Panel A) and the path (Panel B) factors. Lag orders correspond to immediate neighbors (*1st-order*), neighbors of immediate neighbors (*2nd-order*) and so on. Feedback, Home and Foreign Effects are plotted separately. Confidence bounds are given at the 95 % level.

Panel A: Target Factor



Panel B: Path Factor

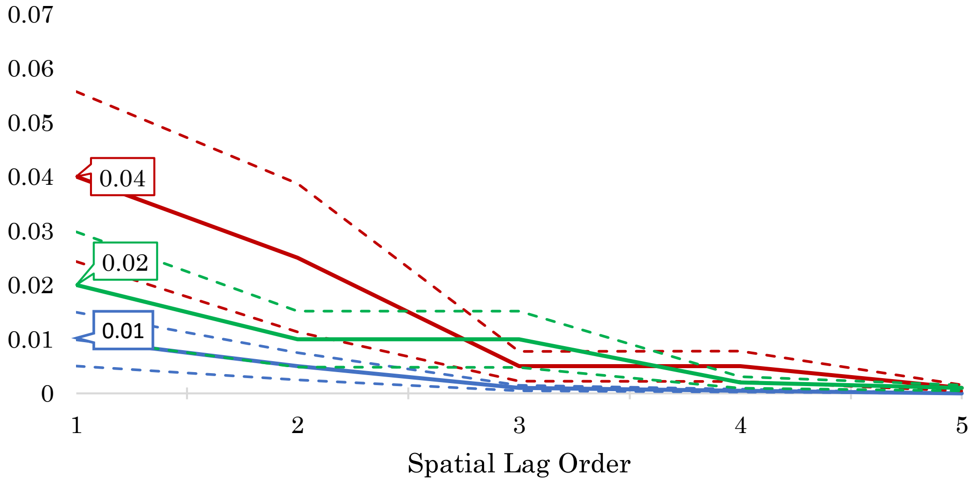


Figure 8: Histogram of Centrality

The figure shows a histogram of supplier centrality. Katz-Bonacich scores are computed using laggeg IO tables during the period 2001-2017.

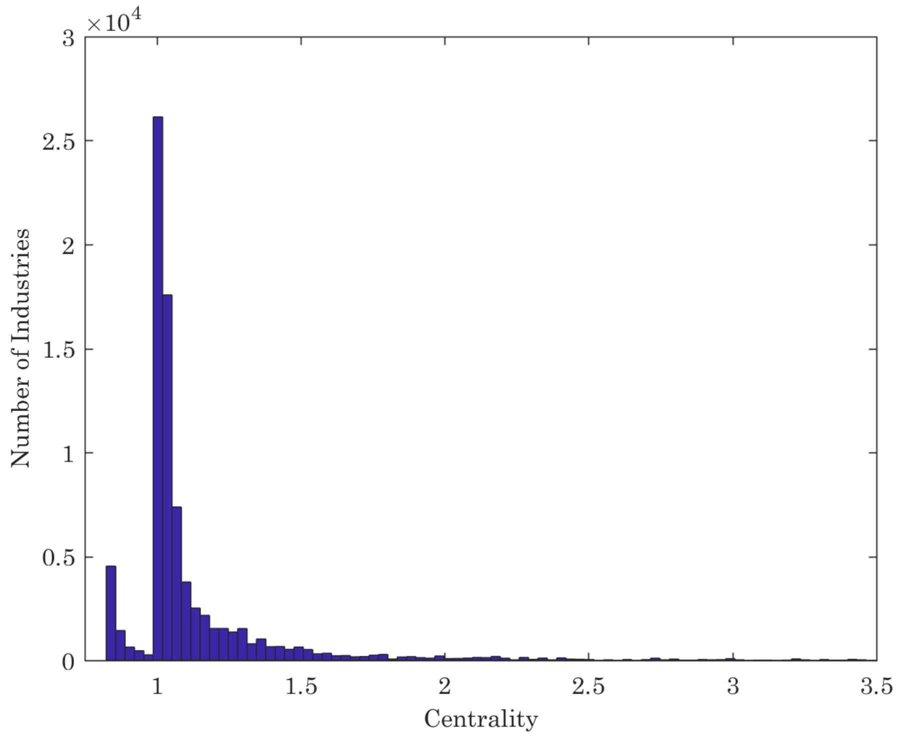


Figure 9: The Global Production Network

The figure draws the Global input-output production network as of 2014. Each vertex corresponds to a sector defined by Table 15: Concordance Table. Directed links represent the share of output of industry i used in the production of industry j in country k. Only links, where the transaction is above 10 % of the total output of a sector are drawn. Shorter links represent stronger connections. The following color code applies: PEACH- Eurozone; GREEN- the rest of the world. Kamada-Kawaii algorithm is used. The figure is produced using the Pajek software package.

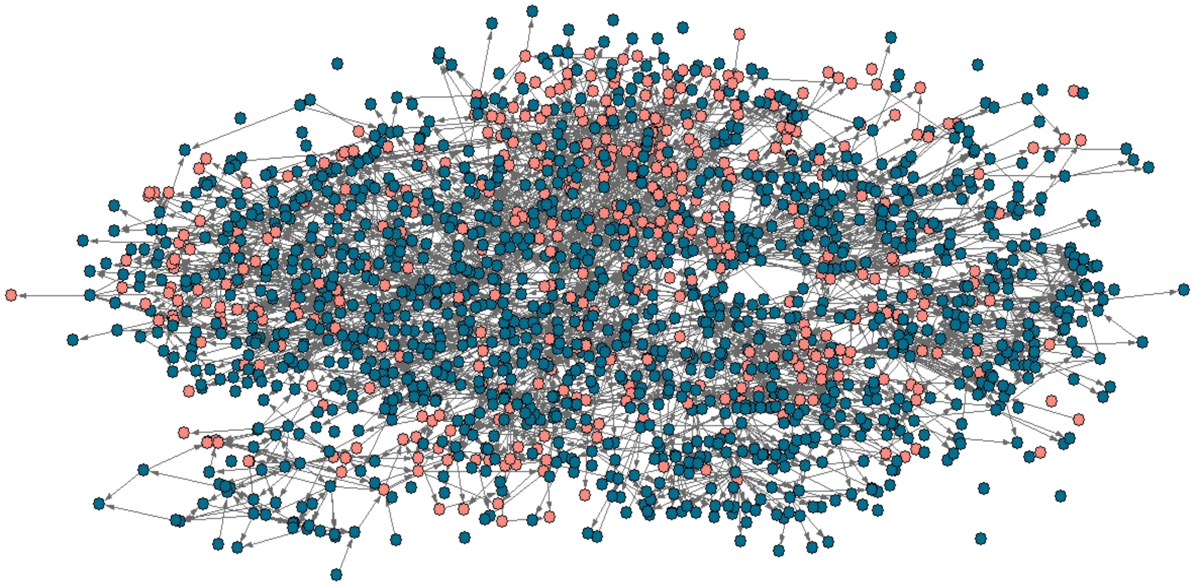
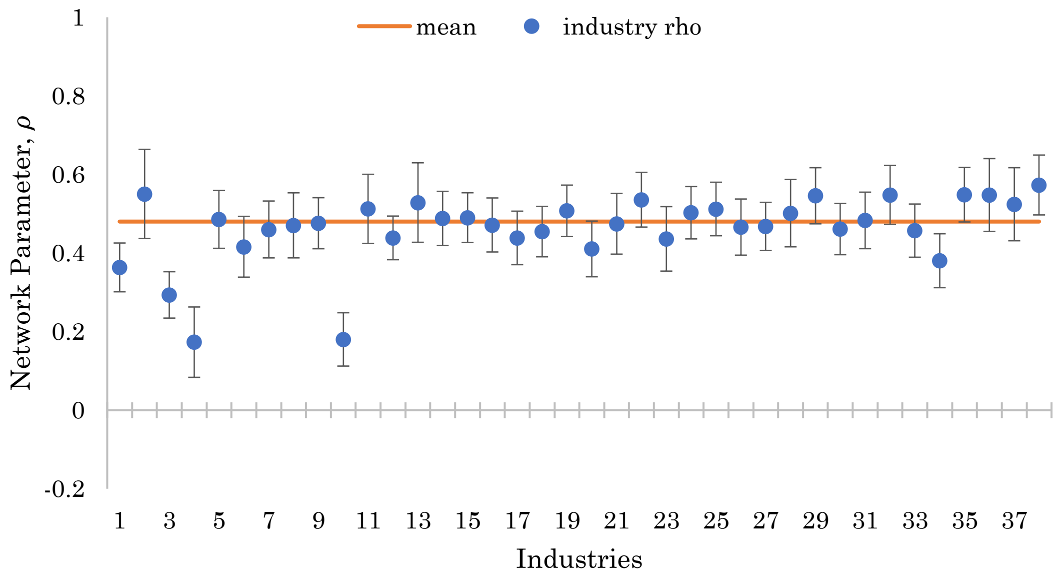


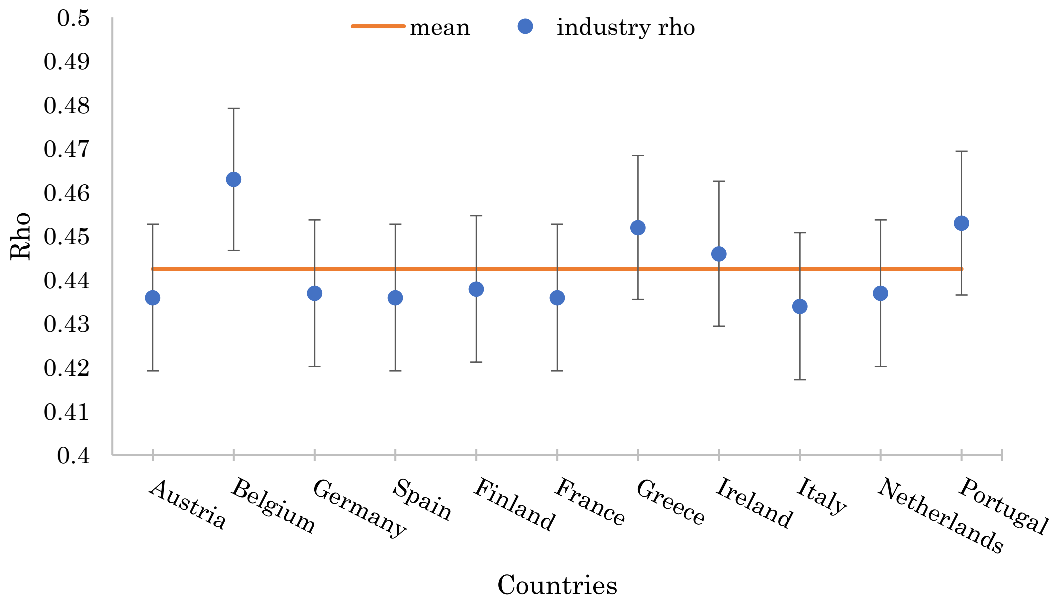
Figure 10: Parameter Heterogeneity

The figure plots the value of the spatial autoregressive parameter ρ and the 95 % confidence interval for each of the 38 industries (Panel A) and the 11 countries (Panel B) in the sample.

Panel A: Industries



Panel B: Countries



List of Tables

Table 1 Rotated Factor Loadings

The table gives the loadings of the GSS2005 factors, *target* and *path*, on a series of forward rates.

	f_1^0	f_1^1	f_3^3	f_3^6	f_3^9	f_6^{12}	f_6^{18}
Target	1.00	1.08	0.83	0.70	0.57	-0.08	-0.34
Path	0.00	0.13	0.29	0.47	0.57	0.50	0.51

Table 2 An Example of a World Input-Output Table

The table gives an example of a world input-output table with 2 countries and 2 industries.

			Use by country industries				**Final use by consumers**		**Total Use**
			Country 1		Country 2		Country1	Country2	
			Ind. 1	Ind. 2	Ind. 1	Ind. 2			
Supply from country industries	Country1	Ind. 1	1.34	8.23	13.40	0.50	4.15	2.00	**29.62**
		Ind. 2	3.90						
	Country2	Ind. 1	2.50						
		Ind. 2	12.42						
	Value added by labor and capital		9.46						
	Gross output		**29.62**						

Table 3: Summary Statistics

The table gives summary statistics for an average supplier industry during the 1-day time interval bracketing monetary policy announcements. Very small customers are those that buy less than 1 % of total supplier sales; small customers between 1-5 %; medium customers between 5-20 % and large customers ≥ 20 %.

	Mean	Standard Deviation	Min	Max
Panel A: Total Network Connections				
Number of non-zero connections	410.81	43.54	19	418
Number of links to very small customers	393.92	43.12	19	415
Number of links to small customers	12.98	5.79	0	30
Number of links to medium customers	3.18	1.73	0	9
Number of links to large customers	0.73	0.65	0	3
Panel B: Home vs. Foreign Connections				
Number of home links ≥ 1 % of sales	14.31	6.21	0	29
Number of foreign links ≥ 1 % of sales	2.58	4.36	0	24
% of foreign sales	15.68	21.72	0	99.97
Panel C: Industry Characteristics				
Log Size	8.30	1.59	2.89	11.89
Returns	0.00	0.061	-0.62	0.04
Pairwise return correlation	0.12	0.07	-0.05	0.29
Pairwise return correlation with biggest counterparty	0.16	0.17	-0.28	0.90

Table 4: Response of industry-level returns to monetary policy shocks

The table reports the results of regressing country-level industry returns in a 2-day time interval [-1,1] bracketing ECB announcements on monetary policy shocks and a spatial lag. Column (1) gives the OLS estimates of the standard model; columns (2)-(5) report maximum likelihood estimates for the spatial autoregressive model (SAR) Regressions controls for *size, value-added, inventory, investment, size composition* and *demand elasticity*. The zero lower bound (ZLB) date is 5th June 2014. Panel A reports point-estimates of the results; Panel B and Panel C offer decomposition of the total reaction into direct and network effects for the two factors. T-stats are reported in parentheses. Significance at the 1%, 5% and 10% is given by ***, ** and * respectively.

$$ret_{ijt} = \beta_0 + \rho W_{ij} ret_{ijt} + \beta_1 Target_t + \beta_2 Path_t + \lambda_i + \delta_j + \mu_{year} + \gamma controls_{ijt} + \varepsilon_{ijt}$$

	(1)OLS Standard Model	(2) SAR: W 2000	(3)SAR: Before the ZLB	(4)SAR: After the ZLB	(5)CAPM	(6)SAR: W^T Downstream
	Panel A: Point Estimates					
$\boldsymbol{\rho}$		**0.44*****	**0.42*****	**0.71****	**0.07*****	**0.22*****
		(77.85)	(65.14)	(59.14)	(21.34)	(33.87)
$Target$	-3.48***	-2.01***	-2.04***	-0.91	-3.45***	-3.02***
	(-19.92)	(-12.76)	(-12.91)	(-0.50)	(-19.71)	(-13.47)
$Path$	0.26***	0.17***	0.20***	-0.62**	0.30***	0.23***
	(5.76)	(4.02)	(4.60)	(-2.09)	(6.60)	(5.08)
Country, Ind. FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Controls	NO	YES	YES	YES	YES	YES
Adj R^2, %	0.51	21.02	19.94	24.82	8.64	6.86
N	418	418	418	418	418	418
T	198	198	172	26	198	198
Observations	82764	82764	71896	10868	82764	82764

	Panel B: Decomposition of Network Effects *Target* Factor					
Indirect Effect	N/A	-1.39***	-1.32***	-2.08	-0.25***	-0.75***
		(-12.54)	(-12.70)	(-0.52)	(-14.68)	(-15.01)
Feedback Effect	N/A	-0.19***	-0.19***	-0.18	-0.0001***	-0.14***
		(-12.64)	(-12.96)	(-0.51)	(-14.68)	(-18.17)
Total	-3.48***	-3.59***	-3.54***	-3.19	-3.70***	-3.91***
	(-19.92)	(-12.65)	(-13.01)	(-0.53)	(-20.09)	(-17.98)
Network/Total	NA	44.01%	42.65 %	70.84 %	6.76 %	22.76 %
	Panel C: Decomposition of Network Effects *Path* Factor					
Indirect Effect	N/A	0.11***	0.13***	-1.40**	0.02***	0.06***
		(3.94)	(4.55)	(-2.10)	(6.21)	(3.98)
Feedback Effect	N/A	0.02***	0.02***	-0.12**	0.0001***	0.01***
		(3.95)	(4.56)	(-2.09)	(6.22)	(3.67)
Total	0.26***	0.30***	0.34***	-2.13**	0.33***	0.30***
	(5.76)	(3.95)	(4.56)	(-2.10)	(6.20)	(4.28)
Network/Total	NA	43.33%	44.12 %	71.36 %	6.06 %	23.33 %

Table 5: Upstream Propagation of Shocks

The table reports the reaction of country-level industry returns, sorted by their proximity to the end-customer into terciles, to monetary policy shocks. See text for details on the sorting procedure. T-stats are reported in parentheses. Significance at the 1%, 5% and 10% is given by ***, ** and * respectively.

$ret_{ijt} = \beta_0 + \rho Wret_{ijt} + \beta_1 Target_t + \beta_2 Path_t + \lambda_i + \delta_j + \mu_{year} + \gamma controls_{ijt} + \varepsilon_{ijt}$

	Target				***Path***			
	Direct	Network	Total	Direct	Direct	Network	Total	Direct
Upstream	-1.53***	-1.93***	-3.46***	44.22	0.22***	0.17***	0.39**	56.41
	(-6.15)	(-6.08)	(-6.84)	.	(3.30)	(3.29)	(2.31)	.
Midstream	-2.02***	-1.68***	-3.70***	54.59	0.06	0.04	0.10	60.00
	(-7.19)	(-6.81)	(-6.91)	.	(0.82)	(0.87)	(0.82)	.
Downstream	-2.69***	-1.23***	-3.92***	68.66	0.22***	0.10***	0.32**	68.75
	(-9.52)	(-9.45)	(-9.54)	.	(3.01)	(3.02)	(3.01)	.

Table 6: Decomposition of Network Effects Across Borders

The table gives a decomposition of the indirect network effect across and within borders. The models estimated in column (2)-(5) correspond to the ones in Table 3. T-stats are reported in parentheses. Significance at the 1%, 5% and 10% is given by ***, ** and * respectively.

	(1) Baseline	(2) W^H and W^F
$\boldsymbol{\rho}$	**0.44*** **	**[0.39***, 0.86***]**
Panel D: Decomposition Across Borders *Target* Factor		
Home	-1.04 ***	-0.66***
	(-12.64)	(-11.54)
Foreign	-0.35***	-0.67***
	(-12.65)	(-10.84)
Foreign/Indirect	25.17 %	49.62 %
Foreign/Total	9.74 %	18.23 %
Panel E: Decomposition Across Borders *Path* Factor		
Home	0.08***	0.07***
	(3.95)	(3.72)
Foreign	0.03***	0.05***
	(3.97)	(3.70)
Foreign/Indirect	27.27 %	41.66 %
Foreign/Total	10.00 %	17.24 %

Table 7: Breakdown of Central Industries

The table presents a breakdown of the industries in the highest tercile of centrality by country and sector.

By Country, %		By Industry, %			
Austria	3.94	*Agriculture*	3.67	*Insurance*	2.03
Belgium	6.33	*Mining*	0.66	*Financial Services*	2.39
Germany	22.37	*Manufacturing*	38.21	*Real Estate*	4.04
Spain	14.12	*Utilities*	2.65	*Professional Services*	3.27
Finland	0.96	*Construction*	3.08	*Administration & Education*	3.59
France	19.27	*Wholesale & Retail Trade*	3.93	*Health*	2.86
Greece	2.62	*Transport, Warehousing, Courier*	15.43	*Arts, Entertainments & Recreation, Others*	3.76
Ireland	0.59	*Accommodation & Services*	0.28		
Italy	18.16	*Telecommunications*	1.42		
The Netherlands	9.48	*Computer & IT*	5.57		
Portugal	2.16	*Banks*	3.12		

Table 8: Heterogeneous Effects

The table reports the results of regressing event returns on monetary policy shocks, spatial lag and centrality/local demand. Bonacich centrality scores are used. T-stats are reported in parentheses. Significance at the 1%, 5% and 10% is given by ***, ** and * respectively.

$$ret_{ijt} = \beta_0 + \rho Wret_{ijt} + \theta Wret_{ijt} \times X_{it} + \beta_1 Target_t + \beta_2 Path_t + \beta_3 X_{it} + \beta_4 Target_t \times X_{ijt} + \beta_5 Path_t \times X_{ijt} + \delta Controls_{ijt} + \varepsilon_{ijt}$$

	(1) $X =$ $centrality$	(2) $X =$ $centrality$	(3) $X = local$ $demand$
$\boldsymbol{\rho}$	**0.46*** **	**0.44*** **	**0.42*** **
	(78.35)	(81.25)	(71.39)
$\boldsymbol{\theta}$	**0.03*** **	**0.029*** **	**0.048*** **
	(4.53)	(4.32)	(8.77)
$Target$	-1.89*** (-11.81)	-2.13*** (-11.87)	-2.18*** (-13.12)
$Path$	0.16*** (3.92)	0.15*** (3.11)	0.17*** (3.95)
X	-0.001 (-0.96)	0.006 (0.55)	-0.02** (-1.96)
$Target \times X$	0.58*** (3.48)	0.59*** (2.64)	-0.13** (-2.01)
$Path \times X$	0.002 (0.45)	0.001 (0.67)	0.027 (0.58)
$Controls$	NO	YES	YES
Year, Industry FE	NO	YES	YES
Adj R^2, %	21.09	20.81	21.17
Observations	82764	82764	82764

Table 9: Network Centrality: Decompositions

The table reports the effect of 1 standard deviation increase in centrality ($\Delta\sigma_{CNTR}$) on the sensitivity of industry returns to monetary policy surprise factor *Target*.

		Target		
$\Delta\sigma_{CNTR}$	$\rho + \theta$	Direct	Network	Total
Mean	0.44	-2.13	-1.67	-3.80
1	0.47	-1.54	-1.36	-2.90
2	0.50	-0.95	-0.94	-1.89
3	0.53	-0.36	-0.40	-0.77
4	0.56	0.23	0.28	0.51
5	0.59	0.82	1.15	1.97

Table 10: Are central industries less risky?

This table presents mean equally-weighted industry monthly returns sorted by Bonacich centrality for the period January 2001-June 2017. α_{CAPM} and α_{3FF} refer to abnormal returns with respect to the CAPM model and the Fama-French 3-factor model. Statistical significance at the 10%, 5%, and 1% level is indicated by *, ** and ***.

	Centrality			
	Low	*2*	*High*	*H-L*
Panel A: Excess Returns				
$R - R_f$	1.14***	1.18***	1.74***	0.60***
	(21.48)	(24.92)	(35.84)	(8.57)
α_{CAPM}	0.85***	0.92***	1.54***	0.69***
	(17.04)	(20.98)	(32.87)	(9.74)
α_{3FF}	0.69***	0.81***	1.46***	0.77***
	(13.83)	(18.14)	(30.90)	(7.75)
Panel B: Industry Characteristics				
Centrality	0.68***	1.21***	3.57***	2.89***
$\times 10^1$	(10.77)	(82.11)	(23.33)	(18.91)
LogSize	14.56***	14.94***	14.97***	0.41***
$\times 10^1$	(9.76)	(11.31)	(12.45)	(2.196)

Table 11: Eurozone Demand By Centrality Tercile

The table reports local (within Eurozone) end-customer and intermediates demand as a percentage of total output by centrality tercile. The measure of centrality used here is Bonacich centrality. T-stats are reported in parentheses. Significance at the 1%, 5% and 10% is given by ***, ** and * respectively.

	Centrality			
	Low	*(2)*	*High*	*L-H*
$\frac{\text{Local Demand}}{\text{Output}}, \%$	38.56***	28.04***	20.90***	17.66***
	(10.53)	(12.84)	(13.06)	(8.88)
$\frac{\text{Local Intermediates}}{\text{Output}}, \%$	53.88***	50.19***	39.17***	14.71***
	(24.56)	(28.56)	(20.47)	(20.01)

Table 12: Robustness: Identification

The table reports the results of regressing country-level industry returns in a 1-day time interval [-1,1] bracketing ECB announcements on monetary policy shocks and a spatial lag. The table reports maximum likelihood estimates for the spatial autoregressive model (SAR). Panel A reports point-estimates of the results; Panels B and C offer decompositions. T-stats are reported in parentheses. Significance at the 1%, 5% and 10% is given by ***, ** and * respectively.

$$ret_{ijt} = \beta_0 + \rho W ret_{ijt} + \beta_1 Target_t + \beta_2 Path_t + \lambda_i + \delta_j + \mu_{year} + \gamma controls_{ijt} + \varepsilon_{ijt}$$

	(1) Initial Jobless Claims	(2) Rigobon IV	(3) Random Events	(4)Industry-demeaned	(5) VW returns
		Panel A: Point Estimates			
$\boldsymbol{\rho}$	**0.48*** **	**0.68*** **	**0.69*** **	**0.42*** **	**0.53*** **
	(53.24)	(72.31)	(46.42)	(67.92)	(86.39)
Target	-1.26***	-0.68***	0.76	-2.09***	-1.62***
	(-4.38)	(-6.02)	(0.45)	(-13.20)	(-10.39)
Path	0.11*	0.11***	1.40	0.18***	0.18***
	(2.03)	(4.06)	(0.18)	(4.22)	(4.89)
Constant	-0.00	-0.00	0.00	-0.00	0.00
	(-1.37)	(-0.53)	(0.01)	(-0.00)	(0.37)
Adj R^2, %	21.64	27.59	28.50	19.68	11.96
N	418	418	418	418	418
T	95	396	198	198	198
Observations	39710	165528	82764	82764	82764

Panel B: Decomposition of Network Effects: *Target*					
Indirect Effect	-1.02***	-1.36***	1.61	-1.31***	-1.62***
	(-4.27)	(-6.12)	(0.47)	(-13.29)	(-10.59)
Feedback Effect	-0.13***	-0.13***	0.15	-0.19	-0.20***
	(-4.28)	(-6.13)	(0.45)	(-13.44)	(-10.64)
Total Effect	-2.41***	-2.16***	2.52	-3.59***	-3.44***
	(-4.29)	(-6.13)	(0.43)	(-13.47)	(-12.60)
Network/Total	47.71 %	68.98%	69.84 %	41.78 %	52.90%
Panel B: Decomposition of Network Effects: *Path Factor*					
Indirect Effect	0.08**	0.22***	2.93	0.11***	0.18***
	(2.02)	(4.15)	(0.17)	(4.25)	(3.67)
Feedback Effect	0.01**	0.02***	0.27	0.02***	0.02***
	(2.02)	(4.14)	(0.18)	(4.25)	(3.67)
Total Effect	0.19**	0.35***	4.60	0.30***	0.39***
	(2.02)	(4.16)	(0.15)	(4.26)	(3.68)
Network/Total	47.36 %	68.57%	65.31%	43.33 %	51.28%

Table 13: Robustnesss: Spatial Weights Matrix

The table reports the results of regressing country-level industry returns in a 1-day time interval [-1,1] bracketing ECB announcements on monetary policy shocks and a spatial lag. See text on details about the construction of different *W*. Panel A reports point-estimates of the results; Panel B and Panel C offer decomposition of the total reaction into direct and indirect (network) effects for the two factors. T-stats are reported in parentheses. Significance at the 1%, 5% and 10% is given by ***, ** and * respectively.

$$ret_{ijt} = \beta_0 + \rho W ret_{ijt} + \beta_1 Target_t + \beta_2 Path_t + \lambda_i + \delta_j + \mu_{year} + \gamma controls_{ijt} + \varepsilon_{ijt}$$

	W does not sum up to 1	Zero-diagonal W	Simulated W	Time-varying W
	(2)	(3)	(4)	(5)
		Panel A: Point Estimates		
$\boldsymbol{\rho}$	**1.11*** **	**0.52*** **	**0.095*** **	**0.48*** **
	(74.94)	(85.10)	(24.12)	(78.34)
$Target$	-1.74***	-1.79***	-3.11***	-1.96***
	(-11.06)	(-10.61)	(-18.06)	(-12.42)
$Path$	0.15***	0.15***	0.26***	0.16***
	(3.67)	(3.38)	(5.21)	(3.91)
Constant	0.00***	0.00**	0.00***	0.00
	(3.18)	(2.52)	(3.90)	(1.24)
Adj R^2, %	20.49	9.11	1.32%	20.81
N	418	418	418	418
T	198	198	198	198
Observations	82764	82764	82764	82764

Panel B: Decomposition of Network Effects: Target				
$Indirect\ Effect$	-1.66***	-1.78***	-0.33***	-1.61***
	(-11.51)	(-11.07)	(-14.59)	(-12.69)
$Feedback\ Effect$	-0.20***	-0.02***	-0.001***	-0.21***
	(-11.63)	(-10.86)	(-18.42)	(-12.59)
$Total$	-3.60***	-3.62**	-3.49***	-3.77***
	(-11.65)	(-11.22)	(-18.49)	(-12.77)
$Network/Total$	51.66 %	49.73 %	9.62 %	48.27%
Panel C: Decomposition of Network Effects: Path				
$Indirect\ Effect$	0.15***	0.16***	0.03***	0.13***
	(3.68)	(3.48)	(5.13)	(4.26)
$Feedback\ Effect$	0.02***	0.001***	0.001***	0.02***
	(3.68)	(3.47)	(5.74)	(3.98)
$Total$	0.32***	0.31***	0.29***	0.31***
	(3.86)	(3.48)	(5.27)	(4.28)
$Network/Total$	53.12%	51.61 %	10.34 %	48.38%

Table 14: Robustness: Centrality Measures

The table reports the results of regressing country-level industry returns in a 2-day time interval [-1,1] bracketing ECB announcements on monetary policy shock, centrality and an interaction term. Column (1) uses eigenvector centrality; column (2) uses closeness centrality; column (3) uses customer centrality. Centrality and all controls are standardized. T-stats are reported in parentheses. Significance at the 1%, 5% and 10% is given by ***, ** and * respectively.

$$ret_{ijt} = \beta_0 + \rho Wret_{ijt} + \theta Wret_{ijt} \times X_{ijt} + \beta_1 Target_t + \beta_2 Path_t + \beta_3 X_{ijt} + \beta_4 Target_t \times X_{ijt} + \beta_5 Path_t \times X_{ijt} + \delta Controls_{ijt} + \varepsilon_{ijt}$$

	Eigenvector (1)	Outdegree (2)	(3) Closeness
$\boldsymbol{\rho}$	**0.42*** **	**0.43*** **	**0.42*** **
	(69.00)	(72.97)	(71.39)
$\boldsymbol{\theta}$	**0.032*** **	**0.038*** **	**0.037*** **
	(5.70)	(6.85)	(8.14)
$Target$	-2.27*** (-12.89)	-2.19*** (-12.57)	-2.29*** (-12.95)
$Path$	0.15*** (3.21)	0.15*** (3.81)	0.16*** (3.30)
$Centrality$	-0.007 (-0.92)	-0.03 (-0.89)	0.007 (0.88)
$Target \times Centr$	0.36** (2.21)	0.70* (1.65)	0.23 (2.26)
$Path \times Centr$	0.02 (0.57)	-0.02 (-0.30)	0.07 (0.77)
Controls	YES	YES	YES
Country, Ind. FE	YES	YES	YES
Year FE	YES	YES	YES
Adj R^2, %	20.14	20.69	19.47
Observations	82764	82764	82764

Table 15: Propagation of Shocks from Influential Industries

The table reports network effects from monetary policy shocks for industries connected to at least one influential customer, where the connection is strong and significant. Influential customers are defined as industries with above-median *centrality* and above-median *Eurozone demand*. T-stats are reported in parentheses. Significance at the 1%, 5% and 10% is given by ***, ** and * respectively.

	Not connected to influential	*Connected to Influential*	*C-NC*
	Panel A: First-order Neighbors		
Target	-1.38***	-3.31***	-1.93***
	(-9.14)	(-8.98)	(-4.91)
Path	0.11***	0.31***	0.20**
	(2.67)	(3.28)	(1.99)
	Panel B: Second-order Neighbors		
Target	-1.22***	-1.83***	-0.61***
	(-8.02)	(-10.12)	(-5.23)
Path	0.10***	0.14***	0.04
	(2.66)	(3.03)	(1.56)
	Panel C: Third-order Neighbors		
Target	-0.95***	-1.01***	-0.06***
	(-5.52)	(-10.32)	(-6.36)
Path	0.08*	0.08***	0.00**
	(1.85)	(3.09)	(1.87)

Table 16: Asset Pricing Factors

The table reports the results of regressing country-level industry returns in a 2-day time interval [-1,1] bracketing ECB announcements on monetary policy shock, network concentration/sparsity and an interaction term. T-stats are reported in parentheses. Significance at the 1%, 5% and 10% is given by ***, ** and * respectively.

$$ret_{ijt} = \beta_0 + \rho Wret_{ijt} + \theta Wret_{ijt} \times X_t + \beta_1 Target_t + \beta_2 Path_t + \beta_3 X_t + \delta Controls_{ijt} + \varepsilon_{ijt}$$

	(1) Concentration	(2)Sparseness
ρ	**0.44*** **	**0.42*** **
	(75.16)	(68.40)
θ	**-0.0469*** **	**0.03**
	(-8.50)	(0.59)
Target	-1.72*** (-10.65)	-2.09*** (-13.13)
Path	0.18*** (4.44)	0.18*** (41.21)
Concentration	0.002 (0.31)	.
Sparseness	.	-0.007 (-1.03)
Controls	YES	YES
Country, Ind. FE	YES	YES
Year FE	YES	YES
Adj R^2, %	20.14	20.69
Observations	82764	82764

Appendix

A. Spatial Autoregressions

To understand how the spatial lag is formed consider the following linkages between industries are given by the matrix A:

$$A = \begin{pmatrix} 0 & 4 & 3 \\ 0 & 1 & 3 \\ 1 & 2 & 1 \end{pmatrix}$$

Row-normalizing A, the spatial weights matrix W is constructed:

$$W = \begin{pmatrix} 0 & 4/7 & 3/7 \\ 0 & 1/4 & 3/4 \\ 1/4 & 2/4 & 1/4 \end{pmatrix}$$

Having defined the spatial weights matrix W, it is easy to see that the spatial lag Wy corresponds to:

$$Wret = \begin{pmatrix} 0 & 4/7 & 3/7 \\ 0 & 1/4 & 3/4 \\ 1/4 & 2/4 & 1/4 \end{pmatrix} \begin{pmatrix} ret_{1t} \\ ret_{2t} \\ ret_{3t} \end{pmatrix} = \begin{pmatrix} (4ret_{2t} + 3ret_{3t})/7 \\ (ret_{2t} + 3ret_{3t})/4 \\ (ret_{1t} + 2ret_{3t} + ret_{3t})/4 \end{pmatrix}$$

B. Monetary Policy Shocks

This section contains a detailed description of the procedure used to derive the monetary policy shock factors *target* and *path* using the methodology by Gürkaynak, Sack and Swanson (2005). Slight modifications have been introduced to fit the research purpose of this paper.

Let $F = [F_1 F_2]$ give the first two principal components and $Z = [Z_1 Z_2]$ denote the two rotated factors. The rotated factors are connected to the original factors by the following relationship:

$$Z = FU \tag{1}$$

where

$$U = \begin{pmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{pmatrix} \tag{2}$$

The matrix U is identified by the following three restrictions:

1) Z_1 and Z_2 are orthogonal:

$$\alpha_1\beta_1 v_1 + \alpha_2\beta_2 v_2 = 0 \tag{3}$$

with $v_1 = E(F_1{}'F_1)$ and $v_2 = E(F_2{}'F_2)$

2) Z_1 and Z_2 have variance equal to one:

$$\alpha_1^2 v_1 + \alpha_2^2 v_2 = 1 \tag{4}$$

$$\beta_1^2 v_1 + \beta_2^2 v_2 = 1 \tag{5}$$

3) Z_2 is orthogonal to Δf_1^0. Let γ_1 and γ_2 denote the loading of Δf_1^0 on F_1 and F_2. It is immediate to see that:

$$\gamma_2\alpha_1 - \gamma_1\alpha_2 = 0 \tag{6}$$

Solving for (3)-(6) for α_2 and α_1, the following expressions obtain:

$$\alpha_2 = \sqrt{\frac{1}{\left(\left(-\frac{\gamma_1}{\gamma_2}\right)^2 v_1 + v_2\right)}}$$

$$\alpha_1 = \left(\frac{\gamma_1}{\gamma_2}\right)\alpha_2$$

Substitute to find β_1 and β_2:

$$\beta_2 = \sqrt{\frac{\alpha_1^2 v_1}{\alpha_2^2 v_2^2 + \alpha_1^2 v_1 2_2}}$$

$$\beta_1 = \frac{-\alpha_2\beta_2 v_2}{\alpha_1 v_1}$$

Finally, to facilitate interpretation, of the rotated factors and Z_2 are rescaled to Z_1^* and Z_2^* such that Z_1^* moves one-to-one with Δf_1^0 and the sensitivity of Δf_3^9 to Z_1^* and Z_2^* are the same. Let ζ_0, ζ_1, φ_0 and φ_1 be coefficients from the following regressions:

$$\Delta f_1^0 = \zeta_0 + \zeta_1 Z_1, t + \varepsilon_{1,t} \tag{7}$$

$$\Delta f_3^9 = \varphi_0 + \varphi_1 Z_1^* + \varphi_2 Z_2 + \varepsilon_{2,t} \tag{8}$$

Then, $Z_1^* = \zeta_1 Z_1$ and $Z_2^* = \frac{\varphi_2}{\varphi_1} Z_2$.

C. Data and Network Construction

I.WIOT Database

The World Input-Output Database (WIOD) provides annual time-series of world input-output tables (WIOT) from 1995 onwards (Timmer, Dietzenbacher, Los, Stehrer, & de Vries, 2015). The project was funded by the European Commission as part of the 7th Framework Programme. Data is publicly available online[30]. It covers 44 countries, including all members of the EU, 13 other major economies: Australia, Brazil, Canada, China, India, Indonesia, Japan, Mexico, Russia, South Korea, Taiwan, Turkey and the USA and a model for the rest of the world. The countries in the sample account for more than 85% of the world gross domestic product (GDP) as of 2008. WIOTs are constructed using data on national supply and use tables (SUTs), merged with national accounts data and connected to each other by bilateral international trade flows. The tables are constructed in accordance with the International System of National Accounts and obey its rules and accounting identities[31]. WIOTs have an industry-by-industry format and provide details about 56 industries, classified according to the International Standard Industrial Classification (ISIC) revision 4[32].

The most important accounting identity in the WIOT is that the gross output of an industry is equal to the sum of the uses of the output from that industry. All values in the WIOTs are given in millions of US dollars. Transaction values are given in basic prices, which reflect costs borne by the producer. International trade flows are expressed as "free-on-board" prices. In addition to input-output data, WIOD provides information on quantity and prices of factor inputs such as data on workers and wages by level of education, capital stocks and value-added. This data is provided by the Socio-economic Accounts.

II. Matching Procedure

[30] http://www.wiod.org The latest release is from 2016.

[31] For a detailed overview of sources and characteristics the reader is advised to Timmer, M. P., Los, B., Stehrer, R. and de Vries, G. J., (2016).

[32] https://unstats.un.org/unsd/cr/registry/regcst.asp?Cl=27

One of the most challenging tasks in the data collection process is how to match the input-output production matrix from WIOD[33] to stock market data on a country-industry level. This is done in three steps. *First*, daily stock prices on a company level are obtained from Datastream from all stock market exchanges in the 11 countries. Each company is associated with a "home country" based on the country where the headquarters is located. In the cases where a company is listed on more than one stock exchange, data from the exchange market in the country of incorporation is used. *Second*, Datastream identifies the industrial sector of a company through its 4-digit Industry Classification Benchmark Code (ICB code). The ICB system uses 41 sectors, which are comparable, but not identical, to the ISIC system used by WIOD. Using a concordance table designed specifically for this project, I match the 56 ISIC Industry Sectors to the 41 ICB sectors. To prevent bias in the data, I require that each sector in each country is populated by at least 2 companies. To fulfill this requirement, the number of sectors is collapsed to 38. *Appendix A: Concordance Table* describes the procedure in detail. Finally, an equally-weighted return index using stock market data for all companies matched to a given country-industry unit is created. When time-series data is poor or not available, returns are proxied by an *aggregate* industry index available from Datastream[34].

III. Concordance Table

This section contains the concordance table used to translate the ICB industry codes to the ISIC codes used by the WIOD. The only ISIC industries that are not matched to ICB codes from DATASTREAM are: *Activities of Households as Employers* (T) and *Activities of Extraterritorial Organizations* (U).

Table 17: Concordance Table

---

[33] The countries in the sample are Austria, Belgium, Germany, Spain, Finland, France, Greece, Ireland, Italy, the Netherlands and Spain. The choice of the countries is dictated by data quality and availability so that a balanced panel can be fed into the SAR model.

[34] In principle, such an aggregate return index can be used for all industries. However, these indices are based on large-cap companies and, hence, it would introduce a measurement error in the data.

This table gives the concordance between the ISIC codes and industries used in WIOD and the ICB codes and industries from DATASTREAM.

ISIC Industry	ISIC Number	Name of the ISIC Industry	ICB Code	Industry Number
A01	1	Farming	3573	1
A02	2	Forestry and logging	1733	1
A03	3	Fishing	.	
B	4	Mining and quarrying	1753	2
B	4	Mining and quarrying	1771	2
B	4	Mining and quarrying	1773	2
B	4	Mining and quarrying	1775	2
B	4	Mining and quarrying	1777	2
B	4	Mining and quarrying	1779	2
C10-C12	5	Manufacture of foods, beverages, tobacco	3577	3
C10-C12	5	Manufacture of foods, beverages, tobacco	3533	3
C10-C12	5	Manufacture of foods, beverages, tobacco	3535	3
C10-C12	5	Manufacture of foods, beverages, tobacco	3537	3
C10-C12	5	Manufacture of foods, beverages, tobacco	3785	3
C13-C15	6	Manufacture of textiles, apparel and leather	3763	4
C13-C15	6	Manufacture of textiles, apparel and leather	3765	4
C16-C17	7:8	Manufacture of paper	1737	5
C18	9	Printing and reproduction of recorded media	5557	6
C19	10	Manufacture of coke and petroleum products	533	7
C19	10	Manufacture of coke and petroleum products	537	7
C20	11	Manufacture of chemical products	1353	7
C21	12	Manufacture of pharmaceuticals	4577	8
C21	12	Manufacture of pharmaceuticals	3767	8
C21	12	Manufacture of pharmaceuticals	4537	8
C22	13	Manufacture of rubber and plastic	1357	9
C23	14	Manufacture of non-metallic mineral products	2353	10
C24	15	Manufacture of basic metals	1755	11
C24	15	Manufacture of basic metals	1757	11
C25	16	Manufacture of fabricated metal products	2717	11
C26	17	Manufacture of computer, electronic and optical products	2737	12
C26	17	Manufacture of computer, electronic and optical products	3743	12
C26	17	Manufacture of computer, electronic and optical products	9572	12
C26	17	Manufacture of computer, electronic and optical products	9574	12
C26	17	Manufacture of computer, electronic and optical products	9576	12
C27	18	Manufacture of electrical products	2733	13
C28	19	Manufacture of machines and equipment	573	14
C28	19	Manufacture of machines and equipment	583	14

C28	19	Manufacture of machines and equipment	2757	14
C28	19	Manufacture of machines and equipment	4535	14
C28	19	Manufacture of machines and equipment	2797	14
C28	19	Manufacture of machines and equipment	9578	14
C29	20	Manufacture of motor vehicles	3353	15
C29	20	Manufacture of motor vehicles	3355	15
C29	20	Manufacture of motor vehicles	3357	15
C29	20	Manufacture of motor vehicles	2753	15
C30	21	Manufacture of other transport equipment (aircraft)	2713	16
C31-C32	22	Manufacture of furniture & others	3726	17
C31-C32	22	Manufacture of furniture & others	3722	17
C31-C32	22	Manufacture of furniture & others	3724	17
C31-C32	22	Manufacture of furniture & others	3745	17
C31-C32	22	Manufacture of furniture & others	3747	17
C33	23	Installation of machinery and equipment	2727	18
D35	24	Electricity, gas, steam and air-conditioning supply	587	19
D35	24	Electricity, gas, steam and air-conditioning supply	7535	19
D35	24	Electricity, gas, steam and air-conditioning supply	7537	19
D35	24	Electricity, gas, steam and air-conditioning supply	7573	19
D35	24	Electricity, gas, steam and air-conditioning supply	7575	19
E36	25	Water collection, treatment and supply	7577	19
E37-E39	26	Sewage and waste collection and management	2799	19
F41-F43	27	Construction	2357	20
F41-F43	27	Construction	3728	20
G45-G47	28:30	Wholesale and retail trade	5333	21
G45-G47	28:30	Wholesale and retail trade	5337	21
G45-G47	28:30	Wholesale and retail trade	5371	21
G45-G47	28:30	Wholesale and retail trade	5373	21
G45-G47	28:30	Wholesale and retail trade	5375	21
G45-G47	28:30	Wholesale and retail trade	5377	21
G45-G47	28:30	Wholesale and retail trade	5379	21
H49	31	Land transport and transport via pipelines	577	22
H49	31	Land transport and transport via pipelines	2775	22
H49	31	Land transport and transport via pipelines	2779	22
H50	32	Water transport	2773	23
H51	33	Air transport	5751	24
H52	34	Warehousing and support activities for transportation	2777	25
H53	35	Postal and courier activities	2771	26
H53	35	Postal and courier activities	2723	26
I55-I56	36	Accommodation and food service activities	5753	27
I55-I56	36	Accommodation and food service activities	5757	27
J58-J60	37:38	Media, publishing and entertainment	5553	28
J58-J60	37:38	Media, publishing and entertainment	5555	28
J61	39	Telecommunications	6535	29

J61	39	Telecommunications	6575	29
J62-J63	40	Computer programming; information service	9533	30
J62-J63	40	Computer programming; information service	9535	30
J62-J63	40	Computer programming; information service	9537	30
K64	41	Finance; banks	8355	31
K65	42	Insurance	8532	32
K65	42	Insurance	8534	32
K65	42	Insurance	8536	32
K65	42	Insurance	8538	32
K65	42	Insurance	8575	32
K66	43	Financial services	8771	33
K66	43	Financial services	8773	33
K66	43	Financial services	8775	33
K66	43	Financial services	8777	33
K66	43	Financial services	8779	33
K66	43	Financial services	8985	33
K66	43	Financial services	8995	33
L68	44	Real estate	8633	34
L68	44	Real estate	8637	34
L68	44	Real estate	8671	34
L68	44	Real estate	8672	34
L68	44	Real estate	8673	34
L68	44	Real estate	8674	34
L68	44	Real estate	8675	34
L68	44	Real estate	8676	34
L68	44	Real estate	8677	34
M69-M75	45:49	Professional, scientific and technical services	2791	35
M69-M75	45:49	Professional, scientific and technical services	4573	35
M69-M75	45:49	Professional, scientific and technical services	5555	35
N77-P85	50:52	Administrative and support services; public administration; education	2793	36
N77-P85	50:52	Administrative and support services; public administration; education	2795	36
N77-P85	50:52	Administrative and support services; public administration; education	2797	36
N77-P85	50:52	Administrative and support services; public administration; education	5759	36
Q86-Q88	53	Health	4533	37
R90-S96	54	Arts, entertainment and recreation; other service activities	5752	38
R90-S97	54	Arts, entertainment and recreation; other service activities	5755	38
T	55	Activities of households as employers	NA	.
U	56	Activities of extraterritorial organizations	NA	.

D. Proximity to the End-customer

To determine "proximity" of an industry to the end-customer, this paper uses the concept of a global value chain (GVC) introduced by Timmer et al. (2013)[35]. The GVC of a final product is the set of all value-adding activities necessary for its production. It is identified by the country-industry, in which the last stage of production was completed. A GVC includes the economic value added in this final production step, as well as in all other industries both in the same country as well as other abroad, where previous stages of production were carried out. It builds on the simple idea that as one moves down the supply chain i.e. from suppliers through manufacturers to distributors value is added at each stage until the final product has been delivered and 100 % of the value has been created. To decompose the production network into tiers, this paper makes use of the Leontief's decomposition method. Let $\mathbf{C}$ be a N x 1 vector of total consumption levels, $\mathbf{Q}$ be N x 1 of output levels and $\mathbf{W}$ is a N x N matrix with intermediate input coefficients, which gives how much intermediates are necessary to produce a unit of output in a given industry. Then, $\mathbf{Q} = (\mathbf{I} - \mathbf{W})^{-1}\mathbf{C}$ where $\mathbf{I}$ is the identity matrix and the term $(\mathbf{I} - \mathbf{W})^{-1}$ is the well-known Leontief inverse. It represents the gross output values that are created in each stage of the production chain of one unit of consumption. To see this, let $\mathbf{C}$ be a column vector of which the first element represents the global production of airplanes in France, and all other elements are zero. Then $\mathbf{C}$ represents the final output of airplanes produced in France (*stage 1*[36]) and $\mathbf{WC}$ is the vector of intermediates (*stage 2*), both French and foreign, needed to assemble the aircraft in France. However, these intermediates need to be produced and then $\mathbf{W}^2\mathbf{C}$ represent the intermediates needed to produce $\mathbf{WC}$ (*stage 3*) . This process continues until the mining of basic metals is taken into account. To derive the gross

[35] An alternative way to reconstruct the production network would be to use data on household consumption spending. This methodology is used in Weber & Ozdagli (2017), Saito et al. (2015) and Su (2016).

[36] Counting backwards from customer to supplier

output created in any country-industry pair that contributes to the manufacturing of aircrafts in France, sum across all stages $\mathbf{C} + \mathbf{WC} + \mathbf{W}^2\mathbf{C} + \ldots$ A geometric series obtains, which can be rewritten as $(\mathbf{I} - \mathbf{W})^{-1}\mathbf{C}$. Finally, the value added in each stage of the production of $\mathbf{C}$ is summarized by the vector $\mathbf{K} = \mathbf{F}(\mathbf{I} - \mathbf{W})^{-1}\mathbf{C}$, where $\mathbf{F}$ is a diagonal matrix of value added to gross output ratios in all country-industry pairs.

Based on the methodology outlined above, each year the supply chain is sliced into tiers. Industries in *tier 1* have produced cumulative value-added greater than the 90th percentile of the *stage 1* production vector. Industries in *tier 2* are those already not classified in *tier 1* and having produced cumulative value-added greater than the 80th percentile of the *stage 2* production vector. Similar rules are used for the other tiers. This procedure with moving thresholds ensures that the supply chain is sliced into equally-sized tiers. The process goes on until approximately 100 % of the value-added has been distributed. In this case 9 stages account for 99.999 % of the value-added. Industries in tiers 1-3 are classified as *downstream*, those in tiers 4-5 are *midstream* and those in tiers 6-9 are *upstream*.

E. Construction of Control Variables

This section describes how each of the control variables used in the regressions is constructed:

Variable	*Construction*
Exchange rate sensitivity	Changes in exchange rates of the euro to 44 major currencies in the 1-day interval bracketing monetary policy announcements are computed. Principal components analysis is applied to the data to extract the most important factor driving exchange rates (eigenvector corresponding to the largest eigenvalue). Returns of industry i in country j are regressed on the factor (*Exrate*) : $$ret_{ijt} = \beta_0 + \beta_j Exrate_t + \varepsilon_{ijt}$$ β_j correspond to country-level sensitivities to exchange rate risk. The results hold of β_{ij} are allowed.
Demand Elasticity	To derive estimates for demand elasticities data on M3 money aggregate estimates from the ECB is used. M3 is calculated as the sum of currency

	in circulation, overnight deposits, deposits with maturity of up to two years, as well as repurchase agreements, money market funds and debt securities with maturity up to two years (ECB definition). Data has been downloaded from ECB Statistical Data Warehouse[37]. Following the money supply equation from the theoretical model, for each industry in each country elasticities are estimated from a linear regression of consumer consumption on money supply (*M*) using the formula $E_D = \frac{dD}{dM}\frac{M}{D}$. In this sense, industries with $abs(E_D) < 1$ are inelastic and those with $abs(E_D) > 1$ are elastic and likely to response more to changes in prices. For example, the *Food and Beverages* industry has average elasticity $E_D = 0.66$ and the *Auto and Motor Vehicles* industry has average elasticity of $E_D = 4.35$.
Size	Logarithm of industrial output
Production structure	$\frac{value-added}{Output}$. Measures to what extent industries are dependent on inputs for their production
Inventory	Net change in inventory
Investment	$\frac{fixed\ capital\ formation}{Output}$. Measures how much of the "value" created by an industry is reinvested back into it in the form of fixed assets.
Debt-to-capital ratio	Industry-level debt-to-capital index is computed by averaging firm-level debt-to-capital ratios in a given industry in a given country.
VSTOXX	Volatility Index
Size composition	$\sum_{i=1}^{N} s_{ij}^2$ *where* s_i are firm market shares in industry *j*. Market shares are calculated as firm-level market capitalization in the year before divided by total ME in the industry. It measures the size concentration and competition within an industry and is also know as a *Herfindhal Index*. High levels of the index indicate a concentrated industry with a small number of large players. Average value of the index is 0.51: high concentration. This high value corroborates to the results in the paper: industries are dominated but large firms, which are not likely to be financially constrained.

[37] http://sdw.ecb.europa.eu

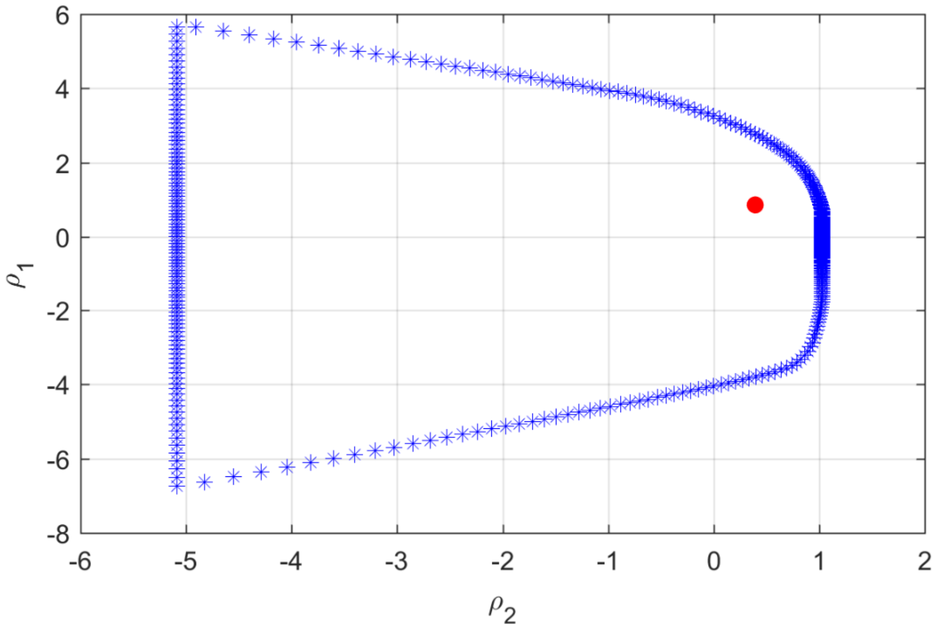
F. Stationarity Region for the Second-order SAR model

The paper uses the approach by Elhorst (2012) to calculate the parameter space over which the matrix $(I_N - \rho_1 W^H - \rho_2 W^F)^{-1}$ exists[38]. **Figure 12** plots the stationarity region (area within the blue figure) and shows that the estimated parameters ρ_1 and ρ_2 (red dot) belong to that region.

Figure 11: Stationarity Region for the Second-order SAR model

The figure plots the stationarity region for the model:

$ret_{it} = \beta_0 + \rho_1 W^H\, ret_{it} + \rho_2 W^F\, ret_{it} + \beta_1 Target_t + \beta_2 Path_t + \varepsilon_{it}$



G. Bonacich Centrality

Definition 1 (Bonacich, 1987). Consider a network g with an $N \times N$ adjacency matrix $W$and a scalar parameter ρ such that $M(g,\rho) = (\mathbb{I}_N - \rho W)^{-1}$ is well-defined. Let $\boldsymbol{\iota}$ is a $N \times 1$ vector of ones. The vector of centralities for all nodes in g is:

$$c(\alpha, g, \rho) = \alpha(\mathbb{I}_N - \rho W)^{-1}.\boldsymbol{\iota}$$

[38] The Matlab program is downloadable from https://spatial-panels.com/software/ under the section "Code Regional Science and Urban Economics 2012"

As can be seen, the parameter α affects only the length of $c(\alpha, g, \rho)$ and is chosen in a way such that the squared length of $c(\alpha, g, \rho)$ equals the number of nodes in the network. Therefore, a node with $c_i(\alpha, g, \rho) = 1$ does not have a too large or small centrality regardless of the number of observations in the network.

H. Identification through Heteroskedasticity

The issues of simultaneity and omitted variables can be described in the following simple set of equations:

$$ret_{it} = \rho W ret_{it} + \beta_1 Target_t + \beta_2 Path_t + \gamma z_t + \varepsilon_{it}$$

$$Target_t = \rho W ret_{it} + \alpha_1 ret_{it} + z_t + \eta_{1t}$$

$$Path_t = \rho W ret_{it} + \alpha_2 ret_{it} + z_t + \eta_{2t}$$

where z_t is a variable that affects both asset prices and interest rates and η_{1t} and η_{2t} are the "true" monetary policy shocks. Let F be a subsample of event (policy dates) and F' a sample of non-policy dates preceding by 1 day those included in F. Following Rigobon (2003), the estimates ρ, β_1 and β_2 can be identified given the following assumptions:

$$\sigma_{\eta_1}^{F} > \sigma_{\eta_1}^{F'} \text{ and } \sigma_{\eta_2}^{F} > \sigma_{\eta_2}^{F'} \tag{1}$$

$$\sigma_{\varepsilon}^{F} = \sigma_{\varepsilon}^{F'} \tag{2}$$

$$\sigma_{z}^{F} = \sigma_{z}^{F'} \tag{3}$$

In words, what these assumptions are saying is that monetary policy shocks are heteroskedastic, asset price shocks are homoscedastic and that parameters are stable over time. In the data, $\sigma_{\eta_1}^{F} = 4.61 > 4.06 = \sigma_{\eta_1}^{F'}$, $\sigma_{\eta_2}^{F} = 17.49 > 17.03 = \sigma_{\eta_2}^{F'}$ and $\sigma_{\varepsilon}^{F} = 2.06 \cong 2.05 = \sigma_{\varepsilon}^{F'}$. Common shocks z_t are unobservable, but given that F and F' are 1 day apart, it is likely that $\sigma_{z}^{F} = \sigma_{z}^{F'}$.

To arrive at the instrumental variable setting proposed by Rigobon and Sack (2004), pull together asset prices and monetary policy shocks on both policy and non-policy dates and define the following variables:

$$r \equiv \{r_{it}, t \in F\} \cup \{r_{it}, t \in F'\}$$

$$Target \equiv \{Target_t, t \in F\} \cup \{Target_t, t \in F'\}$$

$$Path \equiv \{Path_t, t \in F\} \cup \{Path_t, t \in F'\}$$

which are $2T \times N$ and $2T \times 1$ (target and path) vectors respectively with T being the number of policy dates and N the number of industries. The following are valid instruments:

$$w_r \equiv \{r_{it}, t \in F\} \cup \{-r_{it}, t \in F'\}$$

$$w_{target} \equiv \{Target_t, t \in F\} \cup \{-Target_t, t \in F'\}$$

$$w_{path} \equiv \{Path_t, t \in F\} \cup \{-Path_t, t \in F'\}$$

Intuitively, the instrument w_{target}, for example, is correlated with $Target_t$ because the policy subsample outweighs the non-policy subsample due to the heteroskedasticity of the monetary policy shocks. However, the exclusion restriction holds because w_{target} does not correlate with η_{1t} and z_t. These shocks are homoscedastic and the two samples cancel out.

Chapter 2

Systemic Risk and Network Spillovers in the European Sovereign CDS Market: A Spatial Autoregressive Approach

1. Introduction

Sovereign credit default spreads (CDS) in the Euro-area feature a time-varying pattern of comovement, which constitutes a major challenge for econometric modelling and forecasting. During the recent European Sovereign Debt Crisis of 2010-2012 spreads have reached levels that cannot be predicted by standard models, which typically model spreads as a persistent mean-reverting process driven by two factors: a local and a global factor. The local factor is determined by fundamentals, whereas the global factor captures risk aversion i.e. proxies for global market conditions. Predicted spreads from these models cannot match the pattern in the data and are, on average, 100 basis points lower than realized values. This empirical evidence suggests a non-linear relationship between a sovereign's theoretical probability of default and observed credit spreads, a phenomenon dubbed as the "credit spread puzzle" (Amato & Remolona (2003); Chen, Collin-Dufresne, & Goldstein (2008); Longstaff, Pan, Pedersen, & Singleton (2011)).

Until recently, the probability of a developed country defaulting on its sovereign debt was considered to be close to zero. However, with the onset of the Global Liqudity and the subsequent European Sovereign Debt Crisis many governments had to step in and save their financial sectors, as a result of which fiscal deficits reached levels unseen since World War II. This led to a revision of credit markets and raised a discussion about the true probability of sovereign default. Credit rating agencies responded with a series of downgrades, notwithstanding developed countries. For example, Germany was the only country in Europe, which retained its AAA rating.

Furthermore, in the aftermaths of the European Sovereign Debt crisis economists, policymakers and the media have raised concerns over the different forms of contagion in the financial system. One common source of anxiety is that given the interconnectedness of the European financial systems, the default of one country would have spillover effects that would result in higher borrowing costs for other sovereigns, and potentially would trigger a series of other defaults.

This prompted academics to addresses the three points outlined above by considering an extension to the standard models of credit spreads and proposing a *new factor*, which captures time-varying financial linkages among European sovereigns. The factor is constructed as a country-speficic weighted average of CDS spreads, where the weighting scheme is determined by financial network connections and the model is estimated in a *spatial autoregressive framework*. Eder & Keiler (2015) and Blasques et al. (2016) apply the spatial model to European CDS markets and find evidence of considerable spillover effects. The results are motivated by arguing that the new systemic factor reflects important nonlinearities in CDS markets, because it allows that the credit risk of one sovereign depends not only on its own fundamentals, but also on the credit risk of the countries to which it is financially exposed.

However despite its intuitive appeal, one major criticism against the spatial autoregressive model and the choice of weighting scheme of CDS spreads, is that the results are on a *statistical* level and are not motivated by network and economic theory.

This paper addresses this concern by providing a microfoundation for the use of spatial autoregressions in modeling sovereign CDS spreads. *First*, it develops a simple network model of credit risk with asset value interdependencies in the spirit of recent theoretical models (Acemoglu, Ozdaglar, & Tahbaz-Salehi (2015); Elliot, Golub, & Jackson, 2014; Glaserman & Young (2015); Rogers & Veraart

(2013)). It shows that under mild regularity conditions the theoretical network model empirically translates into a system of simultaneous equations, which can be estimated using *spatial autoregressive model* (SAR). *Second,* on the empirical side, the paper extends the work by Keiler & Eder (2015) and Blasques et al. (2016) and shows that the spatial model can be used to predict more accurately sovereign CDS spreads. *Third,* the paper investigates how exogenous financial shocks propagate through sovereign networks and offers a decomposition of the market's reaction into *direct* and *indirect network* effects.

The model is estimated using data on CDS spreads and financial linkages among 10 European sovereigns from 2006 to 2016. To construct the empirical counter-part of the cross-holdings matrix in the theoretical model, I use data from Bank of International Settlements (BIS), which provides detailed information on bilateral lending and borrowing relationships between BIS-reporting countries. A directed link in this network exists if country i holds a claim vis-à-vis country j and the strength of the connection, x_{ij}, is given by dollar value of the outstanding debt to country j, divided by the total amount that country i borrows from all the countries in the sample. It is important to note that this data exists on an aggregate level i.e. how much do all banks of country i borrow/lend to country j. BIS provides a split by sector, which allows me to uncover the total amount borrowed from general governments, but not the identity of the countries. I call this amount D_i^{Gvmt}. Finally, to obtain the link between two *sovereigns* I weigh D_i^{Gvmt} by the strength of the connection x_{ij}.

The theoretical framework considers "balance sheet" mechanism of contagion, where spillovers from a severe financial shock occur via direct losses to assets held by creditors. Using a fixed-point argument, it is possible to show that in the presence of *asset interdependencies* and *discontinuities* in value multiple equilibrium solutions for organization's values are possible. In this context of multiple equilibria, contagion emerges because of linkages and the joint determination of asset prices: organizations experience losses because people *expect* that other connected organizations will incur losses as well and this then becomes *self-fulfilling*.

From an empirical standpoint, the paper reconfirms the results of Keiler & Eder (2015) and Blasques et al (2015) and finds significant evidence for the presence of credit risk spillovers in CDS markets. Introducing the systemic risk factor considerably improves in-sample model fit and explanatory power. The results indicate that network linkages account for 15 % to 20 % of the CDS variance. In out-of-sample predictive tests, the SAR model consistently outperforms standard models. The SAR model is better able to match monthly changes in CDS spreads and leads to 25 % to 35% improvement in predictive accuracy, measured in the root mean squared error (RMSE) sense. Finally, the paper shows that the constructed network of financial linkages between sovereigns is an important mechanism for the propagation of exogenous financial shocks. Using the SAR model, it is possible to decompose the *total* effect of financial shocks into *direct* and *indirect* effects. The results of an event study around announcement dates indicate that as much as 45% of the overall effect of shocks is due to indirect (network) effects. The findings of the paper are robust to the chosen time period, data frequency and alternative specifications of the network links.

2. Related Literature

This paper contributes to three main strands of literature. First, it is related to the literature studying contagion in financial networks[39]. Allen and Gale (2000) and Freixas, Parigi and Rochet (2000) consider a simple interbank lending network, where liquidity shocks arise due to consumers. They study contagion of insolvencies as a result of one bank failing and reach the conclusion that the more connected the network is, the more robust to contagion it is. More recently, Allen, Babus and Carletti (2010) develop a model, where institutions form connections by swapping projects in order to diversify their risks. The authors find that these type of connections lead to two different network topologies: a clustered one, in which

[39] Other related works are: Gai, Hadane, & Kapadia (2011), Gai and Kapadia (2010), Craig and Von Peter (2014), Leitner (2005), Georg (2013), Nier, Yang, Yorulmazer, & Alenton (2007), Battiston, Gatti, Gallegatti, Greenwald, & Stiglitz (2012), Battiston, Puliga, Kaushik, Tasca, & Caldarelli (2012)

institutions hold identical portfolios and fail together, and an unclustered dispersed network. Eisenberg and Noe, (2001) look at default of firms as a part of a clearing mechanism. They develop an computationally efficient algorithm that clears the financial system and, at the same time, provides information about the systemic risk faced by the individual firms. Rogers and Veraart (2013) extend the seminal work of Eisenberg and Noe (2001) by introducing costs of default and offer a rigorous analysis of those situations in which banks will have an incentive to cooperate and bail-out other distressed banks. Elliot, Golub and Jackson (2014) study how failures cascade in a network of interconnected financial organizations and how discontinuous changes in asset value trigger further failures. Whereas Elliot, Golub and Jackson assume that the non-inflated market value of an organization is well-captured by the value of equity held by outside investors, this paper takes a different approach. Building on recent work by Barucca, et al. (2016), the market value of organizations is obtained via a valuation function that takes into consideration at the same time interdependencies and uncertainty. Furthermore, Acemoglu, Ozdaglar and Tahbaz-Salehi (2015) consider an exogenous network and derive results based on the topology of the network. On a related note, Glasserman and Young (2015) derive bounds on the magnitute of network effects of contagion.

Despite this plethora of theoretical works on financial networks and contagion, there is relatively little empirical research drawing on these models[40]. Allen and Gale (2000), Elliot, Jackson and Golub (2014) and Glasserman (2015) suggest that their models can be applied in an empirical setting. Although both Elliot, Jackson and Golub (2000) and Glasserman (2015) use empirical data to illustrate their models, these results are intended as complementing the theoretical results and are not intended as a robust econometric exercise. In this strand of literature the paper is closest to Glover & Richards-Shubik (2016), who use a similar network model of credit risk to make an important contribution. The authors use observed CDS spreads to impute the risk-neutral probability of default. This paper on the

[40] Examples are Cohen-Cole, Patacchini and Zenou (2011);

other hand uses a completely different empirical strategy, which focuses on the simultaneous modelling of sovereign CDS spreads. In particular, this paper uses the network model to microfound the use of SAR. The goal here is to show that the SAR can be used to predict more accurately CDS spreads and to quantify the direct and indirect effects of exogenous financial shocks.

Spatial autoregressive models (SAR) have been traditionally developed in regional and social sciences. Important theoretical contributions in the field include, but are not limited to: Prucha and Kelejian ((1998), (1999), (2004), (2006), (2010); Anselin and Bera (1998); Anselin, Bera, Florax, and Yoon (1996); LeSage and Pace (2006), (2009); Arbia ((2006), (2012); Lee (2004). On the empirical side, Kim, Phipps and Anselin (2003) apply a spatial approach to study the benefits of air quality improvement. In a political economy setting, Bordignon, Cerniglia, & Revelli (2003) how spatial dependencies affect the way property taxes are set. Furthermore, Bloningen et al. (2007) study how foreign direct investment (FDI) into a host country depends on the FDI in proximate countries.

SAR models in finance are not widespread, but have began to gather momentum in the recent years. Examples of recent applications are Ozdagli and Weber (2016), who study how monetary policy shocks propagate through the input-output production network. Fernandez (2011) derives a spatial version of the CAPM model and uses the results to perform value-at-risk simulations. This paper is closest to Blasques et al. (2016) and Eder and Keiler (2015) The former investigates a network of eight European sovereigns and their financial linkages and finds time-varying spatial dependence. However, in contrast to this paper, Blasques et al. (2015) are more concerned with the statistical properties of the model and do not provide a micro-foundation for their results. The article by Eder and Keiler looks at network connections determined by asset correlations on the stock market between systemically important institutions. The paper finds strong empirical evidence of spillovers in CDS markets and concludes that around 10 % to 15 % of the CDS variance can be explained by network connections. The results of this paper are consistent with Eder and Keiler (2015). Differently from Eder and

Keiler (2015), the focus here is on a network of sovereigns, the connections between which are given by cross-border borrowing and lending exposures.

Finally, this paper is related to research that studies modeling and forecasting credit spreads in the euro area. The standard specification adopted for sovereign spreads in the Euro Area models them as a persistent processs reverting to a time-varying mean explained by two factors: a local country-specific factor, related to fiscal fundamentals, and a global factor, which measures market appetite for risk (Favero, Pagano and von Thadden (2010); Beber, Brandt and Kavajecz (2009); Laubach (2009, 2011)).Another common finding in this literature, starting with Codogno, Favero and Missale (2003) and Geier, Kossmeier and Pichler (2004) among others, is that the sovereign spread yields in the Euro area are strongly comoving. This comovement given the heterogenousliquidity of bonds issued by different countries suggests either that credit risk is dominating credit risk and that the two strongly move together (Favero, Pagano and von Thadden (2010); Beber, Brandt and Kavajecz (2009)). Credit risk should theoretically depend on fiscal fundamentals, but the empirical literature suggests that a linear relationship between the credit risk and fundamentals has been largely illusive (Attinasi, Checherita and Nickel (2010); Laubach (2009)). Finally, this paper contributes to the literature on market spillovers. Diebold and Yilmaz (2009,2011) measure spillovers by employing variance decomposition of Vector Autoregressive models (VARs). In particular, they look at the proportion of the (conditional) variance of the returns to an asset, which is explained by the (conditional) variance of other assets. Another approach is the Global Vector Autoregressive model (GVAR) advanced by Peseran, Schuermann and Weiner (2004). The model provides a flexible reduced-form framework, which allows for accommodating time-varying co-movement between local and global country-variables. An important contribution in this field is Favero (2013), who augments the standard GVAR framework by introducing two variables. These variables define for each country a global spread, which is a weighted average of the spreads of other countries, where the weights are given by distances in the fiscal fundamentals (debt and deficit) between countries.

3. Theoretical Model

This section develops a simple input-output model with interbank cross-holdings. The network model builds on earlier work of Eisenberg and Noe (2001), Suzuki (2002) and more recently, Rogers & Veraart (2013), Elliot, Golub and Jackson (2014) and Barucca et al. (2016). The model investigates cascades of failures in a network composed of interdependent institutions and shows that discontinuous changes in the asset value of organizations trigger further failures and that this depends on the properties of the network structure. Contagion is discussed in the context of multiple equilibria, the source of which is interdependencies of asset values: organizations fail because people *expect* that other organizations will fail as well and this then becomes *self-fulfilling*.

The equilibrium concept is investigated in the context of a process, which repeats itself over $t = 1, \dots, T$ number of periods. Each period is treated independently and is assumed to unfold in the following steps:

Step 1: The financial sectors in each country are endowed with bilateral claims to foreign sovereigns, which are established in the previous period

Step 2: The financial sector collects deposits and equity and invests these in primitive assets (e.g. loans, equities, bonds, commodities etc.)

Step 3: Exogenous financial shocks are realized

Step 4: Valuation of claims and investments is performed: *market value* obtains

Step 5: If the market value hits a threshold value (exogenously given), failure occurs and default costs are incurred

Step 6: Claims are established for the next period

Step 7: CDS contracts are traded for credit events in the next period

The following sections describe in detail *Steps 1:7*.

3.1 Primitive Assets, Organizations and Cross-Holdings

Consider a financial system composed of n countries making up a set $N = \{1,..,n\}$. In each country, the financial sector[41] collects deposits (d) and equity (e) and invests these in primitive assets $M = \{1,..,m\}$. To fix ideas, a primitive asset may be thought of as a project that generates net cash flow over time. Let the amount of the primitive asset k of bank i at time t be π_{ikt} and p_{kt} be its price, then $\pi_{ikt}p_{kt}$ is the book value of the primitive assets. One can think of these k assets as different asset classes: e.g. loans to firms and households, equities, bonds, commodities etc., or simply as factors of production. Banks/sovereigns also lend to each other in the interbank market. Let a_{ijt} is the amount that country i lends to the general government of country j. The quantity a_{ijt} represents sovereign debt claims and is of main interest in the model. Both sovereign debt and interbank lending values are determined in the previous period. All magnitudes are expressed in monetary terms, euros, and represent *book value*. The balance sheet of bank i is given by:

$$d_i + e_i + a_{1i} + \cdots + a_{n1} = a_{i1} + \cdots + a_{in} + \sum_k \pi_{ik} p_k \tag{1}$$

It is possible to aggregate the balance sheets of all the n banks and express it in matrix form:

$$d + e + A_M \iota_n = A'_M \iota_n + \Pi p \tag{2}$$

where d, e and p are column vectors, ι_n is column vector of ones, Π is $n \times m$ matrix, A_M is bilateral interbank exposure and everything is expressed in monetary terms. Let v be a vector representing total book value of bank assets:

$$v = d + e + A_M \iota_n = A'_M \iota_n + \Pi p \tag{3}$$

Finally, sovereigns are exposed to exogenous financial shocks, which is the only source of stochasticity in the model.

[41] The terms "financial sector" is understood here as the collection of all BIS-reporting banks in a given country

3.2 Introducing Market Value

A proper valuation of country i, denoted by V in the model, depends on how much the country values its assets. Such valuation can substantially differ from the book value of assets and depends on assets of other nodes in the model, and more precisely, on how much they value their assets. Let T_{ij} denote the maturity of the debt contract between i and j and let t denote the time at which the evaluation of the financial claim takes place. I assume that $t = T_{ij}$ for all i and j i.e. the evaluation takes place at maturity. Capitalizing on recent work by Barucca and co-authors (2016), I introduce the following valuation function:

Definition 1: *Given an integer $q \leq n$, a function $\mathbb{U}: \mathbb{R}^q \to [0,1]$ is called feasible valuation function if and only if:*

1. *It is non-decreasing: $V \leq V' \Rightarrow \mathbb{U}(V) \leq \mathbb{U}(V'), \forall V, V' \in \mathbb{R}^q$*
2. *I is continuous from above*

The intuition behind the definition is that the market value of any asset can be written as the product of its book value multiplied by the valuation function. Thus, the market value of an asset ranges from its face value to zero. I assume that the valuation depends only on banks' asset values. If $\mathbb{U}^{IB}(V)$ and $\mathbb{U}(V)^{PA}$ are the valuation functions for the interbank assets and the proprietary assets respectively, then the market value of assets is given by:

$$V(t) = A'_M \iota_n \mathbb{U}^{IB}(V(t)) + \Pi p \mathbb{U}^{PA}(V(t)) \tag{4}$$

In the interest of readability, in the following sections the explicit dependence on the time at which the valuation is carried out will be dropped.

Let $\hat{V}$ be the corresponding diagonal matrix such that $\hat{V}\iota_n = V$. Then, the right hand side of equation (11) can be represented as:

$$V = \hat{V}^{-1} A'_M \hat{V} \iota_n \mathbb{U}^{IB}(V) + \Pi p \mathbb{U}^{IB}(V) = ZV\mathbb{U}^{IB}(V) + \Pi p \mathbb{U}^{PA}(V) \tag{5}$$

where $Z = \hat{V}^{-1} A'_M$ is a matrix such that each *row* is divided by the total assets of the *lending bank*. The entries of the vector $z = Z\iota_n$ are fractions of unity, which give the proportion of country i's interbank lending to its total assets. Then, the parameter $\psi = \frac{1}{n}\sum z_{ij}$ gives the average contribution of interbank lending to the total value of the organization. This parameter is endogenously determined from the model as it depends on the face value of interbank claims and assets and their market valuations. Furthermore, since $z \in [0,1)$, then by definition $\psi \in [0,1)$. Rewrite equation (12) as:

$$V = \psi W V \mathbb{U}^{IB}(V) + \Pi p \mathbb{U}^{PA}(V) \tag{6}$$

where $W = \frac{1}{\psi} Z$.

Since all valuation functions take values in the interval [0,1], then organization's values V are bounded both from above and from below:

$$V_{min} \equiv 0 \leq V \leq V_{max} \tag{7}$$

3.3 Introducing Discontinuities in Value and Failure Costs

An important aspect of the model is that organizations can lose value in a discontinuous way if their values hit certain critical thresholds. These discontinuities can trigger cascading failures and multiple equilibria. There could be many sources of such discontinuities: for example, if an airline cannot pay for its fuel, then its fleet has to sit idle, which leads to a discontinuous drop in revenues because of taxes paid to the ground operators, lost bookings, failure to meet delivery obligations to courier companies and so forth. When a country's sovereign debt is downgraded, then it experiences a discontinuous jump in the cost of capital. An increase in the borrowing cost could lead to an interruption in the ability to pay or acquire other factors of production, which could then lead to a decrease in revenue and loss of value. The particular source of discontinuity that this paper considers is an exogenous financial shock.

The model assumes that if the value of an organization V falls below some threshold $\underline{V}$, then it is said to fail and incurs a failing cost of γ proportionate to the price of the proprietary asset and expressed as fraction of cents on the euro. The organization incurs this cost, because it needs to liquidate its asset in order to cover its liabilities. Since debt is given priority over equity in this setting, it can be assumed that organizations are *always* able to recover the market value of their interbank claims. Such discontinuities could easily be accommodated in the valuation function $\mathbb{U}(V)$ using the following rule:

1. $\mathbb{U}^{IB} = 1$

2. $\mathbb{U}^{PA} = I_{V>\underline{V}} + (I_{V\leq\underline{V}} - \gamma I_{V\leq\underline{V}})$

The discontinuous drop imposes a loss on the organization and so its value becomes:

$$V = \psi WV + \Pi p\left(I_{V>\underline{V}} + \left(I_{V\leq\underline{V}} - \gamma I_{V\leq\underline{V}}\right)\right) \tag{8}$$

Alternatively, V can be expressed as:

$$V = (\mathbb{I}_n - \psi W)^{-1}\Pi p\left(I_{V>\underline{V}} + \left(I_{V\leq\underline{V}} - \gamma I_{V\leq\underline{V}}\right)\right) = \mathcal{A}\Pi p\left(I_{V>\underline{V}} + \left(I_{V\leq\underline{V}} - \gamma I_{V\leq\underline{V}}\right)\right) \tag{9}$$

The matrix $\mathcal{A}$ determines how the costs of a failing organization j are distributed among other organizations in the system and how this affects their value. Note that the dependency matrix $\mathcal{A}$ is *not* row-stochastic, which is not surprising. In this model, there are two types of assets: interbank claims, expressed as *fractions* of total value, and proprietary assets. The entries of $\mathcal{A}$ describe total value as a fraction of the proprietary asset and, so its rows sum up to more than unity[42].

---

[42] For example, if no organization fails, $V = \begin{pmatrix}1\\1\end{pmatrix}$ and the value of the interbank asset is $\psi WV = \begin{pmatrix}0 & 0.5\\0.5 & 0\end{pmatrix}V$, then the value of the proprietary asset as a fraction of total value is $\Pi p = \begin{pmatrix}0.5 & 0\\0 & 0.5\end{pmatrix}V$. Finally, $V = \mathcal{A}\Pi p = \begin{pmatrix}1.33 & 0.67\\0.67 & 1.33\end{pmatrix}\begin{pmatrix}0.5 & 0\\0 & 0.5\end{pmatrix}V = \begin{pmatrix}0.67 & 0.33\\0.33 & 0.67\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}1\\1\end{pmatrix}$. The obtained matrix $\begin{pmatrix}0.67 & 0.33\\0.33 & 0.67\end{pmatrix}$ is row-stochastic, which means that the values of the underlying assets sum up to the values of the organizations. Hence, no value is lost and no value is created

A final caveat of the model regards the matrix $\mathbb{I}_n - \psi W$. In order to show that it is not singular and that the spatial multiplier $(\mathbb{I}_n - \psi W)^{-1}$ exists, I use *Lemma 1*.

Lemma 1:

If Z is a matrix with $\|Z\| < 1$, where $\|Z\| = \|Z\|_1 = \sup_x = \frac{\|Zx\|_1}{\|x\|_1}$, then $\mathbb{I}_n - Z$ is invertible and

$$(\mathbb{I}_n - Z)^{-1} = \sum_{k=1}^{\infty} Z^k$$

Given that $\|\psi W\| = \|Z\|$ and that the entries of the matrix Z are all fractions of unity, the conditions in **Lemma 1** are fulfilled and the matrix is invertible[43]. *Appendix A* offers a simple proof of the lemma.

3.4 Relationship to Leontief 's Input-Output Model

Equation (13) is reminiscent of the famous input-output analysis of Leontief (1951). In the traditional application, each firm or industry sector i uses z_{ij} units of output from sector $j (j = 1, \dots, n)$, labor and other units of primary inputs to produce one unit of final output. Here, each bank borrows, i.e. "uses" funds from other banks in the model, deposits and equities and invests these in assets (e.g. production factors and/or loans). If the z_{ij} are relatively constant, then the relation between V and the fundamental asset is given by the well-known Leontief Inverse $\mathcal{A} = (\mathbb{I}_n - \psi W)^{-1}$. It is interesting to notice that the literature on social *networking* interprets the vector $b = (\mathbb{I}_n - \psi W)^{-1} \iota_n$ as the centrality of a node (Katz, (1953); Bonacich, (1987)). The vector b is also known as *Katz-Bonacich* centrality, and in the context here, measures the number of direct and indirect connections that a

[43] In practice, in order to ensure that $\mathbb{I}_n - \psi W$ is not singular and that the spatial multiplier $(\mathbb{I}_n - \psi W)^{-1}$ exists, the model imposes that the dependency matrix W is normalized to a row-stochastic matrix W. Kelejian and Prucha (2004; 2010) discuss regularity conditions for spatial models and show that for a row-stochastic spatial weights matrix and for a spatial autoregressive parameter belonging to the interval $(-1,1)$, the spatial multiplier exists and is well defined.

country in the network of bilateral exposures has and the parameter ψ reflects a discount factor, which assigns less influence to distant nodes.

3.5 Equilibrium Existence and Multiplicity

A solution for the values of organizations in equation (16) is an *equilibrium set of values* that takes into consideration dependencies between countries. Invoking Tarski's fixed point theorem (Tarski, 1955), it is possible to show that there always exists a solution to the problem in (16), and moreover, that there is a least and a greatest solution. In fact, the set of solutions forms a complete lattice, which follows from the fact that failures are strategic complements. *Appendix A* offers a simple proof of the result.

The presence of discontinuities and equilibrium multiplicity can come from two distinct sources. The first type of discontinuity arises when the failure of organization i is caused by a drop in the value of its underlying assets. The second type of discontinuity is triggered when another organization j, in which i holds claims, hits the failure threshold: the value of i's assets drops discontinuously, and so does its total value. Consequently, these two sources of discontinuity result in two different types of multiplicities of equilibria. The first is when the values of other organizations and their assets are taken as given: there could be multiple values of organization i consistent with equation (16). There may exist a value V_i that solves equation (16) such that i does not fail and another value V_i satisfying equation (16) such that i fails. This mechanism generates the type of multiple equilibria corresponding to the classical models of self-fulfilling bank runs (e.g. Dybvig and Dybvig, (1983)). The second type of multiple equilibria arises due to interdependencies between organizations: the value of i depends on the value of j and viceverca. There might be two consistent values for i and j: one in which both fail, and another one in which none fail. This source of equilibrium multiplicty is distinct from the bank run story, as in this case organizations fail because people expect that other organizations fail, which then becomes self-fulfilling.

3.6 A Simple Microfoundation

Consider a network composed of two countries only: country 1 and country 2. The dependence relation between the two countries is given by the matrix W, which is assumed to be for simplicity row-stochastic:

$$W = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

The two countries in the model lend each other 1 million. Let $\psi = 0.5$. Then according to Lemma 1 $(\mathbb{I}_n - \psi W)^{-1}$ exists and is equal to:

$$(\mathbb{I}_n - \psi W)^{-1} = \begin{pmatrix} 1.33 & 0.67 \\ 0.67 & 1.33 \end{pmatrix}$$

If the assumption is made that the fundamental asset that each country owns is its fiscal stream, then by exchanging cross-holdings, countries obtain holdings whose value depends not only on the value of their own fiscal stream, but also the fiscal stream of other countries. Thus, $m = n$ and $\Pi = \mathbb{I}$. Finally, assume that countries fail if their value falls below 50 and that in this case they incur a cost of 50[44].

Given the dependence relation, the conditions for a country failing are given by:

Country 1:

$1.33p_1 + 0.67p_2 < 50$

Country 2:

$0.67p_1 + 1.33p_2 < 50$

These inequalities define two failure frontiers FF_1 and FF_2 and four regions (*only Country 1 Fails, only Country 2 Fails, Both fail, None Fail*), which are graphically illustrated in Figure 1 Panel A. In this situation, the prices for which one of the four possible scenarios occur are uniquely determined.

[44] This means that $p\gamma = 50$

It is interesting to discuss what happens to the failure frontier of Country 1, conditional on Country 2 failing and vice versa. In this case the conditions for failing are given by:

Country 1:

$1.33p_1 + 0.67(p_2 - 50) < 50$

Country 2:

$0.67(p_1 - 50) + 1.33p_2 < 50$

The two new *conditional* failure frontiers FF_1' and FF_2' identify a region of multiple equilibria given by the shaded area in Panel B. Multiplicities arise here because the value of Country 1 decreases discontinuously when Country 2 hits the failure threshold and defaults and the value of Country 2 decreases discontinuously when Country 1 fails. It then becomes consistent for both Country 1 and Country 2 *not* to fail, in which case the relevant frontiers are the unconditional ones. However, it is also consistent for both Country 1 and Country 2 to *fail*, in which case the relevant frontiers are the conditional frontiers. This mechanism is of particular importance from a policy-making perspective because for the same price of the fundamental asset, a “good” and a “bad” equilibrium are possible. As discussed above, the source of multiplicities is interdependencies of asset values. Organizations fails because people expect that other connected organizations fail, and this then becomes self-fulfilling.

To summarize, this paper considers a network model with inter-related asset values. Default occurs when the value of a sovereign’s assets hits a critical threshold. In this framework contagion is understood as defaults or other significant losses transmitted via the network of financial linkages. This is a direct, “balance sheet” type of contagion. To be clear, in order to estimate the structural model of spillovers in financial networks outlined in this section, this paper applies a static equilibrium concept on repeated observations within a fixed set of countries[45]. In order to treat each period independently, the paper ignores any dynamic aspect of decision making. Additionally, unobserved shocks are assumed to be independent over the time period. As in Denbee, et al. (2014), this paper treats financial linkages as exogenous if the network is determined by actions in previous periods.

Recall that the first-order condition writes as:

$$V = \psi WV + \Pi p \left(I_{V > \underline{V}} + \left(I_{V \leq \underline{V}} - \gamma I_{V \leq \underline{V}} \right) \right)$$

It is immediate to observe that the equation above has exactly the form of a *spatial autoregression*, where the relevant weighting matrix is given by the bilateral cross-holdings (lending) matrix between countries. Assuming that a natural empirical proxy for the credit riskiness of sovereign's assets is the CDS spread, then ψ captures spillovers in sovereign CDS markets. The model predicts that a sizeable exogenous financial shock reduces assets value i.e. increases riskiness and is empirically reflected in higher CDS spreads. The next section introduces the empirical framework used to test these two model predictions.

4. Empirical Framework

4.1 Spatial Autoregressions

The spatial autoregressive model (SAR) is given by:

$$y = \rho W y + \beta X + \varepsilon \tag{10}$$

with data-generating process:

$$y = (\mathbb{I}_n - \rho W')^{-1}(\beta X + \varepsilon)$$

$$\varepsilon \sim N(0, \sigma^2 \mathbb{I}_n)$$

[45] By applying a static equilibrium concept, the paper follows the existing literature. See Denbee et al. (2014); Cohen-Cole, Patacchini and Zenou (2011); Bonaldi, Hortacsu and Kastl (2014)

where y is a vector of CDS spreads, X is a vector of controls and W is a row-normalized spatial weights matrix. In my application W corresponds to a cross-holdings matrix, which gives the consolidated foreign claims of banks from one country on the debt obligations of the general government of another country. All entries on the main diagonal of W are zero, because I rule out dependence of an observation on its own value. The spatial parameter ρ indicates the relevance of a country's connectedness for its probability of default and in this sense can be regarded as a measure for network spillover effects. Testing for the presence of spillovers is tantamount to the following hypotheses:

$$H_0 : \rho = 0$$

$$H_1 : \rho \neq 0$$

Here, I assume that $abs(\rho) < 1$. I use the term $N(0, \sigma^2 \mathbb{I}_n)$ to denote a zero mean disturbance process with constant variance σ^2 and zero covariance between the observations. This results in a diagonal variance-covariance matrix $\sigma^2 \mathbb{I}_n$ with $\mathbb{I}_n$ representing an $n \times n$-dimensional identity matrix. The term Wy is called a *spatial lag* and is constructed as a linear combination of neighboring values to each observation.

Example 1: Consider the following dependence matrix C:

$$C = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

In this case, there are four countries, the column gives the country whose debt is being held and the row is the country, which holds the debt obligation. Where there exists a dependence the value of the debt is 1 million. The dependence relationship reads like this: country 1 lends to country 2, 3 and 4; country 2 lends only to country 4; country 3 lends to country 1 and 2; country 4 lends to country 1. In order to form a spatial lag from neighboring observations, we can normalize C, such that its rows sum up to unity. This results in a row-stochastic matrix, which I label W:

$$W = \begin{pmatrix} 0 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 0 & 1 \\ 1/2 & 1/2 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

Having defined the spatial weights matrix W, it is easy to see that the spatial lag Wy corresponds to:

$$Wy = \begin{pmatrix} 0 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 0 & 1 \\ 1/2 & 1/2 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} (y_2 + y_3 + y_4)/3 \\ y_4 \\ (y_1 + y_2)/2 \\ y_1 \end{pmatrix}$$

The presence of spatial lags of the dependent variable renders the OLS parameter estimates and standard errors inconsistent. On the other hand, maximum likelihood is consistent for the spatial autoregressive model in (10) (Lee, 2004).

4.2 Spatial Autoregressions: Parameter Interpretation

Parameter estimates in a linear regression have a straightforward interpretation as partial derivatives of the dependent variable with respect to the independent variable. This arises because of assumed linearity and independence of the observations in the model. When the model contains spatial lags of the dependent variable, interpretation of the parameters becomes more complicated, because the model incorporates a richer information set. In a spatial context, a change in any given explanatory variable (e.g. GDP) will have an impact on the CDS spread of the country itself (direct effect) and potentially an impact of the CDS spreads of other related countries (indirect effect). In fact, a large number of researchers have argued that spatial autoregressive models require special attention to the interpretation of the parameters (Anselin & LeGallo (2006); Kelejian, Tavlas and Hondronyiannis (2006); Kim, Phipps and Anselin (2003); Le Gallo, Ertur and Baumont (2003)).

To see more clearly the complication of parameter interpretation, rewrite the model in (10) as:

$$(\mathbb{I}_n - \rho W)y = \beta X + \varepsilon$$

$$y = S(W)X + V(W)\varepsilon$$

where

$$S(W) = V(W)\,\mathbb{I}_n\beta \tag{11}$$

$$V(W) = (\mathbb{I}_n - \rho W)^{-1} = \mathbb{I}_n + \rho W + \rho^2 W^2 + \cdots \tag{12}$$

To illustrate the point, consider again the previous example with four countries, but for simplicity assume that there is only one covariate: *fiscal deficit*. The data-generating process can be expanded to the following:

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} S(W)_{11} & S(W)_{12} & S(W)_{13} & S(W)_{14} \\ S(W)_{21} & S(W)_{22} & S(W)_{23} & S(W)_{24} \\ S(W)_{31} & S(W)_{32} & S(W)_{33} & S(W)_{34} \\ S(W)_{41} & S(W)_{42} & S(W)_{43} & S(W)_{44} \end{pmatrix} \times \begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{pmatrix} + V(W)\varepsilon$$

with $S(W)_{ij}$ indicates the ij^{th} element of the matrix $S(W)$. Focusing on country 1, the following obtains:

$$y_1 = S(W)_{11}X_1 + S(W)_{12}X_2 + S(W)_{13}X_3 + S(W)_{14}X_4 + V(W)_1\varepsilon \tag{13}$$

with $V(W)_1$ referring to the first row of the matrix $V(W)$. An immediate implication from equation (13) is that the CDS spread of country 1 depends not only on changes in its own fiscal deficit, but also on the fiscal deficit of other countries it is connected to. In this sense, $S(W)_{11}$ gives the reaction of country 1's CDS spread to a change in its own fiscal deficit. Similarly, $S(W)_{12}$ denotes the reaction of country 1's CDS spread to a change in the fiscal deficit of country 2. Therefore, $S(W)_{11}$ gives the direct effect of the shock and $S(W)_{12}$, $S(W)_{13}$ and $S(W)_{14}$ give the indirect effect due to country 1's exposure to countries 2, 3 and 4 through the network of debt cross-holdings.

The response of a country's CDS spread to changes in X is determined by the cross-holdings matrix W through its effect on liquidity provision, the spatial autoregressive parameter ρ, which denotes the strength of the network spillover effects, and the parameter β. The own derivative of y with respect to X results in $S(W)_{ii}$ and measures the direct effect. These elements are located on the diagonal of the matrix $S(W)$. It is important to note that this impact takes into consideration *feedback effects*, where observation i affects observation j and j affects i, as well as longer paths i.e. from i to j to k and back to i. The off-diagonal elements of $S(W)$ represent indirect effects. Following LeSage and Pace (2006), it is possible to define three scalars, which summarize the *total, direct* and *indirect* effects:

v. *Average Direct Impact*: the average of the diagonal elements of $S(W)$, which equals $\frac{1}{n}tr(S(W))$ with tr being the trace of a matrix.

vi. *Average Total Impact from an Observation*: the sum down the jth column of $S(W)$ gives the impact on all y as a result of changing the credit rating variable by an amount in the jth observation (e.g. Greece's rating going from A to A-). There are N of these sums given by the row vector $r = \iota_n{}'S(W)$, where $\iota_n{}'$ is a vector of ones. The average of these effects is equal to $\frac{1}{n}r\iota_n$.

vii. *Average Indirect Effect*: the difference between average total impact and average direct impact.

LeSage and Pace (2009) show that for the SAR model (1) with a row-stochastic matrix W and $abs(\rho) < 1$, the summary measure of total impacts, $\frac{1}{n}\iota_n{}'S(W)\iota_n$, takes the simple form of:

$$\frac{1}{n}\iota_n{}'S(W)\iota_n = \frac{1}{n}\iota_n{}'S(W)(\mathbb{I}_n - \rho W)^{-1}\beta\iota_n = (1-\rho)^{-1}\beta \tag{14}$$

It is computationally inefficient to calculate the summary measures defined in *(i)-(iii)*, because this would require inverting the $n \times n$ matrix $(\mathbb{I}_n - \rho W)$ in $S(W)$. It is possible to approximate the infinite expansion in (12) using traces of the powers of W. In this case, the highest power considered has to be large enough in order to ensure that there is approximate convergence.

In order to make inferences about the statistical significance of the direct and indirect effects of changing the explanatory variables, the distribution of the scalars *(i)-(iii)* is required. I produce the empirical distribution of the parameters β, ρ, σ using a Bayesian Markov Chain Monte Carlo (MCMC) estimation method proposed by LeSage (1997). The idea is that since MCMC yields draws from the posterior distribution of the model parameters, these then can be used in *(i)* and *(ii)* to generate the posterior distribution of the summary measures. Importantly, Gelfand & Smith (1990) show that MCMC yields valid inference in the case of a non-linear function of the parameters, such as *(i)* and *(ii)*. The only requirement is the evaluation and storage of non-linear combinations of parameter values.

The SAR model (10) allows for a differential impact of a change in the explanatory variable depending on the order of the neighbors. Approximating the infinite series expansion of $(\mathbb{I}_n - \rho W)^{-1}$ using the first q powers of W, it is possible ro represent $S(W)$:

$$S(W) \approx (\mathbb{I}_n + \rho W + \rho^2 W^2 +, \dots, +\rho^q W^q)\beta \tag{15}$$

Such a representation allows to the study the effect associated with each power of W. The powers in equation (15) correspond to observations themselves (zero-order), immediate neighbors (first-order), neighbors of neighbors (second-order) etc. Consider again *Example 1*: the first-order neighbor of country 4 is country 1. The *second-order* neighbors are neighbors to the *first-order* neighbors. In this sense second-order neighbors to country 4 are the first-order neighbors to country 1: i.e. country 2, 3 and 4.

$$W^1 = \begin{pmatrix} 0 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 0 & 1 \\ 1/2 & 1/2 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, W^2 = \begin{pmatrix} 0.50 & 0.167 & 0 & 0.333 \\ 1 & 0 & 0 & 0 \\ 0 & 0.167 & 0.167 & 0.667 \\ 0 & 0.333 & 0.333 & 0.333 \end{pmatrix}$$

In the case of W^2, positive elements appear on the diagonal. This is so, because, for example, country 4 is a second-order neighbor to itself. This is not surprising because by definition of a second-order neighbor, country 1 is a neighbor to its neighbor. Importantly, given that $abs(\rho) < 1$, the data-generating process assigns decreasing influence to higher-order neighbors, where the decay declines geometrically as the order increases. If $\rho = 0.5$, then $\rho^2 = 0.25$, $\rho^3 = 0.125$ etc.

Stronger dependence reflects bigger values for ρ, which on its term means that more importance will be assigned to distant neighbors (higher-order).

4.3 Example

To illustrate how the SAR model operates in the context of the theoretical network model, consider again *Example 1* and the matrix W. Recall that:

$$V = \psi W V + \Pi p\left(I_{V>\underline{V}} + \left(I_{V\leq \underline{V}} - \gamma I_{V\leq \underline{V}}\right)\right) \Leftrightarrow (\mathbb{I}_n - \psi W)^{-1}\Pi p\left(I_{V>\underline{V}} + \left(I_{V\leq \underline{V}} - \gamma I_{V\leq \underline{V}}\right)\right)$$

Consider the extreme case, in which country 1's value falls below the payment obligation: country 1 fails and all creditor countries are rationed in proportion to V_1 with countries 3 and 4 claiming $w_{31}V_1$ and $w_{41}V_1$ respectively of its value. For simplicity, I assume that all creditors are of equal seniority. If country 1 cannot meet its debt obligations, it is forced to liquidate its fundamental asset, i.e. its fiscal stream, at a cost $\gamma_1 = 0.10$. Let $p_1 = p_2 = p_3 = p_4 = 10$ and let us assume for the moment that countries 2, 3 and 4 are solvent and so $\gamma_2 = \gamma_3 = \gamma_4 = 0$. Countries have only one fundamental asset, hence $\Pi = \mathbb{I}$ and $p\Pi = \vec{p}$. Further, assume $\psi = 0.5$. Equation (14) rewrites:

$$\begin{pmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{pmatrix} \left(\begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 0.1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right)$$

$$= \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} \begin{pmatrix} 9 \\ 10 \\ 10 \\ 10 \end{pmatrix}$$

The immediate loss in asset value equals $p_1\gamma_1 = 1$. Using the infinite series approximation from equation (12), it is possible to study how the costs are shared and transmitted across the network. The direct effect here is given by $\mathbb{I}_n p\gamma$ and equals -1: this is the failure cost that country 1 pays. However, this is not the end of the story: column 1 of ψW gives the effect of a failure of country 1 on the countries that hold its debt; here, these are country 3 and 4. The incurred cost for country 3 is $\psi w_{31}^1 \times (-1) = 0.5 \times 0.5 \times (-1) = -0.25\, million$ and the cost for country 4 is $\psi w_{41}^1 \times (-1) = 0.5 \times 1 \times (-1) = -0.5\, million$. The second-order neighbors of country 1 are country 1 itself and country 2[46]. Hence, the second-order effects of the shock are $\psi^2 w_{11}^2 \times (-1) = -0.125\, million$ for country 1 and $\psi^2 w_{21}^2 \times (-1) = -0.25$ for country 2. Please notice that if a linear model without dependence relationships was assumed the only cost incurred would be given by $\mathbb{I}_n p\gamma$. This would be largely imprecise and misleading: e.g. due to feedback effects, the costs for country 1 are amplified and it incurs second-order costs of substantial magnitude. The spatial autoregressive model relaxes the independence assumption and allows to study network spillover effects in an easy and intuitive way.

5. Data

5.1 CDS Spreads

In this paper financial health of the sovereign is proxied by CDS spreads. CDS spreads offer a hedge against credit risk, in which the protection sellers agrees to compensate the buyer if the underlying defaults before the contract matures. The fee, which the seller charges, is paid up to end of the contract or until the buyer defaults. This fee is denoted as a CDS spread and is usually quoted in basis points. The way CDS contracts are designed makes them a suitable proxy to assess the probability of default of the borrower. Another advantage of CDS spreads is that they are market-based instruments. As such, they are forward-looking and any price changes today reflect anticipated future performance.

Data for 10 euro-zone sovereigns is collected from Credit Market Analytics (CMA) for the period 2006-2016. The sovereigns included in the sample are: Austria, Belgium, France, Germany, Greece, Ireland, Italy, the Netherlands, Portugal and Spain. All spreads are on 5-year contracts and are used in the analysis either on a monthly or on a daily basis. Since Greek CDS spreads are flat

[46] Country 1 borrows from Country 3 and 4; Country 3 borrows from country 1 and Country 4 borrows from Country 1 and 2. The second-order neighbors of Country 1 are the first-order neighbors of its immediate neighbors i.e. Country 1 and Country 2

for a large part of the sample (2010-2012), for the main part of the analysis the paper works with 9 countries and uses a $9x9$ weight matrix. Greek spreads play a crucial role in the section on predictive regressions.

5.2 Financial Linkages

BIS reports consolidated asset holdings of the financial sector vis-à-vis entities in other countries at quarterly frequency. This measure includes *all* financial assets held by the financial sector and offers a breakdown according to the country that issues the claim. This information is contained in Table 9B of the Quarterly BIS Bulletin. A directed link in this network exists if country i holds a claim vis-à-vis country j and the strength of the connection, x_{ij}, is given by dollar value of the outstanding debt to country j, divided by the total amount that country i borrows from all the countries in the sample. The problem is that Table 9B reports *all* financial claims, not only sovereign debt. On the other hand, Table 4B provides a breakdown on a country level by sector of the counterparty: *banks, public sector*[47]*,* *non-bank* and *private sector*. According to the definition by BIS, international public sector claims refer to “claims to the general government”, which matches the empirical purpose of this paper. Table 4B gives the amount of sovereign debt held abroad, but it does not provide the nationalities of the foreign creditors. I call this amount D_i^{Gvmt}. Finally, to obtain the link between two *sovereigns* I weigh D_i^{Gvmt} by the strength of the connection x_{ij}. These weighted directed links are collected in the matrix W, which is the main input into the SAR model. *Appendix B1* gives an example of how the matrix is constructed.

5.3 Identification of Financial Shocks

In order to identify exogenous financial shocks, this paper focuses on the US financial market and uses two approaches. The first approach is to conduct an event study around announcement dates of major US financial events. In particular, changes in *daily* CDS spreads in the interval $[t - \Delta t^{-}, t + \Delta t^{+}]$ around even time t are examined. The use of daily data and a tight event window allows

[47] In more recent reports the term “public sector” is substituted with the term “official sector”

to minimize contamination problems, which might bias the results. Going through newspaper articles and press releases, 18 events are identified (e.g. collapse of Lehman Brothers, etc.). *Appendix B2* provides a detailed list of these events. However, one problem with this approach is that most of these events occur in 2007-2008, which leaves a considerable part of the data on CDS spreads unused. An alternative approach is to use financial shocks, based on a financial conditions index.

In this paper I use data on the National Financial Conditions Index (NFCI), which is published by the Federal Reserve Bank of Chicago. This index provides a comprehensive weekly update on US financial conditions in money markets, debt markets, equity markets and the traditional and “shadow” banking systems. The NFCI is published every week at 08:30 a.m. ET on Wednesday, and reflects information for the time period through Friday. To construct financial shocks, innovations of the NFCI are taken and values between announcement days are linearly interpolated.

5.4 Other data

This section lists all other variables used as controls in regression models:

Fundamentals:

- ***GDP, Debt-to-GDP, Fiscal Deficit***: measured at a quarterly frequency; seasonally adjusted. Values are linearly interpolated to match the frequency of CDS spreads data. Data is obtained from Datastream

Liquidity:

- ***MRO***: the interest rate on Main Refinancing Operations in the euro-zone, which is set by the Governing Council of the European Central Bank (ECB). MROs provide the bulk of liquidity to the banking system. Updates are made on a monthly basis following regular meetings of the Governing Council. Data is obtained from the ECB Data Warehouse

Market conditions:

- ***OIS***: Overnight Indexed Swap, where the floating payment is chained to the Eonia rate. Available at daily frequency. Data is obtained from the ECB Data Warehouse
- ***LIBOR***: London Interbank Offered Rate for short-term loans. Available at daily frequency. Data is obtained from Datastream. LIBOR carries more risk that the "risk-free" OIS rate.
- ***LIBOR-OIS***: the difference between OIS and LIBOR rates
- ***Eonia***: Euro OverNight Index Average. This is the 1-day interbank interest rate for the euro-zone. Available at daily frequency. Data is obtained from Datastream.

Risk aversion:

- ***Baa-Aaa Spread***: the yield spread of Moody's Baa corporate bonds over Moody's Aaa bonds. A wide spread indicates worsening economic conditions. This occurs because more investors switch to safer Aaa bonds, which pushes down the yield. The money flowing into Aaa bonds typically comes from lower-rated Baa bonds, which at the same time increases the rate of Baa bonds. As economic conditions improve, more investors invest into low-rated bonds, which narrows the spread. For example, the highest value of the Baa-Aaa spread in the sample is 3.50 and it obtains on December 3rd 2008. Its lowest value is 0.53 and it obtains on the June 13th 2014. In this paper it is used as a measure of global risk aversion Data is daily and comes from Federal Reserve Bank St. Louis
- ***VSTOXX*:** Measures the implied 30-day volatility of the EURO STOXX 50. It reflects investor sentiment and overall economic uncertainty in Europe. It is used as a measure of local risk-aversion. Data is daily and comes from Datastream.

6. Empirical Results

6.1 Summary Statistics

In line with prior empirical evidence, CDS spreads of euro-zone sovereigns exhibit significant degree of comovement throughout the sample (Figure 2). Prior to the financial crisis, spreads of all countries move closely together. During the peak of the Sovereign Debt Crisis, two groups of countries are noticeable: central countries (France, Germany, Belgium, Austria and the Netherlands) with low CDS spreads and peripheral countries (Portugal, Ireland, Italy and Greece) with high CDS spreads.

Table 1 shows summary statistics of CDS spreads over the sample period from 2006-2016. The results are in line with Figure 2. Countries such as Austria, France, Germany and the Netherlands have mean spreads between 30 and 50 basis points, whereas the spreads of Ireland, Italy, Portugal and Spain exhibit spreads that are on average 5 to 10 times bigger. Germany has the lowest mean and lowest standard deviation. In fact, Germany was the only country, which retained its AAA rating during the Sovereign Debt Crisis of 2010-2012. For this reason, it is reasonable to use German CDS spreads as a proxy for the risk-free asset. In all subsequent sessions, any reference to spreads should be understood to be on an *adjusted* basis i.e. the spread over Germany.

6.2 Unit Root Tests

If the variables in the spatial regressions are not stationary, then standard assumptions on asymptotic behavior and inference are not valid. Table 2 reports the results of panel unit root tests based on Augmented Dickey-Fuller (ADF) regressions. The value of the test-statistic for *CDS* suggests the presence of a unit root. However, the adjusted measure, *CDSadj* is stationary with p-value close to zero. For the rest of the variables, with the exception of *Debt, MRO* and *Eonia,* the null hypothesis of a non-stationarity is rejected. Therefore, regression results, where these three variables are included should be interpreted with caution.

6.3 Spatial Dependence

A useful visual tool for exploratory spatial analysis is the Moran Scatter Plot (Anselin, 1996). It allows to assess how similar an observed value is to its neighboring observations. The horizontal axis gives the values of the observations and is called the *response axis*. The vertical axis is based on the weighted average or spatial lag of the corresponding observations on the response axis. On Figure 3, the horizontal axis gives monthly changes in CDS spreads, the vertical axis gives the spatial lag, where the weighting scheme is defined based on financial linkages. In this particular example, data from 2005 Q3 is used to construct the spatial lag. Two observations are immediate. First, if the spatial weights matrix did not contain any relevant information, observations would be *randomly* scattered in the plot. This is not the case and data appears to be centered. Second, the majority of the data points are concentrated in Quadrant I and in Qudrant III. High values surrounded by high neighboring values are located Quadrant I and low values surrounded by low values are located in Quadrant III. Together, these two quadrants account for roughly 75% of the data points. These results are intuitive: countries that are connected to countries with high credit risk are riskier themselves and this is reflected in a higher CDS spread. Thus, Moran Scatter Plot offers preliminary evidence for the relevance of spatial dependence for credit risk.

6.4 In-sample Results

Table 3 compares the in-sample performance of a standard model of CDS spreads (columns (1)-(2)) and the SAR model (columns (3)-(5)). The specification of the standard model is motivated by the existing literature (*See Section II. Literature Review*): spreads are modeled as a persistent mean-reverting process determined by local factors, driven by fundamentals, and a global factor, driven by risk-aversion. To capture the impact of fundamentals, I use debt-to-GDP ratio (*Debt*) and deficit-to-GDP ration (*Deficit*). Risk aversion is proxied by Moody's Baa-Aaa spread (*spread*). To fulfill stationarity conditions, lags of adjusted CDS spreads and risk aversion are included. Column (1) reports the results of OLS regressions, consistent with prior results in the literature: the higher the indebtedness of a sovereign, the higher the CDS spread; the smaller the fiscal deficit, the smaller the CDS spread; the higher the risk aversion and market uncertainty, the higher is the CDS spread. The standard model is able to explain around 19 % of the total variation, which is markedly low.

In column (2), other variables, which have been found relevant in the empirical literature are included. *GDP* is found to be negatively correlated with CDS spreads. The volatility of the stock market (*VSTOXX*) carries a positive premium, although the magnitude of the coefficient is lower than the one for the global risk aversion. A potential explanation could be that the two are positively correlated. The coefficient on the European stock market index is close to zero and insignificant. The ECB's interest rate (*MRO*) is positively related to CDS spreads and the coefficient is statistically significant. The overnight interbank offered rate (*Eonia*) carries a negative sign, which is not surprising because the rate was considerably lowered during the peak of the crisis, when CDS spreads were at their highest. The spread *LIBOR-OIS* captures marker uncertainty and in this model is negatively related to CDS spreads. Finally, to control for the size of the financial system, the variable *SizeFinSector* is included, which is defined as total banking assets over GDP. Countries, where the financial sector accounts for a bigger fraction of total GDP are more exposed to systemic risk and, hence are characterized by higher levels of CDS spreads. It is interesting to note that even though the model is saturated with many variables, the improvement in explanatory power is low.

Column (3) reports the results of the SAR model estimated by maximum likelihood. To address endogeneity issues, the spatial weights matrix W is calculated using data from BIS 2005 Q3, which is entirely pre-determined with respect to the sample. The spatial autoregressive parameter ρ is positive and strongly statistically significant. This is interpreted as evidence for strong network spillover effects. Since, unlike the OLS estimates, the SAR coefficients have the interpretation of *average* partial derivatives, it is not possible to discuss directly their magnitude, only their sign. Comparing columns (1) and (2) to column (3), note that the direction of the coefficients is the same. Importantly, the explanatory power of the model increases more than two times with an R-squared of 49 %.

One concern could be that the spatial lag does not capture new information, but it is correlated to other variables, previously found in the literature. To discard this doubt, controls are included in the SAR model (column (4)). The parameter ρ continues to be positive and statistically significant and its magnitude is largely unchanged. This leads to the conclusion that the spatial lag is an important factor, which is omitted from standard specifications.

Given the high values for ρ, another concern naturally emerges. It could be that this value is mechanically attributed because changes in CDS spreads are present both on the right and the left hand side of the equation. A random *pseudo* W is generated, whose entries are drawn from an uniform distribution, is used in column(5). The value of ρ is 0.13, which is 5 times lower than the values in the previous two specifications. Therefore, it is safe to assume that the results on network spillovers are not driven by randomness.

In Table 4, the effect of the SAR coefficients is decomposed into direct and indirect effects using the formula from Section IV B. Except for *Debt*, all decompositions are statistically significant. Indirect effects constitute approximately 60% of the overall effect. Comparing *Total* and *OLS*, it is noticed that OLS coefficients are systematically lower than SAR coefficients and the difference is roughly 20%.

To summarize, the in-sample results provide evidence for the systemic nature of credit risk and for strong network spillovers in CDS markets. The SAR model introduces a new factor, which improves significantly the explanatory power of the model. The results are intuitive and are in line with the predictions of the theoretical model outlined in Section III.

Having addressed the first research question, the paper moves on to study how exogenous financial shocks are transmitted through the network of financial linkages.

6.5 Financial Shocks

To answer this *second research question*, the paper uses exogenous financial shocks identified using the methodology described in Section V C: (1) using announcement dates of negative financial events and (2) using innovations of a financial conditions index (NFCI). All results in this section use *daily* CDS spreads or changes in the CDS spreads.

To study the effects of exogenous financial events, I conduct an event study around announcement dates. Changes in CDS spreads in a two-day event window [-1,1] bracketing US financial shocks are regressed on an event dummy (*FinShock*) and a spatial lag. Table 5 tabulates the results of the event study.

Panel A gives point estimates of three models: OLS without a spatial lag, a SAR model with a weight matrix constructed from BIS data 2005 Q3 and a SAR model with a *pseudo W* for robustness. The OLS estimate of the event dummy is positive and significant: a negative financial shock in the US financial market affects positively the European sovereign CDS market. Introducing the spatial lag in column (2), the event dummy remains positive and significant. The spatial autoregressive parameter equals 0.50 and is strongly statistically significant. In the last column a test with a *pseudo spatial weights matrix* is performed, which yields results consistent with the statement that the results of the SAR model are not random.

In order to be able to compare the OLS and SAR results, Panel B offers a decomposition of *FinShock*. For the OLS model, the *Total Effect* equals the point estimate from Panel A. In the case of the SAR model, the *Total Effect* is a sum of the *Indirect Effect* and the *Direct Effect*. It is interesting to note that indirect network effects account for nearly 46% of the overall effect. Comparing the overall effect, SAR model yields effects, which are 25% higher than the OLS estimate. The difference between the two is attributed to propagation effects through the network of sovereign financial linkages.

Although significant, the overall effect of event study is very small in magnitude. One reason could be the small number of relevant events in the sample, which reduces the statistical power of the test. Another approach is to use innovations in the NFCI index computed by the Federal Reserve Bank of St. Louis. This allows to identify a continuum of shocks throughout the entire sample 2006:2016.

The main variable of interest in *ShockNFCI*, defined as weekly innovations in the NFCI index linearly interpolated between announcement days (Thursdays). High values of the NFCI signify economic instability, whereas low values indicate downturn and recession. The mean value of the index is -0.33, the minimum is -0.94 and the maximum is 2.86. The average value of the innovations is 0.001 with occasional positive spikes (positive shocks) and negative spikes (negative shocks). Table 6 reports the average effect of exogenous financial shocks. Changes in CDS spreads react significantly to the unexpected component of the NFCI index. The effect is between 15 to 16 basis points with nearly 45 % of it due to indirect network effects. In column (3), the SAR model uses an *averaged W*. This matrix is computed by averaging the bilateral sovereign exposures from 2006 Q1 to 2016 Q4 and using these values to compute the relative weights. The results are qualitatively unchanged, which I interpret as evidence of the robustness of the results.

6.6 Out-of-sample Prediction

Finally, this paper investigates whether the forecasting performance of models of CDS spreads is improved when a spatial lag is introduced. To test this, I use data from 2006 to 2012. I use monthly data, because data at a higher frequency normally contains a lot of noise. I split the sample into two parts: an estimation sample 2006:2010 and an evaluation sample 2011-2012. The evaluation period is chosen deliberately to cover the Sovereign Debt Crisis. I estimate the standard and the SAR model on data from 2006 to 2010. The estimation includes the Greek CDS spread in the sample and conditions on the spatial lag, fundamentals and risk aversion. Conditioning on the same set of information allows to attribute differences in forecasting performance to the way the two models process the same information. The evaluation sample consists of Austria, Belgium, France, Ireland, Italy, the Netherlands, Portugal and Spain i.e. Germany and Greece are excluded for obvious reasons. The German CDS spread is considered to be the safe asset, whereas in the case of Greece the sharp increase in CDS was due to misrepresentation of fundamentals and, hence, the Greek spread is unlikely to be predicted from a weighted average of neighboring countries[48].

Table 7 reports the results for predicted CDS spreads in changes (Panel A) and in levels (Panel B). Let's focus on first on Panel A. The mean observed change in CDS spread is -70.88. The standard model is unable to match the pattern of comovement in the data with a predicted mean of -6.38, which is nearly 63 basis points away from the observed one. The SAR model predicts -31.83, which is an improvement of 25.45 basis points. Furthermore, the SAR model improves the forecasting accuracy by 20 % in the root-mean-squared-error (RMSE) sense. Turning to predictions in levels (Panel B), the results are even more impressive. The observed mean CDS spread is 237.60 basis points. The prediction of the standard model, 77.52 basis points, misses the true mean by 160 basis points. The SAR model on the other hand predicts spreads, which at mean value of 159.99 are considerably closer to the true data. In terms of accuracy, the SAR model reduces RMSE by around 15%.

The SAR model consistently outperforms the standard model in this out-of-sample prediction exercise both in terms of matched values and accuracy. This suggests that the SAR model with a spatial weights matrix based on sovereign financial linkages offers a powerful, yet intuitive tool for econometric modeling and forecasting of CDS spreads.

7. Robustness

This section discusses additional results and demonstrates that the findings reported in the previous section are robust to alternative specifications of time periods and the spatial weights matrix.

7.1 Structural Breaks

I test for structural breaks in the coefficients by employing a simple Wald Test approach. I split the sample into five periods: 2006:2007, 2008:2010, 2011:2012,

[48] Neighbors in the financial network sense

2013:2014, 2015:2016. These correspond roughly to the *Pre-crisis Period, the Global Liquidity Crisis, the European Sovereign Debt Crisis* and 2 *Post-crisis periods*. I create five dummy variables *Period1, Period2, Period3, Period 4* and *Period5* that take values equal to one if one of the five periods applies. Then variables are interacted with the dummies. I run the following regression:

$$\begin{aligned}\Delta CDS_t = \beta_0 + \sum_{j=1}^{5} \rho W \Delta CDS_t \times Period_j \\ + \sum_{j=1}^{5} \beta_j Debt_t \times Period_j + \sum_{j=1}^{5} \beta_{2j} Deficit_t \times Period_j \\ + \sum_{j=1}^{5} \beta_{3j} RiskAv_t \times Period_j + \sum_{j=1}^{5} \beta_{4j}\, CDS_{t-1} \times Period_j \\ + \sum_{j=1}^{5} \beta_{5j} RiskAv_{t-1} \times Period_j + \varepsilon_t\end{aligned} \tag{2}$$

Then, I conduct Wald tests for each set of variables. For example, $\beta_{11} = \beta_{12} = \beta_{13} = \beta_{14} = \beta_{15}$ etc. Judging from the values of the F-statistic and the corresponding p-values, there is significant evidence for structural breaks in the coefficients induced by the Global Financial Crisis and the European Sovereign Debt Crisis (Table 8).

7.2 Alternative Time Periods

The results of the tests in Section A suggest changes in the time-series of the variables from one period to the other. One concern could be that the presence of network spillovers in CDS markets is an artefact of the long sample or that it is driven by very strong spillovers during a sub-sample of the data. It will be useful to see whether the spatial autoregressive parameter remains significant when the data is split into subsamples and how its values change.

Table 9 documents significant differences in the spatial autoregressive parameter throughout the samples. It is lowest during 2006:2007, when CDS spreads are low and the levels of risk are small. It skyrockets during 2008:2010 to 0.67 at the peak of the Liquidity Crisis and the onset of the Sovereign Debt Crisis. During 2011:2012, the value of the parameter is reduced nearly in half, which is a result of government interventions and macroeconomic policies aiming at decoupling the financial system. During the next two periods after the crisis, values of the parameter increase, which reflects the strong comovement and narrowing down of the differences in CDS spreads between countries in the sample. These results are consistent with the graphical evidence in Figure 2.

7.3 Alternative Specifications of the Spatial Weights Matrix

This section establishes the robustness of the results using different specifications of the spatial weights matrix.

In column (1) of Table 10, I use a categorical spatial matrix, which is computed in the following way. Data for financial linkages is averaged over all periods from 2006 Q1 to 2016 Q4 and collected in a matrix *Average*. Then the following rule is applied:

$$w_{ij}\begin{cases}1 & if\ average_{ij} \leq p.25\\ 2 & if\ p.25 < average_{ij} \leq p.50\\ 3 & if\ average_{ij} \geq p.75\end{cases}$$

where $p.25, p.50$ and $p.75$ stand for the 25^{th} ,50^{th} and 75^{th} percentile of the matrix *Average*. In order to ensure that $abs(\rho) < 1$, row-normalization is applied. The results indicate significant spillover effects, but the magnitude of the spatial autoregressive parameter is lower, because some information is lost when the categorical scheme above is carried out.

In column (2) of Table 10, I use data from BIS 2010 Q4 to capture changes in the dynamics of cross-border borrowing and lending. The network parameter remains significant and positive.

Another potential source of bias could arise due to the assumptions of the construction method. Recall that the entry $w_{ij} = x_{ij}D_i^{Gvmnt}$, $x_{ij} = \frac{a_{ij}}{\sum_j a_{ij}}$, a_{ij} is claims of i vis-à-vis j in 2005 Q3 and $j = 1,..,9$. This *assumes* a closed system i.e. no lending and borrowing outside of the network of countries in the sample. Since, it is reasonable to believe that in reality countries in the sample hold claims vis-à-vis countries outside of the euro-zone, the weighting scheme x_{ij} might *overstate* the effect. To check whether this is true, the following is introduced:

$$x_{ij}^{BIS} = \frac{a_{ij}}{\sum_{j=1}^{BIS} \sum_j a_{ij}}$$

where *BIS* stands for the total number of BIS-reporting banks. Then, let $W^{BIS} = x_{ij}^{BIS} D_i^{Gvmnt}$. The results of column (3) alleviate this concern: the sign and magnitude of all coefficients in the model remain almost the same.

8. Conclusion

To summarize, using a simple network model of sovereign credit risk, this paper provides an economic motivation for the use of spatial autoregressions to model and predict sovereign CDS spreads. In the spirit of recent theoretical work on networks, this paper develops a network model with asset interdependencies. A "balance sheet" mechanism of contagion is considered, where spillovers following an exogenous financial shock (e.g. a US financial shock) occur via direct losses to assets held by creditors. Using a fixed-point argument, it is possible to show that in the presence of *asset interdependencies* and *discontinuities* in value multiple equilibrium solutions for organization's values are possible. In this context of multiple equilibria, contagion emerges because of linkages and the joint determination of asset prices: organizations fail because people *expect* that other connected organizations will fail as well and this then becomes *self-fulfilling*.

Next, the paper shows that the theoretical model can be empiricaly operationalized via a spatial autoregressive (SAR) model. Using methods from spatial econometrics, the paper makes two empirical findings. First, the paper shows that the constructed network of financial linkages between sovereigns is an important mechanism for the propagation of exogenous financial shocks. Using the SAR model, it is possible to decompose the *total* effect of financial shocks to *direct* and *indirect* effects. The paper finds that as much as 45% of the overall effect of shocks is due to indirect (network) effects. Second, in out-of-sample predictive tests, the SAR model consistently outperforms standard models. The SAR model is better able to match monthly changes in CDS spreads and leads to 25 % to 35% improvement in predictive accuracy, measured in the root mean squared error (RMSE) sense.

References

Acemoglu, D., Ozdaglar, A., & Tahbaz-Salehi, A. (2015). Systemic Risk and Stability in Financial Networks. *American Economic Review, 105*, 564-608.

Afonso, A., Furceri, D., & Gomes, P. (2012). Sovereign Credit Ratings and Financial Market Linkages. *Journal of International Money and Finance, 31*(3), p. 606-638.

Allen, F., & Gale, D. (2000). Financial contagion. *Journal of Political Economy, 108*, 1-33.

Allen, F., Babus, A., & Carletti, E. (2010). Financial connections and systemic risk. *National Bureau of Economic Research*.

Amato, D., & Remolona, E. (2003). The credit spread puzzle.

Anselin, L., & Bera, A. (1998). Spatial dependence in linear regression models with an introduction to spatial econometrics. *Statistics Textbooks and Monographs, 155*, 237-290.

Anselin, L., & LeGallo, J. (2006). Interpolation of Air Quality Measures in Hedonic House Price Models. *Spatial Economic Analysis, 1*(1), p. 31-52.

Anselin, L., Bera, A., Florax, R., & Yoon, M. (1996). Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics, 26*(1), 77-104.

Arbia, G. (2006). *Spatial econometrics: statistical foundations and applications to regional convergence.* Springer Science and Business Media.

Attinasi, M., Checherita, C., & Nickel, C. (2010). What explains the surge in euro area sovereign spreads during the Financial Crisis 2007-2009? *Public Finance and Management, 10*(4), 595-645.

Barucca, P., Bardosia, M., Caccioli, F., D'Errico, M., Visentin, G., Battiston, S., & Caldarelli, G. (2016). Network Valuation in Financial Systems. *Working Paper*.

Battiston, S., Gatti, D., Gallegatti, M., Greenwald, B., & Stiglitz, J. (2012). Liaisons dangereuses: Increasing connectivity, risk sharing, and systemic risk. *Journal of Economic Dynamics and Control, 36*(8), 1121-1141.

Battiston, S., Puliga, M., Kaushik, R., Tasca, P., & Caldarelli, G. (2012). Debtrank: Too central to fail? financial networks, the fed and systemic risk. *Scientific Reports*.

Beber, A., Brandt, M., & Kavajecz, K. (2009). Flight-to-quality or flight-to-liquidity? *Review of Financial Studies, 22*(3), 925-957.

Bernanke, B., & Kuttner, K. (2005). What explains the stock market's reaction to Federal Reserve Policy. *Journal of Finance, 60*(3), p. 1221-1257.

Blasques, F., Koopman, S., Lucas, A., & Schaumburg, J. (2016). Spillover dynamics for systemic risk measurement using spatial financial time series models. *Journal of Econometrics, 195*(2), 211-223.

Bloningen, B., Davies, R., Waddell, G., & Naughton, T. (2007). FDI in space: Spatial autoregressive relationships in foreign direct investment. *European Economic Review, 51*(5), 1303-1325.

Bonacich, P. (1987). Power and centrality: a family of measures. *American Journal of Sociology, 92*, p. 1170-1182.

Bordignon, M., Cerniglia, F., & Revelli, F. (2003). In search of yardstick competition: a spatial analysis of Italian municipality property tax setting. *Journal of Urban Economics, 54*(2), 199-217.

Chen, L., Collin-Dufresne, P., & Goldstein, R. (2008). On the relation between the credit spread puzzle and the equity premium puzzle. *The Review of Financial Studies, 22*(9), 3367-409.

Codogno, L., Favero, C., & Missale, A. (2003). Yield spreads on EMU government yields. *Economic Policy, 18*(37), 503-532.

Cohen-Cole, E., Patacchini, E., & Zenou, Y. (2011). Systemic risk and network formation in the interbank market. *CERP Discussion Paper*.

Cook, T., & Hahn, T. (1989). The effects of changes in the federal funds rate target on market interest rates. *Journal of Monetary Economics, 24*(3), p. 331-351.

Craig, B., & Von Peter, G. (2014). Interbank tiering and money center banks. *Journal of Financial Intermediation, 23*(3), 322-347.

Diebold, F., & Yilmaz, K. (2009). Measuring financial asset return and volatility spillovers, with application to global equity markets. *Economic Journal, 119*, 158-171.

Diebold, F., & Yilmaz, K. (2011). Equity market spillovers in Americas. (R. Alfaro, A cura di) *Financial Stability, Monetary Policy and Central Banking*, 199-214.

Dybvig, D., & Dybvig, P. (1983). Bank runs, Deposit Insurance and Liquidity. *Journal of Political Economy, 91*, p. 401-419.

Eder, A., & Keiler, S. (2015). CDS Spreads and Contagion Amongst Systemically Important Financial Institutions–A Spatial Econometric Approach. *International Journal of Finance & Economics, 20*(4), 291-309.

Eisenberg, L., & Noe, T. (2001). Systemic Risk in Financial Systems. *Management Science, 47*, p. 236-249.

Elliot, M., Golub, B., & Jackson, M. (2014). Financial Networks and Contagion. *American Economic Review, 104*(10), p. 3115-3153.

Favero, C. (2013). Modeling and forecasting government bond spreads in the euro area: a GVAR model. *Journal of Econometrics, 177*(2), 343-356.

Favero, C., Pagano, M., & von Thadden, E.-L. (2010). How does liquidity affect government bond yields? *Journal of Financial and Quantative Analysis, 45*(1), 107-134.

Fernandez, V. (2011). Spatial linkages in international financial markets. *Quantative Finance, 11*(2), 237-245.

Freixas, X., Parigi, B., & Rochet, J. (2000). Systemic risk, interbank relations and liquidity provision by the central bank. *Journal of Money, Credit and Banking*, 611-638.

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

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

Gande, A., & Parsley, D. (2005). News spillovers in the sovereign debt market. *Journal of Financial Economics, 75*(3), p. 691-734.

Geier, A., Kossmeier, S., & Pichler, S. (2004). Measuring systematic risk in EMU government yield spreads. *Review of Finance, 8*(2), 171-197.

Gelfand, A., & Smith, A. (1990). Sampling-based Approaches to Calculating Marginal Densities. *Journal of the American Statistical Association, 85*, p. 398-409.

Georg, C. (2013). The effect of the interbank network structure on contagion and common shocks. *Journal of Banking and Finance, 37*(7), 2216-2228.

Glaserman, P., & Young, P. (2015). How likely is contagion in financial networks? *Journal of Banking and Finance.*

Glasserman, P., & Young, H. (2015). How likely is contagion in financial networks? *Journal of Banking and Finance, 50*, 383-399.

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

Kelejian, H., & Prucha, I. (1998). A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *The Journal of Real Estate Finance and Economics, 17*(1), 99-121.

Kelejian, H., & Prucha, I. (1999). A generalized moments estimator for the autoregressive parameter in a spatial model. *International Economic Review, 40*(2), 509-533.

Kelejian, H., & Prucha, I. (2004). Estimation of simultaneous systems of spatially interrelated cross sectional equations. *Journal of Econometrics, 118*, 27-50.

Kelejian, H., & Prucha, I. (2010). Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of Econometrics, 157*, 53-67.

Kelejian, H., Tavlas, G., & Hondronyiannis, G. (2006). A Spatial Modeling Approach to Contagion Among Emerging Economies. *Open Economies Review, 17*(4/5), p. 423-442.

Kim, C., Phipps, T., & Anselin, L. (2003). Measuring the Benefits of Air Quality Improvement: A Spatial Hedonic Appriach. *Journal of Environmental Economics and Management, 45*, p. 24-39.

Laubach, T. (2009). New Evidence on the interest rate effect of budget deficits and debt. *Journal of the European Economic Association*, 858-885.

Laubach, T. (2011). Fiscal policy and interest rates: the role of sovereign default risk. (R. Clarida, & F. Giavazzi, A cura di) *NBER International Seminar on Macroeconomics*, 7-29.

Le Gallo, J., Ertur, C., & Baumont, C. (2003). A spatial econometric analysis of convergence across European regions, 1980-1995. *European Regional Growth*.

Lee, L.-F. (2004). Asymptotic Distributions of Quasi-Maximum Likelihood Estimators for Spatial Econometric Models. *Econometrica, 72*, p. 1899-1926.

Leitner, Y. (2005). Financial networks: Contagion, commitment, and private sector bailouts. *Journal of Finance, 60*(6), 2925-2953.

Leontief, W. (1951). *Input-Output Economcis.* New York: Oxford University Press.

LeSage, J. (1997). Bayesian Estimation of Spatial Autoregressive Models. *International Regional Science Review, 20*(1&2), p. 113-129.

LeSage, J., & Pace, R. (2006). Interpreting Spatial Econometric Models. *Regional Science Association International North American Meetings.*

LeSage, J., & Page, R. (2009). *Introduction to Spatial Econometrics.* CRC Press Taylor & Francis.

Longstaff, F., Pan, J., Pedersen, L., & Singleton, K. (2011). How sovereign is sovereign credit risk? *American Economic Journal, 3*(2), 75-103.

Nier, E., Yang, J., Yorulmazer, T., & Alenton, A. (2007). Network models and financial stability. *Journal of Economic Dynamics, 31*(6), 2033-2060.

Peseran, M., Schuermann, T., & Weiner, S. (2004). Modeling Regional Interdependencies using a Global Error-correcting macroeconometric Model. *Journal of Business and Economic Statistics, 22*(2), 129-162.

Rogers, L., & Veraart, L. (2013). Failure and rescue in the interbank network. *Management Science, 59*, p. 882-898.

Suzuki, T. (2002). Valuing corporate debt: the effect of cross-holdings of stock and debt. *Journal of the Operations Research Society, 45*, p. 123-144.

Tarski, A. (1955). A lattice-theoretical fixpoint theorem and its applications. *Pacific Journal of Mathematics, 5*(2), p. 285-309.

Weber, M., & Ozdagli, A. (2016). Monetary Policy through Production Networks. *Working Paper*.

List of Figures

Figure 1: Conditional and Unconditional Failure Frontiers

The figure plots the example from section *3.6 A Simple Microfoundation*.

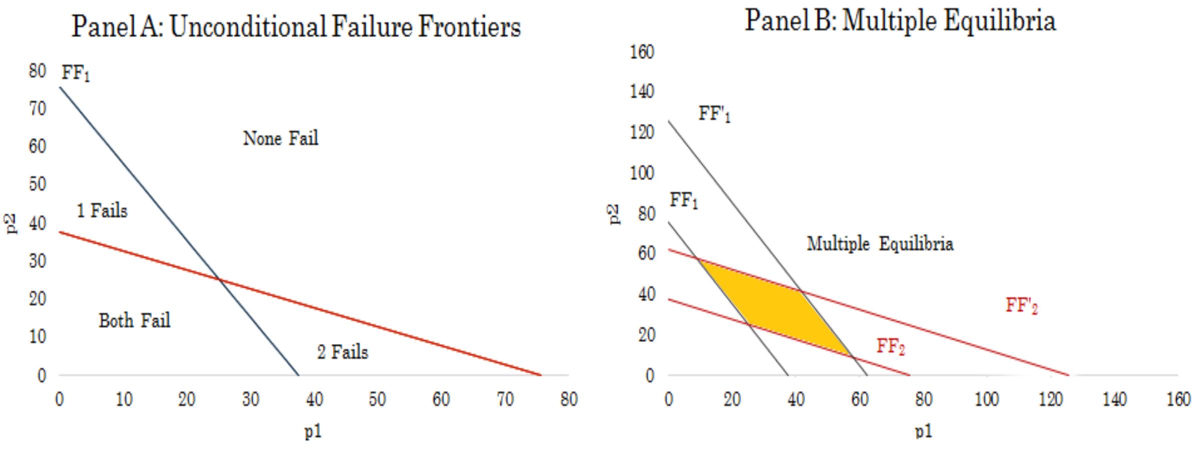


Figure 2: CDS spreads of Euro-zone Countries

The figure shows monthly CDS spreads of Euro-zone countries: Austria, Belgium, France, Germany, Greece, Ireland, Italy, The Netherlands, Portugal and Spain. Sample period is 2006-2016.

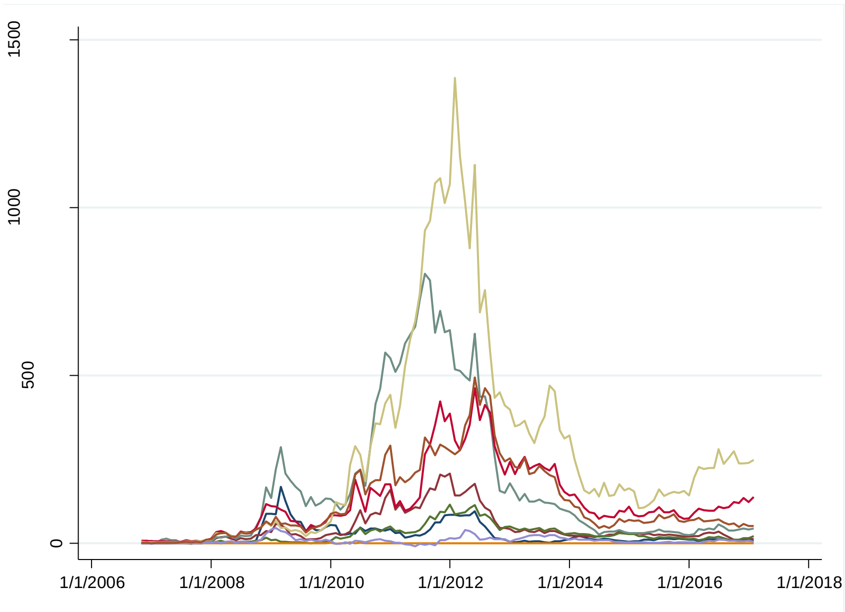


Figure 3: Moran Scatter Plot

The figure shows Moran Scatter Plot. The Spatial lag is constructed using a weighting matrix from 2005 Q3 BIS Data.

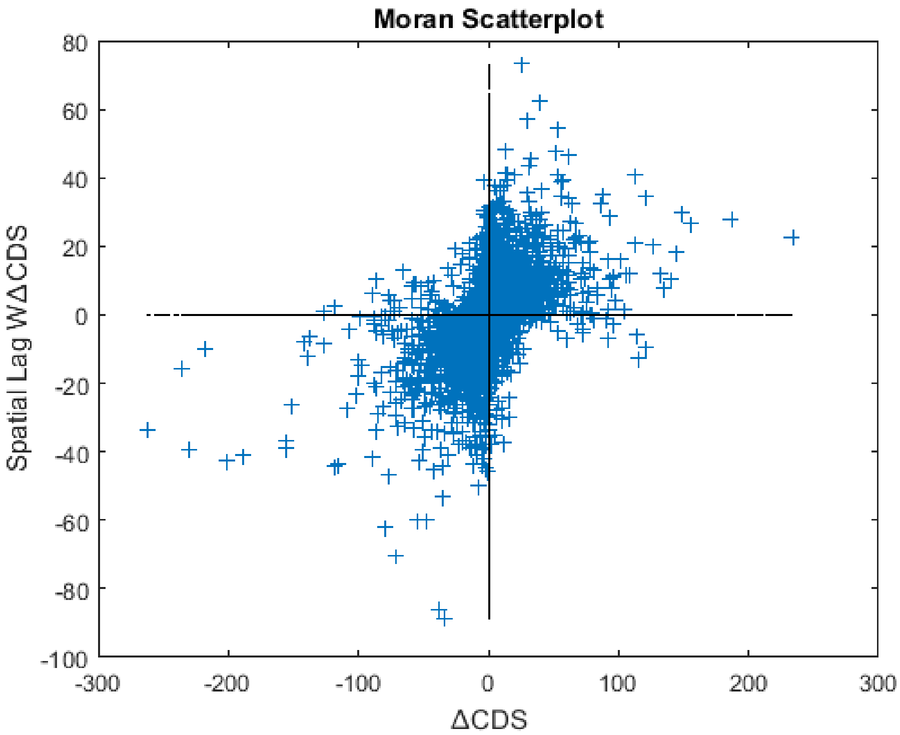
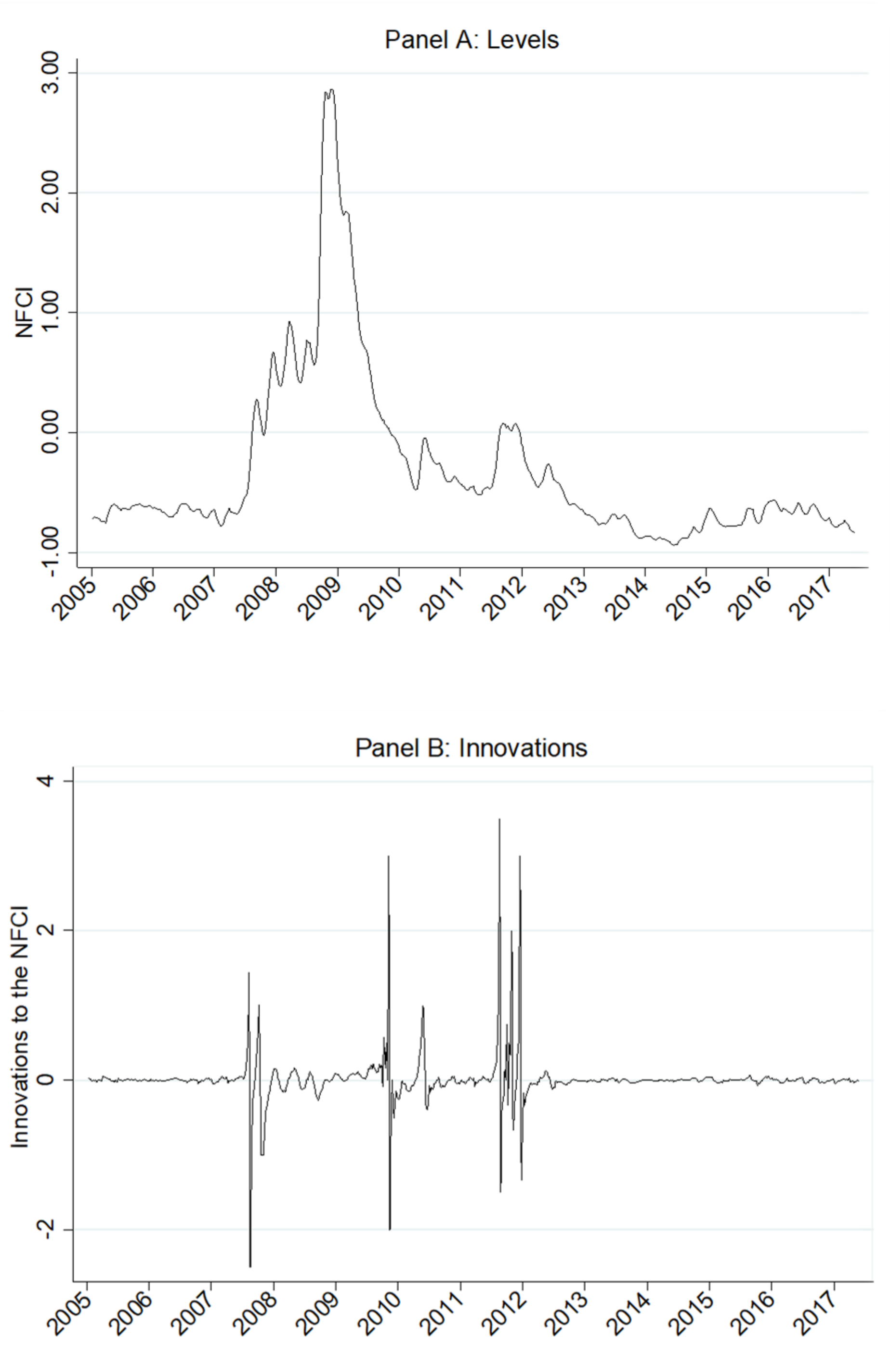


Figure 4: The NFCI Index

The figure plots the NFCI Index: Panel A in levels, Panel B innovations



List of Tables

Table 1: Summary Statistics of CDS Spreads

Country	Mean	Min	Max	SD
Austria	55.81	1.4	255.81	51.90
Belgium	76.51	1.7	307.41	73.02
France	57.08	1.5	214.86	49.16
Germany	30.58	1.5	114.35	25.11
Greece	10504.24	4.5	37030.49	15899.13
Ireland	202.50	2	866.19	232.82
Italy	158.85	5.6	563.40	125.92
The Netherlands	38.53	1.3	128.27	30.55
Portugal	304.35	3.4	1471.74	323.74
Spain	152.34	2.5	595.93	133.00

Table 2: Unit Root Tests

The Table reports results of Panel Unit Root Tests (Levin-Lin-Chu).Values of the Augmented Dickey-Fuller (ADF) statistic and p-value are reported. ADF regressions are performed with one lag.

Variable	ADF Statistic	p-value
CDS	17.34	0.63
CDSadj	60.00	0.00
GDP	-1.66	0.04
Debt	-1.21	0.11
Deficit	-5.75	0.00
spread	-3.00	0.00
VSTOXX	-3.17	0.00
MRO	-0.68	0.77
Libor-OIS	-2.82	0.05
Eonia	-0.77	0.82
STOXX	-1.35	0.08

Table 3: In-sample Regression Analysis

The table reports the results of regressing changes in CDS spreads on innovations of the National Financial Conditions Index (NFCI) published by the Federal Reserve Bank of St. Louis (FRED), a set of controls and a spatial lag (column 3). The sample ranges from 2006 to 2017. T-stats are reported in parentheses. Significance at the 1%, 5% and 10% is given by ***, ** and * respectively.

$$\Delta CDS_t = \beta_0 + \rho W \Delta CDS_t + \beta_1 Debt_t + \beta_2 Deficit_t + \beta_3 RiskAv_t + \beta_4 CDS_{t-1} + \beta_5 RiskAv_{t-1} + \varepsilon_t$$

	OLS Standard Model	OLS Standard Model	SAR: 2005Q3	SAR: 2005Q3	SAR: Pseudo W
	(1)	(2)	(3)	(4)	(5)
ρ			**0.64*** (28.93)	**0.61*** (25.88)	**0.13*** (18.09)
$Debt$	0.08 (1.39)	0.34*** (5.29)	0.035 (0.80)	0.111** (2.14)	0.124** (2.26)
$Deficit$	-1.18*** (-3.49)	-1.25*** (-5.11)	-1.12*** (-4.16)	-1.32*** (-4.58)	-1.42*** (-4.67)
$RiskAv$	45.00*** (5.80)	6.09 (0.33)	19.46*** (3.10)	0.40 (0.04)	2.78 (0.289)
CDS_{t-1}	-0.11*** (-12.80)	-0.14*** (-12.99)	-0.06*** (-8.47)	-0.06*** (-8.75)	-0.06*** (-8.67)
$RiskAv_{t-1}$	-56.16*** (-7.84)	-32.98*** (-3.66)	-23.75*** (-3.76)	-13.15* (-1.75)	-15.46** (-1.96)
GDP		-0.82*** (-2.74)		-0.40* (-1.64)	-0.206 (-0.78)
$VSTOXX$		1.07*** (4.07)		0.60*** (2.72)	0.51*** (2.19)
$StockInd$		0.000 (0.33)		0.000 (0.03)	0.001 (0.20)
MRO		2.47*** (4.19)		0.84 (1.58)	1.27 (0.80)
$Eonia$		-3.95 (-1.44)		-1.55 (-0.67)	-1.94*** (-2.27)
$Libor - OIS$		-2.06*** (-2.95)		-0.95 (-1.55)	-0.97 (-1.51)
$Size\ FinSector$		0.23*** (4.75)		0.20*** (3.75)	0.15 (3.00)
Adj R^2, %	18.64	26.42	48.70	49.50	44.52
Observations	1080	1080	1080	1080	1080

Table 4: Decomposition: Direct, Indirect and Total Effects

This table reports the decomposition to and indirect effects of the coefficients of the SAR model in column (3) of Table 2

	$Debt$	$Deficit$	$RiskAv$	ΔCDS_{t-1}	$RiskAv_{t-1}$
$Direct$	0.04	-1.28***	22.42***	-0.06***	-27.26***
	(0.85)	(-4.14)	(3.24)	(-8.34)	(-3.85)
$Indirect$	0.05	-1.82***	31.605***	-0.09***	-38.43***
	(0.84)	(-3.89)	(3.24)	(-6.91)	(-3.85)
$Total$	0.10	-3.10***	54.02***	-0.15***	-65.70***
	(0.85)	(-4.04)	(3.26)	(-7.71)	(-3.90)
OLS	0.08	-1.18***	45.00***	-0.11***	-56.16***
	(1.39)	(-3.49)	(5.80)	(-12.80)	(-7.84)

Table 5: Event study: Response of CDS Spread to Exogenous Negative Financial Events

The table reports the results of regressing changes in CDS spreads in a two-day event window [-1, 1] bracketing US financial shocks on an event dummy and a spatial lag. The sample ranges from 2006 to 2017. T-stats are reported in parentheses. Significance at the 1%, 5% and 10% is given by ***, ** and * respectively.

$$\Delta CDS_{[t-1,t+1]} = \beta_0 + \rho W \Delta CDS_{[t-1,t+1]} + \beta_1 FinShock_t + \varepsilon_t$$

	OLS No Spatial Lag	SAR: 2005Q3	SAR: Pseudo W
	(1)	(2)	(3)
Panel A: Point Estimates			
ρ		0.50*** (88.96)	0.13*** (67.95)
FinShock	**1.49***** (2.36)	**0.89**** (1.89)	**0.75**** (1.96)
Constant	0.11** (5.32)	0.005 (0.02)	0.006 (0.06)
Adj R^2, %	0.40	17.00	17.3
Observations	27408	27408	27408
Panel B: Decomposition of *FinShock*			
Direct Effect		**0.98** (1.88)**	**0.76** (1.95)**
Indirect Effect		**0.86** (1.84)**	**0.11** (1.88)**
Total Effect	1.49*** (3.23)	**1.84** (1.96)**	**0.87** (1.96)**

Table 6: Response of CDS Spreads to Exogenous Financial Shocks

The table reports the results of regressing changes in CDS spreads on innovations of the National Financial Conditions Index (NFCI) published by the Federal Reserve Bank of St. Louis (FRED), a set of controls and a spatial lag (column 3). The sample ranges from 2006 to 2017. T-stats are reported in parentheses. Significance at the 1%, 5% and 10% is given by ***, ** and * respectively.

$$\Delta CDS_t = \beta_0 + \rho W \Delta CDS_t + \beta_1 FinShock_t + \beta_2 Debt_t + \beta_3 Deficit_t + \beta_4 Spread_t + \beta_5 \Delta CDS_{t-a} + \beta_6 Spread_{t-1} + \varepsilon_t$$

	OLS No Spatial Lag	SAR: 2005Q3	SAR: Average
	(1)	(2)	(3)
Panel A: Point Estimates			
ρ		0.49*** (77.58)	0.40*** (54.08)
$\boldsymbol{ShockNFCI}$	**15.37*** (3.80)**	**8.40*** (2.28)**	**9.15*** (2.43)**
$Debt$	0.32 (0.115)	0.75 (0.37)	0.72 (0.34)
$Deficit$	-0.067*** (-4.72)	-0.067*** (-4.75)	-0.06*** (-4.39)
$Spread$	-0.89 (-0.44)	-0.16 (-0.08)	-0.48 (-0.25)
ΔCDS_{t-1}	-0.002*** (-5.87)	-0.002*** (-4.85)	-0.002*** (-5.57)
$Spread_{t-1}$	0.88 (0.43)	0.14 (0.07)	0.47 (0.25)
Constant	-0.026 (-0.17)	-0.032 (-0.25)	-0.038 (-0.26)
Adj R^2, %	00.25	17.16	14.02
Observations	23490	23490	23490
Panel B: Decomposition of $\boldsymbol{FinShock}$			
Direct Effect		**8.95*** (2.25)**	**9.49*** (2.45)**
Indirect Effect		**7.73*** (2.24)**	**5.80*** (2.45)**
Total Effect	15.37*** (3.23)	**16.68*** (2.24)**	**15.30*** (2.45)**

Table 7: Predictive Regressions

The table reports the results of out-of-sample predictive performance of the standard model and the SAR model during the period 2011 to 2012. Parameters have been estimated using data from 2006:2010. The standard model and SAR model simulations are conditioned to the same information set consisting of the behavior of the Greek CDS spread, fiscal fundamentals in all euro area countries and the Baa-Aaa spread. The weight matrix used in the SAR model is constructed using BIS data from 2005 Q3. p-values are reported in parentheses. Significance at the 1%, 5% and 10% is given by ***, ** and * respectively

	Observed Values	Standard Model	SAR: 2005Q3
	Panel A: Spreads Changes ΔCDS_t		
$\boldsymbol{Mean}$	**-70.88*** **	**-6.38*** **	**-31.83*** **
	(<0.00)	(<0.00)	(<0.00)
SD	72.17	9.15	84.02
Min	-541.72	-28.51	-140.00
Max	215.99	5.026	43.00
$\boldsymbol{RMSE}$		**96.11**	**76.47**
	Panel B: Spreads Levels		
$\boldsymbol{Mean}$	**237.60**	**77.52*** **	**159.99*** **
		(<0.00)	(<0.00)
SD	281.45	44.40	70.36
Min	-9.36	-18.47	18.035
Max	1386.55	203.14	358.45
$\boldsymbol{RMSE}$		**295.86**	**253.35**

Table 8: Structural Breaks

The Table reports the results of tests of structural breaks in the regression coefficients. The data sample is split into four parts, separate SAR regressions are performed on each and then the coefficients are compared using a series of Wald tests.

Variable	F-statistic	p-value
Spatial Lag	0.01	0.98
Debt	0.14	0.96
Deficit	2.23	0.10
Risk aversion	0.16	0.84

Table 9: SAR Model: Sub-samples

The table reports the results of regressing changes in CDS spreads on a spatial lag and a set of controls over four sub-samples of the data. The full sample ranges from 2006 to 2017. T-stats are reported in parentheses. Significance at the 1%, 5% and 10% is given by ***, ** and * respectively.

	2006:2007	2008:2010	2011:2012	2013: 2014	2015:206
	(1)	(2)	(3)	(4)	
ρ	**0.25****	**0.67*****	**0.38*****	**0.52*****	**0.46*****
	(2.08)	(18.11)	(4.07)	(13.46)	(6.39)
$Debt$	-0.3	0.026	0.17	0.06	0.09*
	(-0.46)	(0.42)	(0.69)	(0.94)	(1.71)
$Deficit$	-0.002	-1.08***	-1.83	-0.22	-0.006
	(-0.02)	(-4.27)	(-1.09)	(-0.54)	(-0.01)
$RiskAv$	3.42	16.93***	43.25	14.68	4.43
	(1.40)	(3.34)	(1.05)	(0.72)	(0.49)
CDS_{t-1}	-0.18***	-0.003	-0.04**	-0.06***	-0.02
	(-2.96)	(-0.14)	(-1.99)	(-5.69)	(-1.02)
$RiskAv_{t-1}$	6.33	-21.11***	-132.32***	-9.97	1.79
	(1.39)	(-4.02)	(-3.01)	(-0.42)	(0.19)
Adj R^2, %	17.49	54.04	29.98	41.85	18.38
Observations	117	327	216	216	207

Table 10: Robustness Results for the Spatial Weights Matrix

The Table reports robustness checks using different spatial weight matrices. See the text for details on how the matrices (1)-(4) are constructed.

	Categorical W	2010 Q4	W^{BIS}
	(1)	(2)	(3)
ρ	**0.44***** (16.20)	**0.59***** (27.03)	**0.62***** (29.28)
$Debt$	1.10** (2.28)	0.03 (0.81)	0.04 (0.88)
$Deficit$	-1.08*** (-3.61)	-1.07*** (-3.95)	-1.12*** (-4.14)
$RiskAv$	25.09*** (3.69)	20.02*** (3.15)	16.83*** (2.67)
CDS_{t-1}	-0.09*** (-12.23)	-0.05*** (-8.29)	-0.06*** (-8.73)
$RiskAv_{t-1}$	-31.02*** (-4.53)	-24.52*** (-3.84)	-21.24*** (-3.36)
Adj R^2, %	39.93	47.62	47.83
Observations	1080	1080	1080

Appendix A: Theoretical Model

A1. Tarski's Fixed Point Theorem

Tarski's Fixed Point Theorem (1955): Let X be a non-empty complete lattice. If $\Phi: X \rightarrow X$ is non-decreasing, then the set of fixed points of Φ is a non-empty complete lattice. Moreover, there exists a least fixed point $\underline{x}$ and a greatest fixed point $\overline{x}$ such that for any solution x^*, $\underline{x} \leq x^* \leq \overline{x}$.

Proof:

Recall that by the definition of a valuation function, organizations' values are bounded from above and from below. Introduce the following map:

$$\Phi: [V_{min}, V_{max}]^n \rightarrow [V_{min}, V_{max}]^n$$

$$\Phi(\mathrm{V}) = \psi W V \mathbb{U}^{IB}(V) + \Pi p \mathbb{U}^{PA}(V)$$

In order to prove that Tarski's Fixed Point Theorem applies, we must show that the function Φ maps a complete lattice onto itself (i) and that the function Φ is a non-decreasing function (ii). To prove (i), notice that if the valuation functions are feasible, then:

$$V_{min} \leq \Phi(\mathrm{V}) \leq \psi W V + \Pi p \leq V_{max}$$

and so $X = [V_{min}, V_{max}]^n$ is a complete lattice such that $\Phi: X \rightarrow X$. Since Φ is a linear combination of monotonic non-decreasing functions in V, then $\forall V, V'$ if $V < V'$, then $\Phi(V) \leq \Phi(V')$. Since both (i) and (ii) hold, Tarski's Fixed Point Theorem applies.

A2. Invertibility of the dependency matrix $\boldsymbol{\mathcal{A}}$

Lemma 1:

If Z is a matrix with $\|Z\| < 1$, where $\|Z\| = \|Z\|_1 = \sup_x = \frac{\|Zx\|_1}{\|x\|_1}$, then $\mathbb{I}_n - Z$ is invertible and

$$(\mathbb{I}_n - Z)^{-1} = \sum_{k=1}^{\infty} Z^k$$

Proof:

First, note that if $\vec{v} \neq \vec{0}$, then $\|Z\vec{v}\| < \|\vec{v}\|$ and so

$$\|(\mathbb{I}_n - Z)\vec{v}\| \geq \|\mathbb{I}_n\vec{v}\| - \|Z\vec{v}\| > 0$$

This means that $\mathbb{I}_n - Z$ has a trivial kernel, and therefore it is invertible. Letting $S = \sum_{k=0}^{N} Z^k$, then $S(\mathbb{I}_n - Z) = \mathbb{I}_n - Z^{N+1}$. Therefore:

$$\sum_{k=0}^{N} (\mathbb{I}_n - Z)^{-1}(\,\mathbb{I}_n - Z^{N+1})$$

Letting $N \to \infty$ and observing that $Z^{N+1} \to 0$, the desired result obtains.

Appendix B: Data

B1. Construction of the Financial Linkages and the Spatial Weights Matrix

Financial Linkages are constructed in three simple steps. This example uses BIS data from 2005 Q4

Step 1: Obtain raw bilateral links from BIS Quarterly Bulletin. Entries a_{ij} give the dollar amount of claims of banks headquartered in country i borrow from banks headquartered in country j. Values are in millions of US dollars. Entries on the diagonal are zero (Table 1).

Table 1: Bilateral claims in millions of US dollars, BIS Table 9B 2005 Q4

a_{ij}	AUS	BEL	FRA	GER	IRE	ITA	NET	PRT	SPA
AUS	0	6033	7187	145059	13381	4808	11539	882	2359
BEL	1318	0	64537	53323	5417	21053	95870	1661	10737

FRA	5156	82700	0	173779	25304	25935	90181	7309	33202
GER	39772	47890	101727	0	127759	36161	195813	3166	38637
IRE	4803	33095	31208	108124	0	14403	35659	2592	14996
ITA	16870	56114	125249	183973	42058	0	79915	4214	28407
NET	8635	165238	68538	136317	7532	19182	0	2773	15667
PRT	1212	6709	14200	30334	7049	6284	8568	0	48114
SPA	3379	20967	84677	133490	21892	12587	81661	13448	0

Step 2: Row-normalize to obtain fractions: $x_{ij} = \frac{a_{ij}}{\sum_j aij}$. Note[49]

Step 3: Weigh sovereign debt by the strength of the connection $w_{ij} = x_{ij} D_i^{Gvmt}$. The first column of Table 2 gives the dollar amount that the banks in country i borrow from the general government of foreign countries (Table 2).

Step 4: Transpose the matrix: focus on lending relationships $W = W'$

Step 5: Construct spatial weights matrix: row-normalize W. This ensures that the spatial multiplier exists

Table 2: Sovereign Debt Matrix, BIS 2005 Q4

a_{ij}	D_i^{Gvmt}	AUS	BEL	FRA	GER	IRE	ITA	NET	PRT	SPA
AUS	48522	0	1531	1823	36803	3395	1220	2928	224	599
BEL	58060	301	0	14757	12193	1239	4814	21921	380	2455
FRA	143311	1666	26719	0	56146	8175	8379	29136	2361	10727

[49] Data for borrowing from Portugal to Ireland are not available in the BIS Bulletin 2005 Quarter 4 and the reported value is missing. I approximate the entry by using the first data available, which is from 2008 Quarter 1.

GER	310758	20915	25185	53497	0	67186	19016	102975	1665	20319
IRE	6956	136	940	886	3071	0	409	1013	74	426
ITA	333049	10467	34815	77709	114143	26094	0	49582	2615	17625
NET	64282	1310	25058	10394	20673	1142	2909	0	421	2376
PRT	40950	405	2243	4748	10143	2357	2101	2865	0	16088
SPA	101180	919	5701	23025	36298	5953	3423	22205	3657	0

B1. Financial Shocks

The Table below lists financial events used in the empirical analysis to identify financial shocks.

Table 18: List of Event Dates

Date	Event Description
2/27/2007	Mortgage giant Freddie Mac says it will no longer buy the most risky subprime loans
4/2/2007	Subprime mortgage lender New Century Financial files for bankruptcy-court protection
7/31/2007	Investment bank Bear Stearns liquidates two hedge funds that invested in risky securities backed by subprime mortgage loans
8/6/2007	American Home Mortgage Investment, which specializes in adjustable-rate mortgages, files for bankruptcy protection
8/16/2007	Fitch Ratings cuts the credit rating of giant mortgage lender Countrywide Financial to its third-lowest investment-grade rating
1/11/2008	Bank of America, the biggest U.S. bank by market value, agrees to buy Countrywide Financial for about $4 billion
3/16/2008	The Federal Reserve agrees to guarantee $30 billion of Bear Stearns' assets in connection with the government-sponsored sale of the investment bank to JPMorgan Chase
7/11/2008	Federal regulators seize IndyMac Federal Bank after it becomes the largest regulated thrift to fail
9/7/2008	Mortgage giants Fannie Mae and Freddie Mac are taken over by the government
9/15/2008	Lehman Brothers files for bankruptcy-court protection

9/16/2008	American International Group, the world's largest insurer, accepts an $85 billion federal bailout that gives the government a 79.9% stake in the company
9/21/2008	Goldman Sachs and Morgan Stanley, the last two independent investment banks, will become bank holding companies subject to greater regulation by the Federal Reserve
9/25/2008	Federal regulators close Washington Mutual Bank and its branches and assets are sold to JPMorgan Chase in the biggest U.S. bank failure in history
9/29/2008	Congress rejects a $700 billion Wall Street financial rescue package, known as the Troubled Asset Relief Program or TARP, sending the Dow Jones industrial average down 778 points, its single-worst point drop ever
10/3/2008	Congress passes a revised version of TARP and President Bush signs it. Wells Fargo & Co., the biggest U.S. bank on the West Coast, agrees to buy Wachovia for about $14.8 billion
11/18/2008	Ford, General Motors and Chrysler executives testify before Congress, requesting federal loans from TARP
11/23/2008	The Treasury Department, Federal Reserve and Federal Deposit Insurance Corp. agree to rescue Citigroup with a package of guarantees, funding access and capital. Citigroup will issue preferred shares to the Treasury and FDIC in exchange for protection against losses on a $306 billion pool of commercial and residential securities it holds.
12/19/2008	The U.S. Treasury authorizes loans of up to $13.4 billion for General Motors and $4.0 billion for Chrysler from TARP

Chapter 3

Firm Returns and Network Centrality

1. Introduction

Systemic risk can be broadly defined as any event that threatens the stability of or public confidence in the financial system. Although the risk of such an event is unlikely to be captured by a single metric, what is conceptually more important is the idea of contagion: how failures and shocks propagate from one institution, market or system to another one. Indeed, this is one of the lessons of the global financial crisis of 2007-2008: a shock to the subprime mortgage industry, which accounted for only 3 % of the US financial sector, spread quickly through linkages to the rest of the finance industry, and consequently to the entire economy. Parts of the literature argue that systemic risk is closely related, or even part of systematic risk, and as such there is no need to distinguish between the two. The main argument of this paper is that in order to better understand the sources and dynamics of market risk, it is important to disentangle spillover effects from exposures to common factors. This calls for the need to study the connections between market participants.

The standard macroeconomic diversification argument discards the possibility that network interconnections could have an impact on aggregate volatility or asset prices. The idea there is that at high levels of aggregation in the economy, *individual* shocks average out and their idiosyncratic effect will be negligible (Lucas, (1977)). However, one problem with this argument is that it ignores the fact that firms do not function in isolation but are imbedded in intricate supply chain networks. Firms in such a network are economically related: they receive shocks from their business partners through the network, and as a result they tend to move together. I call this propensity to co-move with related stocks *network risk*.

Recent work by Gabaix (2011) and Acemoglu et al. (2012) shows that if the firm-size distribution or intersectoral input-output linkages are sufficiently heavy-tailed, network granularity renders firms imperfect aggregators of idiosyncratic risk and shocks to individual firms could be transmitted across the economy. For example, many of the components that car manufacturers use are highly specialized and are produced by a few leading companies. *Bose Corporation* is a leader in high-quality audio equipment and supplies simultaneously to automotive giants such as *General Motors*, *Ford*, *Renault*, *Daimler-Chrysler* and *Volkswagen* (*VW*). A negative productivity shock (e.g. a strike) at *Bose Corporation* is going to adversely affect production rates of its customers. As a result, those car manufacturers, which are competitors and whose supply chains might seem unrelated at first sight, become affected through their link to a common supplier, which is central to the automotive industry as a whole.

If idiosyncratic shocks are potential drivers of the volatility of the economic system, then companies with greater exposure to idiosyncratic shocks will be characterized by higher levels of market risk. Furthermore, since *network risk* emanates from idiosyncratic shocks transmitted through a network, *network risk* constitutes a fraction of total idiosyncratic risk. Modelling the economy empirically, as a network of co-moving returns, exposure to *network shocks* is given by the firm's centrality in the network. To capture the network of business relationships, I estimate the asset correlations after having controlled for common market factors. Given this framework, I investigate three main questions in this paper: *(1) Does network centrality have predictive power to identify and characterize systemic events?; (2) How does the network position of a firm relate to its stock returns? and (3) What implications for asset pricing models does network centrality have?*

To address these questions, I borrow methods from graph theory and network analysis. In my model firms are connected to form networks through business relationships. These could be customer-supplier relationships or links to consultancies, financial services, marketing companies, logistics service providers etc[50]. Although I am agnostic to the exact nature of the linkages, what is more important for the analysis is that network links are reflected in asset prices. In this paper I *infer* connections from market data, and more precisely from the variance-covariance (VCV) matrix of *residual* ie. *idiosyncratic returns*.

I use data from January 2000 to December 2015. This is a relevant sample, because it reflects the recent trend of supply networks integration and encompasses both tranquil, boom and crisis periods. To construct the VCV matrix of residuals in any given month τ, I use the past 36 monthly returns in excess of the risk-free rate and regress them on common risk factors. This way I obtain an undirected weighted fully-connected network, where each stock represents a node and the strength of the connections between any two nodes is given by the covariance between them. To evaluate how much each stock is connected to the entire *system*, I follow the approach by Billio et al. (2012) and use principal components analysis. The connectivity (centrality) score for each stock, *PCAS,* is given by the loadings on the first most significant eigenvectors, computed from the VCV residuals matrix. The advantage of using such a measure is that it is recursively defined and allows for feedback effects. *PCAS* measures a node's power to influence other nodes in the network both directly and indirectly through its neighboring nodes. Links to other nodes that are themselves better connected to the rest of the economy are given higher scores than links to nodes that are poorly connected.

Given its systemic nature, one important application of *PCAS* is to provide early warning signals to regulators. In univariate rank regressions I find that stocks that were more exposed to network risk, *i.e.* with larger *PCAS* loadings, are more likely to suffer considerable losses during the recent financial crisis. The beta-coefficients are significant at the 5 % level, which suggests that the centrality measure correctly identifies stocks that will be more affected during crisis periods.

[50] Following the strategic management literature, I define a *supply network* or a *value-added network* as *"a network used to deliver products and servces from raw materials to end customers through an engineered flow of information, physical goods and cash"* (Schönsleben, 2007)

In univariate portfolio sorts on *centrality* I find that there is a positive correlation between centrality and expected future stock returns. Firms in the highest quintile of centrality have average equal-weighted monthly returns over December 2002 to December 2015 that are between 0.09 to 0.47 % higher than the returns of companies in the lowest centrality decile. The result is economically and statistically significant.

As a next step, I test whether the positive relationship between centrality and expected stock returns is related to other firm characteristics. In double-sorted portfolios that control for firm size, the positive relationship between centrality and expected stock returns remains: within the second to fourth quintiles of size, the stocks in the highest quintile of centrality earn higher returns. For example, the *H-L* average equally-weighted monthly return for stocks in second quintile of size is 1.71%, which is statistically significant at the 1% level. The returns in the *H-L* portfolio decrease from small to big firms, which suggests that small stocks in the highest quintile of centrality earn on average higher returns. One possible interpretation could be that small central firms are less capable of absorbing idiosyncratic disturbances than large firms, hence they are more affected by idiosyncratic risks and are accordingly compensated. In equally-weighted double sorts by centrality and book-to-market, the relationship between centrality and market returns remains positive and strongly statistically significant. Within each quintile of book-to-market, more central returns earn higher returns. Moreover, within each centrality quintile, growth stocks (low book-to-market) outperform value stocks (high book-to-market).

I use centrality scores as weights and construct a *mimicking factor* of network risk, which I call *CNTR*. Following the conventional practice in the literature, I use a two-pass procedure to determine whether *CNTR* is priced in the cross-section of returns. In a univariate regression I obtain that, *CNTR* earns a positive risk premium of 0.479%. Controlling for market beta, size, book-to-market and liquidity, *CNTR* positive risk premium remains statistically significant. This finding supports the hypothesis that a stock's true market risk is partly influenced by its relative position in the network of business relations.

Finally, I demonstrate that centrality is related to macroeconomic variables. I define the *centrality spread* as the difference between the returns of firms in the highest quintile of centrality minus the returns of stocks in the lowest quintile. Controlling for unemployment, a recession variable and a linear time trend, I obtain that the *centrality spread* is significantly positively correlated to future consumption in non-durable goods. This finding confirms once again that central firms are more exposed to aggregate risks.

The primary contribution of this paper is empirical. Using methods from graph theory and network analysis, I identify, visualize and analyze the network formed by residual stock returns. First, I develop an econometric measure of network connectedness extending the approach by Billio et al. (2012) in two ways. I apply the measure in a more general economic context and explicitly differentiate between exposures due to a network and due to common factors. Second, I use the connectivity scores to construct a novel factor mimicking network risk and show that it carries a positive risk premium. The paper extends our understanding of how idiosyncratic shocks transmit through the economy and provides a micro foundation for market risk by emphasizing the role of firm-level network connections.

2. Relationship to the Literature

This paper speaks to three main strands of literature. First, by analyzing asset return commonalities, the strength of linkages between individual stocks and the economy as a whole and the sensitivities of these connections to changing economic conditions, the paper broadly relates to the literature on systemic risk and identification of systemic events. Three measures have been developed recently to capture linkages between financial institutions: conditional value-at-risk (Adrian & Brunnermeier, (2011)), systemic expected shortfall (Acharya et al. (2010)) and distressed insurance premium (Huag et al.(2011)). This paper measures connectedness in the economy directly and unconditionally through principal components analysis. Although the paper does not focus exclusively on the financial sector, the paper is relevant for the systemic risk literature because it offers insights into network risk and how idiosyncratic shocks spread in the aggregate economy.

Second, the paper relates to the literature studying the asset pricing implications of networks. One of the biggest challenges in this type of analysis is the availability of data to identify network linkages. Using Bureau of Economic Analysis (BEA) industry input-output tables, Ahern (2013)[51] discovers that a factor mimicking portfolio of returns long in the highest quintile of centrality and short in the lowest quintile of centrality, is positively priced in the cross-section of returns. Buraschi & Porchia (2013), who define network connectivity as the ability to transfer distress states in a directed and timely manner, find that central stocks have lower price-to-dividend ratio and earn a positive centrality premium[52]. In contrast to these papers, which look at within-industry connections or between industrie, I conduct the analysis on a firm-level and use PCA from the variance-covariance matrix of residual firm returns using the most recent 36 months. This approach has several advantages. First, it allows to compute a centrality measure for each firm from the CRSP-COMPUSTAT universe, which could be attractive for investors, who do not invest on an industry-level. Second, it offers a simple, intuitive and easy-to-calculate measure of connectedness. Third, it allows to create a continuous time-series of centrality for each firm and, thus, follow its dynamics over time.

Finally, the paper is also related to the literature that studies correlation networks in finance. Following the pioneering work of Mantegna and Stanley (1999), which introduced concepts of statistical physics in the description of financial systems, a number of studies have used correlations between asset returns to *infer* network connections. Some examples are: Tse et al. (2010), Bonnano et al. (2004), Namakia et al. (2011), Lee & Djauhari, 2012, Takayuki et al. (2006) etc. Whereas these studies are typically limited to analyzing the

[51] Using a similar dataset, Ahern & Harford (2012) study the propagation of merger waves through production networks.
[52] Other related works, which link economic fundamentals to market risk are (Gomes, Kogan, & Zhang, (2003); Carlson, Fisher, & Giammario, (2004); Hou & Robinson,(2006); Hong, Torous, & Valkanov, (2007)).

topological properties of the emerging network, this paper adds an additional layer of analysis by investigating the asset pricing implications of the network structure. Following the idea of granularity in networks, Kelly, Lustig, & Nieuwerburgh (2013) show that the firm size distribution influences the volatility of returns and sales growth rates. Barigozzi and Brownlees (2016) propose a new network analysis technique for high-dimensional multivariate time-series: a two-step LASSO regression that allows to estimate large sparse long-run partial correlation networks. Diebold and Yilmaz (2014) use vector autoregressive models to decompose the variance of a stock into different components contributed by other stocks. The former two studies are primarily concerned with idiosyncratic volatility and its decomposition, whereas I focus on stock returns and interactions between centrality and other market forces.

3. Economic Mechanism

The main hypothesis of this paper is that a stock's true market riskiness is, in part, influenced by its relative position in the network of stock returns. There are five important identification assumptions. First, I assume that firms are connected in networks through their business relationships. Second, shocks that originate at the firm level are transmitted through business relationships. Third, I assume that firm-level shocks can aggregate to form economy-wide shocks. Fourth, business relationships are reflected in the correlation matrix of returns. Fifth, idiosyncratic shocks do not cancel out through diversification.

First, I assume that two stocks are connected to each other if there exists a business relationship between the two. Note that I am agnostic with respect to the nature of this business relationship: it could be a direct relationship as in the case of a supplier or a customer, or an indirect link such as a marketing consultancy, a financial audit firm or a logistics service provider. What is important is that this business partner is crucial for the operations of the reference company. In the next chapter I discuss this aspect in more detail. To clarify: I acknowledge that business relationships are *not* the only way through which a network could be identified, but rather one of the many ways to model stocks from a network perspective. Other approaches include production networks or networks of ownership shares.

Second, there are many ways in which systematic shocks can originate. The most immediate example is a macro shock, such as a change in the policy of the central bank. Such a shock is likely to have a simultaneous effect on all of the industries at the same time. Another way in which a systematic shock can originate is from a firm-level shock. An illustrative example is a shock to technology. Imagine energy company XYZ develops a new technology that allows it to extract gas from deep shale gas reserves, previously unattainable with existing technology. Such a shock is likely to affect gas prices (assuming the reserve is big enough to disrupt the market, demand stays constant and supply increases in the aggregate) and transmit to other firms in XYZ's immediate supply chain. Third, it is also reasonable to believe that changes in the prices of gas would spill over to companies from other sectors, such as transportation and manufacturing, which are highly sensitive to fluctuations in the price of natural gas and oil.

Fourth, I assume that business relationships are reflected in the correlation matrix of returns. Previous research such as Hendricks & Singhal (2005) and Lang & Stulz (1992) discover that stocks in the same industry segment tend to move together because their businesses are tightly connected to each other. In a related vein of research, Hou (2007), Cohen & Frazzini (2008) and Menzly & Ozbas (2010) examine how shocks to one firm translate into the linked firm in both real quantities, as well as stock prices, and find returns predictability arising from economic links.

The fifth underlying assumption of this paper is that random idiosyncratic shocks do not cancel out. The long-standing macro argument against the importance of idiosyncratic shocks for aggregate volatility has been that a positive shock in one industry is balanced by a negative shock in another, and so, on average the economy is unaffacted. However, recent work by Acemoglu, Carvalho, Ozdaglar, & Tahbaz-Salehi (2012) disputes this view. The diversification argument assumes that connections are uniformly distributed and random. In a theoretical model, Acemoglu et al. (2012) relax this assumption and show that in the presence of asymmetric sectors, the diversification argument breaks down and idiosyncratic shocks could indeed be translated into aggregate ones. From an intuitive standpoint, the diversification argument relies on the Law of Large Numbers: if connections between sectors are asymmetric, then the law no longer holds. Supporting empirical evidence is given in Acemoglu et al. (2012), Ahern & Harford (2014) and Ahern (2013). The authors show that the U.S. intersectoral trade is characterized by a large number of industries with very few connections and a small number of industries with many connections. Furthermore, Ahern & Harford (2014) show that economy-wide merger waves are formed by the accumulation of industry-level merger waves that ripple through product market relationships. In essence, the asymmetry in the connections creates merger cascades, which are similar to the volatility cascades that Acemoglu et al. (2012) discover in their work.

To summarize, if idiosyncratic shocks can be accumulated to form aggregate shocks, then such local shocks will affect asset prices. If shocks are transmitted through the business network, then those companies that are more connected to the economy are more exposed to such shocks. Moreover, stocks in business networks are economically related, their asset prices tend to move together and such commonality will be reflected in the correlation matrix. Finally, companies cannot fully protect themselves against these shocks through diversification. Consequently, network risk becomes undiversified and the risk should be priced in the cross-section of returns. My main premise is that central companies earn higher returns because they are compensated for higher exposure to network risk.

4. A Network Representation of the Stock Market

4.1 Network Identification

One of the challenges in this work is the availability of data to identify the network of stock returns. Previous research has used industry input-output linkages, provided by the U.S. Bureau of Economic Analysis (BEA). Although such an approach provides clear identification of customer-supplier relationships, it poses the disadvantage that information in BEA tables is only available at an industry level and is updated every five years. On a firm-level, studies such as Kelly, Lustig, & Nieuwerburgh (2013), Atalay et al. (2011) and Barrot & Sauvagnat (2016) among others, use Compustat data of firm's own reports of major customers. Financial Accounting Standards N.131 mandates that companies report customers accounting for more than 10 % of the seller's revenues. Such a network is illustrative, but data is available for a small subset of companies over relatively short spells of time and the 10% threshold could potentially introduce a bias.

The approach that this paper takes is to measure connections directly and unconditionally: through principal components analysis applied to residual returns. Using market data has the following advantages: market data are easily available, have higher frequency (allows to update the links on a continuous basis) and connections constructed from market data are forward looking in cotrast to actual connections extracted from accounting data, which provide "snapshots" and could be considered as backward looking. Such a forward looking interpretation is consistent with the idea that market prices reflect all information available to market participants, and, in equilibrium correspond to the discounted value of future dividends. For example, consider a silicon wafer manufacturer and its accounting firm. News of abuse of accounting practices by the accounting firm will not only have an impact on it, but also on the wafer manufacturer. Assuming efficient markets, new information will be reflected in the stock prices of the manufacturer and the accounting company.

The next section formalizes the concept of a network from a graph theory point of view and introduces the measure of network centrality.

4.2 Network Representation of the Stock Market

A network, or a graph object in mathematics, is a set of nodes (vertices) linked by edges. Nodes represent the objects of interest, whereas edges contain information on how the nodes are linked. For example, on *Google Scholar* each academic paper is a node and connections are given by citations between works. In a social network, such as *Facebook*, each personal account is a node and edges are "friendships" between people. If edges are undirected (such as on *Facebook*), the network is *undirected*; if edges have a direction (such as citations) the network is *directed*. In this paper vertices are represented by assets traded on the stock market and edges give possible undirected connections between them. Networks can be graphically represented by a square *adjacency matrix* $\boldsymbol{A}$. More formally, let $\boldsymbol{A}$ be the adjacency matrix of a graph $\boldsymbol{G}(\boldsymbol{V},\boldsymbol{E})$, where the set of N nodes is given by $\boldsymbol{V} = \{v_1, v_2, \ldots, v_N\}$ and the set of edges is given by $\boldsymbol{E}$. Each entry a_{ij} in the adjacency matrix $\boldsymbol{A}$ represents a possible link between nodes i and j:

$$= a_{ij} \begin{cases} 1 & if\ asset\ i\ is\ connected\ to\ asset\ j \\ 0 & otherwise \end{cases}$$

By convention, entries on the main diagonal of $\boldsymbol{A}$ are set equal to zero i.e. self-loops are excluded. Since, I use the VCV matrix of residual returns as an adjacency matrix, no entry is *exactly* equal to zero and so I assume that the network is fully connected, but the strength of these connections is different: the value of entry a_{ij} is given by the covariance between nodes i and j. Also, because the VCV matrix is symmetric, the network is undirected. This is a slight departure from the conventional way of modeling networks, but on a conceptual level the intuition is the same. One could think of values of a_{ij} close to 0 as very weak connections and values of a_{ij} close to 1 (in absolute terms) as very strong connections.

4.3 Network Centrality

A number of measures have been developed in the literature to characterize centrality in networks. Some examples include in-degree and out-degree, closeness, betweenness, diameter and eigenvector centrality[53]. I follow the approach by Billio et al. (2012), who suggest that changes in the correlations between assets could be captured using an econometric measure based on Principal Components Analysis (PCA). The authors apply PCA in the context of *systemic risk of the financial system* and use it to analyze the returns of hedge funds, banks, brokers/dealers and insurance companies.

[53] The reader is referred to Borgatti (2005) for a detailed discussion of the use of these measures and the assumptions underlying them.

The rationale behind using PCA to identify network correlations is the following. PCA describes the covariance structure of a given set of variables by identifying the primary sources of variation. By identifying the sources, PCA reduces the dimensionality of the data to a few common orthogonal factors of decreasing explanatory power. In order to do so, PCA computes the eigenvalues and eigenvectors of the covariance matrix. The biggest eigenvalue corresponds to the first principal component and the associated eigenvectors gives the direction that constitutes the biggest source of variation in the data. More formally, let r_i be the stock return of firm $i, i = 1,..,N$, let $R^s = \sum_i r_i$ be the system's aggregate return, and let $E[r_i] = \mu_i$ and $Var[r_i] = \sigma_i^2$. Then, the variance of the system is given by: $\sigma_S^2 = \sum_i^N \sum_j^N \sigma_i \sigma_j E[z_i z_j]$

where $z_k = (r_k - \mu_k)/\sigma_k$, for $k = i, j$and z_k is the standardized return of firm k.

If we introduce N zero-mean uncorrelated variables ς_k, for which

$$E[\varsigma_k \varsigma_l] = \begin{cases} \lambda_k \; if \; k = l \\ 0 \; if \; k \neq l \end{cases}$$

and λ_k is the k-th eigenvalue, then we can express the z's as linear combinations of ς_k's in the following way:

$$z_i = \sum_{k=1}^N L_{ik} \varsigma_k$$

where L_{ik} is a factor loading for firm i. It follows that

$$E\left[z_i z_j\right] = \sum_{k=1}^N \sum_{l=1}^N L_{ik} L_{jk} \, E[\varsigma_k \varsigma_l] = \sum_{k=1}^N L_{ik} L_{jk} \, \lambda_k$$

and that the total variance of the system is given by the following expression

$$\sigma_S^2 = \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^N \sigma_i \sigma_j L_{ik} L_{jk} \, \lambda_k$$

PCA decomposes the variance-covariance matrix of returns of the N firms into an orthonormal matrix of loadings L, given by the eigenvectors of the variance-covariance matrix, and a diagonal matrix of eigenvalues Λ. Since the first eigenvalues usually explain most of the variance of the system, it is possible to focus on a subset $K < N$ of them. Conditional on a common component, Billio et al. (2012) formally show that a univariate measure of the centrality of each firm to the system is given by

$$PCAS_{i,k} = \sum_{k=1}^{K} \frac{\sigma_i^2}{\sigma_S^2} L_{ik}^2 \lambda_k \,|K < N$$

Intuitively, the idea is that when the system is highly interconnected, a small number of principal components will explain most of the volatility of the system. Thus, by following the time variation of the eigenvalues and the percentage of cumulative variance explained by them, it will be possible to detect increasing correlations between different firms/ and or industry sectors, as well as to detect common sources of risk exposure.

In this paper I propose to apply the *PCAS* measure in a more general context that involves not only the financial system but the economy as a whole. Such an application is possible for three main reasons.

First, Billio et al. (2012) model the financial system as a collection of interconnected institutions that engage in mutually beneficial business relationships, through which illiquidity and financial distress can quickly propagate. Such a definition is very similar to my understanding of supply networks, which are collections of business partners connected through an engineered flow of information, physical goods and cash. In this sense, by means of induction, one could think of the financial system of Billio et al (2012) as a special instance of a supply network, where the parties involved are hedge funds, banks, brokers/dealers and insurers. Second, in the original paper the authors capture changes in the correlations of financial assets by analyzing stock returns. In this paper I focus on *residual stock return*, because I want to isolate the effect of network risk, which requires that I net out common risk factors. In essence, both approaches empirically detect connections between assets, but the difference is that they identify different types of risk.

Third, from a technical point of view *PCAS* is well suited for my research purposes, because it is recursively defined. The measure allocates relative scores to nodes based on the number and strength of their connections to their neighbors, as well as those neighbors' centrality scores. Consequently, *PCAS* allows feedback effects. The relevance of feedback effects for supply networks could be illustrated using the following example. Imagine, a strike at a second-tier supplier *X* could disrupt deliveries to a first-tier supplier *Y*, which in turn could compromise production at the manufacturer. To improve reliability the manufacturer decides to introduce an additional supplier *Y'* and allocates input orders on a 50-50 percent basis between old and new supplier. This means that supplier *Y'* would have to adjust production levels accordingly (50% reduction) and so do its up-stream partners (*X* in this case). This means that just because the shock originated at *X* does not imply that the firm is immune from subsequent feedback shocks. All in all, *PCAS* is suited to measure a node's power to influence other nodes in the network both directly and indirectly through its neighboring nodes. Links to other nodes that are themselves better connected to the rest of the economy are given higher scores than links to nodes that are poorly connected.

To summarize, by filtering the effect of common risk factors, I construct the matrix of residual returns, which I treat as an adjacency matrix. I obtain a fully connected weighted network. The idea of using residual returns is simple: when idiosyncratic shocks spread across the supply network, residual returns tend to move together. However, not every stock is equally connected to the rest of the firms and so some firms will be located in the center of network, whereas others will be in the periphery. I argue that that central stocks are more likely to be hit by idiosyncratic shocks than peripheral ones. Consequently, higher exposure to network risk should be reflected in higher returns for central firms.

4.4 Econometric Model

4.4.1 Decomposition of Returns

A key step in studying the network of stocks is to decompose returns into a component that is attributed to common market factors and into an idiosyncratic component. Idiosyncratic returns are constructed on a rolling 36-month basis. To construct a measure of *PCAS* within each month τ, a factor model using the past

36, *excluding* month τ, is estimated. Such a procedure results in 157 rolling windows.

$$R_t - R_t^f = \alpha + \beta' F_t + \varepsilon_t \qquad (4)$$

where t denotes a monthly observation from $t = 1,..,36$, R_t denotes an N-dimensional set of time t asset returns and ε_t is an N-dimensional vector of asset idiosyncratic returns. The factor model that I consider specifies F_t as a *3 × 1* vector of the three Fama-French factors: *Mktrf* (the return on the market portfolio), *SMB* (the size factor) and *HML* (book-to-market factor).

Then, I compute the VCV of the residuals $\varepsilon_{i,t}$ for each of the 157 rolling windows. I apply the *PCAS* measure outlined in the previous section and calculate a centrality score for each stock i in month t. This allows me to identify the stocks with the highest and lowest centrality scores within a given month. Next, I proceed to create a factor mimicking portfolio of network risk. Let X be an N × *1* vector of centralities in month t, normalized such that $X'X = 1$. Then, it follows that $CNTR_t = X'\varepsilon_t$ is a common factor: a mimicking portfolio of network risk. The procedure is simple and intuitive: the network risk factor is derived out of residual returns, using network centralities as weights. To obtain exogeneity of the weights, I lag network centralities by 12 months. The factor that obtains is a non-traded factor. It has to be noted that this is not the only way to construct a mimicking factor: one alternative could be to create a traded factor by employing long-short portfolio strategy.

Regressing ε_t on $CNTR_{t-12}$ gives the following:

$$\varepsilon_t = \gamma + \beta_{CNTR} CNTR_{t-12} + \eta_t \qquad (5)$$

The vector of coefficients β_{CNTR} is the loading on the network risk factor and the central object of interest. If the hypothesis that central firms earn higher returns than peripheral ones is supported in the data, then I expect a positive exposure to the network risk factor.

4.4.2 Estimation and Assumptions for the Factor Model

Since the unobserved network factor $CNTR_t$ is essentially derived using PCA, it is important to discuss the properties of the principal components estimator. Throughout the paper I make the following assumptions[54]:

Assumption A: Factor

(A.1): $E[CNTR_t CNTR_t'] = \sigma^2_{CNTR}$

$(A.2): \frac{1}{T}\sum_{t=1}^{T} CNTR_t CNTR_t' \xrightarrow{p} \sigma^2_{CNTR}$

Assumption B: Factor Loadings

$(B.1): \left\|\beta^i_{CNTR}\right\| \leq \bar{\beta}_{CNTR} < \infty$

$(B.2): \left\|\frac{1}{N}\beta_{CNTR}{}'\beta_{CNTR} - \sigma^2_N\right\| \to 0$

Assumption C: Time and Cross-sectional Dependence and Heteroscedasticity

$(C.1): E(\eta_{it}) = 0, E[\eta_t \eta_t'] = \Sigma_\eta$

$(C.2): \frac{1}{T}\sum_{t=1}^{T} \eta_t \eta_t{}' \xrightarrow{p} \Sigma_\eta$

Assumption D: Orthogonality Condition

$(D.1): E[\eta_t\ CNTR_t'] = 0$

Assumptions A-D are common in the literature and consistent with the discussion in Bai (2003). *Assumption A* allows factors to be serially correlated, but does not allow the dynamics of $CNTR_t$ to enter into ε_t, so that the relationship between ε_t and $CNTR_t$ remains static. This rules out stochastic trends, unit roots and other processes with non-constant second moments. *Assumption B* assumes that the factor has a non-trivial contribution to the variance of ε_t[55]. *Assumption C.1*

[54] $\|.\|$ refers to the Euclidean norm
[55] For simplicity, I only assume nonrandom factor loadings. The results still hold when $\beta^t_{CNTR}$are random under the condition that they are independent of the CNTR factor and the idiosyncratic disturbances η_t and that $E\left\|\beta^i_{CNTR}\right\|^4 \leq M$

guarantees that Σ_η exists and is finite and *Assumption C.2* ensures that the sample second moments converge. *Assumption D* makes sure that η_t and $CNTR_t$ are independent, so that the factor loading β_{CNTR} can be consistently estimated.

4.4.3 Determining the Number of Principal Components in the Factor Model

Central to the theoretical and the empirical validity of the factor model outlined in section 3.4.1 is the correct specification of the number of principal components, the loadings on which are used to derive *PCAS*. This is particularly important, because this paper works with a large dataset, where the cross-sectional dimension N is greater than the time dimension T, which constitutes a departure from classical factor models. It is worthwhile noting that while the *true* covariance can have rank N even if $N > T$, the rank of the $\sqrt{T} - consistent\ N \times N$ *sample* covariance matrix is always less than or equal to $\min\{N, T\}$ (Lucia, Barigozzi, & Capasso, 2009). Here, this gives $T = 36$ and allows me to obtain 36 eigenvectors.

There are a number of ways to determine how many of these 36 eigenvectors to retain. One informal way would be to impose a certain threshold on the cumulative variance explained and retain only those eigenvectors that explain total variance up to the threshold. However, since there is no existing economic theory on which the network risk factor is based, postulating a variance threshold of e.g. 50% is not particularly informative. Another way that I consider is the Kaiser criterion (1960), according to which only factors with eigenvalues greater than 1 are retained. Intuitively, the idea is that those factors with eigenvalues less than 1 account for less variability than does a single variable. A third alternative could be to use a formal statistical procedure. I rely on the method of Bai & Ng (2002), who developed information criteria based on which the number of factors can be consistently estimated, even when both N and T are large. Central to their approach is the idea that overfitting the model involves a penalty, which is a function of *both* N and T. Let $V\left(k, \widehat{F_k}\right)$ be the residual variance from regressing the residuals in any rolling time window on the first k principal components. Then, the following three information critera (IC) are proposed by Bai and Ng (2002)[56]:

$$IC_{p1}(k) = \ln\left(V\left(k, \widehat{F_k}\right)\right) + k\left(\frac{N+T}{NT}\right) ln\left(\frac{N+T}{NT}\right)$$

$$IC_{p2}(k) = \ln\left(V\left(k, \widehat{F_k}\right)\right) + k\left(\frac{N+T}{NT}\right) ln(C_{NT}^2)$$

$$IC_{p3}(k) = \ln\left(V\left(k, \widehat{F_k}\right)\right) + k\left(\frac{\ln(C_{NT}^2)}{C_{NT}^2}\right)$$

I compute the three criteria and choose k such that the ICs are minimized.

5. Data and Variables

I obtain monthly stock prices from the CRSP database for January 2000 to December 2015. A firm-month observation is included if the stock has a CRSP share code of 10, 11 or 12 and the stock has no missing monthly observations in the following 36 months. The requirement for non-missing observations is quite stringent and a relevant concern could be that it introduces a selection bias. However, the monthly set of stocks in the sample changes relatively slowly, with the average replacement rate (the ration between the number of new companies, from a month to the next, and total number of companies) being less than 5%. Economic fundamentals are obtained from the COMPUSTAT Database. I merge on company identifier *permno* and time variable *date*, which results in 481,609 firm-month observations.

To avoid look-ahead bias, when calculating size and book-to-market I follow the matching procedure by Fama and French (1992), which imposes a minimum gap of 6 months between fiscal year-ends and firm returns. I match monthly returns from July of calendar year T to June of calendar year $T+1$. Throughout the paper I use the natural logarithm of market equity and book-to-market value as controls. I calculate market equity (ME) as share price times number of shares outstanding. The variable is measured in June of year T, for the returns between

[56] $C_{NT} = \min\{\sqrt{N}, \sqrt{T}\}$

July of year *T* and June of year *T+1*. Book-to-market (BM) is computed as the ratio between book equity and market equity in December of year *T-1*. Book equity is calculated by summing shareholders' equity, balance sheet deferred taxes and investments and subtracting the book value of preferred stock. The book value of preferred stock is depending on availability of data using the following (in this order): redemption, liquidation or par value. In the regressions to follow I use the natural logarithm of *ME* (*LogME*), and the natural logarithm of *BM* (*LogBM*). All variables used in the subsequent analyses, including *logME* and *LogBM*, are winsorized at the 1st and 99th percentiles level to avoid the effect of outliers.

Individual stocks are re-assigned to portfolios on a monthly basis using the same market equity or book-to-market value Fama and French's factors would otherwise employ. I am aware that from an investor's point of view such a monthly rebalancing could involve higher turnover, and hence costs. However, such a slight departure from the classical Fama and French approach is necessitated by the dynamics of the centrality measure, which changes substantially from month to month. If portfolios were formed once a year, in July as in the Fama-French approach, a considerable amount of information would be lost.

5 The Empirical Relationship between Centrality and Returns

First, I present the results of *PCAS*, on which the centrality analysis is based. Then I proceed to analyze centrality. To test the hypothesis that central returns earn higher returns, I assign stocks into equal-weighted and value-weighted portfolios sorted by centrality. In order to assess how firm characteristics, such as size and book-to-market relate to centrality, I form double-sorted portfolios. Residual returns used in this chapter are obtained by the Fama-French 3 Factor Model.

5.1 Network Visualization

A useful starting point in the analysis is to visualize the network formed by residual returns. To uncover the topology, I use a method from theoretical graph theory and represent the network of residual returns as a spanning tree of shortest length or a *minimum spanning tree* (MST)[57]. Such a tree is a graph without loops, which connects all *n* nodes with *n-1* links. One begins with an empty tree and starting from an arbitrary node *i* adds a link to the node closest to it in the sense of Euclidean distance. One proceeds by adding one link at a time, until all nodes are connected in the graph. *Therefore*, the MST retains the *n-1* shortest (strongest) links that span all the nodes in the network. *Apendix A.* gives a detailed description of *MST* and the precise algorithm to extract the links.

Figure 1 shows an example of a network comprised of the 500 largest stocks as of December 2002. I choose to focus on a sub-sample and not the entire sample of 4, 732 companies in order to facilitate visual comprehension. However, the results and intuition do not change when one turns to the full sample. Panel A shows stocks (nodes) divided into 11 Fama-French Industries with different colors representing different industries. Two messages are immediate from the graph. First, the tree-like structure with different branches suggests that there is not a single "central node" in the economy. This calls for a centrality measure, which considers multiple features of the data unlike *eigenvector centrality*, which rates one node as central and assigns scores to others that drop very sharply in magnitude. Second, different clusters can be seen in the graph: for example, the **yellow** region of **financials**, **orange** region of **energy** companies, **blue** area of **non-durables**. This supports existing evidence that stocks in the same industry move together because their businesses are related (Lang & Stulz, (1992); Hendricks & Singhal (2005) etc.). However, not all industries are grouped into clear economic clusters.

In Panel B, I "zoom" in the network and look at the position of the three biggest automakers in my sample: **General Motors**, **Daimler-Chrysler** and **Ford Motor Company**. These are given in **purple**. In a testimony to the Senate Committee on Banking, Housing and Urban Affairs on December 4th 2008, Alan

[57] Theoretically, one could use the full VCV of the residuals to visualize the topology. However, the problem is that this way one obtains a very dense network, which is hard to visualize in an easy to comprehend way.

Mulally, CEO of Ford said: "*The collapse of one or both of our domestic competitors would threaten Ford because we have 80 percent overlap in supplier networks and nearly 25 percent of Ford's top dealers also own GM and Chrysler franchises."* The graph confirms these statement: *GM, Daimler-Chrysler* and *Ford* are topologically close and, indeed, share a lot of common links to **manufacturing** companies and **equipment** providers. Furthermore, visual inspection of the graph shows the financial company **Toronto Dominion** is located very "close" to **Daimler Chrysler** and **Citibank** is close to Ford. It turns out that Toronto Dominion has a division, which offers financial services related to the auto industry: dealer programs, car loans for consumers, car-insurance etc.[58] On the other hand, **Ford** and **Citibank** share a connection through a VISA and Mastercard program launched back in the 1990s that entitles card-holders to a discount on purchases of new Ford vehicles.

These examples serve two purposes. First, they demonstrate that using the definition of SNs is a reasonable assumption to identify business links. And, second, correlation matrices of residual returns are rich source of information from which meaningful economic links can be inferred.

5.2 Components Analysis

In this section, I implement the measures defined in Chapter 3 using firm-level monthly *residual* returns. The time-series graph of cumulative variance explained by the first 1, 5, 10, 15, 10 and 36 principal components is presented in Figure 2 and means of the cumulative variance broken down by periods are given in Table 1. The first principal component captures on average roughly 7% of the total variation; the first 5 PCs account for approximately 23% of it; the first 10 PCs capture approximately 40%; the first 15 PCs capture around 54% and the first 20 PCs give around 67% of the total variability. I do not find any sizeable differences across time periods.

To demonstrate that cumulative variance explained by the PCs is statistically robust, I bootstrap the estimated cumulative variances using 1,000 samples and

[58] In fact in 2010 TD bought Chrysler Financial.

compute confidence intervals from the resulting empirical distribution. Figure 3 plots the results for the first 10 PCs. As it is clear to see from the graph, the estimated cumulative variance always lies with the confidence bounds. The results for PC1, PC1-5, PC1-10, PC1-15, PC1-20 and PC1-36 are qualitatively the same.

The fact that the first 10-15 PCs explain nearly half of the variation is suggestive of a factor structure behind residual returns. One way to explain such common variation would be comovement among factor model residuals, for example due to omitted common factors. However, average pairwise correlations among residuals are around 2%, and never exceed more than 10% in a year. This finding persists even after the inclusion of the momentum factor. Indeed, the Fama-French 3 Factor Model and the 4 Factor Model seem to be able to absorb all commonality in returns.

After having shown results for the entire sample of 36 PCs, I now turn to the discussion in Chapter 3.4.3 and determine the number of PCs I am going to use derive measures of network centrality. Table 2 gives information on the mean values of the ICs and the eigenvalues. This is possible, since no sizeable differences are noticed in different estimation windows. The value of the ICs grows monotonically with the inclusion of more PCs and, therefore to conserve space only results up to PC 15 are presented. Minimum values are given in red. IC_1 and IC_3 suggest the usage of one factor, whereas IC_2 suggest using the first 2 PCs. Turning to eigenvalues, the first eigenvalue has an average magnitude of 2.625; the values of the subsequent eigenvalues decline monotonically. Technically, the last eigenvalue, which has magnitude greater than 1 is λ_{14}. However, $\lambda_{11} - \lambda_{14}$ have values very close to 1, and I decide not to include them.

Based on these results, I calculate network centrally using the loadings of the first 10 PCs. I obtain *PCAS1, PCAS5* and *PCAS10.* In the case when my centrality measure is calculated using only one PC, then it collapses to a scaled version of eigenvector centrality, which is a measure commonly used in network analysis[59].

---

[59] Let $x(i)$ be the eigenvector centrality (EVC) of a node v_i. Then, $x(i) = \frac{1}{\lambda}\sum_{j\epsilon\Gamma(v_i)} x(j) = \frac{1}{\lambda}\sum_{j=1}^{N} a_{ij}x(j)$ $\Gamma(v_i)$ denotes the neighbourhood of of v_i

5.3 Summary Statistics of Centrality

In Panel A of Figure 5 I plot the empirical densities of the three centrality measures. My estimates are based on a normal kernel function, which is evaluated at a 100 equally-spaced points. It is immediate to observe that the distribution is heavily skewed to the right, which suggests that there is a small number of firms that are highly connected to the rest of the economy i.e. are central in the network. To further characterize such heavy-tailed distributions, in Panel B I plot the empirical counter-cumulative distribution function (i.e. one minus the empirical cumulative distribution) on a log-log scale. The panel shows the right tail of the distribution can be approximated by a power-law distribution as indicated by the nearly linear relationship. The presence of a fat right tail offers evidence that the diversification argument could break down. Furthermore, this result is fully consistent with the main hypothesis of the paper that idiosyncratic shocks could be translated into the aggregate.

Since the centrality measures *PCAS1, PCAS5, PCAS10* are highly skewed, I take the logarithm and use the transformed variables in Table 3 and henceforth. *LogPCAS1* is negatively skewed with extra kurtosis of 1.061 compared to the normal distribution. *LogPCAS5* and *LogPCAS10* look similar to each other: they are positively skewed and have kurtosis less than 3. In terms of volatility, *PCAS1* scores highest with standard deviation equal to 1.134, followed by *PCAS5* with 0.606 and *PCAS10* with 0.547. Please note that from now on, any reference to centrality considers its logarithmic transformation.

Figure 5 shows as an example a histogram of *LogPCAS10* with the density of the normal distribution superimposed over it. The figure visually confirms the results of Table 3. It is important to note that even after the log-transformation the empirical distribution of the centrality measure exhibits a slightly thicker

i.e. the set of nodes to which it is connected and λ is the largest eigenvalue of the adjacency matrix $\boldsymbol{A}$. Rewritten in matrix form with $\boldsymbol{x} = \{x(1), x(2), \dots, x(N)\}'$ being the vector of EVCs of all nodes, it obtains: $\boldsymbol{x} = \frac{1}{\lambda}\boldsymbol{A}\boldsymbol{x} \ \rightarrow \ \lambda\boldsymbol{x} = \boldsymbol{A}\boldsymbol{x}$

(compared to a normal cdf) right tail, although the result is far less pronounced than that in Figure 5.

5.4 Single-sort Results

In Table 4, I sort stocks into five quintiles based on centrality. Portfolios are rebalanced monthly. To avoid look-ahead bias, for each stock in my sample I match its returns in month τ to its centrality measure in month $\tau - 1$. I show equally-weighted and value-weighted returns, where I calculate value-weighted returns using ME/average ME as weights.

The difference between the average equally-weighted monthly returns of the stocks in the highest quintile of *LogPCAS1* and the returns in the lowest quintile (*H-L*) is 0.09%. However, the *t-stat* indicates that the result is not statistically significant. Turning to the other centrality measures, the equally-weighted *H-L* return is 0.39% and 0.47% for *LogPCAS5* and *LogPCAS10* respectively. The result is economically meaningful and statistically significant at the 95 percent level. For portfolios of value-weighted returns, the relationship is reversed. Stocks in the highest quintile of centrality earn lower returns than those in the lowest quintile. This indicates that a negative relationship between firm size and centrality. In fact, the cross-sectional correlation between size and centrality is -12.44 %. One explanation for this is that larger central firms are less adversely exposed to idiosyncratic shocks because they are better able to diversify or because they possess more resources that allow them to cope with such shocks. Another possibility, suggested by Ahern (2012), is that return estimates for value-weighted portfolios are affected by the presence of multi-divisional firms, for which centrality is likely to be less precisely calculated.

5.5 Double Sort Results

To account for possible interactions between size, book-to-market and centrality, I double sort stocks into 25 portfolios. Table 5 Panel A shows equally-weighted returns and panel B shows value-weighted returns. shows the results for double-sorted equally-weighted portfolios on *LogPCAS10 centrality* and *size*[60]. The positive relationship between centrality and stock returns remains: within the second to fourth quintile of size, the stocks in the highest quintile of centrality earn higher returns. For example, the *H-L* average equally-weighted monthly return for stocks in second quintile of size is 1.71%, which is statistically significant at the 1% level. The returns in the *H-L* portfolio decrease from small to big firms, which suggests that small stocks in the highest quintile of centrality earn on average higher returns. Central firms are likely to be more exposed to shocks than firms in the periphery, but small firms are more likely to be affected than large firms. Kelly, Lustig, & Nieuwerburgh (2013) show that small firms are characterized by higher total and idiosyncratic volatility than bigger firms.One possible interpretation could be that small central firms are more volatile because they are exposed to higher market risk, and hence are compensated for such risk.

Turning to equally-weighted double sorts by centrality and book-to-market, the relationship between centrality and market returns remains positive and strongly statistically significant. Within each quintile of book-to-market, more central returns earn higher returns. Moreover, within each centrality quintile, growth stocks (low book-to-market) outperform value stocks (high book-to-market). This offers a new interesting insight into the value anomaly, which can be explained in the following way. Growth stocks are stocks that are expected to generate future earnings at a higher-than-average rate, by retaining earnings and investing into capital projects. One typical example of growth companies are those that invest heavily in research and introduce new products or services. Being at the forefront of industry developments, such companies are unlikely to be located in the periphery of the network. This could be because they need partners to cooperate with (technology companies creating clusters) or because a large number of customers use their innovative products as input materials. Another example of growth stocks are companies that have a competitive advantage in manufacturing scale: by reducing per-unit costs, they are able to achieve higher earnings, which

[60] Qualitatively similar results hold for the other centrality measures and are available upon request.

could be reinvested and translated into higher growth rates. In this case, growth stocks need a large customer base, and so, it not likely that they share few connections with other companies and are located in the periphery. On the other hand, results on value-weighted returns indicate that within each book-to-market quintile, central stocks earn a lower premium. This is a consequence of the negative relationship between size and centrality and size and book-to-market value in my sample.

These findings suggest that there is a positive relationship between network centrality and stock returns, although the effect is influenced by size and book to-market.

5.6 Two-pass Cross-Sectional Regression

I would like to investigate the implications of the above results for asset pricing models. In particular, I would like to test whether the introduction of a new factor based on centrality is priced in the cross-section of firm returns. In the spirit of Arbitrage Pricing Theory (APT) of Ross (1976), if I find that *CNTR* is priced, then I could interpret the result as centrality premium consistent with no-arbitrage in the market.

Following the conventional methodology in the literature developed by Black, Jensen and Scholes (1972) and Fama and Macbeth (1973), I run two-pass cross-sectional regressions (CSR). In the first step, I estimate the betas of the test assets using ordinary least squares (OLS) time series regression of *residual* returns on the *centrality* factor *CNTR* (equation 5) i.e. I run:

$$\varepsilon_{i,t} = \alpha_i{}' + \beta_{i,CNTR} CNTR_t + \eta_{i,t} \quad t = 1, \dots, T \qquad (5)$$

In the second pass, excess returns are regressed on β_{CNTR} estimated from the time-series regression and a set of controls:

$$r_{it} - r_{ft} = \lambda_{0t} + \boldsymbol{\lambda_{CNTR,t}} \beta_{CNTR,i,t} + \lambda_{kt} \sum_{k=1}^{K} X_{kit} + \nu_{it} \quad i = 1, .., N, t = 1, \dots T \qquad (6)$$

The vector X_k contains the following variables: beta, size, book-to-market, turnover and idiosyncratic volatility (IVOL). I use beta to proxy for systematic risk and estimate it by regressing excess returns on the return of the market using rolling

36-month observations. Size and book-to-market are calculated as in Fama-French (1992). Turnover is obtained by dividing trading volume by numbers of shares outstanding. Idiosyncratic volatility (IVOL) is calculated as the standard deviation of residuals from the Fama-French 3 Factor Model, again using rolling 36-month observations. The logarithm of all variables is used in regressions and all controls are lagged one period in order to avoid bias and to ensure that information is available to the public before returns. Running this second-pass cross-sectional regression on a period-by-period basis, I obtain time series of the intercept and the slope coefficients. The values of the intercept and the slope coefficients are averaged out to obtain estimates for the factor premia, where standard errors are also calculated from these estimates. Although the Fama-Macbeth procedure (1973) produces standard errors robust to correlations across firms at given point in time, it does not address correlations between observations in different months[61]. To address time-series dependence, I still use the Fama-Macbeth estimator, but correct for autocorrelation in the standard errors using the Newey-West procedure using lag length of 1.

Panel A of Table 6 shows summary statistics of the variables used in the Fama-Macbeth regressions. Excess returns have a mean on 1.2 %, a standard deviation of 15% and a positive skew. The average of beta is approximately 1. The network beta, β_{CNTR}, has a lower mean than the market beta and is considerably more positively skewed. Size has a mean of 3.82 and a standard deviation of 2.161. *LogBM, LogTrnv* and *LogIVOL* have negative means and very low values of skewness, which is a result of the logarithmic transformation. All in all, the summary statistics of Table 6 do not show striking patterns and are consistent with existing research

Table 7 shows the main result of this paper. Model 1 shows a regression of monthly excess returns on market beta, size, BM and IVOL. As documented in Fama and French (1992), there is a flat relationship between returns and beta, a negative relationship between returns and size is a positive one between returns

[61] In fact, in the original application Fama and Macbeth find that the residuals are virtually serially uncorrelated

and BM and returns and turnover. Model 2 shows that the network risk factor is priced in the cross-section of returns and earns a monthly risk-premium of 0.479%. The result is statistically significant and economically meaningful. Furthermore, pricing errors are reduced almost in half compared to Model 1. Model 3 adds beta, size and book-to-market as controls and obtains that β_{CNTR} still bears a positive premium, although the coefficient is reduced. This is because there is a negative correlation between size and the network beta. Model 4 adds turnover and idiosyncratic volatility as controls. The premium on the centrality factor remains positive and significant. It is slightly higher than in Model 3 because of the positive interaction with turnover.

This empirical evidence supports the hypothesis that firms that are central in the network of business relationships earn higher returns as a compensation for higher exposure to network risk. This result is consistent with previous studies in the field, such as Ahern (2013) and Buraschi & Porchia (2013), both of which find that centrality bears a positive monthly premium. I acknowledge that a positive loading on *CNTR* could be reconciled with multiple hypotheses. The most natural explanation could be that it captures systematic risk that is not accounted by the other standard factors. Alternatively, as Ahern (2013) argues, it could be the case that the ex-post beta coefficient is measured poorly, and thus, *CNTR* gives a better estimate of ex-ante exposure to market risk.

5.7 Risk decomposition

The model in (5) in combination with Assumptions *A-D* allows for the following decomposition of idiosyncratic variance $\mathbb{V}(\varepsilon_t)$:

$$\mathbb{V}(\varepsilon_t) = \beta_{CNTR}\Sigma_{CNTR}\beta'_{CNTR} + \Sigma_\eta \qquad (6)$$

The first component gives the effect of interconnections on idiosyncratic risk, or the contribution of network exposure to total idiosyncratic risk. Another way to look at this component is as *systemic network risk,* which reflects the transmission of idiosyncratic shocks and their endogenous amplification. I call this component network variance (NVAR). The second component refers to pure or residual idiosyncratic variance, which is independent of the network connections. I call this component (RVAR). Then, network volatility, $NVOL = \sqrt{NVAR}$ and $RVOL = \sqrt{RVAR}$. Panel B of Table 6 shows summary statistics of the two variables. $RVOL$ is nearly 10 times bigger than $NVOL$, however the network component is considerably more skewed. I proceed to investigate how the two idiosyncratic risk components are priced in the cross-section of expected returns.

In an influential study Ang, Hodrick, Xing and Zhang (2006) find that stocks with higher idiosyncratic return volatility have, on average, lower expected future returns. The effect has been documented using US data (Ang et al. (2006)), as well as internationally (Ang et al. (2009)) and is found to be robust to size, value, momentum and liquidity effects. This presents a puzzle, because from a theoretical point of view idiosyncratic volatility should not be priced, or if it is, it should earn a positive premium at the most.

To investigate what are the implications of network risk for the idiosyncratic volatility puzzle, I use a daily version of IVOL: to obtain an estimate of IVOL in month τ, I calculated the standard deviation of residuals from regressions of daily excess returns on the three Fama-French factors within this month τ, multiplied by $\sqrt{business\ days}$. I similarly use daily observations for the construction of *CNTR, NVOL* and *RVOL*. In Model 1 of Table 8, consistent with previous research, I document that after controlling for market risk, size, value and liquidity effects, high idiosyncratic volatility in month $\tau - 1$ is predictive of lower future returns in month τ. In Model 2, I split *IVOL* into its two components and find that *NVOL* is positively priced, whreas *RVOL* is negatively priced. Also, the magnitude of the coefficient of *NVOL* is considerably higher than the magnitude of the coefficient of *RVOL*. After the inclusion of controls, the signs of coefficients of the two risk components remain unchanged, however, *NVOL* loses its statistical significance. In Model 4, I use the ratio of network risk to residual idiosyncratic risk i.e. $\frac{NVOL}{RVOL}$ as an explanatory variable and find a positive coefficient equal to 0.0385% per month.

All in all, these results give an important insight into the idiosyncratic volatility puzzle. However, they should not be intepreted as a solution to the anomaly. Rather, the message is that once network risk has been separated from total idiosyncratic risk, then regression results indicate a positive cross-sectional premium, which is consistent with theoretical predictions.

5.8 Out-of-sample prediction exercise

To evaluate the out-of-sample performance of the network measures PCAS, I follow the approach by Billio et al (2012). First, I rank stocks based on *PCAS1*, *PCAS5* and *PCAS10*. Then, I compute the maximum percentage loss (*Maximum % Loss*) suffered by each stock in my sample during the recent financial crisis period from July 2007 to December 2008. The maximum percentage lost for a firm is defined to be the difference between market capitalization in the end of June 2007 and the minimum market capitalization during the period from July 2007 to December 2008, divided by the market capitalization in the end of June 2007. Finally, I rank all companies according to their *Maximum % Loss*. I estimate univariate regressions of *Maximum % Loss* on the stock's PCAS rankings. Table 8 reports the results for 2 samples: October 2002 – September 2005 and July 2004 – June 2007. For each regression, I report the coefficients along with t-stats, p-values and Kendall τ (1938) rank-correlation coefficient.

I find that companies that are more exposed to idiosyncratic shocks i.e. those that have higher loadings on the *PCAS* measures, were more likely to suffer losses during the financial crisis. I register a difference in the predictive ability between the two samples: in the second sample, July 2004- June 2007 the size of the coefficients is higher and the Kendall τ indicates a positive relationship between *PCAS* and financial loss during the crisis. The results from the sample from October 2002- September 2005 are significant, but considerably lower. In this sample, I the value of Kendall's τ indicates that hypothesis of independence between the two ranks cannot be rejected.

These results suggest that the *PCAS* measure is able to identify systemic events and can serve as an early-warning signal.

5.9 Relationship to Macroeconomic Variables

In this section I would like to study the relationship between centrality and network risk and macroeconomic variables. I define *centrality spread, CSPR,* as the difference between the monthly returns of the firms in the highest quintile of *logPCAS10* and the monthly returns of those in the lowest quintile of *logPCAS10*. I look at three macroeconomic variables: unemployment, recessions and consumption growth. The recession variable is based on NBER business cycle dates and equals 1 during troughs and 0 during peaks. During the time period that I look at, there are 4 announcement dates: March 2001 (peak), November 2001 (trough), December 2007 (peak) and June 2009 (trough). In the months between these dates, I fill the recession variable with linearly interpolated values. Consumption growth is defined as the growth in nondurable consumption per capita in the future 6 months. I compare the observed relationship between the macro variables and *CSPR* to the relationship between the macro variables and five additional risk factors. Those factors are: *Mktrf* (the return spread of the value-weighted CRSP portfolio minus the risk-free rate); *SMB* (the return spread of small minus big stocks, i.e. size effect); *HML* (the return spread of high/cheap minus low/expensive Book-to-Market stocks, i.e. value effect), *CMA* (the return spread of firms that invest conservatively minus aggressively, i.e. investment effect) and *RMW* (the return spread of firms with robust minus weak operating profitability) (Fama & French, 2015). Panel A of Table 9 shows summary statistics. *CSPR* has a mean of 0.83% and a standard deviation of 5.24%. This compares very well to the statistics of *Mktrf,* which has a mean of 0.69% and a standard deviation of 4.16.

Panel B shows correlation coefficients between the variables. *CSPR* is positively correlated to *HML* and *CMA*, but the result is insignificant at the 10 % level. *CSPR* is significantly positively correlated to the market excess return, which I interpret as evidence in favor of the hypothesis that centrality is related to market risk: companies located more centrally in the network co-vary more with the market compared to companies in the periphery. Also, *CSPR* shows positive correlation to small stocks (which usually have high market betas) and negative correlation to more profitable stocks (which traditionally have lower market betas).

With respect to market variables I observe that unemployment and recessions are uncorrelated with the factors. On the other hand, consumption growth is positively correlated to the market premium (*Mktrf)* and the centrality spread (*CSPR).*

Table 10 gives the results of cross-sectional regressions of each of the risk factors defined above and the *centrality spread* on *unemployment, recession, consumption growth* and a time trend. Consumption growth is positively related to *Mktrf* and *CSPR* with coefficients of nearly equal magnitudes. Unfortunately, the other two macroeconomic outcomes are not related to any of the factors. In the case of *recession*, one potential explanation could be the limited amount of announcement dates, which makes the interpolation procedure less precise and reduces the power of these tests.

The results are consistent with the hypothesis that central firms face more economic risks than peripheral ones. When *CSPRP* is high i.e. the spread between the returns of central minus peripheral industries is higher, then future consumption growth increases as well. Furthermore, the tests that were carried out in the previous chapter's sections provide evidence that *CNTR*, the network risk mimicking factor, bears similarities to the market premium, which means that centrality could enrich our understanding of the driving forces of market risk.

6 Robustness

In this section, I provide robustness checks in support of the asset pricing tests in Chapter 5. In Chapter 6.2, I construct a *traded* factor mimicking portfolio of network risk and using a wide range of test assets I show in Fama-Macbeth regressions that this factor is priced in the cross-section of returns. In Chapter 6.3 I provide additional cross-sectional regressions that study the relationship between centrality variables and stock characteristics.

6.1 Construction of a Traded Network Factor

I construct the factor mimicking portfolio in the following way: every month I subtract the mean return of the firms in the lowest quintile of *logPCAS10* from the mean return of the firms in the highest quintile of *logPCAS10*. I call this long-short trading strategy *CMP*. To test whether *CMP* helps explain the cross-section of returns I use the same two-step procedure outlined in Chapter 5. Betas in the second-pass cross-sectional regression are lagged by one period in order to avoid bias and to consistently show that the model is able to predict future stock returns. Additionally, to account for the fact that the regressors in the cross-sectional Fama-Macbeth regression are not fixed, but estimated in a time-series regression, I correct t-statistics using the Shanken method. For further details, please refer to *Appendix A.3*. The advantage of the so-constructed traded factor is that I can evaluate its performance on a variety of test assets. Such a portfolio testing approach was not possible in the case of *CNTR* because it was asset-specific, which did not allow to estimate exposures on a portfolio level. Here, I use four sets of assets: 25 portfolios sorted on size and book-to-market, 25 portfolios sorted on centrality and size, 25 portfolios sorted on centrality and book-to-market and firm-level returns.

Table 11, Model 1 presents results that are consistent with prior research: there is a negative market beta and a positive constant. The factor premium on *CMP* is positive, but not significant and hence, *CMP* does not help explain variation in these portfolios. In contrast, when I use portfolios sorted on size and centrality (Model 2) and BM and centrality (Model 3), *CMP* remains positive and the factor premium becomes significant, but the other factors become insignificant. Also, the factor premium is considerably higher than in Model 1 and in the models in Table 7. One possible explanation could be that due to the inherent factor structure of the *size* and *book-to-market* portfolios, many additional factors could be correlated with these portfolios (Daniel & Titman, 2012) (Lewellen, Nagel, & Shanken, 2010). Actually, *CMP* exhibits a positive correlation with both *Mktrf* (0.49) and *SMB* (0.44). In Model 4 I conduct the most conservative tests using firm-level returns. The factor premia for all factors are statistically significant. *CMP* offers a premium of 0.33%, which is remarkably close to the factor premium on the market factor. This offers additional evidence that the network risk factor, constructed using centrality scores, bears similarities to the market factor. This means that centrality could enrich our understanding of the driving forces of market risk.

6.2 Additional Cross-sectional Regressions

I would like to investigate if and how network variables relate to stock returns, controlling for a range of firm characteristics. To do this, I carry cross-sectional regressions in the spirit of Fama-Macbeth (1973). The results are presented in Table 6.

In columns 1-3 I find that the measures of centrality: *log(PCAS1), log(PCAS5)* and *log(PCAS10)* are positively and significantly related to stock returns, where the effect for *log(PCAS10)* is the strongest. Consistent with prior research, I find that the logarithm of size (column 4) and share turnover (column 5) are positively related to firm returns.

In column 6 I regress stock returns on the Herfindahl index[62] and find a negative and significant relationship. The Herfindahl index measures the degree of industry concentration and proxies for the amount of competition firms in a

[62] Herfindahl Index $= \frac{1}{N}(1 + V^2)$, where N equals the number of firms in a given industry and V is the variance of the firm shares. Industry definitions follow the description by Fama and French and are downloaded from Kenneth French' website. Herfindahl Index $\epsilon\ [\frac{1}{N}, 1]$ with $\frac{1}{N}$ being the case when all companies have equal market share

given industry are exposed to. The negative loading on the Herfindahl Index suggests that firms in more concentrated industries earn lower stock returns. The finding is consistent with the results of Ahern (2013), who finds a negative but insignificant relationship.

In column 7, I obtain that average industry market equity is negatively related to stock returns. In column 8, I regress stock returns on the Fama-French and Momentum factors. I obtain positive coefficients for the *Mktrf, SMB* and *HML* factors and a negative one for the momentum factor.

In column 9 I include all variables used in the previous specifications. I find that the coefficient on *log(PCC10)* is reduced in half, but remains highly statistically significant. The reason for this reduction could be the strong correlation with *log size*. However, even after the inclusion of size, the effect of centrality remains and this suggests that centrality captures a dimension of market risk not entirely captured by firm size.

In conclusion, even after accounting for a wide range of stock and industry characteristics, such as size, share turnover, industry concentration and average industry size, there is still a strong positive cross-sectional relationship between centrality in the network and firm returns.

6. Conclusion

In this paper I investigate the empirical relationship between centrality and firm returns. I find that firms located in the highest quintile of centrality have higher expected returns than firms in the lowest quintile of centrality. I construct a factor mimicking portfolio of network risk using centrality scores as weights. In Fama-Macbeth cross-sectional regressions, I obtain that the factor is priced and earns a positive risk premium of 0.47%. The premium remains significant even after the inclusion of relevant control variables. This finding supports the hypothesis that a stock’s true market risk is partly influenced by its relative position in the network of business relations. Furthermore, in univariate rank regressions I find that stocks with higher scores of *PCAS,* were more likely to suffer higher losses during the recent financial crisis of 2007-2008. This confirms that

PCAS has out-of-sample predictive power. Finally, I demonstrate that the *centrality spread*, i.e. the difference between the returns of firms in the highest quintile of centrality minus the returns of stocks in the lowest quintile, is significantly positively correlated to future consumption in non-durable goods. By identifying links between stock market firms, their strength and sensitivity to changing market conditions and quantifying the individual contribution of a stock to the idiosyncratic volatility of the system, the paper offers insights into network risk. A better understanding of network risk is crucial because it has implications for corporate policy decisions (e.g. hiring, compensation, investment) as well as for firm value and asset prices.

One limitation of this paper is that estimating centrality scores using PCA precludes the possibility of assigning directionality to the inferred connections. In fact, PCA allows to establish whether there is a connection between firm A and firm B, but not to establish the direction of the link. This prevents me from being able to estimate precisely *close* and *distant* firms (in terms of the number of links between them). A potential extension for future research could be to use Granger-causality to establish the direction of links and then to use the information to study the dynamic propagation of shocks from *close* to *distant* firms in the network.

References

Acemoglu, D., Carvalho, V., Ozdaglar, A., & Tahbaz-Salehi, A. (2012). The network origins of aggregate fluctuations. *Econometrica, 80*, 1977-2012.

Acharya, V. V., Pedersen, L. H., Philippon, T., & Richardson, M. P. (2010). Measuring systemic risk. *Working Paper*.

Adrian, T., & Brunnermeier, M. (2011). CoVaR. *No. w17454. National Bureau of Economic Research*.

Ahern, K. (2013). Network Centrality and the Cross-Section of Stock Returns. *Working Paper*.

Ahern, K., & Harford, J. (2014). The importance of industry links in merger waves. *Journal of Finance, 69*(2), 527-576.

Anderson, R. R. (2004). Founding-Family Ownership and Firm Performance:Evidence from the S&P 500. *Evidence from the S&P 500" Journal of Finance 58, 1301-1328, 58*, 1301-1328.

Ang, A., Hodrick, R., & Xing, Y. Z. (2009). High idiosyncratic volatility and low returns. *Journal of Financial Economics, 91*(1).

Ang, A., Hodrick, R., Xing, Y., & Zhang, X. (2006). The cross-section of volatility and expected returns. *Journal of Finance, 61*(1), 259-299.

Arellano, M. (1987). Computing robust standard errors for within-groups estimators. *Oxford Bulletin of Economics and Statistics, 49*, pp. 431–434.

Atalay, E., Hortaçsu, A., Roberts, J., & Syverson, C. (2011). Network structure of production. *Proceeding of the National Academy of Sciences in the United States of America.*

Bai, J. (2003). Inferential theory for factor models or large dimensions. *Econometrics, 71*(1), 135-171.

Bai, N., & Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica, 70*(1), 191-221.

Barigozzi, M., & Brownlees, C. (2016). Network estimation for time series. *Working Paper*.

Barrot, J.-N., & Sauvagnat, J. (2016). Input specificity and the propagation of idiosyncratic shocks in production networks. *Quarterly Journal of Economics*.

Billio, M., Caporin, M., Panzica, R., & Pelizzon, L. (2015). Network Connectivity and Systematic Risk. *Working Paper*.

Billio, M., Getmansky, M., W.Lo, A., & Pelizzon, L. (2012). Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *Journal of Financial Economics, 104*, 535-559.

Black, F., Jensen, M., & Scholes, M. (1972). The Capital Asset Pricing Model:Some Empirical Findings. *Studies in the Theory of Capital Markets*. New York: Praeger.

Bloom, N., Bond, S., & Van Reenen, J. (2007). Uncertainty and Investment Dynamics. *The Review of Economic Studies, 74*(2), 391-415.

Bonnano, G., Caldarelli, G., Lillo, F., Miccich`e, S., Vandewalle, N., & Mantegna, R. (2004). Networks of equities in financial markets. *The European Physical Journal B, 38*, 363–371.

Borgatti, S. (2005). Centrality and network flow. *Social networks, 27*, 55-71.

Buraschi, A., & Porchia, P. (2013). Dynamic networks and Asset Pricing. *AFA 2013 San Diego.*

Cameron, A., & Trivedi, P. (2005). *Microeconometrics: Methods and Applications.* New York: Cambridge University Press.

Campbell, J., & Vuolteenaho, T. (2004). Bad beta, good beta. *American Economic Review, 94*, 1249-1275.

Carhart, M. (1995). Survivor bias and mutual fund performance. *Working Paper, University of Chicago.*

Carlson, M., Fisher, A., & Giammario, R. (2004). Corporate investment and asset price dynamics: implications for the cross-section of returns. *Journal of Finance, 59*, 2577-2603.

Chudik, A., & Pesaran, M. (2013). Large Panel Data Models with Cross-Sectional Dependence:A Survey. *Federal Reserve Bank of Dallas*.

Cohen, L., & Frazzini, A. (2008). Economic links and predictable returns. *The Journal of Finance, 63*(4), 1977-2011.

Daniel, L., & Titman, S. (2012). Testing factor-model explanations of market anomalies. *Critical Finance Review*, 103-139.

Diebold, F., & Yilmaz, K. (2014). On the network topology of variance decompositions: measuring the connectedness of financial firms. *Journal of Econometrics, 182*(1), 119-134.

Eicker, F. (1967). Limit theorems for regressions with unequal and dependent errors. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* (pp. 59-82). Berkely, CA: University of California Press.

Fama, E., & French, K. (2015). A five-factor asset-pricing model. *Journal of Financial Economics, 116*(1), 1-22.

Fama, E., & Macbeth, J. (1973). Risk, Return, and Equilibrium: Empirical Tests. *The Journal of Political Economy, 81*(3), 607-636.

Froot, K. (1989). Consistent covariance matrix estimation with cross-sectional dependence and heteroskedasticity in financial data. *Journal of Financial and Quantitative Analysis, 24*, 333-355.

Gomes, J., Kogan, L., & Zhang, L. (2003). Equilibrium cross-section of returns. *Journal of Political Economy, 111*, 693-732.

Gross, D., & Souleles, N. (2004). An Empirical Analysis of Personal Bankruptcy and Delinequency. *Review of Financial Studies, 15*, 319-347.

Hendricks, K., & Singhal, V. (2005). An empirical analysis of the effect of supply chain disruptions on long-run stock price performance and equity risk of the firm. *Production and Operations Management, 14*(1), 3552.

Herskovic, B., Kelly, B., & Van Nieuwerburgh, S. (2016). The common factor in idiosyncratic volatility: Quantitative asset pricing implications. *Journal of Financial Economics, 119*(2), 249-283.

Hoechle, D. (2007). Robust standard errors for panel regression. *The STATA Journal, 7*(3), 281-312.

Hong, H., Torous, W., & Valkanov, R. (2007). Do industries lead stock markets? *Journal of Financial Economics, 83*, 367-396.

Hou, K. (2007). Industry information diffusion and the lead-lag effect in stock returns. *Review of Financial Studies, 20*(4), 1113-1138.

Hou, K., & Robinson, T. (2006). Industry concentration and average stock returns. *The Journal of Finance, 61*, 1927-1956.

Huang, X., Zhou, H., & Zhu, H. (2011). Systemic Risk Contribution. *Board of Governors of the Federal Reserve System*.

Huber, P. (1967). The behavior of maximum likelihood estimates under nonstandard. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* (pp. 221-233). Berkely, CA: University of California Press.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20*, 141-151.

Kelly, B., Lustig, H., & Nieuwerburgh, S. (2013). Firm Volatility in Granular Networks. *University of Chicago, University of Los Angelis and University of New York Working Paper*.

Kendall, M. (1938). A new measure of rank correlation. *Biometrica 30*, 81-89.

Lang, H., & Stulz, R. (1992). Contagion and competitive intra-industry effects of bankruptcy announcements. *Journal of Financial Economics, 32*(1), 45-60.

Lee, G., & Djauhari, M. (2012). An Overall Centrality Measure: The Case of U.S. *International Journal of Basic & Applied Sciences, 12*(06).

Lewellen, J., Nagel, S., & Shanken, J. (2010). A skeptical appraisal of asset pricing tests. *Journal of Financial Economics, 96*, 175-194.

Lucas, R. (1977). Understanding business cycles. *Carnegie-Rochester conference series on public policy*. Elsevier.

Lucia, A. .., Barigozzi, M., & Capasso, M. (2009). A robust criterion for determining the number of factors in approximate factor models. *ECARES Working Papers*.

Mantegna, R., & Stanley, H. (1999). *Introduction to econophysics: correlations and complexity in finance.* Cambridge University Press.

Menzly, L., & Ozbas, O. (2010). Market segmentation and cross-predictability of returns. *The Journal of Finance, 65*(4).

Muhammad, U., & Radha, H. (2010). A KLT-inspired node centrality for identifying influential neighborhoods in graphs. *Information Sciences and Systems*, 1-7.

Namakia, A., Shirazi, A., Raei, H., & Jafari, G. (2011). Network analysis of a financial market based on genuine correlation and threshold method. *Physica A, 390*, 3836-3841.

Newey, W., & West, K. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica, 55*, 703-708.

Petersen, M. (2009). Estimating standard errors in finance panel datasets: Comparing. *Review of Financial Studies, 22*(1), 435-480.

Ross, S. (1976). The Arbitrage Theory of Capital Asset Pricing. *Journal of Economic Theory, 13*, 341-360.

Schönsleben, P. (2007). *Integral Logistics Management.* Auerbach Publications.

Shanken, J. (1992). On the Estimation of Beta-Pricing Models. *Review of Financial Studies, 5*(1), 1-33.

Stokey, N. (2012). Uncertainty and Investment Delays. *Working Paper-University of Chicago*.

Takayuki, M., Takayasu, H., & Takayasu, M. (2006). Correlation networks among currencies. *Physica A: Statistical Mechanics and its Applications, 364*, 336-342.

Tse, C. K., Liu, J., & Lau, F. C. (2010). A network perspective of the stock market. *Journal of Empirical Finance, 17*(10), 535-559.

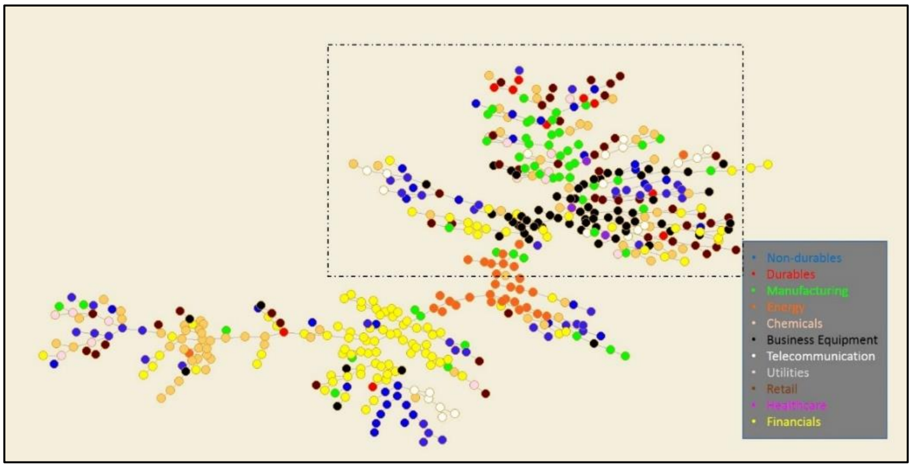
White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica, 48*, 817-838.

List of Figures

Figure 1: A Spanning Tree of the 500 Largest Stocks in December 2002

Panel A represents the network of the 500 Largest Stocks in the sample as of December 2002. Links between nodes are obtained through *a minimum spanning tree* and the strength of the connection between node *i* and *j* is determined by the absolute value of the correlation between two nodes. Companies belonging to different industries are color-coded as shown on the legend. *Panel B* shows a subgraph, where the three biggest automotive manufacturers are given in purple: **GM** (General Motors), **CHR** (Daimler-Chrysler) and FORD (Ford Motor Company). Additionally, two financial companies are shown for illustrative purposes in **yellow**: **TD** (Toronto Dominion) and **CITI** (Citibank Inc). The graph is produced using the Pajek Software for network analysis.

Panel A



Panel B

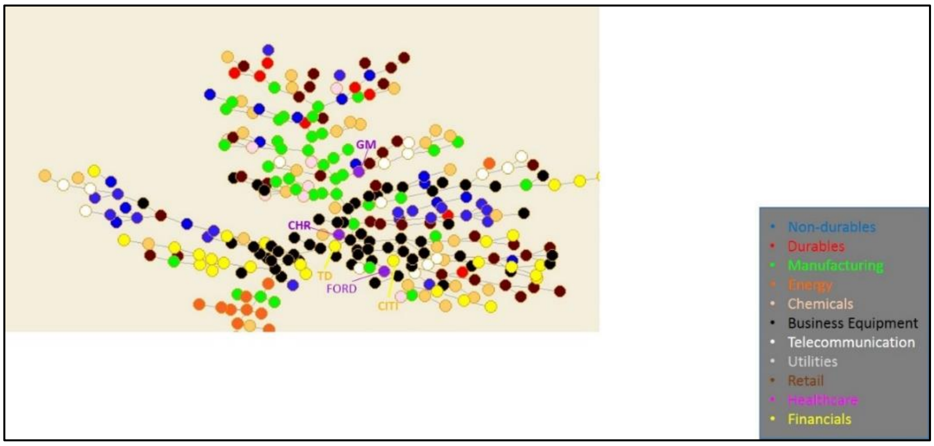


Figure 2: Principal Components Analysis of residual returns

The figure shows 36-month rolling-window estimates for the cumulative fraction of total variance explained by the first 1, 5, 10, 15, 20 and 36 principal components over the period from December 2002 to December 2015.

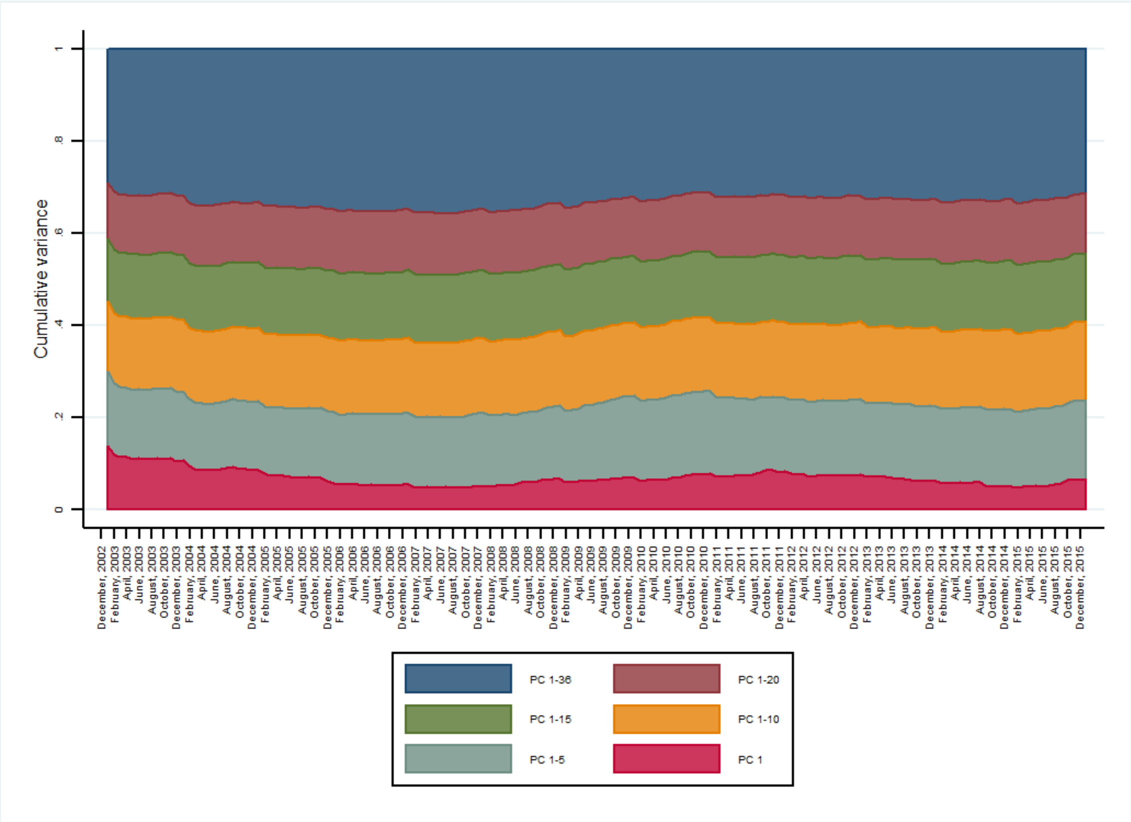


Figure 3: Confidence Intervals for Cumulative Variance

The figure plots upper and lower confidence bounds at the 95% level for the cumulative variance explained by PC 1-10. Confidence intervals are estimated using the empirical distribution from 1,000 bootstrapped samples.

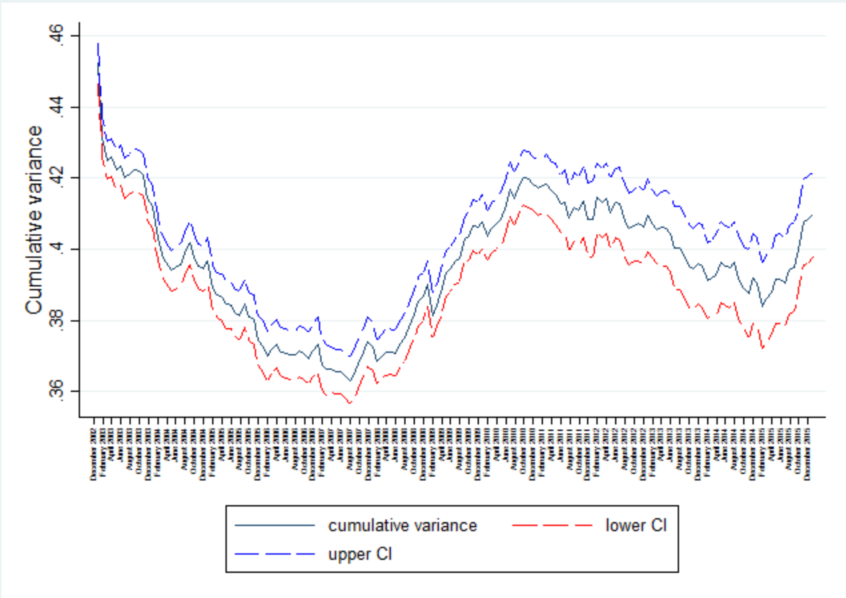


Figure 4: Empirical Density of Centrality

Panel A shows the empirical density function of PCAS, PCAS5 and PCAS10. The estimate is based on a normal kernel function, and is evaluated at a 100 equally-spaced points. Panel B gives the counter-cumulative distribution on a log-log scale.

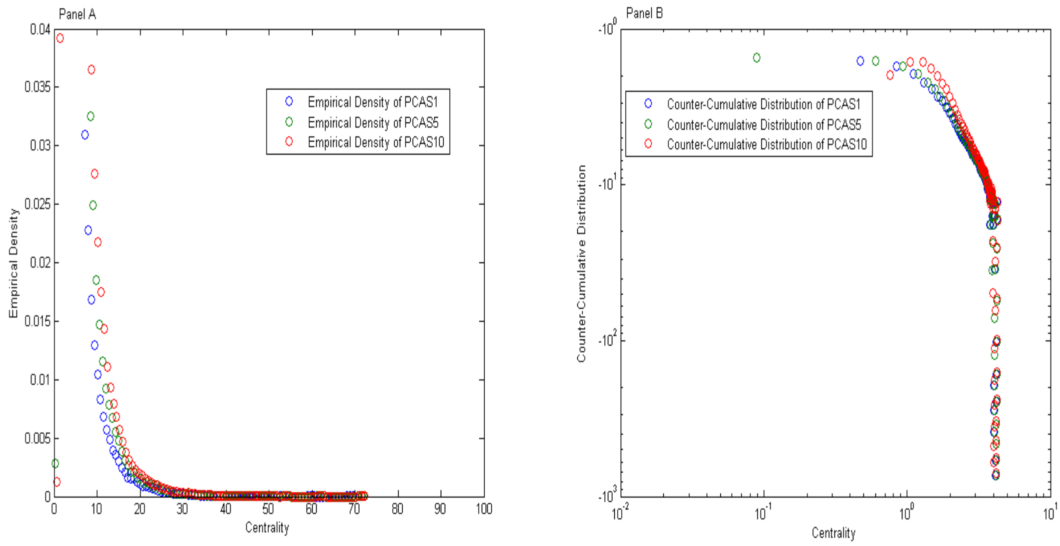
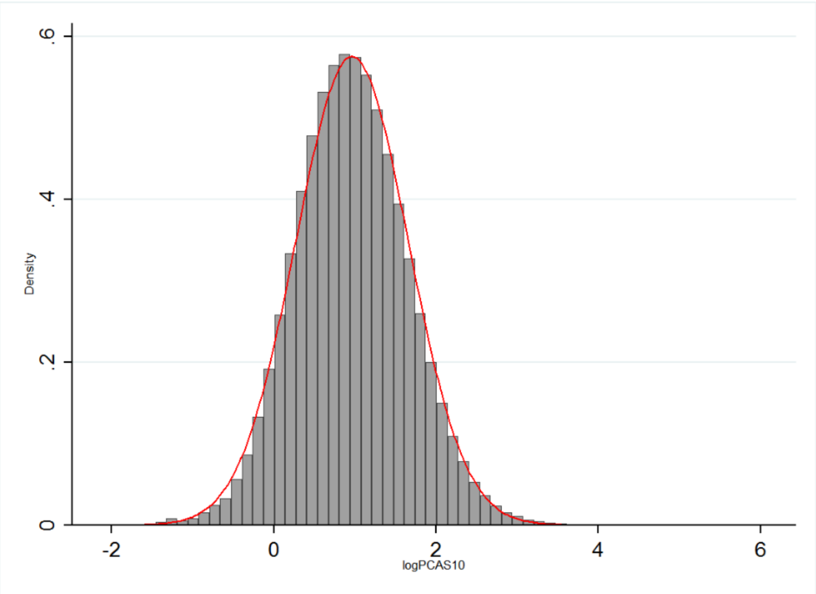


Figure 5: Histogram of LogPCC10

The figure plots the histogram of *logPCC10*. Frequencies are given in percentages. The density of the normal distribution is superimposed over the histogram to facilitate visual comparison of the statistical properties of the centrality measure.



List of Tables

Table 1: Cumulative Variance from PCA

The table presents the fraction of total returns variance explained by the first 1, 5, 10, 15, 20 and 36 Principal Components. PCs are calculated using rolling 36-month periods.

Sample Period	PC1	PC1-5	PC1-10	PC1-15	PC1-20	PC1-36
Dec 2002-Dec2005	9.4	24.5	40.1	54.0	67.1	≈100
Jan 2006-Dec2009	5.8	21.5	37.7	52.1	65.4	≈100
Jan 2010-Dec2013	7.5	24.5	41.0	54.9	67.9	≈100
Jan2014-Dec2015	5.8	22.4	39.4	54.0	67.3	≈100

Table 2: Information Criteria for PCA

Panel A of Table 2 shows values for the three information criteria IC_1, IC_2 and IC_3 developed by Bai and Ng (2002). By minimizing the value of the information criteria, the number of factors/ principal components to be used in factor models can be consistently estimated. Panel B gives the values of the eigenvalues. To conserve space, only information about the first 15 PCs is presented. The full table is given in Appendix A. 2

A: Information Criteria				B: Eigenvalues
PCs	IC_1	IC_2	IC_3	λ_i
1	**-3.139**	-15.889	**-2.976**	**2.625**
2	0.375	**-25.123**	0.702	**1.67**
3	11.545	11.569	0.801	**1.40**
4	15.226	15.258	0.900	**1.32**
5	18.903	18.946	0.998	**1.27**
6	22.587	22.635	1.098	**1.22**
7	26.268	26.324	1.197	**1.18**
8	29.948	30.012	1.296	**1.15**
9	33.629	33.701	1.396	**1.12**
10	37.309	37.390	1.495	**1.10**
11	40.990	41.079	1.594	1.05

12	44.671	44.768	1.694	1.03
13	48.352	48.457	1.793	1.01
14	52.033	52.146	1.893	1.00
15	55.714	55.834	1.992	0.99
16	59.395	59.523	2.091	0.98
…	….	…..	…	< 0.98

Table 3: Summary Statistics of Centrality

	logPCAS1	*logPCAS5*	*logPCAS10*
Mean	0.710	1.372	1.530
Standard Dev.	1.134	0.606	0.547
Variance	1.280	0.367	0.299
Skewness	-0.911	0.154	0.325
Kurtosis	4.061	2.839	2.857

Table 4: Mean Portfolio Returns sorted by Centrality

This table presents mean equally-weighted and value-weighted returns sorted by centrality measures for the period Dec 2002-Dec 2015. For each stock, returns in moth τ are matched to its centrality measure in month $\tau - 1$. Centrality measures PCAS1, PCAS5 and PCAS10 are calculated using the loadings on the first one, five and ten eigenvalues of the variance-covariance matrix of residual returns. For the estimation 36-month rolling windows are used. Value-weighted returns are calculated using ME/average ME as weights. Statistical significance between the mean returns in the highest and lowest centrality quintiles is reported by t-tests from Newey-West standard errors. Statistical significance at the 10%, 5%, and 1% level is indicated by *, ** and ***.

	Centrality						
	Low	2	3	4	High	H-L	t-stat
Panel A: Mean Monthly Returns, %							
EW PCC1	1.38	1.32	1.25	1.25	1.47	0.09	(1.39)
VW PCC1	0.72	0.79	0.75	0.69	0.56	-0.16	(-1.27)
EW PCC5	1.21	1.25	1.3	1.31	1.6	0.39***	(4.97)
VW PCC5	0.76	0.75	0.85	0.75	0.24	-0.52***	(-3.29)
EW PCC10	1.17	1.22	1.33	1.32	1.64	0.47***	(5.86)
VW PCC10	0.92	1.01	0.66	0.56	0.37	-0.55***	(-2.99)
Panel B: Firm Characteristics by Log PCAS10							
Log PCAS10	0.847	1.240	1.494	1.772	2.297	1.449***	(23.49)
Log Size	14.14	13.60	13.11	12.58	11.83	-2.31***	(-59.85)

Table 5: Mean Portfolio Returns Double-Sorted by Centrality and Size

This table presents mean equally-weighted and value-weighted returns of double-sorted portfolios for the period Dec 2002-Dec 2015. Centrality here is measured by PCC10, which is calculated using the loadings on the first ten eigenvalues of the variance-covariance matrix of residual returns. For the estimation 36-month rolling windows are used. Value-weighted returns are calculated using ME/average ME as weights. Panel A shows double-sorts on centrality and size, whereas Panel B shows double-sorts on centrality and Book-to-Market. Portfolios are rebalanced monthly. Statistical significance between the mean returns in the highest and lowest centrality quintiles is reported by t-tests from Newey-West standard errors. Statistical significance at the 10%, 5%, and 1% level is indicated by *, ** and ***.

Panel A: Equally-Weighted Returns, %

	Centrality						
Size	L	2	3	4	H	H-L	t-stat
S	1.4	1.28	1.1	1.42	1.23	-0.17	(-0.58)
2	1.27	1.12	1.23	1.61	1.71	0.44***	(2.21)
3	1.09	1.34	1.26	1.52	1.57	0.48***	(2.81)
4	1.22	1.37	1.45	1.46	1.52	0.30**	(1.79)
B	1.19	1.14	1.09	1.45	1.40	0.21	(1.18)
S-B	0.21	0.14	0.01	-0.03	-0.17	-0.38	(-1.05)
tstat	(0.79)	(0.64)	(0.08)	(-0.17)	(-0.74)	(-1.05)	
BM	L	2	3	4	H	H-L	t-stat
L	1.21	1.52	1.63	1.66	1.63	0.42***	(2.66)
2	1.25	1.28	1.51	1.42	1.54	0.29*	(1.72)
3	1.25	1.2	1.27	1.36	1.71	0.46**	(2.53)
4	1.04	1.16	1.1	1.26	1.62	0.58***	(3.30)
H	1.16	1.06	1.06	1.24	1.51	0.35*	(1.70)
L-H	0.05	0.46***	0.57***	0.42**	0.12	0.07	(0.29)
tstat	(0.34)	(2.99)	(3.35)	(2.06)	(0.58)	(0.29)	

Panel B: Value-Weighted Returns, %

	Centrality						
Size	L	2	3	4	H	H-L	t-stat
S	0.79	0.67	0.49	0.52	0.09	-0.70	(-3.9)
2	1.02	0.86	0.96	1.29	1.04	0.02	(0.12)
3	0.87	1.04	0.97	1.23	0.94	0.07	(0.32)
4	0.97	1.14	1.07	1.00	1.13	0.16	(0.86)
B	1.10	1.0	0.75	0.90	0.60	-0.5**	(-2.1)
S-B	-0.31	-0.3**	-0.26	-0.38*	-0.50*	-0.20	(-0.6)
tstat	(-1.7)	(-2.2)	(-1.5)	(-1.8)	(-1.9)	(-0.62)	
BM	L	2	3	4	H	H-L	t-stat
L	1.35	1.75	1.04	0.62	0.33	-1.0***	(-3.0)
2	1.26	1.2	1.04	0	0.27	-0.99**	(-2.8)
3	1.46	0.38	0.54	0.65	0.55	-0.91*	(-1.7)
4	1.22	1.03	0.40	1.28	0.58	-0.64	(-1.2)
H	0.9	-0.17	0.41	0.59	-0.41	-1.31**	(-2.0)
L-H	0.45	1.92***	0.63	0.03	0.74	0.29	(0.44)
tstat	1.20	(3.74)	(1.33)	(0.05)	(1.34)	(0.44)	

Table 6: Summary Statistics of Variables

Table 6 shows summary statistics related to stock variables used in cross-sectional regressions. Estimation period is Dec 2002-Dec 2015.

	Mean	St. Deviation	Q1	Q3	Skew
Panel A: Stock Characteristics					
Exret	0.012	0.15	-0.055	0.064	4.00
Beta	1.05	0.96	0.47	1.49	1.08
β_{CNTR}	0.71	0.45	0.43	0.85	2.61
LogSize	3.82	2.161	2.22	5.3	0.18
LogBM	-0.58	0.86	-1.05	-0.09	-0.40
LogTrnv	-0.16	1.28	-0.88	0.71	-0.73
LogIVOL	-2.13	0.52	-2.50	-1.70	0.31
Panel B: Network Variables					
NVOL	0.010	0.017	0.002	0.0111	17.15
RVOL	0.120	0.114	0.058	0.144	10.41

Table 7: Fama-Macbeth Cross-sectional Regression

This table presents results of two-stage monthly cross-sectional regressions over December 2002 to December 20015 of average excess firm-level returns on centrality beta and a set of controls. The CNTR factor is calculated using residuals weighted by LogPCAS10 centrality. Market betas are estimated from time-series regressions stock returns on Fama-French 3 Factors (Mktrf, SMB, HML); logIVOL is the standard deviation of the residuals from this model. LogSize and LogBM are calculated as in Fama and French (1992). LogTrnv measures liquidity and is calculated as trading volume divided by number of shares outstanding. Coefficients estimates are in percentages. T -statistics are computed using Newey-West procedure with lag length 1. R^2s give average values from the Fama-Macbeth regressions. Values in the parentheses are p-values. Statistical significance at the 10%, 5%, and 1% level is indicated by *, ** and ***.

	(1)	(2)	(3)	(4)
betaMktrf	0.04		0.125**	0.05
	(<0.18)		(0.03)	(0.127)
LogSize	-0.252***		-0.172***	-0.261***
	(<0.00)		(<0.00)	(<0.00)
LogBM	0.491***		0.455***	0.481***
	(<0.00)		(<0.00)	(<0.00)
LogTrnv	0.311***			0.317***
	(<0.00)			(<0.00)
LogIVOL				-0.02
				(0.75)
betaCNTR		0.479***	0.165*	0.282**
		(<0.00)	(<0.08)	(<0.07)
Constant	1.52***	0.864***	1.89***	2.37***
	(<0.00)	(<0.00)	(<0.00)	(<0.00)
R^2	0.400	0.021	0.070	0.077

Table 8: Risk decomposition

This table presents results of Fama-Macbeth cross-sectional regressions over December 2002 to December 20015 of average expected excess firm-level returns on idiosyncratic risk components and stock characteristics. IVOL, NVOL and RVOL are calculated using daily observations in a given month τ and lagged one period. Market betas are estimated from time-series regressions stock returns on Fama-French 3 Factors (Mktrf, SMB, HML. LogSize and LogBM are calculated as in Fama and French (1992). LogTrnv measures liquidity and is calculated as trading volume divided by number of shares outstanding. T -statistics are computed using Newey-West procedure with lag length 1. Values in the parentheses are p-values. Statistical significance at the 10%, 5%, and 1% level is indicated by *, ** and ***.

	(1)	(2)	(3)	(4)
betaMktrf	0.080		0.080	0.04
	(0.18)		(0.03)	(0.266)
LogSize	-0.029**		-0.028**	0.025**
	(0.03)		(<0.03)	(0.031)
LogBM	0.631***		0.633***	0.645***
	(<0.00)		(<0.00)	(<0.00)
LogTrnv	0.174***		0.173***	0.120***
	(<0.00)		(<0.00)	(<0.00)
IVOL	-0.007**			
	(0.04)			
NVOL		**0.074***	**0.04**	
		(0.092)	(0.32)	
RVOL		**-0.009***	**-0.013****	
		(0.097)	(0.028)	
$\frac{NVOL}{RVOL}$				0.0384***
				(<0.00)
Constant	1.77***	1.24***	1.79***	2.37***
	(<0.00)	(<0.00)	(<0.00)	(<0.00)
R^2	0.400	0.43	0.070	0.077

Table 9: Predictive Power of PCAS Measures

The table shows coefficients, t-stats, p-values and Kendall (1938) τ rank-correlation coefficients of regressions of Maximum % Loss on PCAS 1, PCAS 1-5 and PCAS 1-10. The maximum percentage lost for a firm is defined to be the difference between market capitalization in the end of June 2007 and the minimum market capitalization during the period from July 2007 to December 2008, divided by the market capitalization in the end of June 2007. Estimates are shown for two samples: October 2002- September 2005 and July 2004 to June 2007. Statistical significance at the 10%, 5%, and 1% level is indicated by *, ** and ***.

	Maximum % Loss			
	Coeff	t-stat	p-value	Kendall τ
	October 2002-Septermber 2005			
PCAS 1	0.01	1.14	0.256	0.01
PCAS 1-5	0.12***	7.00	0.00	0.08
PCAS 1-10	0.17***	9.49	0.00	0.11
	July 2004-June 2007			
PCAS 1	0.05***	3.22	0.001	0.03
PCAS 1-5	0.19***	10.92	0.00	0.12
PCAS 1-10	0.24***	14.19	0.00	0.16

Table 10: Summary Statistics and Correlation Coefficients of Risk Factors

This table presents summary statistics (Panel A) and pairwise correlation coefficients (Panel B) for the five Fama-French Factors (2015): *Mktrf, SMB, HML, CMA* and *RMW* and the centrality spread. In Panel B, statistical tests for the significance of the pairwise correlations use the Bonferroni correction to account for the fact that multiple comparisons are jointly tested. *P-values* are given in parentheses. Statistical significance at the 10%, 5%, and 1% level is indicated by *, ** and ***.

Panel A: Summary Statistics

	Obs	Mean	SD	Min	Max	t-stat
Mktrf	157	.69	4.16	-17.23	11.35	(1.94)
SMB	157	.23	2.32	-4.25	5.79	(1.3)
HML	157	.07	2.27	-9.67	7.65	(0.33)
CMA	157	5.43	135.21	--316	344	(0.46)
RMW	157	17.31	173.35	-712	570	(1.15)
CSPR	157	0.83	5.24	-12.03	20.8	(1.92)

Panel B: Correlation Coefficients

	Mktrf	SMB	HML	CMA	RMW	CMP
SMB	0.409***					
	(<0.000)					
HML	0.251**	0.123				
	(0.025)	(>0.999)				
CMA	0.146	0.153	0.412***			
	(>0.999)	(0.832)	(<0.000)			
RMW	-0.518***	-0.484***	-0.1045	-0.227*		
	(<0.000)	(<0.000)	(>0.999)	(0.062)		
CSPR	0.572***	0.554***	0.138	0.221	-0.689***	
	(<0.000)	(<0.000)	(>0.999)	(0.19)	(<0.000)	
Recession	-0.159	-0.144	-0.025	-0.152	0.066	-0.209
	(>0.999)	(>0.999)	(>0.999)	(>0.999)	(>0.999)	(0.383)
Consumption	0.250**	0.015	-0.053	-0.102	-0.215	0.244*
	(0.05)	(>0.999)	(>0.999)	(>0.999)	(0.244)	(0.075)
Unemployment	0.136	0.088	-0.021	0.143	-0.024	0.132
	(>0.999)	(>0.999)	(>0.999)	(>0.999)	(>0.999)	(>0.999)

Table 11: Cross-Sectional Regression with Macro Variables

The table presents results of regressions of 5FF Factors and Centrality on macroeconomic outcomes over the period December 2002 to December 2015. Recession is based on NBER business cycle announcement dates and equals 1 for troughs and 0 for peaks, with values in the intervening months being linearly interpolated. Consumption growth is defined as the growth rate in nondurable consumption per capita in the future 6 months. T-statistics are given in brackets. Statistical significance at the 10%, 5%, and 1% level is indicated by *, ** and ***.

	Mktrf	*SMB*	*HML*	*CMA*	*RMW*	*CSPR*
Unemployment	-0.235	-0.057	-0.000	0.138	0.200	-0.503
	(-0.57)	(-0.23)	(-0.22)	(1.00)	(1.00)	(-0.62)
Recession	-3.192	-2.010	-0.010	-0.269	1.387	-7.41
	(-1.31)	(-1.54)	(-0.82)	(-0.29)	(1.02)	(-1.62)
Consumption	0.405**	-0.043	-0.001	-0.084**	-0.144**	0.311**
	(2.34)	(-0.59)	(-1.42)	(-2.47)	(-2.82)	(2.15)
Time Trend	Yes	Yes	Yes	Yes	Yes	Yes
Observations	157	157	157	157	157	157
R^2	0.093	0.047	0.044	0.062	0.058	0.119

Table 12: Robustness: Fama-Macbeth Cross-sectional Regression

This table presents results of two-stage monthly cross-sectional regressions over December 2002 to December 20015 of average excess firm-level returns on the exposures to the three Fama-French Factors and CMP. CMP is constructed by subtracting the mean returns of stocks in the lowest quintile of centrality from the mean returns of stocks in the highest quintile of centrality. Betas are estimated from time-series regressions. Model 1 uses 25 Size and BM Portfolios; Model 2 uses 25 Size and Centrality Portfolios; Model 3 uses 25 BM and Centrality Portfolios; Model 4 uses firm-level returns. Centrality is based on PCAS10. Coefficients estimates are in percentages. T - statistics are computed using Newey-West procedure with lag length 1 and corrected for errors-in-variables problem using the Shanken method (1992). R^2s give average values from the Fama-Macbeth regressions. Values in the parentheses are p-values. Statistical significance at the 10%, 5%, and 1% level is indicated by *, ** and ***.

	(1) 25 Size and BM Portfolios	(2) 25 Size and Centrality Portfolios	(3) 25 BM and Centrality Portfolios	(4) Firm-level Returns
betaMktrf	-0.62	-0.037	-0.19	0.32***
	(0.613)	(0.650)	(0.861)	(<0.00)
betaSMB	0.19	0.155	0.41	0.18***
	(0.415)	(0.640)	(0.717)	(<0.00)
betaHML	0.019	0.37	0.35	-0.17***
	(0.715)	(0.650)	(<0.434)	(<0.00)
betaCMP	0.44	0.93**	1.14**	0.33***
	(0.675)	(0.03)	(0.022)	(<0.001)
Constant	1.50	0.73	1.89***	0.77***
	(0.283)	(0.353)	(<0.00)	(<0.00)
R^2	17.85	19.93	20.12	0.44

Table 13: Cross-sectional Regression of Firm-Level Monthly Returns

This table presents the results of cross-sectional regressions of firm-level monthly returns on centrality measures, industry characteristics and risk factors. The coefficients are calculated using the Fama-Macbeth (1973) two-step approach, t-statistics are given in parentheses. The dependent variable is measured in percentages. Centrality measures (PCC1, PCC5, and PCC10) are calculated as the loadings on the first one, five and ten eigenvalues of the variance-covariance matrix of returns. Log Sh. Trnv equals trading volume divided by number of shares outstanding. Based on their SIC codes, firms are divided into 49 industry groups using the definitions by Fama and French. The Herfindahl Index measures industry concentration and is calculated as the sum of squared market shares within an industry. Industry ME gives average market equity. Statistical significance at the 10%, 5%, and 1% level is indicated by *, ** and ***.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
LogPCC1	0.34***								
	(15.51)								
LogPCAS5		1.13***							
		(23.77)							
LogPCAS10			1.33***						0.78***
			(23.10)						(16.67)
Log Size				0.162***					0.110***
				(12.03)					(7.34)
Log Sh.Trnv					0.573***				0.598***
					(25.43)				(21.34)
Herfindahl						-0.073***			-0.031
						(-3.64)			(-1.39)
Industry ME							-0.168***		-0.254***
							(-5.18)		(-8.17)
Mktrf								0.963***	0.974***
								(149.58)	(149.47)
SMB								0.766***	0.747***
								(64.31)	(62.67)
HML								0.135***	0.125***
								(11.75)	(17.72)
UMD								-16.94***	-16.28***
								(-19.48)	(18.91)

Constant	1.270***	0.240***	-0.268***	-0.707***	1.416***	1.32***	2.088***	0.451***	0.052
	(55.60)	(4.87)	(-3.54)	(-4.65)	(52.91)	(53.96)	(50.85)	(20.12)	(0.74)
R^2	0.001	0.002	0.002	0.001	0.003	0.000	0.001	0.128	0.134

Appendix A.1 Minimum Spanning Trees and Prim's Algorithm

Given the correlation matrix C of dimension *NxN*, where the correlation coefficient between stocks i and j is:

$$c_{ij} = \frac{\langle r_i r_j \rangle - \langle r_i \rangle \langle r_j \rangle}{\sqrt{(\langle r_i^2 \rangle - \langle r_i \rangle^2)(\langle r_j^2 \rangle - \langle r_i \rangle^2)}}$$

and $\langle r_i \rangle$ is the statistical average of $r_i(t)$ for all t. The distance matrix D can be defined by transforming C in the following way:

$$d_{ij} = \sqrt{2(1 - c_{ij})}$$

The distance d_{ij} defined in this way is the Euclidean distance norm because it satisfies the three axioms for a norm:

$(i)\ \ d_{ij} \geq 0\ and\ d_{ij} = 0\ iff\ i = j$

$(ii)\ \ d_{ij} = d_{ji}$

$(iii)\ \ d_{ij} \leq d_{ik} + d_{kj}$

I apply Prim's Algorithm to obtain a *Minimum Spanning Tree* (MST) from the distance matrix D. Such a tree is a graph without loops, which connects all n nodes with *n-1* links. Therefore, the MST retains the *n-1* shortest (strongest) links, that span the all the nodes in the network. Prim's Algorithm can be described in the following steps:

1. Start with an empty tree
2. Initialize the tree with a vertex chosen at random
3. Grow the tree by one edge: choose the edge that connects the tree to a vertex not yet in the added with the minimum distance.
4. Repeat step 2 until all vertices are in the tree.

Appendix A.2 Information Criteria

Table 1: Information Criteria and Eigenvalues

The table shows the full set of results for the Bai and Ng (2002) information criteria and for the magnitude of the eigenvalues for all PCs.

	Information Criteria			Eigenvalues
PCs	IC_1	IC_2	IC_3	λ_i
1	**-3.139**	-15.889	**-2.976**	**2.625**
2	0.375	**-25.123**	0.702	**1.67**
3	11.545	11.569	0.801	**1.40**
4	15.226	15.258	0.900	**1.32**
5	18.903	18.946	0.998	**1.27**
6	22.587	22.635	1.098	**1.22**
7	26.268	26.324	1.197	**1.18**
8	29.948	30.012	1.296	**1.15**
9	33.629	33.701	1.396	**1.12**
10	37.309	37.390	1.495	**1.10**
11	40.990	41.079	1.594	**1.08**
12	44.671	44.768	1.694	**1.05**
13	48.352	48.457	1.793	**1.03**
14	52.033	52.146	1.893	**1.01**
15	55.714	55.834	1.992	**1.0**
16	59.395	59.523	2.091	0.99
17	63.076	63.212	2.191	0.98
18	66.757	66.901	2.290	0.96
19	70.437	70.589	2.389	0.94

20	74.118	74.278	2.488	0.92
21	77.799	77.967	2.588	0.91
22	81.479	81.656	2.687	0.89
23	85.160	85.344	2.786	0.87
24	88.841	89.033	2.885	0.85
25	92.521	92.722	2.984	0.84
25	96.202	96.4111	3.084	0.82
27	99.883	100.100	3.183	0.80
28	103.564	103.788	3.2831	0.79
29	107.245	107.477	3.382	0.77
30	110.925	111.166	3.481	0.75
31	114.606	114.855	3.581	0.73
32	118.287	118.544	3.680	0.71
33	121.967	122.232	3.778	0.69
34	125.648	125.921	3.878	0.66
35	129.328	129.609	3.977	0.63
36	133.009	133.298	4.076	0.50

Appendix A.3 Shanken Correction

When applying the standard OLS formulas to a cross-sectional regression, it is assumed that the regressors β are fixed. The β in the cross-sectional regression are not fixed, but are estimated in the first-pass time series regression. To address the errors-in-variables (EIV) problem, I use a multiplicative correction due to Shanken (1992).

$$\sigma^2(\hat{\lambda}_{OLS}) = \frac{1}{T}\left[(\beta'\beta)^{-1}\beta'\Sigma\beta(\beta'\beta)^{-1}\left(1+\lambda'\Sigma_f^{-1}\lambda\right)\right]$$

where Σ_f is the variance-covariance matrix of the factors λ is a vector of factors. Here, the Shanken correction ≈ 1.06. To obtain correct estimates, standard errors are multiplied by $\sqrt{1.06}$ and t-statistics are divided by $\sqrt{1.06}$. In my case, the correction does not change the results substantially, which is not surprising given the monthly frequency of the data. Consider for comparison annual data: since both the mean and the variance scale with horizon, at an annual frequency $\lambda'\Sigma_f^{-1}\lambda \approx 0.06 * 12 \approx 0.72$, which is too big a value to be ignored.