Text-Based Linkages and Local Risk Spillovers in the Equity Market*

Shuyi Ge †

October 24, 2020

Abstract

This paper uses extensive text data to construct firms' links via which local shocks transmit. Using the novel text-based linkages, I estimate a heterogeneous spatial-temporal model which accommodates the contemporaneous and dynamic spillover effects at the same time. I document a considerable degree of local risk spillovers in the market plus sector hierarchical factor model residuals of $S\&P$ 500 stocks. The method is found to outperform various previously studied methods in terms of out-of-sample fit. Network analysis of the spatial-temporal model identifies the major systemic risk contributors and receivers, which are of particular interest to microprudential policies. From a macroprudential perspective, a rolling-window analysis reveals that the strength of local risk spillovers increases during periods of crisis, when, on the other hand, the market factor loses its importance.

Keywords— Excess comovement, local risk spillovers, networks, textual analysis, big data, systemic risk, weak and strong cross-sectional dependence, heterogeneous spatial autoregressive model (HSAR)

---

*The author would like to thank Oliver Linton, Hashem Pesaran, Cheng Hsiao, Koen Jochmans, Julius Vainora, Michael Leung, Sampad Mohanty, Shaoran Li, Weiguang Liu, Chong Shu, Ekaterina Gavrilova, Marco Valerio Geraci, Merrick Li, Chen Wang, Yimeng Xie, Lidan Tan, Grigory Franguridi, Chris Yoo and all the participants at the University of Cambridge Econometrics workshop and the University of California Econometrics workshop for their useful comments.

†The author is at the Faculty of Economics, University of Cambridge. Her email address is sg751@cam.ac.uk

1 Introduction

According to asset pricing theories such as the classical capital asset pricing model (CAPM) developed by Sharpe (1964), and the arbitrage pricing theory (APT) by Ross (1976), asset returns have a common factor representation with a strong pervasive component driven by a few common factors and an idiosyncratic component that is weakly correlated. Many studies have found that the APT models with the factors in the existing literature seem to be insufficient to capture all the significant interdependencies in asset returns. Local risk spillovers still play a non-negligible role (see for example Gabaix (2011), Acemoglu et al. (2012), Israelsen (2016), Barigozzi and Hallin (2017), Kou et al. (2018), Bailey et al. (2019), and Barigozzi and Brownlees (2019)).

To distinguish the two sources of dependencies, imagine a group of people sitting in a room on a chilly winter day. People might catch a cold because the heater is broken (common factor), or because someone sitting close to them are ill (local interactions). The network architecture of firms, like the sitting plan in the previous example, is key to studying local risk spillovers. However, such linkage data are usually unavailable to researchers, which hinders the study of local dependencies. This paper uses extensive text data to construct firms' links via which local shocks transmit. Using the novel text-based network, I estimate a heterogeneous spatial-temporal model which accommodates the contemporaneous and dynamic spillover effects at the same time. I document a considerable degree of local risk spillovers in the market plus sector hierarchical factor model residuals of $S\&P$ 500 stocks.

I obtain a novel text-based linkages dataset using extensive news data from LexisNexis Academic. LexisNexis Academic has a collection of news from a wide range of sources. Company names and tickers that are mentioned in each piece of news are tagged. Using this feature, I identify firms that share business links by common news coverage. The maintained assumption is that two companies share a link if they are the only two that get mentioned in the same piece of news. Although I am using a special feature of LexisNexis Academic, the method can be generalized to any other text by applying an additional named-entity-recognition step (see Figure 6 for a demo of named-entity-recognition). Using all the business news from the Business Wire from 2006-2013, the estimated full sample network is plotted in Figure 7. The network has a core-periphery structure. Big banks including JPMorgan Chase (JPM), Citi (C), Goldman Sachs (GS) and Bank of America (BAC), and big hitech firms including Microsoft (MSFT), Apple (AAPL), Intel (INTC) and Oracle (ORCL), and big manufacturers and conglomerates including General Electric (GE) and Procter & Gamble (PG) are the most connected companies in the $S\&P$ 500 universe, occupying the centre of the graph. Companies within the same sector appear as clusters, implying a lot of intra-sector relationships. The method also uncovers many cross-sector relationships. Table 11 presents links aggregated at the sector level. Intra-sector links account for a high percentage of total links for every sector. In particular, the hi-tech sector has the highest percentage of intra-sector links as such companies are very well connected with each other in order to develop products together and form strategic partnerships. On the other hand, the consumer sector has the lowest percentage of intra-sector links. This feature is obvious from Figure 7, as the red nodes are more scattered.

The novel dataset complements existing network datasets in several ways. While existing network datasets are usually lagged, incomplete, and cover only certain types of links for certain types of firms[1], the textual analysis method complements these datasets by identifying additional types of links that facilitate risk spillovers. In addition to interbank relationship and customer-supplier links, the method also finds strategic partnerships, business lines acquisitions, investment banking relationships, funding relationships, similar legal and regulatory exposures, and $M\&A$ relationships, etc. As a comparison, Figure 9 plots the customer-supplier network among $S\&P500$ firms using Compustat segments data. It is visible that far fewer links are identified. In response to the lack of network data, there has been a strand of literature using pure statistical methods to estimate links from a panel of equity returns/volatilites. Examples include Diebold and Yılmaz (2014), Hale and Lopez (2019), Barigozzi and Hallin (2017), Barigozzi and Brownlees (2019), and Demirer et al. (2018). Figure 10 plots the long-run variance decomposition network (LVDN), long-run Granger causality network (LGCN) and partial correlation network (PCN) among $S\&P500$ companies estimated from the idiosyncratic returns (market plus sector hierarchical factor model residuals) using the high-dimensional methodology of Barigozzi and Hallin (2017). The links identified are very different from period to period, showing strong temporal variation. And very few links from the crisis period long-run variance decomposition network (LVDN) appear in the pre-crisis LVDN. The links that turn out to be important for risk transmissions in periods of crisis are like the submerged icebergs that are hard to detect ex-ante and reveal themselves only when large shocks hit the system. Additional sources of information could be fruitful in aiding the link detection. Our text-based network constructed using pre-crisis news outperforms the high-dimensional vector autoregression (VAR) estimated using pre-crisis sample in terms of detecting long-run variance decomposition network (LVDN) links in the crisis period.[2] This is due to the fact that our text-based links are more persistent than the Long-run variance Decomposition network (LVDN) links implied by the high-dimensional vector autoregression (VAR) model. Eighty-two percent of the links are identified in more than two different monthly windows. And on average, 59.32% of the linked pairs identified in a year continue to get identified in later years, showing that the method identifies long-lived economic links among companies. Taken together, it can be seen that the text-based network complements alternative network datasets and can be viewed as a promising alternative to other datasets.

Utilizing the novel text-based linkages data, I quantify the strength of local risk spillovers in the equity market using a heterogeneous spatial autoregressive model (HSAR) studied in Bailey et al. (2016) and Aquaro et al. (2019). The model captures temporal dependence as well as spatial-temporal dependence. It is flexible, and the individual-specific parameters can be consistently estimated for any N as long as T is large. Since the equity returns comovement reflects both exposures to common risk factors (the broken heater) and local risk spillovers (the sick neighbours), I first remove the strong cross-sectional dependence (CSD) by de-factoring equity returns. I show that after removing the 6 Fama French common risk factors and the Fama French industry risk factors, there is still a considerable degree of local risk spillovers via the links identified by our textual analysis method.

[1] For example, interbank network data only cover the lending relationships among banks, and they are not publicly available. The Compustat segments data are available to researchers, but they only cover customer-supplier links.

[2] Table 15 shows the percentages of crisis period Long-run variance Decomposition network (LVDN) links that get identified using alternative pre-crisis network information. Different hard thresholds are applied to the LVDN given the network implied by LVDN is very dense (the link densities for pre-crisis and crisis sample are 77.5% and 95.3%, respectively). We do not need to apply thresholding to text-based network since it is already very sparse (the link density of the full sample network is 4.5%, and for the short pre-crisis sample the density is even smaller). For any non zero thresholds applied, the text-based networks consistently outperform that of pre-crisis LVDN in terms of detecting out-of-sample links.

Thanks to the flexible framework, we have found a substantial degree of sectoral heterogeneity. In particular, manufacturing firms and financial firms are more sensitive to the shocks of their neighbours. It is also worth noticing that the dynamic spillover effect for financial firms is more pronounced as for any lag order, as the percentage of significant individual-specific spatial-temporal coefficients is about twice as large as that of other sector groups. The spatial-temporal framework allows us to analyse a complicated diffusion pattern of local shocks over time and space. The decay of shock along the spatial dimension is slower than that along the time dimension. By constructing spatial-temporal spillover matrices using the estimated parameters, we identify the major systemic risk contributors and receivers, which are of particular interest to microprudential policymakers. The firms that contribute the most to the systemic risks are the large cap financial institutions and manufacturers. Apart from systemic risk contributors, companies that are particularly sensitive to others' shocks are also found. It is worth noticing that the well-connected systemic risk contributors themselves are not necessarily the major risk receivers. Rather, they are the periphery firms that are exposed to a lot of risks from the core.

To examine how the strength of local risk spillovers evolves over time, I consider a rolling window analysis. The estimation results reveal that the local dependencies intensify during periods of financial crisis and turmoil. The surge in local risk spillovers could be a signal of rising systemic risk, which is useful for macroprudential purposes. Previous studies have documented that asset returns depart from fundamentals during times of financial crisis, and stocks dis-connect from the market factor (see Bailey et al. (2019) and Bailey et al. (2020)). Our analysis tracks the evolution of strong cross-sectional dependence (CSD) and weak cross-sectional dependence (CWD) at the same time and documents an interesting fact: local risk spillovers intensify when the market factor loses its importance during times of financial crisis and turmoil, which is evidence of market decoupling.

This paper contributes to two strands of literature. The first strand of literature is textual analysis and its application in the financial market; and in particular, how to quantify the soft information contained in news articles. Textual analysis is a useful tool to construct novel datasets. It fills the gaps in data availability induced by limited disclosure and slow update, thus complementing traditional economic datasets. For example, there has been a steep rise in the number of studies on sentiment analysis recently (see Garcia (2013), Price et al. (2012), and Ke et al. (2019) among others). Similar to the sentiment index, an economic policy uncertainty index (EPU) has been developed by Baker et al. (2016), which is based on newspaper coverage frequency of political words. Textual analysis has also been used for link mining. Hoberg and Phillips (2016), and Hoberg and Phillips (2018) construct peer links using text-based analysis of firm 10-K product descriptions. Scherbina and Schlusche (2015), and Schwenkler and Zheng (2019) both identify firm links from business news.

The second strand of literature is local risk spillovers in equity returns. Equity returns comovement reflects both exposure to common risk factors and local interactions that generate spillovers. While exposure to common factors gives rise to strong cross-sectional dependence (CSD), local risk spillovers represent weak cross-sectional dependence (CWD), with the latter form of interdependence receiving much less attention compared with the former. Since local shocks transmit among economically-linked firms, a key reason for the lack of empirical work is the lack of linkage information. Existing network datasets are limited, because much of a company's data are considered highly proprietary. Facing the challenge, this paper contributes to the literature by uncovering a wide range of business links that facilitate local risk spillovers from publicly available text data. Using the text-based network, the paper documents the existence of "excess-comovement" in linked stocks beyond what is predicted by standard asset pricing models. Another contribution is that the econometric framework applied in this paper addresses contemporaneous spillovers and dynamic cross spillovers at the same time. To the best of the author's knowledge, previous studies focus either on contemporaneous dependence or dynamic cross spillovers but not both at the same time.

To assess the performance of the method, I compare the in-sample and out-of-sample mean squared errors (MSE) of it with various alternative methods. The literature uses publicly available proxies for interfirm linkages include intra-industry links (Moskowitz and Grinblatt (1999), Fan et al. (2016), Engelberg et al. (2018)), geographic links (Pirinsky and Wang (2006), and Parsons et al. (2020)), and customer-supplier links (Cohen and Frazzini (2008)). I consider the spatial-temporal model estimated using the above links as our alternatives. Another competing method is the high-dimensional vector autoregression (VAR) approach used by Barigozzi and Hallin (2017), Barigozzi and Brownlees (2019)) and Demirer et al. (2018). The method shrinks, selects and estimates high-dimensional network when no explicit links are observed. In terms of in-sample fit, the high-dimensional VAR approach has the smallest MSE. This is not surprising, given the method selects the model by minimizing the Bayesian information criterion. However, when we look at out-of-sample fit, which is more important practically, the heterogeneous spatial-temporal model estimated with the text-based network outperforms all other methods. Comparing with the existing linkage datasets that each only cover a particular type of relatedness, our text-based links provide an integrated measure of relatedness. And comparing with the high-dimensional VAR, the textual analysis approach identifies persistent economically-meaningful links. These reasons explain why our method has superior performance. To evaluate the robustness of superior performance, I conduct the model comparison for different sub-samples and the result is robust across all sub-samples.

The rest of the paper is organized as follows: section 2 describes the data and link identification strategy and shows some key properties of the estimated linkages. Section 3 introduces the modelling of strong and weak cross-sectional dependence using a factor plus spatial two-stage procedure. The main focus in on local risk spillovers among linked stocks. Section 4 provides full sample estimation results and the construction of spatial-temporal spillover matrices using estimated parameters. And this section also presents the model comparison results. Section 5 provides a rolling window analysis and characterizes the evolution of local risk spillovers over time. Section 6 considers an extensive alternative specifications and robustness checks. Section 7 gives some concluding remarks.

2 Data and Link Identification

All the stock market related data are from the Center for Research in Security Prices (CRSP). Since our econometric framework requires large T for consistent estimation, I use the daily stock file. Sector classification is based on the Standard Industrial Classification (SIC) code from the CRSP/Compustat Merged database and I modify Fama French classification criteria provided on Kenneth French's homepage. As I will elaborate in section 4, to obtain mean group (MG) estimates of each sector group's parameters, one requires that the number of stocks within each group to be big enough. Due to that consideration, I build the sector classification on top of the FF5 industry definitions where they classify all stocks according to their SIC code into 5 broad groups:

"Consumer", "Health", "Hi-tech", "Manufacturing" and "Others". For the first four categories, I keep the same definitions as Fama and French. Since there are a large proportion of financial companies in the $S\&P500$ universe and our sample period covers the financial crisis, it would be interesting to separate financial firms from the those in the "Others" category. Among the stocks that fall into "Others", I categorize the stocks with a SIC in the range $6000 - 6799$ as "Finance" and put the remaining in the "Others" category. Daily Fama-French factors returns and industry portfolios' returns are taken from Kenneth French's homepage.

As for the text data, I download all the full-text business news from Business Wire that tagged $S\&P500$[3] companies from January 2006 to December 2013 on LexisNexis Academic[4]. A news item contains a title, date, body and classification. A typical business news item in the dataset is in shown in Figure 5. This example news item reports on the strategic partnership between American Express and Regis Corporation. The main subject of the news item is summarized by some key words. And in the classification section, the relevant companies are tagged with their tickers listed. Although I am using a special feature of the news available on LexisNexis Academic, the method can be generalized to any other text by using an additional Named-Entity-Recognition step (see Figure 6 for a demo of Named-Entity-Recognition). There are $345,880$ distinct business news items that have tagged sample companies during the whole sample period, and each sample month has around $3,200$ distinct business news items. This section will mainly focus on the identification of links from our text database and some key properties of those links.

2.1 Identification of Links

Common news coverage reveals information about linkages among companies. In this paper, links are identified by common business news coverage. The identification assumption is that if a piece of business news reports two companies together, then the two firms have some sort of business relationship/link. Although news that mention multiple companies together may carry potential information about links, they are noisier (for example, analyst recommendations, ratings changes, and index movements might stack multiple companies together when they actually do not have real links). Due to those concerns, I discard news that tag more than two firms.

I use a $N \times N$ adjacency matrix $W = (w_{ij})$ to store all the links identified in the sample news. N is the number of sample companies and a typical entry w_{ij} is the number of times i and j are co-mentioned in different news items. The link estimation procedure is as follows. For each piece of distinct news item in the sample, (1) we firstly extract the tickers tagged; (2) keep the news item if only two distinct[5] tickers are tagged; (3) match the tagged tickers with sample companies; (4) if both tagged companies are successfully matched, say if the

[3] The composition of $S\&P500$ index changes over time. All stocks that have stayed on the list and have no missing return observations for more than one year during the sample period are considered.

[4] LexisNexis Academic is a database of full text online news, legal cases and company from information. News from hundreds of source are available. After entering the company names and narrowing down the subject to "business news", Business Wire is always among the top sources list. To maximize the number of relevant business news during sample period and to avoid duplications, only the news from Business Wire is used. The python code of data scraping is available upon request

[5] A same company listed on different stock exchanges may have different tickers. For example, Citi used to list on both the New York Stock Exchange (NYSE) and the Tokyo Stock Exchange (TSE) at the same ticker with different ticker names: C(NYSE) and 8710 (TSE). To avoid double counts of a same company, only tickers associated with the New York Stock Exchange (NYSE), National Association of Securities Dealers Automated Quotation System (Nasdaq) and American Stock Exchange (AMEX) are kept.

they are matched to the companies correspond to the ith and the jth row/column, we add 1 to both w_{ij} and w_{ji}. Process (1)-(4) is repeated for every piece of distinct news item in the sample.

2.2 Estimated Links

The links identified using all the business news from Business Wire from 2006 to 2013 are plotted in Figure 7. Only companies with links are plotted in the figure. Given the long sample period and huge amount of business news, only a small number of sample companies were not co-mentioned with another[6]. In the figure, the nodes represent companies and two nodes are connected by an edge if there is a link between them. The size of a node is proportional to the number of neighbours it has (i.e., its degree) and the colour of a node indicates which sector it is in.

The estimated full sample network has a core-periphery structure. The most connected companies in the network graph include big banks, big hi-tech companies and big manufactures. These big banks include JP-Morgan Chase (JPM), Citi (C), Goldman Sachs(GS) and Bank of America (BAC), the big hi-tech firms include Microsoft (MSFT), Apple (AAPL), Intel (INTC) and Oracle (ORCL) and the big manufacturers and conglomerates include General Electric and Procter & Gamble (PG). They occupy the centre of the graph. Table 9 provides the link validation results for the most frequently mentioned pairs. Big banks engage in a variety of business relationships with other companies including financing, joint ventures, strategic partnerships, joint investment banking, acquisition of business lines and competition. Hi-tech giants are very well connected with one another to form strategic partnerships and develop new products together. Supplier-customer relationships and business lines acquisitions are found among big manufacturers and conglomerates. Companies within the same sector appear as clusters, indicating that there are dense intra-sector linkages. Most of the hi-tech firms lie on the third quadrant (bottom left corner) of the graph, while most of the health companies show up in the first quadrant (top left corner) of the graph. The method also uncovers a lot of cross-sector relationships. Table 11 presents links aggregated at the sector level. Manufacturing and consumer companies have a relatively lower percentages of intra-sector links. This is consistent with the feature from Figure 7 that red and orange nodes are more scattered.

Table 10 gives the summary statistics of the news-based links estimated using full sample. In total, $40,185$ links are identified in the full sample period, among which there are $6,742$ unique pairs of companies that share links. The former number is much larger than the latter one since most pairs of firms are mentioned together multiple times in different news items. The link density of the full sample network is 4.5%. Those links are discovered over time and the yearly networks plotted in Figure 8 are sparser. Over the full sample period, each company is connected to around 24 other companies in the $S\&P500$ universe on average. The network shows a substantial level of degree heterogeneity with a small number of firms being highly-connected, which is shown by the $90th$ percentile of degree. This feature is consistent with the core-periphery structure of most empirical networks. To have a more detailed understanding of the features of the links, I further break down the links into intra-sector and inter-sector links. Our method identifies a lot of inter-sector links as well as intra-sector links.

[6] For the balanced panel of 413 sample companies, only 5 out of 413 never got co-mentioned with others. There are 546 firms that have stayed on the $S\&P500$ list and had no missing return observations for at least a year, and they are included in the one year rolling sample. Among 546 firms that are included in the rolling sample, only 26 of them never got co-mentioned with others.

Although the full sample network is sparse, the intra-sector link densities are high. For example, the intra-sector link density for hi-tech companies reaches 16%. Compared with the peer link mining method from Hoberg and Phillips (2016) and Lee et al. (2015), the method used in this paper has the advantage of discovering not only peer links but also inter-sector (industry) links. Over the full sample, the number of distinct inter-sector pairs identified is larger than the number of distinct intra-sector pairs identified. And this is still true even if we look at different sectors separately.

The full sample period is long and different links are identified in different years. The network graphs and summary statistics for each year from $2006-2013$ are given in Figure 8 and Table 12, respectively. For the links identified using only one year's news, the statistics are much smaller than that those in Table 10. This implies that new links get identified over time and they carry timely information about the interconnectedness among companies. To roughly gauge the percentages of "new" and "stale" links, I calculate the percentage of linked pairs identified in one year that were identified in the previous year. On average, 37.82% of the linked pairs identified in a year are "stale" links and 62.18% are "new" links. And 82% of the links are identified in more than two different monthly windows. However, the high percentage of "stale" information is not necessarily a bad thing, as it implies the news-based links are persistent. I calculate the percentage of linked pairs identified in one year that continue to get identified in later years. On average, 59.32% of the linked pairs identified in a year continue to get identified in later years, showing that the method identifies long-lived economic links among companies.

2.3 Comparison with Other Networks

The novel dataset complements existing network datasets in several perspectives. While existing network datasets are usually lagged, incomplete, and cover certain types of links for certain types of firm, the link mining method complements these information sources by identifying additional types of links that facilitate the transmissions of local shocks. Figure 9 plots the customer-supplier network of $S\&P500$ firms using Compustat segments data. Consumer companies such as Walmart and McKesson are well-connected given they have a wide range of suppliers and customers. Apart from several consumer companies, there is no apparent star in the network. Very few links of financial firms are uncovered. On the other hand, the link mining approach applied in this paper is more successful at identifying links of financial firms.

Instead of turning to existing limited network datasets, one strand of literature uses purely statistical methods to estimate links from a high-dimensional time series (see Barigozzi and Hallin (2017), Barigozzi and Brownlees (2019), and Demirer et al. (2018)). Figure 10 plots the long-run variance decomposition network (LVDN), the long-run Granger causality network (LGCN) and the partial correlation network (PCN) among $S\&P500$ companies estimated from the our sample of idiosyncratic returns (the market plus sector hierarchical factor model residuals) using the high-dimensional methods from Barigozzi and Hallin (2017). Although we are using idiosyncratic returns while they use idiosyncratic volatilities, two prominent features remain true. The first feature is that the Financial Crisis has blown up the interconnectedness in the system. From Figure 10, it is clear that for all three types of networks considered, the network in the crisis period is much denser than that of others[7]. The second feature is that the links identified are very different from period to period. Table 15 shows the percentages of thresholded crisis period LVDN links that also appear in the thresholded pre-crisis LVDN. Expect the results for no thresholding, where the link densities for pre-crisis and crisis sample are 77.5% and 95.3% respectively, for other thresholds applied, very few links from the crisis LVDN appear before the crisis. From those two features, the links that turn out to be important for risk transmissions in the crisis period are like the submerged icebergs that are hard to detect ex-ante and reveal themselves only when large shocks hit the system. As a result, such high-dimensional link estimation method alone is not so useful for policymakers to monitor systemic risk. Rather, additional sources of information, could be fruitful in aiding the link detection. For example, if we apply a 5% hard threshold to the both the pre-crisis and crisis LVDN, then only 4% of the thresholded LVDN links identified from the crisis period are also identified from the pre-crisis period. On the other hand, our text-based links identified from the same pre-crisis period reveals the 34% of the thresholded LVDN links identified from the crisis period. For other non-zero thresholds, the text-based links consistently outperform that of pre-crisis LVDN in terms of predicting crisis period links. This is due to the fact that our text-based links are much more persistent than the LVDN constructed using high-dimensional VAR estimates. On average, 59.32% of the linked pairs identified in a year continue to get identified in later years, showing that the method identifies long-lived economic links among companies. Taken together, it can be seen that the text-based network complements alternative network datasets and can be veiwed as a promising alternative to other datasets.

3 Local Risk Spillovers Among Linked Stocks

Equity returns comovement reflects both exposure to common risk factors and local risk spillovers where the latter source of comovement receives much less attention compared with the former one. However, the models that focus on strong cross-sectional dependence such as CAPM and ATP fail to capture all the cross-sectional dependence in the equity returns. Many studies show that the local dependence in the idiosyncratic component is non-negligible (see Gabaix (2011), Acemoglu et al. (2012), Barigozzi and Hallin (2017), Kou et al. (2018) among others), thus it is important to examine the role played by local interactions. Adopting the econometrics framework in Bailey et al. (2016), I remove the strong cross-sectional dependence using a factor approach and then use spatial models to examine the local risk spillovers (weak cross-sectional dependence) remaining in the idiosyncratic returns. Unlike spatial interactions in the geographical systems, where there exist a natural network structure, for a panel of equity returns there is no natural network structure. The text analysis approach in this paper helps to identity the business links among listed firms and thus allows us to construct the channels through which local shocks transmit. It is found that there is significant local dependence among linked firms' idiosyncratic returns.

[7] Table 13 shows the number of links from the thresholded LVDN for pre-crisis, crisis and full sample periods (different thresholds applied).

3.1 De-factoring Equity Returns

To disentangle weakly correlated idiosyncratic returns from the strongly correlated returns driven by pervasive factors, one could use factor models. To be specific, I apply the below hierarchical factor model:

$$r_{it} - r_{ft} = \alpha_i + \mathbf{b}_i'\mathbf{f}_t + \gamma_i'\mathbf{f}_{g,t} + \epsilon_{it} \tag{1}$$

where r_{it} denotes the return of stock i at time t and subtracting the risk-free rate r_{ft} gives the excess return. $\mathbf{f}_t$ is the K_1 vector of common risk factors that affect every stock in the market. Since a large proportion of the links identified are intra-sector links, to avoid spurious spillovers that are actually caused by sectoral common factors, we add the K_2 sectoral risk factors $\mathbf{f}_{gt}$ that affect members of sector g but not others. $\mathbf{b}_i$ and γ_i are the loadings of common risk factors and sectoral risk factors, respectively. For the choice of factors, we can either use observed factors like Fama-French factors or statistical factors extracted using principal component analysis.

Our analysis needs the number of members to be large within each sector group g, so we consider six broad sectoral categories that I will elaborate upon in details next. For the choice of $\mathbf{f}_t$, I use five Fama French factors (Fama and French (2015)) plus the momentum factor (Carhart (1997)). And as for the sectoral factors $\mathbf{f}_{g,t}$, I use the Fama French industry portfolios. As an alternative to the observed market and sectoral factors, one could use unobserved factors, which I will use it as a robustness check.

3.2 Local Risk Spillovers: a Heterogeneous Coefficient Spatial-temporal Model

3.2.1 Heterogeneous Coefficient Spatial-temporal Model

After removing the strongly pervasive component driven by common risk factors, the remaining dependence is weak (local). Spatial econometrics methods are natural tools to address the weak (local) cross-sectional dependence in the idiosyncratic component, where entities interact locally. Conventional homogeneous spatial models restrict the spatial response parameter to be the same across all units. While such restriction is necessary for small T panels, it need not to be imposed when T is large. For a panel dataset with sufficiently large T, one can exploit the data along the time dimension to estimate individual-specific parameters for all N units.

One might reasonably suspect that the sensitivity to neighbours' risks is different from firm to firm. Since the stock market data set usually covers a long time period, we can utilize this nice feature to explore the heterogeneity in the strength of local dependency. The local risk spillovers in the idiosyncratic component is modelled using a heterogeneous coefficient spatial-temporal model (Bailey et al. (2016), LeSage and Chih (2016) and Aquaro et al. (2019)) that is written as follows:

$$\boldsymbol{\epsilon}_t = \mathbf{a}_\epsilon + \underbrace{\sum_{k=1}^{L_1} \boldsymbol{\Lambda_k} \boldsymbol{\epsilon}_{t-k}}_{\text{temporal dependence}} + \underbrace{\sum_{k=0}^{L_2} \boldsymbol{\Psi_k} W \boldsymbol{\epsilon}_{t-k}}_{\text{spatial temporal dependence}} + \boldsymbol{\upsilon}_t \tag{2}$$

where $\boldsymbol{\epsilon}_t$ is the $N \times 1$ vector of de-factored returns and $\mathbf{a}_\epsilon = (\alpha_{\epsilon,1}, \ldots, \alpha_{\epsilon,N})$ is the $N \times 1$ vector of intercepts. $\boldsymbol{\lambda}_k = diag(\lambda_{k,1}, \ldots, \lambda_{k,N})$ gives the autoregressive parameters of the kth lag for $k = 1, \ldots, L_1$, and $\boldsymbol{\Psi}_0 = diag(\psi_{0,1}, \ldots, \psi_{0,N})$ gives the contemporaneous spatial coefficients, and $\Psi_k = diag(\psi_{k,1}, \ldots, \psi_{k,N})$ gives the spatial-temporal parameters of the of the kth lag for $k = 1, \ldots, L_2$. Notice that for the individual specific spatial coefficients $\boldsymbol{\psi}_i = (\psi_{0,i}, \ldots, \psi_{L_2,i})'$ to be identifiable, company i has to have non-zero number of neighbours. For unconnected i, we need to restrict their spatial related coefficients $\boldsymbol{\psi}_i = \mathbf{0}$. The error variance $\boldsymbol{\sigma}_{\upsilon^2} = var(\upsilon_{it})$ is allowed to differ for diffent i. W is the $N \times N$ adjacency matrix that specifies the channels from which shocks transmits. As a convention in spatial econometrics, the diagonal elements are set to zero ($w_{ii} = 0$ for all $i = 1, \ldots, N$), and all other entries are assumed to be non-negative ($w_{ij} >= 0$). Also, the weights are row-normalized so that $\sum_j^N w_{ij} = 1$ for all $i = 1, \ldots, N$.

The model can be consistently estimated using the QML procedure proposed in Bailey et al. (2016) and Aquaro et al. (2019). We collect all the parameters in the $(N*(L_1+L_2+3))$ by 1 vector $\boldsymbol{\theta} = (\mathbf{a}_\epsilon{}', \boldsymbol{\lambda}_1', \ldots, \boldsymbol{\lambda}_{L_1}', \boldsymbol{\Psi}_0', \ldots, \boldsymbol{\Psi}_{L_2}', \boldsymbol{\sigma}_{\upsilon^2}')'$ and the log-likelihood function of (2) is written as follows:

$$\mathcal{L}_T(\boldsymbol{\theta}) = -\frac{NT}{2} ln(2\pi) - \frac{T}{2}\sum_i^N ln(\sigma_i^2) + \frac{T}{2} ln \mid \boldsymbol{S}'(\boldsymbol{\psi_0})\boldsymbol{S}(\boldsymbol{\psi_0}) \mid - \frac{1}{2}\sum_{t=1}^T [\boldsymbol{S}(\boldsymbol{\psi_0})\boldsymbol{y_t} - \boldsymbol{B}\boldsymbol{x_t}]'\Sigma^{-1}[\boldsymbol{S}(\boldsymbol{\psi_0})\boldsymbol{y_t} - \boldsymbol{B}\boldsymbol{x_t}] \quad (3)$$

where $\boldsymbol{S}(\boldsymbol{\psi_0}) = I_N - \boldsymbol{\Psi}_0 W$, and $\boldsymbol{y_t} = (y_{1t}, \ldots, y_{Nt})$. We stack the constant and all weakly exogeneous variables for i at t in $x_{it} = (1, \epsilon_{i,t-1}, \ldots, \epsilon_{i,t-L_1}, W\epsilon_{i,t-1}, \ldots, W\epsilon_{i,t-L_2})$. $\boldsymbol{x_t} = (x_{1t}', \ldots, x_{Nt}')'$ is the $((1 + L_1 + L_2) * N)$ by 1 vector. $\boldsymbol{B}$ is the N by $((1 + L_1 + L_2) * N)$ block diagonal matrix with elements $\boldsymbol{\beta}_i' = (a_{\epsilon_i}, \lambda_{1,i}, \ldots, \lambda_{L_1,i}, \psi_{1,i}, \ldots, \psi_{L_2,i})'$ on the main diagonal and zeros elsewhere. Finally, $Var(\boldsymbol{\upsilon}) = \Sigma$.

The quasi maximum likelihood estimator $\hat{\boldsymbol{\theta}}_{QMLE}$ maximizes (3). The error terms need not to be Gaussian. But when it is, $\hat{\boldsymbol{\theta}}_{QMLE}$ is the maximum likelihood estimator of $\boldsymbol{\theta}$. For further details of computationally cheaper estimation procedure and inference, interested reader could refer to Aquaro et al. (2019).

3.2.2 Spatial-temporal Responses to Local Risk

The spatial-temporal framework allows us to analyse a complicated diffusion pattern of local shocks over time and space. The parameter estimates of equation (2) only shows a part of the picture. To fully understand how $\epsilon_{i,t}$, a local shock arising from firm i at time t affects $\epsilon_{j,t+h}$, one need to trace the time profile of shocks over time and space. To examine the dependence across time and space implied by (2), we first rewrite it in a vector autoregression (VAR) form that we are familiar with:

$$\boldsymbol{\epsilon_t} = \sum_{\tau=1}^{max\{L_1,L_2\}} \Phi_\tau \boldsymbol{\epsilon_{t-\tau}} + R\boldsymbol{\upsilon_t} \quad (4)$$

where $R = (I_N - \Psi_0 W)^{-1}$, and $\Phi_\tau = R\Lambda_\tau + R\Psi_\tau W$. The lag order of the VAR depends on the number of AR lags and spatial-temporal lags in equation (2). Under the assumptions that $E(\boldsymbol{e_t}) = \mathbf{0}, E(\boldsymbol{\upsilon_t}\boldsymbol{\upsilon_t}') = \Sigma_\upsilon = \{\sigma_{ij}, i, j = 1, \ldots, N\}$, which is a positive definite matrix, $E(\boldsymbol{\upsilon_t}\boldsymbol{\upsilon_\tau}') = \mathbf{0}$ for $t \neq \tau$, and the stability of the process, the VAR can be as a vector moving average (VMA) process,

$$\boldsymbol{\epsilon_t} = \sum_{p=1}^{\infty} A_p \boldsymbol{\eta}_{t-p} \quad \text{for } t = 1, \ldots, T \quad (5)$$

where $\boldsymbol{\eta}_t = R\boldsymbol{\upsilon_t}$. A_p can be obtained recursively by

$$A_p = \Phi_1 A_{p-1} + \Phi_2 A_{p-2} + \cdots + \Phi_m A_{p-m} \quad (6)$$

where $m = max\{L_1, L_2\}$ and $A_0 = I_N$. Given that $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_N)$ is a hypothetical primitive shock hitting the economy at t, the generalized impulse response function (Pesaran and Shin (1998), Koop et al. (1996)) at horizon h is written as

$$GI(h, \boldsymbol{\delta}, \Omega_{t-1}) = E(\boldsymbol{\epsilon}_{t+h} \mid \boldsymbol{\upsilon_t} = \boldsymbol{\delta}, \Omega_{t-1}) - E(\boldsymbol{\epsilon}_{t+h} \mid \Omega_{t-1}) = A_h R \boldsymbol{\delta} \tag{7}$$

The primitive idiosyncratic shock $\boldsymbol{\upsilon_t}$ is allowed to be correlated. To look at the effect of a primitive shock from firm k on the whole system, we integrate out the effect of all other primitive shocks using the historically observed distribution of $\boldsymbol{\upsilon_t}$. The generalized impulse response function of the effect of a primitive shock from firm k at time t on the system h period in the future is given by

$$GI(h, \delta_k, \Omega_{t-1}) = E(\boldsymbol{\epsilon}_{t+h} \mid \upsilon_{k,t} = \delta_k, \Omega_{t-1}) - E(\boldsymbol{\epsilon}_{t+h} \mid \Omega_{t-1}) = \delta_k A_h R(\frac{\Sigma_\upsilon \boldsymbol{e_k}}{\sigma_{kk}}) \tag{8}$$

where $\boldsymbol{e_k}$ is a $N \times 1$ selection vector with 1 as its kth element and zeros elsewhere. $\frac{\Sigma_\upsilon \boldsymbol{e_k}}{\sigma_{kk}}$ is the adjustment due to potentially correlated primitive shocks $\boldsymbol{\upsilon_t}$. When Σ_υ is a diagnoal matrix, $\frac{\Sigma_\upsilon \boldsymbol{e_k}}{\sigma_{kk}} = \boldsymbol{e_k}$, and $GI(h, \delta_k, \Omega_{t-1})$ can be simplified to $\delta_k A_h R \boldsymbol{e_k}$.

4 Full Sample Estimation

In this section, I estimate the local risk spillovers in the weakly correlated idiosyncratic returns using the heterogeneous spatial-temporal model (2) discussed above. Using the estimated parameters, I then compute the spatial-temporal responses and construct the spatial-temporal spillover matrix D_h for each horizon h. Based on the spatial-temporal spillover matrices, we are able to find important systemic risk contributors and receivers. In the end, to assess the performance of the proposed method, I compare the in-sample and out-of-sample mean squared error (MSE) of the spatial-temporal model (2) estimated using alternative W and the high-dimensional vector autoregressive (VAR) model from Barigozzi and Hallin (2017).

4.1 De-factored (Idiosyncratic) Returns

Our full sample spans from 03/01/2006 to 31/12/2013 ($T = 2014$ days). To obtain a balanced panel, we end up with $N = 413$ stocks. We first estimate the hierachical factor model (1) by running time series regression for each company $i = 1, \ldots, N$.

Our analysis needs the number of members to be large within each sector group g,[8] so we adopt the six broad sectoral categories discussed in section 2. For the choice of $\mathbf{f}_t$, I use five Fama French factors (Fama and French (2015)) plus the momentum factor (Carhart (1997)). And as for the sectoral factors $\mathbf{f}_{g,t}$, I use the Fama French industry portfolios. As an alternative to the observed market and sectoral factors, one could use the hierarchical PCA and the statistical factors, which I will use it as a robustness check.

Table 1 summarizes the share of variance explained by the factors (regression R^2) for N cross-sections. The R^2 varies from 13.2% to as high as 77.2%. On average, these factors explain 49.1% of the variation of the excess returns of $S\&P500$ stocks and the R^2 is higher than 40% for three-fourths of the stocks.

[8]In this paper, we are particularly interested setoral heterogeneity, which we turn to Mean-Group estimation for each sector. The consistency of Mean-Group estimator requires large N within each group, which motivates the broad sector classification

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Hierarchical factor model	0.132	0.401	0.498	0.491	0.586	0.772

Table 1: Summary statistics for corss-sectional regression R^2 for the hierarchical factor model

The de-factored (idiosyncratic) returns can be obtained by estimating the above factor model (1) and collecting the residuals $\hat{\epsilon}_{it}$. Below I list some stylized features of the de-factored returns, which motivates our choice of the spatial-temporal modelling approach. First of all, $\hat{\epsilon}_{it}$ is serially correlated for around half of the sample $S\&P500$ stocks. Table 2 shows the summary statistics of the Q_m statistics of the Ljung-Box test and the corresponding P-value for the sample stocks. Four lag orders are considered. $m = 1, 5, 10, 22$ corresponds to the number of trading days in a day, a week, two weeks, and a month, respectively. If we consider the significance level $\alpha = 0.05$, we then reject the hypothesis of white noise for half of the stocks in the sample for all three lag orders except $m = 1$, as the P-value for the median is smaller than 0.05 for $m = 5, 10, 22$. This results shows that there are predictability in terms of estimated idiosyncratic returns. However, examining the estimates of the correlation coefficients (I will not report here) shows that correlations are in general very small economically, rendering the predictability unprofitable given the trading cost.

		Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
m=1	Q_m	0.001	0.451	2.512	5.806	6.050	104.088
	P-value	0.000	0.014	0.113	0.272	0.502	0.985
m=5	Q_m	1.000	6.501	11.183	15.864	18.014	130.446
	P-value	0.000	0.003	0.048	0.174	0.260	0.962
m=10	Q_m	2.951	13.122	19.308	25.598	27.570	189.538
	P-value	0.000	0.002	0.037	0.151	0.217	0.983
m=22	Q_m	11.48	28.19	37.67	47.81	52.30	273.00
	P-value	0.000	0.000	0.019	0.126	0.169	0.967

Table 2: Summary statistics of the Q_m statistics of the Ljung-Box test and the corresponding P-value for the sample stocks. Note: m is the lag order of the test. $Q_m = T(T-1)\sum_{j=1}^{m} \frac{1}{T-j}\hat{\rho_j}^2 \sim \chi^2_m$.

To choose the lag order L_1 for the autoregressive term, one could apply information criterion such as Akaike information criterion (AIC), Bayesian information criterion (BIC), etc. In this paper, since the estimation of a large heterogeneous spatial temporal model is time consuming, applying model selection techniques on the full model (2) is computationally burdensome, although theoretically possible. Thus, I pre-select the lag order L_1 of the model by examining the maximum number of lags included in the autoregressive (AR) model for each individual stock. I select the optimal number of autoregressive lags for each stock i using BIC criterion since AIC criterion usually selects a bigger model than BIC, and we hope to keep the model parsimonious given that the number of parameters need to be estimated is $N * (L_1 + L_2 + 3)$. [9] Among all sample stocks, 95% of them have optimal lag order smaller or equal to 5. according to that, I pre-specify $L_1 = 5$, which is the number of trading days in a week. The spatial temporal part is specified to have the same lag order $L_2 = 5$, and according to the estimation results they are sufficient to capture the spatial-temporal relationships.

[9] For each i, there are $L1$ AR parameters, $(L2 + 1)$ spatial temporal parameters, 1 intercept parameter and 1 scale parameter.

In addition to the temporal correlation, the cross-sectional dependencies in the idiosyncratic returns are of major interest. The de-factoring process removes the strong cross-sectional dependence by reducing the average pairwise correlation coefficient $\hat{\rho_N} = \frac{2}{N(N-1)} \sum_{i=1}^{N} \sum_{j=i+1}^{N} \hat{\rho}_{ij}$ from as large as $\hat{\bar{\rho}}_{N,r} = 0.4308$ to as small as $\hat{\bar{\rho}}_{N,\epsilon} = 0.008$. Then I go on to test the null of cross-sectionally uncorrelated idiosyncratic returns $H_0 = E(\epsilon_{it}, \epsilon_{jt}) = 0$ for all t and $i \neq j$. I compute a scaled version of Breusch and Pagan (1980) LM test statistics, which has an asymptotically standard normal distribution when N and T are both large.

$$CD_{LM} = \sqrt{\frac{1}{N(N-1)}} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} (T\hat{\rho}_{ij}^2 - 1) \tag{9}$$

Using $\hat{\epsilon}_{it}$ estimated from equation (1), $CD_{LM} = 1653.40$, which strongly rejects the null that idiosyncratic returns are cross-sectionally uncorrelated. Consistent with arbitrage pricing theory (APT), the interconnectedness in stock idiosyncratic returns, although being weak, is non-negligible and needs to be accounted for. Spatial models are natural tools for addressing local dependencies among neighbouring units, and with the novel business links constructed using text analysis, we can model the channel through which local shocks transmit and quantity its strength.

4.2 Adjacency Matrix

For the full sample estimation, W contains all the links that are identified within the sample period. Here is how we calculate W: (1) if there are T_m months in the sample, then for the tth $(1 \leq t \leq T_m)$ month in the sample, $W_t = (w_{ij,t})$ with $w_{ij,t}$ being a 1/0 dummy indicating whether company i and j are ever co-mentioned in the news items published in this month. We add up all the monthly observed adjacency matrix $W_1 + \cdots + W_t + \cdots + W_T$ to get a non-normalized adjacency matrix W_{raw}. (2) we row-normalize W_{raw}, as a convention in spatial econometrics, to get W. Notice that W_t is an unweighted adjacency matrix, while on the other hand W is weighted. This is because news tends to report the development of one issue for consecutive days and we may thus observe two companies getting co-mentioned several times within that period. This "multiple co-mentions within a short period of time" does not imply the relationship between two companies is stronger. However, if two companies get co-mentioned consistently in different monthly windows, there is reason to believe their links are stronger or the public are more aware/pay more attention to their links. That is why we add up unweighted monthly matrices W_t for $t = 1, \ldots, T_m$ and then apply row normalization to get weighted W. In section 6, I consider and compare various alternative specifications of W including different weighting schemes, and narrower definitions of links.

4.3 Spatial-temporal Model Estimation Results

4.3.1 Parameter Estimates

Equation (2) is estimated using quasi maximum likelihood (QML) and it is assumed that $\upsilon_{it} \sim IID(0, \sigma_i^2)$ for $i = 1, \ldots, N$. Since it is a heterogeneous coefficient model, we can only identify the spatial coefficients of those units with at least one link. We need to restrict the spatial parameters of companies without any links to zero.

If we apply the full sample adjacency matrix W discussed above, only 5 out of $N = 413$ companies don't have any links. We require large T for consistent estimation.

Given the huge amount of parameters in the model, here I only report some summary statistics of the estimates in Table 3.[10] For a heterogeneous coefficient panel model, what is often of the interest to empirical researchers is the average estimates across all entities (or all entities within a sub-group). If we assume that individual specific coefficients are randomly distributed around their common means as follows:

$$\begin{aligned}&\lambda_{k1,i} = \lambda_{k1} + \zeta_{k1,i}, \psi_{k2,i} = \psi_{k2} + \varsigma_{k2,i} \ \text{ for } k1 = 1, \ldots, L_1,\ k2 = 1, \ldots, L_2 \text{ and } i = 1, \ldots, N \\ &\boldsymbol{\eta}_i = (\boldsymbol{\zeta}_i', \boldsymbol{\varsigma'}_i')' \sim IID(\mathbf{0}, \Omega_\eta)\end{aligned} \tag{10}$$

The common mean parameters λ_{k1} and ψ_{k2} for $k1 = 1, \ldots, L_1$, $k2 = 1, \ldots, L_2$ are the the objects of interest and they can be consistently estimated with the following mean group (MG) estimator given N and T are large.[11] The mean group (MG) estimates are provided in Table 3 with standard errors in the parenthesis.

$$\hat{\lambda}_{k1}^{MG} = \frac{1}{N}\sum_{i=1}^{N}\hat{\lambda}_{k1,i} \ \ and \ \ \hat{\psi}_{k2}^{MG} = \frac{1}{N}\sum_{i=1}^{N}\hat{\psi}_{k2,i} \tag{11}$$

	(1)AR terms					(2) spatial-temporal terms						(3) σ
	λ_1	λ_2	λ_3	λ_4	λ_5	ψ_0	ψ_1	ψ_2	ψ_3	ψ_4	ψ_5	σ
Median	-0.026	-0.013	-0.013	-0.009	-0.004	0.252	0.024	0.010	-0.009	0.008	0.008	1.470
MG Estimates	-0.028	-0.015	-0.016	-0.010	-0.006	0.292	0.036	0.010	-0.007	0.007	0.011	1.561
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.021)	(0.007)	(0.006)	(0.005)	(0.005)	(0.006)	(0.028)
% Sig (at 5%)	38.7%	22.5%	19.9%	17.2%	19.4%	77.0%	21.1%	20.6%	15.2%	16.4%	15.7%	-
Non-zero coef.	413	413	413	413	413	408	408	408	408	408	408	413

Table 3: **QML estimation results of heterogeneous spatial-temporal model (2) using full sample.** Note: The median and mean group (MG) estimates are computed using unrestricted parameters only. The standard errors of the MG estimates are in the parenthesis. The second last row show the percentage of significant parameters at 5% out of the unrestricted parameters and the last row of the table shows the number of unrestricted entities out of N. Panel (1), (2), (3) report the results of autoregressive parameters, spatial-temporal parameters and standard deviations of errors, respectively.

$\hat{\psi}_0^{MG} = 0.292(0.021)$ shows that the contemporaneous local dependency is not only statistically significant but also economically significant. Among 408 unrestricted contemporaneous spatial coefficients $\psi_{0,i}$, 77% of them are individually significant. This high ratio implies we are successful at mining economic links among firms. If our text-based data contain a lot of spurious links, we will more likely to see the spatial parameters to be insignificant for many individuals. Some dynamic spatial terms are also statistically significant, although smaller in economic magnitude. Some general conclusions can be drawn here. After removing the common risk factors and sectoral risk factors there is still a considerable degree of local spatial-temporal risk spillover among $S\&P$ 500 firms.

[10] Full estimation results can be requested from the author.

[11] See Pesaran and Smith (1995) for proofs of the consistency when individual specific coefficients are independently distributed. Recent development by Chudik and Pesaran (2019) prove the consistency under weakly correlated individual specific estimators. In both cases, T and N are required to be big enough. Intuitively, big T is required for the consistent estimation of individual specific coefficients and N needs to be big enough for the consistent estimation of the means. To see how the MG estimators behave in the context of heterogeneous spatial-temporal model, see Aquaro et al. (2019).

It is reasonable to suspect that the mean sensitivities to local risk spillovers are different for different sectoral groups. To explore the sectoral heterogeneity, here we adopt the random coefficient assumptions at sector level.

$$\begin{aligned}
&\lambda_{k1,i,g} = \lambda_{k1,g} + \zeta_{k1,i,g}, \quad \psi_{k2,i,g} = \psi_{k2,g} + \varsigma_{k2,i,g} \\
&\text{for } k1 = 1, \dots, L_1,\ k2 = 1, \dots, L_2 \text{ ,and } i = 1, \dots, N,\ g = 1, \dots, G \\
&\boldsymbol{\eta}_{i,g} = (\boldsymbol{\zeta}'_{i,g}, \boldsymbol{\varsigma}'{}'_{i,g})' \sim IID(\mathbf{0}, \Omega_\eta)
\end{aligned} \tag{12}$$

Given that the consistency of mean group estimator requires large N, one consideration when doing sector classification is that the number of members of each group needs to be sufficiently big. Thus, I adopt a broad sector classification scheme described in section 4.1, which guarantees large N condition to be satisfied for each sector. Table 4 presents the estimation results grouped by sector. It reveals the considerable level of heterogeneity among different sectors. In particular, the size of mean contemporaneous spatial effect for manufacturing firms is largest, with $\hat{\psi}^{MG}_{0,manufacturing} = 0.446(0.033)$. Manufacturing firms are closely connected with other firms via supplier-customer linkages, and it is well documented (eg. Cohen and Frazzini (2008)) that a shock to one firm has sizeable effects on its linked partners along the supply chain. Financial firms are also exposed to quite a large contemporaneous spatial effect with $\hat{\psi}^{MG}_{0,finance} = 0.345(0.039)$. Apart from the large contemporaneous spatial coefficient, it is also worth noticing that the lead-lag effect in risk spillovers for financial firms is more pronounced as the the percentage of significant spatial-temporal coefficients $\hat{\psi}_{k_2,i,finance}$ is about twice as large as that of other sector groups for any lag order $k2 = 1, \dots, 5$. We need to interpret the mean group estimates of these spatial-temporal parameters with care. The individual parameters $\hat{\psi}_{k_2,i,finance}$ are quite dispersed, with some firms having significantly positive spatial temporal terms and some having significantly negative ones. That is why $\hat{\psi}_{k2,finance}$ for $k2 = 2, 3, 4$ are not statistically significant although high percentages of individual coefficients are significant —there is too much heterogeneity! Firms from the consumer sector and the hi-tech sector are also significantly exposed to the local risks of their economic neighbours, although with slightly smaller sensitivities. Health firms are least sensitive to shocks elsewhere and the mean group estimate $\hat{\psi}^{MG}_{0,health} = 0.061(0.061)$ is not statistically significant. However, we should interpret that result with care since the number of companies from the health sector is relatively small thus the mean group estimate is likely to be imprecise.

4.3.2 Spatial-temporal Responses to Local Shocks

For any horizon h, we can summarize the own response and cross-response implied by equation (8) in a similar way that LeSage (2008) and LeSage and Chih (2016) summarize direct and indirect partial effects of a change in the kth explanatory variable. As an illustration, consider a simple example where Σ_v is diagonal, and firm k receives a unit shock at time t, equation (8) can be simplified as $GI(h, \delta_k = 1, \Omega_{t-1}) = A_h R \boldsymbol{e_k}$. $A_h R$ is a $N \times N$ matrix with N own responses and $N(N-1)$ cross-responses at horizon h on the diagonal and off-diagonal, respectively. For $h = 0, A_0 = I_N$,

$$A_h R = R = (I_N - \Psi_0 W)^{-1} = I_N + \Psi_0 W + \Psi_0^2 W^2 + \Psi_0^3 W^3 + \dots \tag{13}$$

R is an infinite series expansion that adds the own effect I_N, first order neighbour effect $\Psi_0 W$, second order neighbour effect $\Psi_0^2 W^2$ and so on. Ψ_0 is a diagonal matrix that every entry is upper-bounded by 1 in absolute

	(1)AR terms					(2) spatial-temporal terms						(3) σ
	λ_1	λ_2	λ_3	λ_4	λ_5	ψ_0	ψ_1	ψ_2	ψ_3	ψ_4	ψ_5	σ
Panel A: Consumer												
Median	-0.020	-0.015	-0.009	-0.009	-0.008	0.236	0.035	0.013	-0.004	0.001	0.001	1.373
MG Estimates	-0.025	-0.015	-0.011	-0.012	-0.008	0.232	0.033	0.026	-0.001	-0.002	0.005	1.456
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.039)	(0.009)	(0.011)	(0.010)	(0.010)	(0.012)	(0.054)
% Sig(at 5%)	29.9%	15.6%	16.9%	18.2%	18.2%	79.2%	15.6%	11.7%	14.3%	11.7%	7.8%	-
Non-zero coef.	77	77	77	77	77	77	77	77	77	77	77	77
Panel B: Finance												
Median	-0.037	-0.013	-0.017	-0.014	-0.005	0.350	0.017	0.000	-0.018	0.001	0.033	1.616
MG Estimates	-0.039	-0.020	-0.021	-0.024	-0.005	0.345	0.056	-0.010	-0.018	0.023	0.050	1.785
	(0.008)	(0.006)	(0.005)	(0.005)	(0.005)	(0.057)	(0.026)	(0.019)	(0.020)	(0.017)	(0.017)	(0.073)
% Sig(at 5%)	57.3%	38.7%	36.0%	32.0%	33.3%	82.7%	32.0%	34.7%	30.7%	30.7%	29.7%	-
Non-zero coef.	75	75	75	75	75	74	74	74	74	74	74	75
Panel C: Health												
Median	-0.007	-0.010	-0.004	0.001	0.008	0.119	0.024	-0.004	0.014	0.027	0.004	1.368
MG Estimates	-0.014	-0.005	-0.010	-0.005	0.006	0.061	0.020	0.001	-0.001	0.029	0.041	1.479
	(0.006)	(0.005)	(0.005)	(0.005)	(0.004)	(0.061)	(0.016)	(0.015)	(0.013)	(0.016)	(0.020)	(0.105)
% Sig(at 5%)	25.7%	20.0%	14.3%	14.3%	8.6%	68.6%	14.3%	11.4%	8.6%	5.7%	14.7%	-
Non-zero coef.	35	35	35	35	35	34	34	34	34	34	34	35
Panel D: Hitech												
Median	-0.036	-0.019	-0.014	-0.009	-0.004	0.212	0.016	-0.004	0.006	0.009	-0.013	1.459
MG Estimates	-0.032	-0.018	-0.012	-0.010	-0.007	0.229	0.018	-0.004	-0.001	0.004	-0.014	1.576
	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)	(0.048)	(0.011)	(0.013)	(0.009)	(0.010)	(0.014)	(0.062)
% Sig(at 5%)	49.3%	19.2%	8.2%	13.7%	16.4%	72.6%	11.0%	13.7%	6.8%	12.3%	11.0%	-
Non-zero coef.	73	73	73	73	73	73	73	73	73	73	73	73
Panel E: Manufacturing												
Median	-0.011	-0.004	-0.018	-0.001	-0.005	0.468	0.022	0.028	-0.011	0.000	0.005	1.249
MG Estimates	-0.019	-0.005	-0.017	-0.002	-0.010	0.446	0.032	0.018	-0.008	0.004	0.005	1.303
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.033)	(0.009)	(0.008)	(0.008)	(0.007)	(0.007)	(0.041)
% Sig(at 5%)	30.0%	20.9%	17.3%	20.9%	20.0%	85.5%	18.2%	18.2%	13.6%	16.4%	13.0%	-
Non-zero coef.	110	110	110	110	110	108	108	108	108	108	108	110
Panel F: Other												
Median	-0.036	-0.015	-0.013	-0.007	-0.002	0.227	0.045	-0.017	-0.005	0.013	-0.017	1.488
MG Estimates	-0.031	-0.016	-0.020	-0.007	-0.001	0.315	0.072	0.010	-0.019	-0.002	-0.007	1.635
	(0.007)	(0.005)	(0.005)	(0.004)	(0.004)	(0.075)	(0.024)	(0.019)	(0.016)	(0.017)	(0.017)	(0.083)
% Sig(at 5%)	46.5%	23.3%	27.9%	7.0%	11.6%	76.7%	32.6%	20.9%	9.3%	9.3%	11.9%	-
Non-zero coef.	43	43	43	43	43	42	42	42	42	42	42	43

Table 4: **QML estimation results of heterogeneous spatial-temporal model (2) using full sample, parameters summarized by sector.**

value, so that higher powers of Ψ_0 assigns smaller impact to higher order neighbours. The main diagonal elements of R gives the own responses to a unit shock, which is in general different from 1 since they are the sums of own effects and feedback from others. The off-diagonal elements of R, on the other hand, are the sums of neighbour effects of different orders. For $h >= 1$, $A_h R$ gives the combination effects of spatial dependence and temporal dependence. In general, when Σ_v is not diagonal, we need to adjust for correlated $\boldsymbol{v_t}$ using equation (8).

For each horizon h, I compute the $N \times 1$ vector $GI(h, \delta_k = 1, \Omega_{t-1})$ for each $k = 1, \ldots, N$ using the estimated parameters. For the diagonal matrices $\Lambda_k, k = 1, \ldots, 5$, the i's diagonal element is $\lambda_{k,i}$ if it is statistically significant at 5% level, otherwise it is replaced by zero. The same is true for the construction of $\Psi_k, k = 1, \ldots, 5$. We denote the spatial-temporal spillover matrix at h as D_h, where $GI(h, \delta_k = 1, \Omega_{t-1})$ is the kth column of it. $D_h = [d_{ij}^h]$ gives the pairwise directional spillovers at horizon h.

Since N is large in our analysis, it is not feasible to report spillovers at the pairwise level. I adopt the scalar summary measure used in LeSage (2008) and LeSage and Chih (2016). For each horizon h, I derive individual level own response, which is the diagonal elements of D_h. As for the individual level indirect effect, two measures are used, which are in-degree (C^h_{in}) and out-degree(C^h_{out}). They are defined as follows:

$$C^h_{i,in} = \sum_{j \neq i}^{N} d^h_{ij} \tag{14}$$

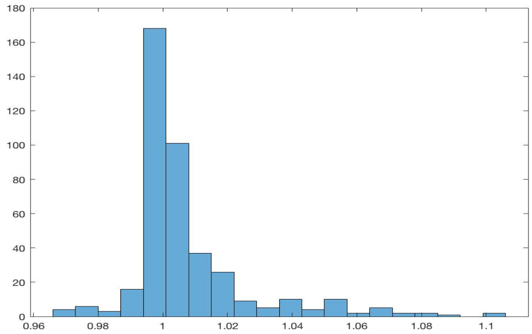
$$C^h_{j,out} = \sum_{i \neq j}^{N} d^h_{ij} \tag{15}$$

The in-degree measures the shocks a firm receives from other firms, and the out-degree, on the other hand, measures the shocks a firm spreads to others.

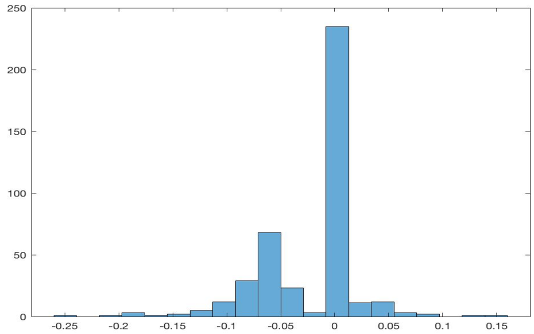
Figure 1 plots the histogram for own response, in-degree and out-degree at horizon $h = 0, 1$. The figures for further horizon are in Figure 11. The two sub-figures on the first row correspond to the contemporaneous responses. When a firm receives one unit primitive shock at t, its contemporaneous own response it not necessarily 1 as the result of the complicated feedback relationships. There are stark differences between two indirect effect measures if we compare the two graphs on the second row with the two graphs on the third row. For $h = 0$, while there is a small proportion of firms that respond negatively to neighbours' shocks, almost all firms are positive spreader of risks (in graph (e), there is only a tiny bin with negative out-degree). Also, it is worth noticing that the out-degree has a heavy right tail, with some companies contributing a lot of risk to the system. The right column with $h >= 1$ corresponds to the dynamic responses, which combine the effects of both temporal and spatial dependencies. From Table 3 and Table 4, the estimates of dynamic parameters (both the pure temporal and the spatial-temporal) are small relative to the contemporaneous spatial effect parameter, which is reflected in Figure 1. Local shocks travel over time and space with decays. One interesting feature is that the decay along the spatial dimension is slower than that along the time dimension. It is noticeable that the current analysis focuses on how a unit shock to one firm affects the system, which explains why the shocks die out quickly. However, in financial crisis episodes where a larger number of firms receive negative shock jointly, local shocks could have a larger and more long-lasting effect.

Firms with high in-degree are vulnerable as they are particularly sensitive to shocks elsewhere and firms with high out-degree are dangerous since their own primitive shocks are widespread. Therefore, it is of interest from a microprudential (firm-specific) perspective to identify these two types of firms. Table 5 shows the 20 firms with the highest in-degree and out-degree for $h = 0, 1$.[12] The firms that contribute the most to the systemic risks are the large cap financial institutions and manufacturers, and the findings are reasonably in line with the systemic risk contributors found in others including Hautsch et al. (2015), Barigozzi and Hallin (2017), Barigozzi and Brownlees (2019). Apart from systemic risk contributors, companies that are particularly sensitive to others' shocks are also found. It is worth noticing that the well-connected systemic risk contributors themselves are not necessarily the major risk receivers; rather these are the periphery firms that receive a lot of risks from the core.

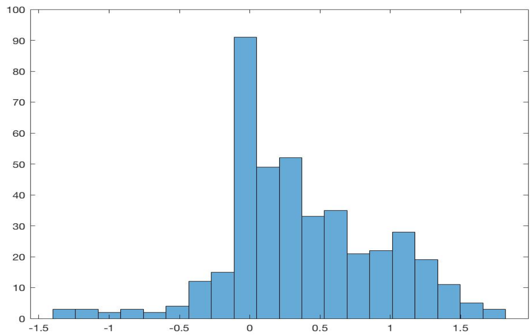
[12] Higher order results are not shown in the main text since the shocks decay along time dimension quickly.



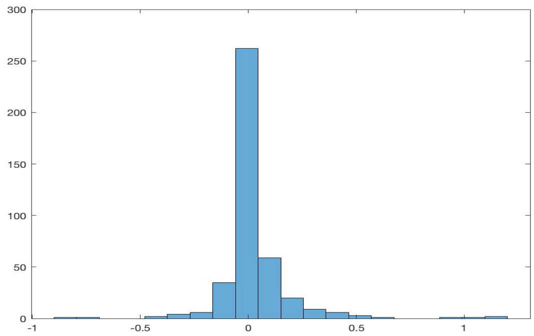
(a) Own response ($h = 0$)



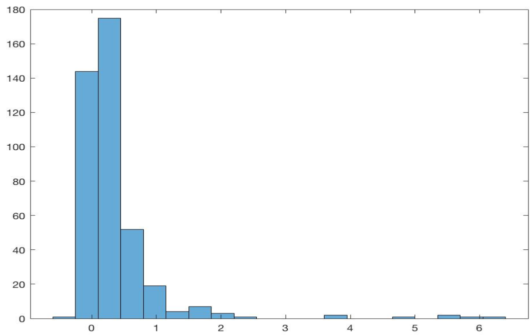
(b) Own response ($h = 1$)



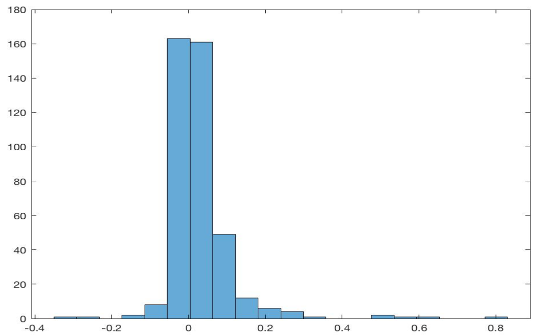
(c) In-degree ($h = 0$)



(d) In-degree ($h = 1$)



(e) Out-degree ($h = 0$)



(f) Out-degree ($h = 1$)

Figure 1: Histogram for own response, in-degree and out-degree at horizon $h = 0, 1$.

		Company Ticker
h=0	In-degree	LEN, EIX, PCG, DUK, DHI, NOC, GD, RIG, RTN, LNC,
		ETR, ALTR, SO, LRCX, CSX, PNW, UNH, HBAN, PFG, POM
	Out-degree	BAC, MSFT, GE, GS, JPM, XOM, C, CVX, LNC, WFC
		APPL, USB, BA, FITB, VZ, JNJ, PG, AET, UNH, PFE
h=1	In-degree	GNW, FITB, HBAN, GE, WY, STT, LEN, LNC, CI, COF
		FLR, PG, ATI, AES, RIG, JEC, PH, CAG, HD, HUM
	Out-degree	BAC, MSFT, GE, GS, JPM, XOM, LM, C, CVX, LNC,
		DUK, WFC, APPL, USB, BA, FITB, BZ, JNJ, PG, HCP

Table 5: The 20 firms with highest in-degree and out-degree for $h = 0, 1$.

4.4 Comparisons with alternative methods

To assess the performance, I compare the in-sample and out-of-sample mean squared error (MSE) of the spatial-temporal model (2) estimated using different adjacency matrices Ws considered in previous studies and the high-dimensional vector autoregressive (VAR) model from Barigozzi and Hallin (2017). The comparison results are shown in Table 6.

The first column of the table corresponds to the benchmark Naive estimator where the predicted de-factored returns are zero all the time. The second column presents the results of the high-dimensional vector autoregressive (VAR) model from Barigozzi and Hallin (2017), or BH-VAR for short. Column three to column seven present the results of the spatial-temporal model (2) estimated using five alternative adjacency matrices. The first candidate W is the empty adjacency matrix where there are no links. Fan et al. (2016) document a a sector-based block diagonal pattern in the factor model residual covariance of $S\&P$ 500 stocks. Inspired by that, our second candidate W is the sectoral network where within each sector, companies are completely connected, and there are no inter-sector links. The third candidate W is the geographic network motivated by Pirinsky and Wang (2006), and Parsons et al. (2020). For $W_{geographic}$, companies whose headquarters are in the same Metropolitan Statistical Area (MSA) are completely connected. The fourth candidate W is the Compustat customer-supplier network. The fifth W is the news-based network.

The spatial-temporal model (2) allows cross-sections to have heterogeneous coefficients. While the highly flexible model promises a better in-sample fit, some might suspect the model does not guarantee a better out-of-sample fit as a result of potential over-fitting. To examine the above issue, for each candidate W, I compute the in-sample and out-of-sample MSE using three alternative specifications, given by row (1)-(3) of each panel. Row (1) corresponds to the heterogeneous coefficient model. Row (2) corresponds to the setoral heterogeneous coefficient model where companies within the same sector have homogeneous coefficients. Row (3) corresponds to the homogeneous coefficient model. The training sample spans from 03/01/2006 to 31/12/2013 (2014 days) and the testing sample spans from 03/01/2014 to 31/12/2014 (252 days).

	Naive	BH-VAR	W_{empty}	W_{sector}	$W_{geographic}$	$W_{compustat}$	W_{news}
In Sample MSE							
(1)Heterogeneous coef	-	-	2.907	**2.829**	2.876	2.903	**2.764**
(2)Sectoral-heterogeneous coef	-	-	2.912	2.902	2.921	2.929	2.863
(3)Homogeneous coef	-	-	2.804	2.918	2.920	2.926	2.865
(4)	2.935	**2.211**	-	-	-	-	-
Out-of-Sample MSE							
(1)Heterogeneous coef	-	-	1.353	1.332	1.371	1.353	**1.287**
(2)Sectoral-heterogeneous coef	-	-	1.350	1.336	1.368	1.347	**1.302**
(3)Homogeneous coef	-	-	1.351	1.338	1.370	1.348	**1.309**
(4)	1.348	1.423	-	-	-	-	-

Table 6: In-sample and out-of-sample MSE (in basis point) of alternative models. Note: for each panel, the best 3 (smallest MSE) cases are in bold.

In terms of in-sample fit, BH-VAR has the smallest MSE. This is not surprising, given the method selects the model by minimizing the Bayesian information criterion. The heterogeneous coefficient spatial-temporal model with news-based network and sectoral network rank second and third, respectively. However, when we look at out-of-sample fit, BH-VAR loses its advantage with its MSE being even larger than that of the Naive predictor. The spatial-temporal model with news-based network, under any of the three parameter heterogeneity assumptions, outperforms the rest of the specifications.

The strength of local risk spillovers via news-based linkages exhibits a high level of heterogeneity. As a result, the heterogeneous coefficient specification improves not only the in-sample fit but also the out-of-sample fit. Although the spatial-temporal model with news-based network underperforms the BH-VAR model in term of in-sample fit, but it beats the BH-VAR when we compare the out-of-sample fit, which is more important practically. It is also worth noticing that W_{news} beats all other alternative Ws in terms of both in-sample and out-of-sample fit. Comparing with the existing linkage datasets that each only cover a particular type of relatedness, our text-based links provide an integrated measure of relatedness. And comparing with the high-dimensional VAR, the textual analysis approach identifies persistent economically-meaningful links. These reasons explain why our method has superior performance. To evaluate the robustness of superior performance, I conduct the model comparison for different sub-samples the results are shown in Table 16. The superior performance of our method is robust across all sub-samples.

5 Dynamic Estimation

Equity returns comovement reflects both strong and weak cross-sectional dependence. It has been documented that asset returns depart from fundamentals during times of financial crisis and stocks dis-connect from the market factor (see Bailey et al. (2019), Bailey et al. (2020)). Our two-stage factor plus spatial approach captures both sources of comovement separately and thus provides an avenue to examine how weak cross-sectional dependence evolves over time. In this section, I consider a rolling window analysis with 251-day (the average number of trading days in a year) rolling samples from 03/01/2006 to 31/12/2013. In total, there are 1761 windows.

5.1 Time Evolution of Weak Cross-sectional Dependence

The composition of $S\&P$ 500 index changes periodically in response to acquisitions and the growth or shrinkage of company values. We update the list of sample companies on a yearly basis and include the securities that stay in the $S\&P$ 500 list that have no missing observations for that year. On average, there are 447 stocks on the list for each update. Then we use a rolling estimation with a 251-day window to gauge the time variations in local dependencies. For the estimation window $[t, t+251]$, we conduct the two-stage procedure. The W_t used for the estimation of the spatial-temporal model is constructed using all the news published one year during the year. In the end, 1761 sets of estimates are obtained.

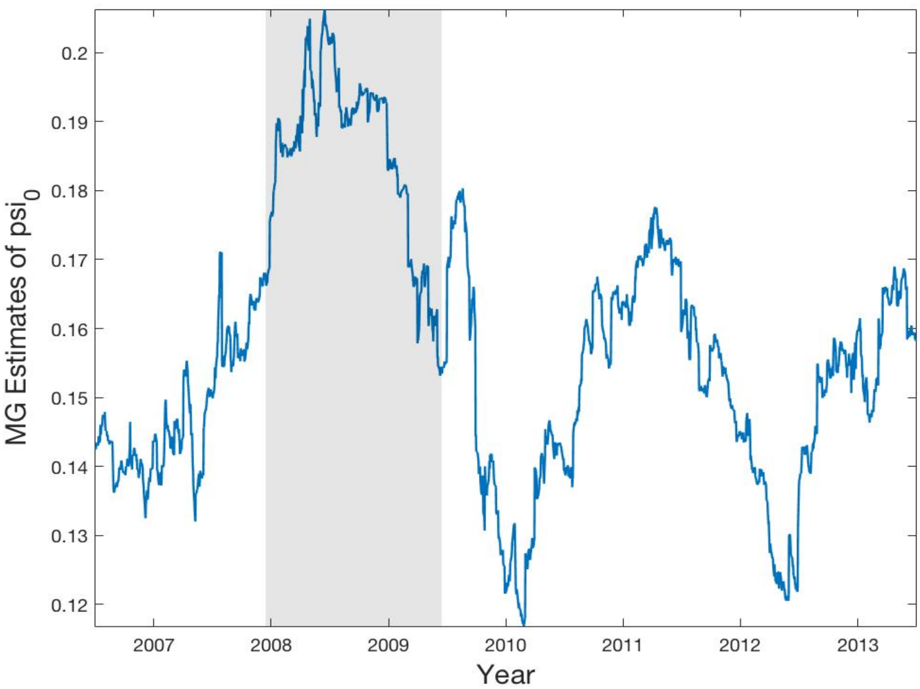
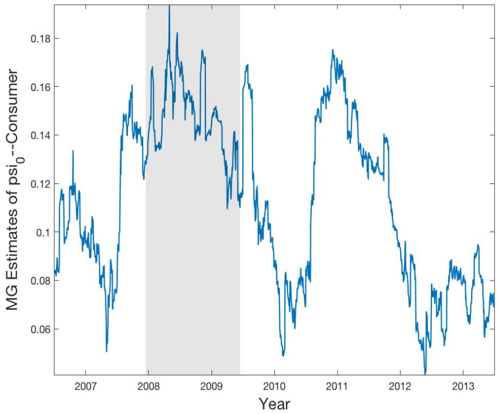
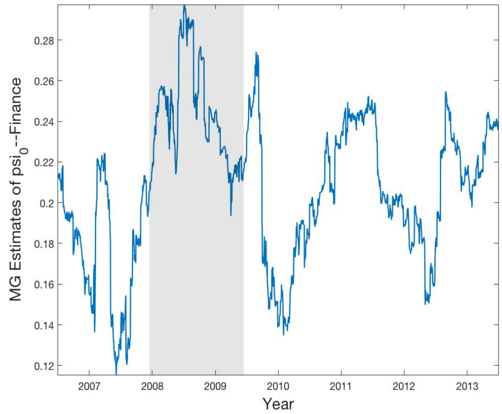


Figure 2: 251-day rolling $\psi_{0,t}^{MG}$ from 03/01/2006 to 31/12/2013. The Grey bar is the NBER recession indicator. Note: for window $[t-125, t+126]$, we use the middle date of the window to denote the the mean group estimate $\hat{\psi}_{0,t}^{MG}$, which explains why the x axis spans from 30/06/2006 to 01/07/2013.

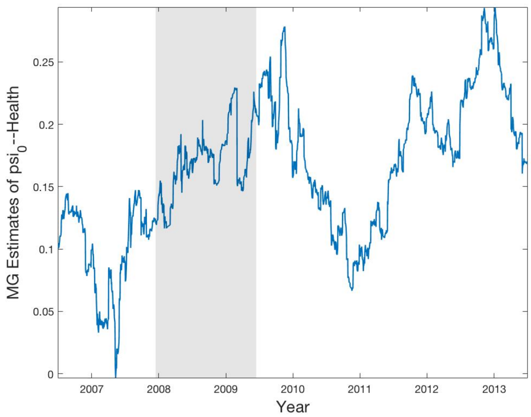
Figure 2 plots $\hat{\psi}_{0,t}^{MG}$, the 251-day rolling mean group estimates of the the contemporaneous spatial parameter. For the window $[t-125, t+126]$, the mean group estimate of the contemporaneous spatial parameter is calculated as $\hat{\psi}_{0,t}^{MG} = \frac{1}{N}\sum_{i=1}^{N}\hat{\psi}_{0,i,t}$. 1761 rolling samples from 03/01/2006 to 31/12/2013 give rise to 1761 sets of estimates from 30/06/2006 to 01/07/2013. The figure reveals the increase in the intensities of local dependencies during times of financial turmoil. $\hat{\psi}_{0,t}^{MG}$ was low in the 2006 and early 2007, and has increased gradually since 2007 following the liquidity crisis. By the end of the year, the public started to be more aware that big US banks might write off a huge amount of losses and that a global financial crisis is unfolding. $\hat{\psi}_{0,t}^{MG}$ skyrocketed afterwards, peaking around the time of the bankruptcy of Lehman Brothers. Months after, with a massive direct capital injection by the US government, the market calmed down and $\hat{\psi}_{0,t}^{MG}$ gradually recovered to the pre-crisis level. Instead of staying low, the several waves of the European Debt Crisis raised $\hat{\psi}_{0,t}^{MG}$ again, although by a smaller magnitude.



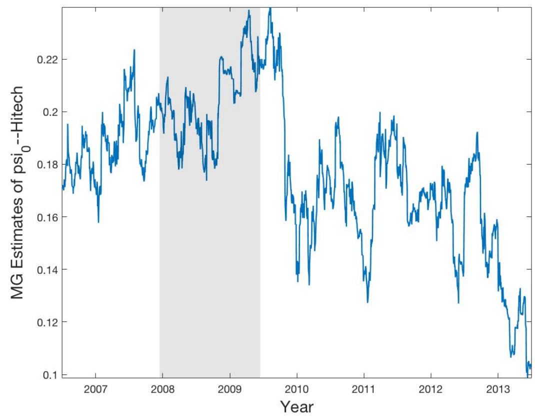
(a) Consumer



(b) Finance

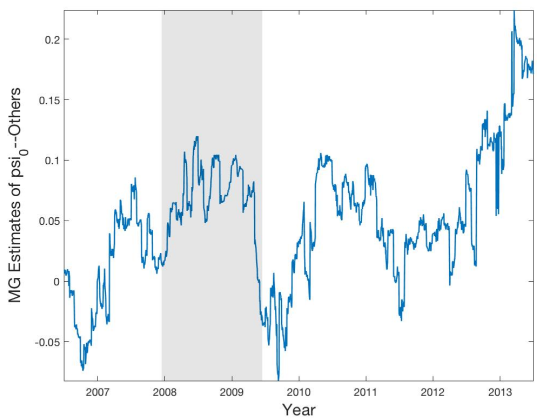


(c) Health



(d) Hitech

(e) Manufacturing



(f) Other

Figure 3: 251-day rolling $\psi_{0,g,t}^{MG}$ from 03/01/2006 to 31/12/2013 for different sectors.

To examine the sectoral heterogeneity of time variations in the strength of local spillovers, I plot the rolling mean group estimates of the contemporaneous spatial parameter for different sectors in Figure 3. For each window, the mean group estimate of the contemporaneous spatial parameter for sector g is calculated as the sample average of individual specific contemporaneous spatial parameter from that sector at t, namely, $\hat{\psi}^{MG}_{0,g,t} = \frac{1}{N}\sum_{i\in g}\hat{\psi}_{0,i,t}$. The time series pattern of $\hat{\psi}^{MG}_{0,g,t}$ shows a considerable degree of heterogeneity. Table 7 presents the correlation coefficients matrix of $\hat{\psi}^{MG}_{0,t}$ and $\hat{\psi}^{MG}_{0,g,t}$ for g = Consumer, Finance, Health, Hi-tech, Manufacturing and Others. $\hat{\psi}^{MG}_{0,consumer,t}$, $\hat{\psi}^{MG}_{0,finance,t}$ and $\hat{\psi}^{MG}_{0,manufacturing,t}$ exhibit similar patterns and they all have two obvious humps around the time of the Great Financial Crisis and the European Debt Crisis episodes. While hi-tech sector also experienced a rise in local risk spillovers during the Great Financial Crisis, it was not overly affected by the European Debt Crisis. Health care stocks belong to the non-cyclical group and $\hat{\psi}^{MG}_{0,health,t}$ moves in opposite directions with others.

	$S\&P500$	Consumer	Finance	Health	Hitech	Manufacturing	Other
$S\&P500$	1	0.77	0.79	0.04	0.39	0.65	0.3
Consumer	0.77	1	0.52	-0.18	0.45	0.5	-0.11
Finance	0.79	0.52	1	0.18	0.14	0.25	0.28
Health	0.04	-0.18	0.18	1	-0.16	-0.48	0.21
Hitech	0.39	0.45	0.14	-0.16	1	0.21	-0.46
Manufacturing	0.65	0.5	0.25	-0.48	0.21	1	0.14
Other	0.3	-0.11	0.28	0.21	-0.46	0.14	1

Table 7: Correlation coefficients of $\hat{\psi}^{MG}_{0,t}$ and $\hat{\psi}^{MG}_{0,g,t}$ for g =Consumer, Finance, Health, Hi-tech, Manufacturing and Others.

5.2 Time Evolution of Market Factor Strength

While weak CSD intensifies during periods of financial crisis and turmoil, strong CSD, as documented in Bailey et al. (2019) and Bailey et al. (2020), loses its power. According to asset pricing theories like the capital asset pricing model (CAPM), all stocks should load significantly on market factor. In these papers, they propose an estimator of factor strength based on the number of statistically significant factor loadings, taking account of the multiple testing problem. For a factor model with $\mathbf{f}_t = (f_{1t}, \ldots, f_{kt})$ being the vector of factors.

$$r_{it} - r_{ft} = \alpha_i + \mathbf{b}_i'\mathbf{f}_t + \epsilon_{it} \text{ for } i = 1, \ldots, N \tag{16}$$

Their proposed an estimator of the factor strength for the jth factor $\hat{\alpha_j}$, which is calculated as

$$\hat{\alpha_j} = 1 + \frac{log(\hat{D}_j/N)}{log(N)} \text{ if } \hat{D}_j > 0 \tag{17}$$

where $\hat{D}_j$ is the total number of statistically significant loadings of factor j out of N cross-sectional regressions. The critical value of the test is adjusted for the multiple testing problem.

According to capital asset pricing model (CAPM), the market factor is a strong factor and all stocks load significantly on the market factor as the number of stocks N grows large. This implies that the market factor have $\alpha_{market} = 1$. I re-do their exercise and conduct a rolling estimation of α_{market}. Figure 4 plots the rolling estimate of the strength of the market factor. The time series is more volatile than that in Bailey et al. (2020) since I am using daily 251-day rolling windows while they are using monthly 10-year rolling windows. As is found in their work, the market factor is pretty strong with its strength being very close to 1 all the time except for a short period during the Great Financial Crisis. This result, together with the time series patterns of the local risk spillovers shows that the strength of strong and weak CSD tend to move in opposite directions. The correlation coefficient of $\hat{\psi}_{0,t}^{MG}$ and $\hat{\alpha}_{market}$ is -0.6. When the market factor loses its importance during the financial crisis, weak cross-sectional dependence gains its power with the strength of local risk spillovers increasing.

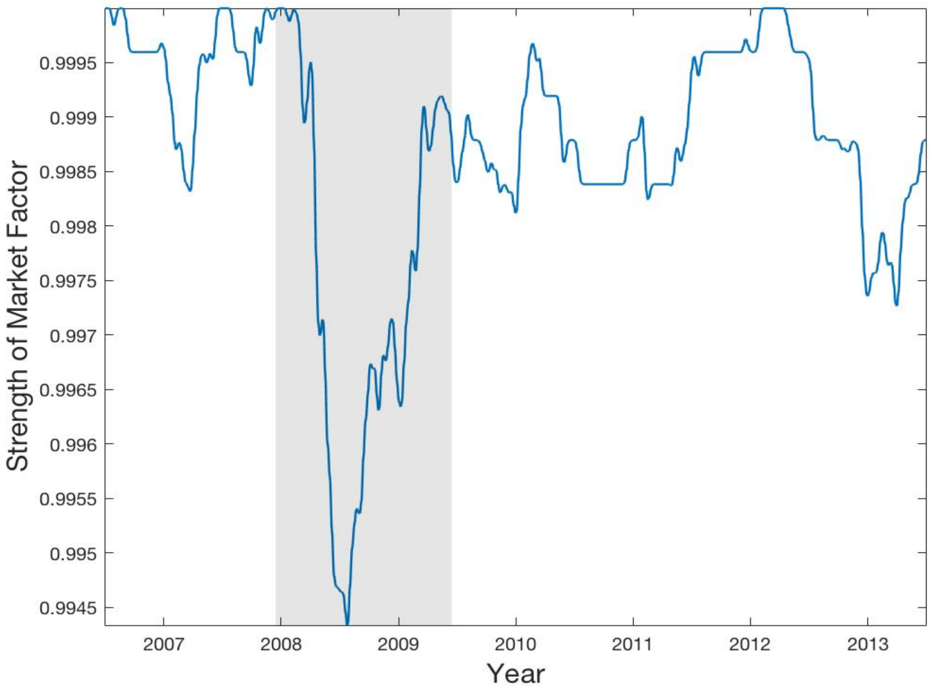


Figure 4: 251-day rolling of α_t from 03/01/2006 to 31/12/2013. Note: The factor strength parameter α is calculated as in Baily et al. (2020).

6 Robustness Check and Placebo Test

In this section, I test whether the results are sensitive to de-factoring procedures and specifications of the adjacency matrix W.

6.1 Robustness Check

As a first robustness check, I de-factor by using unobserved factors instead of observed ones. A hierarchical principal components (PCA) procedure is applied to remove both principal components at the market level and the sector level. Such a hierarchical can be written as:

$$r_{it} - r_{ft} = \alpha_i + \mathbf{b}_i'\mathbf{f}_t + \gamma_i'\mathbf{f}_{g,t} + \epsilon_{it} \tag{18}$$

where $\mathbf{f}_t$ is the vector of market factors that affect every stock i and $\mathbf{f}_{g,t}$ is the vector of sectoral factors that affect every stock in the sector g. Applying the information criteria in Bai and Ng (2002), we select the five market principal components and one sectoral principal component. I then run regression (18) and obtained the residuals $\hat{\epsilon}_{it}^{pca}$. Table 17 summarizes the regression R^2 for N cross sections. On average, unobserved factors have slightly better explanatory power. Estimating the heterogeneous spatial-temporal model (2) using $\hat{\epsilon}_{it}^{pca}$, the estimates of spatial-temporal parameters are smaller than that from Table 3 and Table 4, given that the hierarchical PCA removes a larger degree of commonality in the first stage. The estimation results using $\hat{\epsilon}_{it}^{pca}$ are given in Table 18 and Table 19.

Next, I test whether our results are valid under alternative specifications of the adjacency matrix W, sticking with the hierarchical observed factor model as our first stage. The in-sample and out-of-sample fit of these different specifications are shown in Table 8, where panel (1) considers several alternative adjacency matrices Ws under the one-W specification as in equation (2) and panel (2) considers the two-W specification as in equation (19) described below. The estimation results of alternative specifications are shown in Table 20, Table 21, Table 22, Table 23 and Table 24 from the Appendix.

In the baseline specification $W_{baseline}$ described in section 4.2, I add up monthly unweighted adjacency matrices W_t to account the "multiple co-mentions within a short period of time" issue. Row 2 to row 3 in panel (1) of Table 8 consider two alternative weighting schemes. I firstly consider an equal weight adjacency matrix $W_{unweighted}$, where entry w_{ij} is either $1/N_i$ or 0, depending on whether i and j are ever co-mentioned through the whole sample and N_i is the number of neighbours i has. Then I consider $W_{weighted}$, where its entries are weighted by the number of common news that link the two companies through the whole sample. Both $W_{unweighted}$ and $W_{weighted}$ underperform $W_{baseline}$, showing that if two companies get co-mentioned consistently in different monthly windows, there is reason to believe their links are stronger or the public are more aware/pay more attention to their links. However, frequent co-mentions within monthly windows does not indicate a stronger relationship as news tends to report the development of one issue for consecutive days.

Row 4 to row 7 in panel (1) of Table 8 investigate narrower definitions of links. For row 4, I remove competitor links by throwing away news that contain variations of the keyword "competition" in the headlines. All the non-competitive links are stored in $W_{noncompetitive}$ and the number of identified pairs is reduced by 3%. For row 5, I remove transitory links by keeping links that get identified in at least two different monthly windows. All the persistent links are stored in $W_{persistent}$ and the number of identified pairs is reduced by 18%. Both $W_{noncompetitive}$ and $W_{persistent}$ perform slightly worse than $W_{baseline}$. For row 6, $W_{intersector}$ is constructed by removing linked pairs in the same sector. For row 7, $W_{interindustry}$ is constructed by removing linked pairs in the same industry, which is defined by the four-digit Standard Industrial Classification (SIC) code. Both $W_{intersector}$ and $W_{interindustry}$ have the worset performance in panel (1), showing that intra-sector and intra-industry links contain very important information and can not be discarded. The estimation results using the mentioned alternative Ws are presented in Table 20.

To investigate the different role played by inter and intra-sector /industry links, in Panel (2) of Table 8 I consider a two-W specifications as follows:

$$\boldsymbol{\epsilon}_t = \mathbf{a}_\epsilon + \sum_{k=1}^{L_1} \boldsymbol{\Lambda_k} \boldsymbol{\epsilon}_{t-k} + \sum_{k=0}^{L_2} \boldsymbol{\Psi_{1,k}} W_1 \boldsymbol{\epsilon}_{t-k} + + \sum_{k=0}^{L_2} \boldsymbol{\Psi_{2,k}} W_2 \boldsymbol{\epsilon}_{t-k} + \boldsymbol{\upsilon}_t \tag{19}$$

where W_1 and W_2 are different adjacency matrices for different risk transmission channels. Fan et al. (2016) documents a sector-based block diagonal pattern in the factor model residual covariance of $S\&P500$ stocks. To show that the W_{news} contains additional information, I let W_1 be $W_{intersector}/W_{interindustry}$ defined above, and W_2 be sector/industry based block diagonal matrix. The estimation results using $W_1 = W_{intersector}$ and $W_2 = W_{sector}$ is presented in Table 21 and Table 22. The estimation results using $W_1 = W_{interindustry}$ and $W_2 = W_{industry}$ is presented in Table 23 and Table 24. After controlling for all the intra-sector/industry links, the inter-sector/industry links identified using the text-mining method continue to statistically significant channels of risk spillovers. In general, the intra-sector/industry risk spillovers are more intense than the inter-sector/industry risk spillovers. The consumer and manufacturing companies have stronger inter-sector/industry spillovers than others. From Table 8, the two-W model with $W_1 = W_{interindustry}$ and $W_2 = W_{industry}$ has the best in-sample and out-of-sample fit among all alternatives.

	In-Sample MSE	Out-of-Sample MSE
Panel(1): Spatial-temporal model with one W		
$W_{baseline}$	**2.764**	**1.287**
$W_{unweighted}$	2.793	1.301
$W_{weighted}$	2.770	1.292
$W_{noncompetitive}$	**2.769**	**1.290**
$W_{persistent}$	2.782	1.292
$W_{intersector}$	2.884	1.353
$W_{interindustry}$	2.850	1.337
Panel(2): Spatial-temporal model with two Ws		
$W_{intersector} + W_{sector}$	2.809	1.332
$W_{interindustry} + W_{industry}$	**2.725**	**1.265**

Table 8: In-sample and out-of-sample MSE (in basis point) of specifications . Panel (1) considers several alternative adjacency matrices W under the one-W specification as in equation (2) and Panel (2) considers the two-W specification as in equation (18). For each column panel, the best 3 (smallest MSE) cases are in bold.

6.2 Placebo Test

In this section, I conduct a placebo test by checking whether randomly generated networks would give rise to significant local dependencies. Our full sample news-based network has 6742 linked pairs out of $N*(N-1)/2 = 148785$ pairs of firms. So the linking probability is 4.5%. I generate 100 random graphs using $G(N, p)$ model, which is one version of the Erdős–Rényi (ER) random graph models. In the $G(N, p)$ model, a graph is constructed by connecting nodes randomly. There are N edges and each edge is included in the graph with probability p independent from every other edge. To have the same level of sparsity as our full sample news-based network, I let $p = 4.5\%$.

For each one of the randomly simulated E-R networks, I use it as the adjacency matrix W in equation (2) and estimate the spatial-temporal model. None of them produces significant spatial parameters. The placebo test confirms that the textual analysis approach does help us to identify the economically important links among firms that facilitate transmissions of local shocks.

7 Conclusion

This paper investigates the local dependencies in idiosyncratic asset returns, which is an area less explored in empirical finance studies due to data availability issues. Utilizing the novel text-based linkage data, I construct the channels through which the local shocks transmit. I found that stocks linked via news items co-mentioning exhibit excess comovement beyond that is predicted by standard asset pricing models. By constructing spatial-temporal spillover matrices, we identify the major systemic risk contributors and receivers, which are of the interest to microprudential polices. From a macroprudential perspective, by separately addressing both strong and weak cross-sectional dependencies, I found that the strength of strong and weak CSD tend to move in opposite directions. When equities dis-connect from the market factor during periods of financial turmoil, the strength of local risk spillovers increases.

The text-based network is not intended to replace traditional economic datasets. Rather, it can be viewed as a promising alternative to existing network data. Our empirical study shows that it is competitive in the modelling of local risk spillovers. The author believes that it can be applied to a wider context, such as the modelling of the volatility spillovers and using text-based links as prior information in estimating links from large panel, etc.

References

D. Acemoglu, V. M. Carvalho, A. Ozdaglar, and A. Tahbaz-Salehi. The network origins of aggregate fluctuations. *Econometrica*, 80(5):1977–2016, 2012.

M. Aquaro, N. Bailey, and M. H. Pesaran. Estimation and inference for spatial models with heterogeneous coefficients: an application to us house prices. *USC-INET Research Paper*, (19-07), 2019.

J. Bai and S. Ng. Determining the number of factors in approximate factor models. *Econometrica*, 70(1): 191–221, 2002.

N. Bailey, S. Holly, and M. H. Pesaran. A two-stage approach to spatio-temporal analysis with strong and weak cross-sectional dependence. *Journal of Applied Econometrics*, 31(1):249–280, 2016.

N. Bailey, G. Kapetanios, and M. H. Pesaran. Exponent of cross-sectional dependence for residuals. *Sankhya B*, 81(1):46–102, 2019.

N. Bailey, G. Kapetanios, and M. H. Pesaran. Measurement of factor strenght: Theory and practice. 2020.

S. R. Baker, N. Bloom, and S. J. Davis. Measuring economic policy uncertainty. *The quarterly journal of economics*, 131(4):1593–1636, 2016.

M. Barigozzi and C. Brownlees. Nets: Network estimation for time series. *Journal of Applied Econometrics*, 34 (3):347–364, 2019.

M. Barigozzi and M. Hallin. A network analysis of the volatility of high dimensional financial series. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 66(3):581–605, 2017.

T. S. Breusch and A. R. Pagan. The lagrange multiplier test and its applications to model specification in econometrics. *The review of economic studies*, 47(1):239–253, 1980.

M. M. Carhart. On persistence in mutual fund performance. *The Journal of finance*, 52(1):57–82, 1997.

A. Chudik and M. H. Pesaran. Mean group estimation in presence of weakly cross-correlated estimators. *Economics Letters*, 175:101–105, 2019.

L. Cohen and A. Frazzini. Economic links and predictable returns. *The Journal of Finance*, 63(4):1977–2011, 2008.

M. Demirer, F. X. Diebold, L. Liu, and K. Yilmaz. Estimating global bank network connectedness. *Journal of Applied Econometrics*, 33(1):1–15, 2018.

F. X. Diebold and K. Yılmaz. On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics*, 182(1):119–134, 2014.

J. Engelberg, A. Ozoguz, and S. Wang. Know thy neighbor: Industry clusters, information spillovers, and market efficiency. *Journal of Financial and Quantitative Analysis*, 53(5):1937–1961, 2018.

E. F. Fama and K. R. French. A five-factor asset pricing model. *Journal of financial economics*, 116(1):1–22, 2015.

J. Fan, A. Furger, and D. Xiu. Incorporating global industrial classification standard into portfolio allocation: A simple factor-based large covariance matrix estimator with high-frequency data. *Journal of Business & Economic Statistics*, 34(4):489–503, 2016.

X. Gabaix. The granular origins of aggregate fluctuations. *Econometrica*, 79(3):733–772, 2011.

D. Garcia. Sentiment during recessions. *The Journal of Finance*, 68(3):1267–1300, 2013.

G. Hale and J. A. Lopez. Monitoring banking system connectedness with big data. *Journal of Econometrics*, 212(1):203–220, 2019.

N. Hautsch, J. Schaumburg, and M. Schienle. Financial network systemic risk contributions. *Review of Finance*, 19(2):685–738, 2015.

G. Hoberg and G. Phillips. Text-based network industries and endogenous product differentiation. *Journal of Political Economy*, 124(5):1423–1465, 2016.

G. Hoberg and G. M. Phillips. Text-based industry momentum. *Journal of Financial and Quantitative Analysis*, 53(6):2355–2388, 2018.

R. D. Israelsen. Does common analyst coverage explain excess comovement? *Journal of Financial and Quantitative Analysis*, 51(4):1193–1229, 2016.

Z. T. Ke, B. T. Kelly, and D. Xiu. Predicting returns with text data. Technical report, National Bureau of Economic Research, 2019.

G. Koop, M. H. Pesaran, and S. M. Potter. Impulse response analysis in nonlinear multivariate models. *Journal of econometrics*, 74(1):119–147, 1996.

S. Kou, X. Peng, and H. Zhong. Asset pricing with spatial interaction. *Management Science*, 64(5):2083–2101, 2018.

C. M. Lee, P. Ma, and C. C. Wang. Search-based peer firms: Aggregating investor perceptions through internet co-searches. *Journal of Financial Economics*, 116(2):410–431, 2015.

J. P. LeSage. An introduction to spatial econometrics. *Revue d'économie industrielle*, (123):19–44, 2008.

J. P. LeSage and Y.-Y. Chih. Interpreting heterogeneous coefficient spatial autoregressive panel models. *Economics Letters*, 142:1–5, 2016.

T. J. Moskowitz and M. Grinblatt. Do industries explain momentum? *The Journal of finance*, 54(4):1249–1290, 1999.

C. A. Parsons, R. Sabbatucci, and S. Titman. Geographic lead-lag effects. *The Review of Financial Studies*, 33 (10):4721–4770, 2020.

H. H. Pesaran and Y. Shin. Generalized impulse response analysis in linear multivariate models. *Economics letters*, 58(1):17–29, 1998.

M. H. Pesaran and R. Smith. Estimating long-run relationships from dynamic heterogeneous panels. *Journal of econometrics*, 68(1):79–113, 1995.

C. Pirinsky and Q. Wang. Does corporate headquarters location matter for stock returns? *The Journal of Finance*, 61(4):1991–2015, 2006.

S. M. Price, J. S. Doran, D. R. Peterson, and B. A. Bliss. Earnings conference calls and stock returns: The incremental informativeness of textual tone. *Journal of Banking & Finance*, 36(4):992–1011, 2012.

S. Ross. The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, 13(3):341–360, 1976.

A. Scherbina and B. Schlusche. Economic linkages inferred from news stories and the predictability of stock returns. *Available at SSRN 2363436*, 2015.

G. Schwenkler and H. Zheng. The network of firms implied by the news. *Available at SSRN 3320859*, 2019.

W. F. Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *The journal of finance*, 19(3):425–442, 1964.

American Express and Regis Corporation Announce Strategic Partnership; Hair Care Industry's Global Leader to Roll-out Card Acceptance at all of its U.S. Locations

Business Wire

February 24, 2005 Thursday 2:00 PM GMT

Distribution: Business Editors

Length: 438 words

Dateline: NEW YORK Feb. 24, 2005

Body

American Express and Regis Corporation today announced a plan for nationwide card acceptance at Regis' U.S. salons. With thousands of locations currently accepting the American Express Card, the companies expect all corporate owned Regis U.S. locations to be accepting American Express by the end of calendar year 2005.

"Our partnership with American Express is in direct response to our customers. Over the last several years we have seen an increasing demand by our customers to accept American Express," commented Kyle Didier, vice president, finance at Regis Corporation. "In addition, American Express' willingness to extend the partnership benefits to our franchisees demonstrates their commitment to drive value throughout the entire Regis Corporation network."

"Working with the world's largest operator of hair salons reinforces our commitment to the hair salon industry overall and demonstrates our ability to drive value to hair salon owners," said Elizabeth Langwith, vice president, American Express Establishment Services. "We're delighted to provide our Cardmembers with another opportunity to earn rewards, cash or miles for their everyday purchases."

The companies will work together to develop marketing programs that deliver value to consumers using American Express-branded cards. In addition, Regis franchise owners will qualify for special discounts on a variety of business expenses ranging from shipping, technology, car rentals and cellular phone service through the American Express Business Savings Program.

Classification

Language: ENGLISH

Publication-Type: Newswire

Subject: FRANCHISING (85%); PRESS RELEASES (75%); CONSUMERS (69%); FRANCHISEES (62%); Contract/Agreement (%)

Company: REGIS CORP (94%); AMERICAN EXPRESS CO (94%); HAIR CLUB FOR MEN INC (52%); NY-AMEX/REGIS

Ticker: RGS (NYSE) (94%); ***AXP*** (NYSE) (94%); RGS (NYSE)

Figure 5: A typical business news in the dataset.

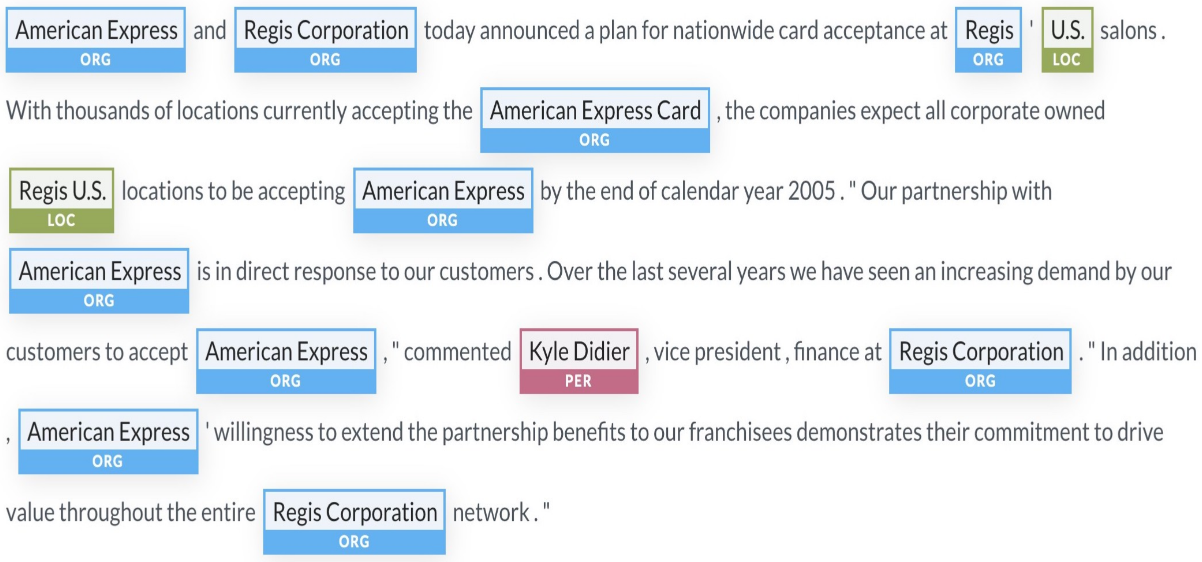


Figure 6: Named-Entity-Recognition(NER) demo. Note: if we feed a text into the NER algorithm, it will tag organizations, locations, person, etc. Given a list of companies with company names and tickers, we can either match by names (by calculating distance between strings, a match corresponds to distance being smaller than some threshold) or by match tickers.

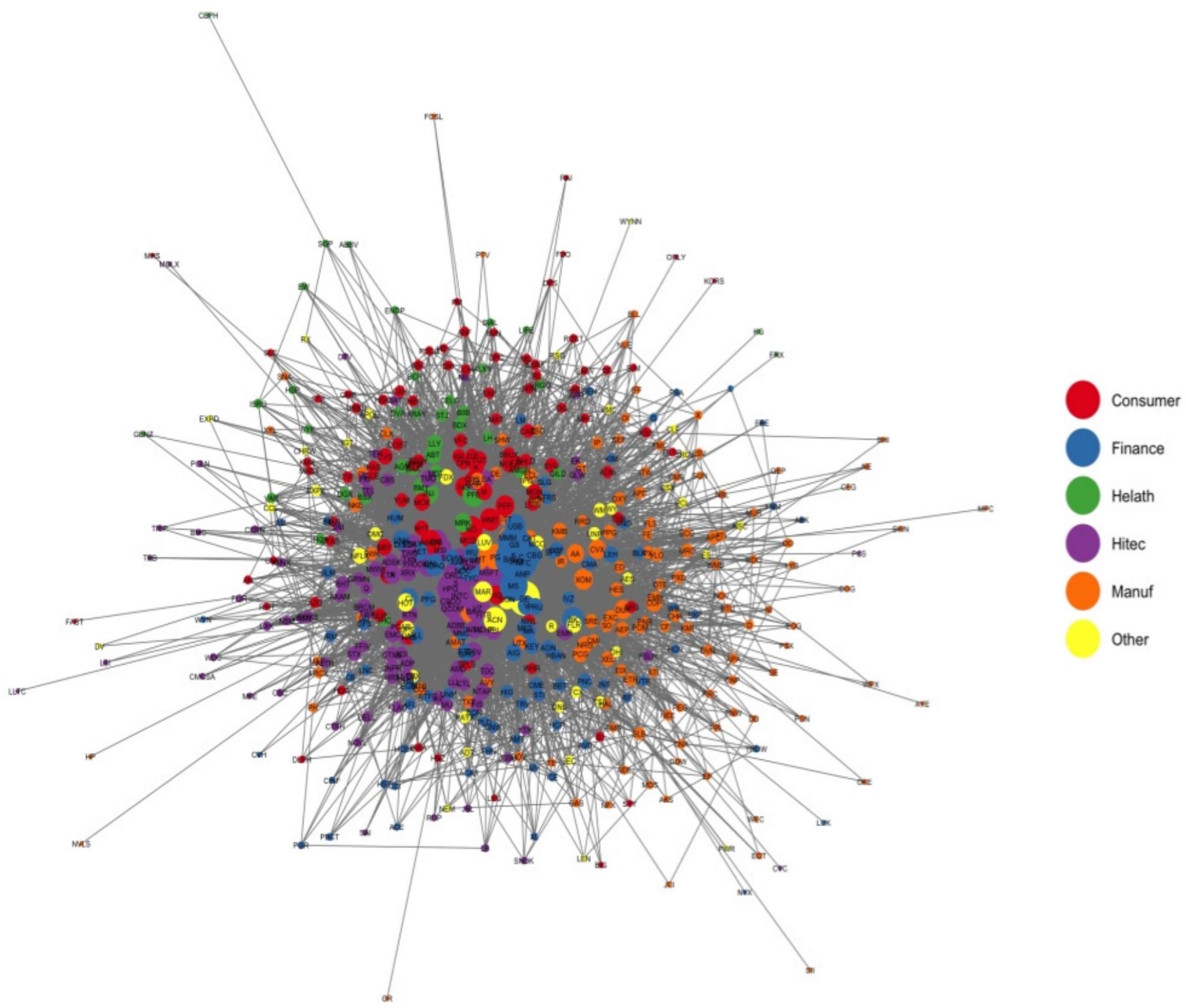


Figure 7: **News-based networks of $S\&P500$ companies identified using all the business news from Business Wire from 2006 to 2013**. The figure plots all the links identified in the sample period. For visualization, only companies with links are plotted. The color of a node indicate which sector the company is in. The sector classification is given in section 4.1. The data source of SIC code for each sample company is from CRSP/Compustat Merged. The size of a node is proportional to the network degree of the company (how many other companies is the node linked with). If there is an edge between two nodes, this indicates there is a link between the nodes.

Linked pair	Link type
(MSFT,INTC)	Strategic partnerships, product developments
(ACE,CB)	Strategic partnerships,merger and acquisition
(C,LM)	Outsourcing, Strategic partnerships
(MSFT,APPL)	Products developments, competition
(BAC,WFC)	Joint venture, strategic partnerships
(CCE,KO)	Partners, common major owners
(MSFT,HPQ)	Strategic partnerships, products developments, competition
(MSFT,ORCL)	Strategic partnerships, products developments
(AXP,V)	Competition, legal
(PEP,KO)	Competition
(BAC,V)	Strategic partnerships, joint venture, products developments
(JPM,GS)	Joint investment banking, competition
(MS,AXP)	Strategic partnerships, joint venture, products developments
(WFC,JPM)	Joint venture, strategic partnerships
(C,JPM)	Joint venture, strategic partnerships
(T,VZ)	Competition
(C,MS)	Joint venture, strategic partnerships
(MSFT,CSCO)	Strategic partnerships, products developments
(NVDA,INTC)	Strategic partnerships, products developments, competition
(Q,CTL)	Competition
(MSFT,ADBE)	Strategic partnerships, products developments
(JPM,GS)	Joint venture, strategic partnerships, joint investment banking, competition
(BA,MSFT)	Strategic partnerships, product developments
(PFE,BMY)	Joint reserach and development
(MSFT,ACN)	Strategic partnerships, products developments
(AMD,INTC)	Supplier-customer
(C,BAC)	Competition
(GE,BA)	Supplier-customer
(BK,JPM)	Business-swap, acquire business lines
(MSFT,INTU)	Strategic partnerships, products developments
(MSFT,CTXS)	Strategic partnerships, products developments
(DISCA,HAS)	Joint venture, supplier-customer
(BSX,STJ)	Legal settlement, competition, joint development effort
(LLY,PFE)	Joint reserach and development
(GS,C)	Strategic partnerships, joint financing
(NOC,BA)	Strategic partnerships, supplier-customer
(K,PG)	Acquire business lines
(AMZN,APPL)	Strategic partnerships, products developments
(GE,JPM)	Strategic partnerships, joint financing
(VZ,MSI)	Alliance, products developments

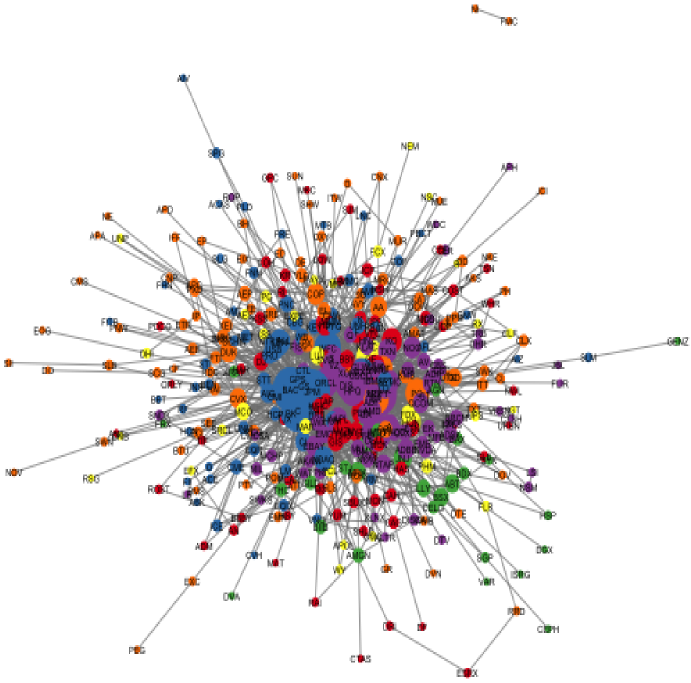
Table 9: Link validation. Note: The table shows the type of economic linkages that the article co-mentioning imply. Since those pairs were co-mentioned quite frequently, for each pair, we randomly read 5 news that have co-mentioned the two firms and infer link type from the news. Thus the listed link types are representative but not exhaustive. Due to space limitations, we only show the validation results for the most frequently co-mentioned pairs.

Table 10: **Summary statistics of news-based links from 2006 to 2013**. I collect all distinct business news within the sample period that tagged the $S\&P500$ companies. Firm links are constructed using the methodology mentioned in section 2. Most pairs of firms are co-mentioned multiple times within the sample period, and we consider both weighted links and unweighted links. For the weighted version, a typical entry w_{ij} of W gives the number of times i and j that are co-mentioned in the sample (co-mentioning multiple times in the same month count only once). On the other hand, for the unweighted version, the entries of W are 0/1 dummies. $w_{ij} = 1$ if i and j are co-mentioned at least once $w_{ij} = 0$ if i and j are never mentioned together in any articles. The unweighted version of the statistics are given in the parentheses below the corresponding weighted statistics. The number of total links identified for sector g is $\sum_{i \in g}^{N} d_i$, where d_i is the links of firm i from sector g. The average and the 90th percentile of d_i are given in the two following rows. I further break down links to intra-sector and inter-sector links. The sector definitions are given in section 4.1

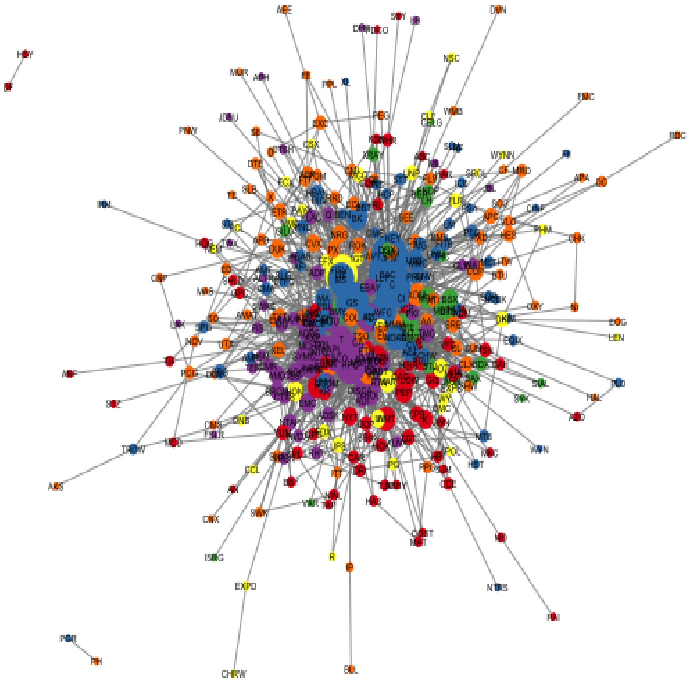
	Finance	Consumer	Manuf	Hitech	Health	All
#of total links identified	11114.00	4843.00	5287.00	13447.00	2821.00	40185.00
	(3403.00)	(2115.00)	(2350.00)	(3616.00)	(870.00)	(13485.00)
Average degree	110.04	47.02	36.46	123.37	70.53	73.60
	(33.69)	(20.53)	(16.21)	(33.17)	(21.75)	(24.70)
90th percentile of degree	169.00	104.80	74.20	247.20	166.90	154.00
	(66.00)	(43.80)	(34.00)	(71.40)	(50.10)	(52.50)
#of intra sector links	6582.00	2020.00	2232.00	8802.00	1660.00	21660.00
	(1372.00)	(724.00)	(864.00)	(1624.00)	(316.00)	(5028.00)
#of inter sector links	4532.00	2823.00	3055.00	4645.00	1161.00	18525.00
	(2031.00)	(1391.00)	(1486.00)	(1992.00)	(554.00)	(8457.00)
%of firms with inter sector links	0.89	0.94	0.87	0.89	0.92	0.90
	(0.89)	(0.94)	(0.87)	(0.89)	(0.92)	(0.90)

Table 11: Links aggregated at sector level (2006-2013). Row normalizations in percentages are given in the parenthesis.

	Consumer	Finance	Health	Hitech	Manuf	Other
Consumer	724	433	97	420	296	145
	(34%)	(20%)	(5%)	(20%)	(14%)	(7%)
Finance	433	1372	183	686	445	284
	(13%)	(40%)	(5%)	(20%)	(13%)	(8%)
Health	97	183	316	149	74	51
	(11%)	(21%)	(36%)	(17%)	(9%)	(6%)
Hitech	420	686	149	1625	442	294
	(12%)	(19%)	(4%)	(45%)	(12%)	(8%)
Manuf	296	445	74	442	864	229
	(13%)	(19%)	(3%)	(19%)	(37%)	(10%)
Other	145	284	51	294	229	128
	(13%)	(25%)	(5%)	(26%)	(20%)	(11%)

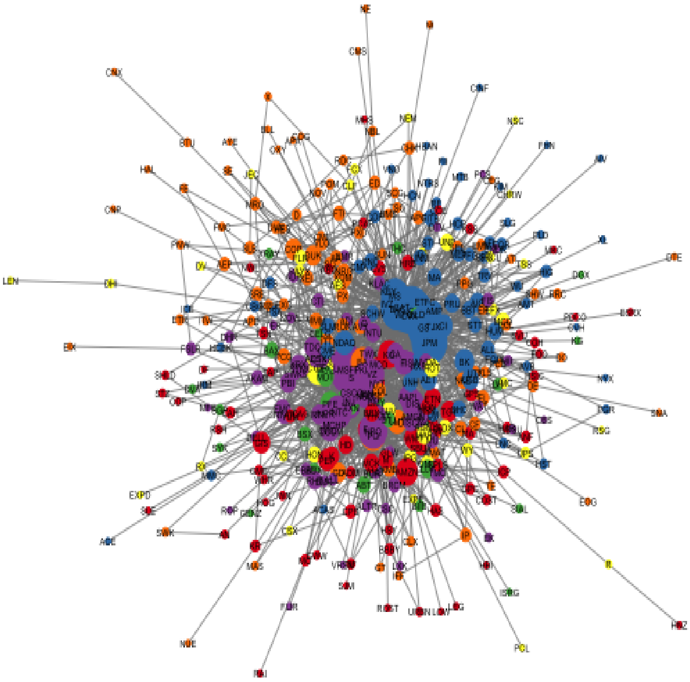


2006

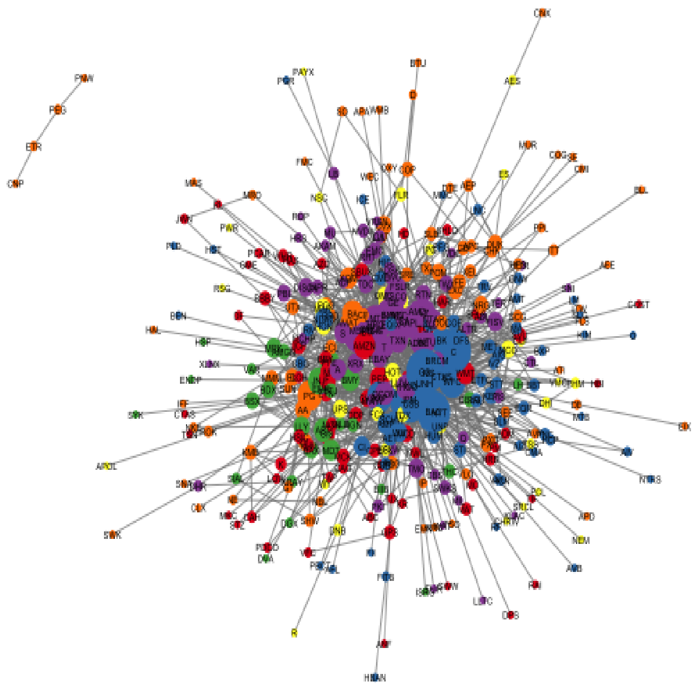


2007

Figure 8: Yearly news-based networks of $S\&P500$ companies. For each year, all the business news from Business Wire that have mentioned sample companies are used to identify links across companies. Only companies with links are plotted. The color code is the same as the aggregate graph and node size is proportional to the network degree of the company (how many other companies is the node linked with).

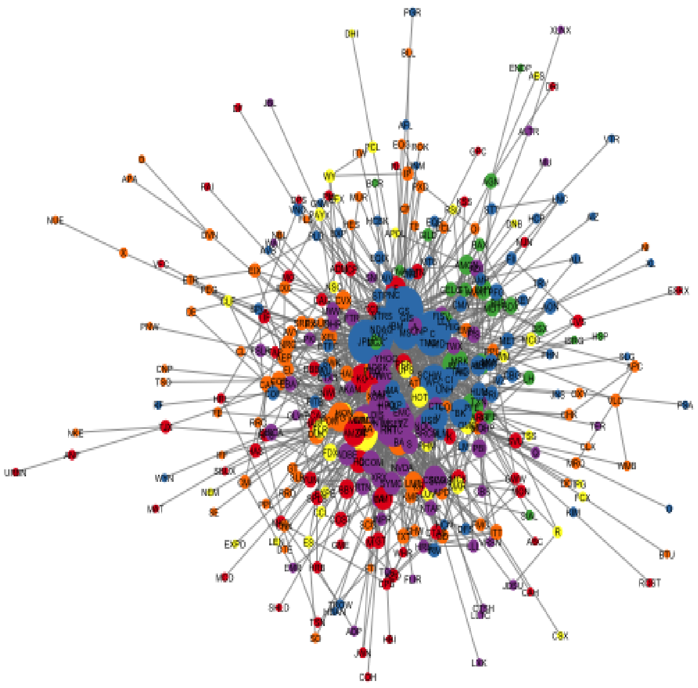


2008

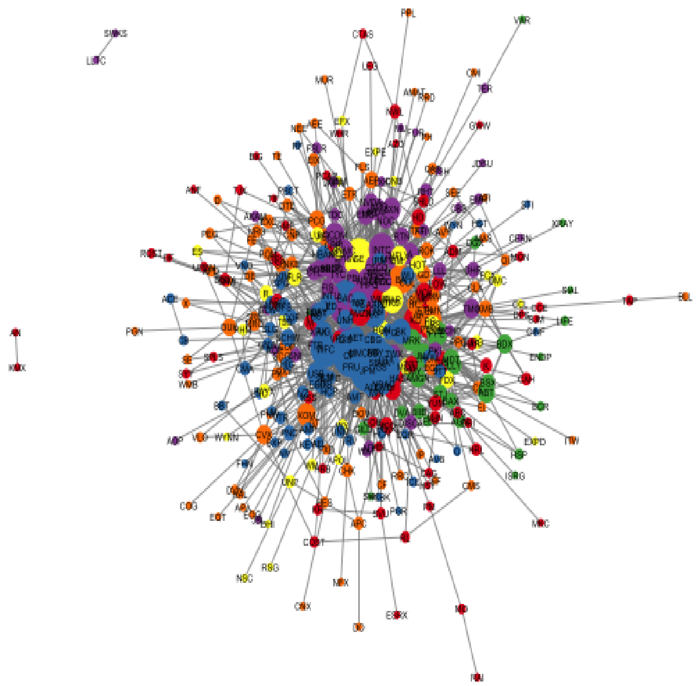


2009

Figure 8: Yearly news-based networks of $S\&P500$ companies. For each year, all the business news from Business Wire that have mentioned sample companies are used to identify links across companies. Only companies with links are plotted. The color code is the same as the aggregate graph and node size is proportional to the network degree of the company (how many other companies is the node linked with).

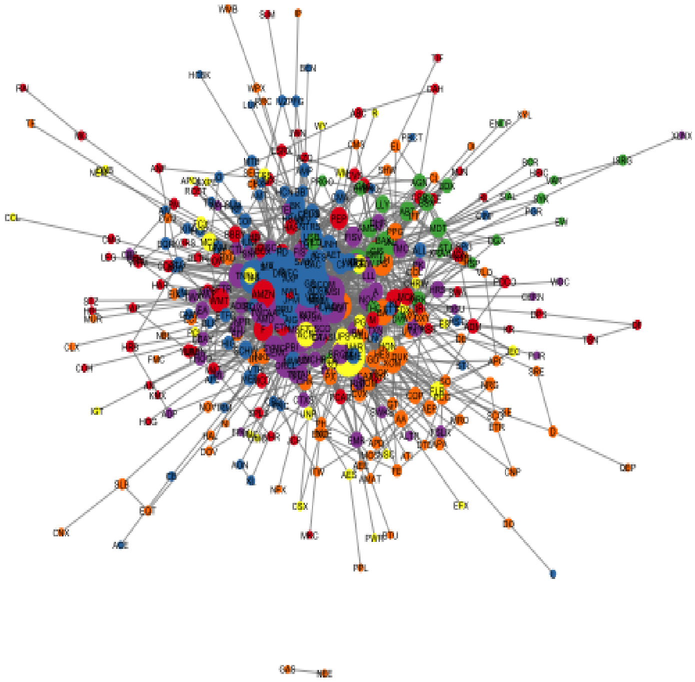


2010

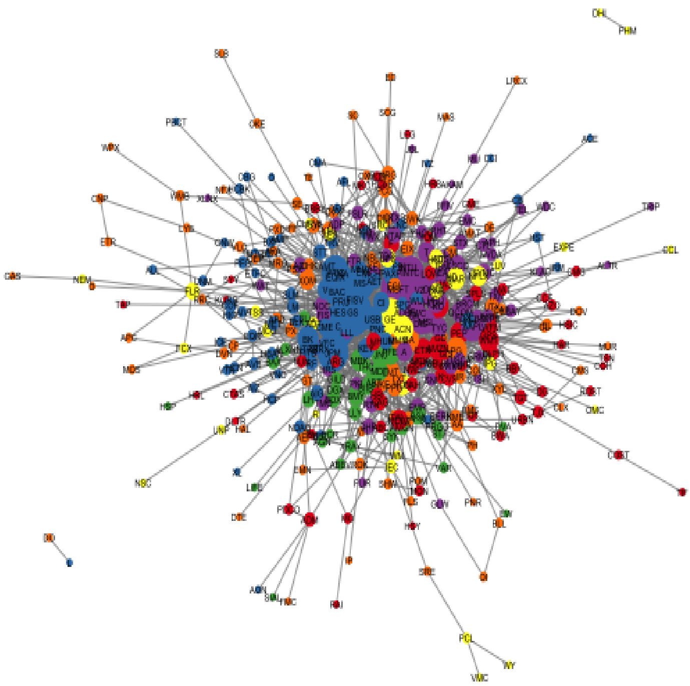


2011

Figure 8: Yearly news-based networks of $S\&P500$ companies. For each year, all the business news from Business Wire that have mentioned sample companies are used to identify links across companies. Only companies with links are plotted. The color code is the same as the aggregate graph and node size is proportional to the network degree of the company (how many other companies is the node linked with).



2012



2013

Figure 8: Yearly news-based networks of $S\&P500$ companies. For each year, all the business news from Business Wire that have mentioned sample companies are used to identify links across companies. Only companies with links are plotted. The color code is the same as the aggregate graph and node size is proportional to the network degree of the company (how many other companies is the node linked with).

Table 12: **Summary statistics of news-based links from 2006 to 2013**. For each year from 2006 to 2013, I collect all distinct business news within that year and construct firm links using the methodology mentioned in section 2. The summary statistics of that year's news-based links are given in sub-tables. For each year, most pairs of firms are co-mentioned multiple times, and we consider both weighted links and unweighted links. For the weighted version, a typical entry w_{ij} of W gives the number of times i and j that are co-mentioned in the sample (co-mentioning multiple times in the same month count only once). On the other hand, for the unweighted version, the entries of W are 0/1 dummies. $w_{ij} = 1$ if i and j are co-mentioned at least once $w_{ij} = 0$ if i and j are never mentioned together in any articles. The unweighted version of the statistics are given in the parentheses below the corresponding weighted statistics. The number of total links identified for sector g is $\sum_{i \in g}^{N} d_i$, where d_i is the links of firm i from sector g. The average and the 90th percentile of d_i are given in the two following rows. I further break down links to intra-sector and inter-sector links. The sector definitions are given in section 4.1

2006

	Finance	Consumer	Manuf	Hitech	Health	All
#of total links identified	1208.00	544.00	644.00	1762.00	276.00	4671.00
	(739.00)	(409.00)	(486.00)	(954.00)	(190.00)	(2959.00)
Average degree	15.10	6.80	5.24	20.97	8.1	10.57
	(9.24)	(5.11)	(3.95)	(11.35)	(5.58)	(6.69)
90th percentile of degree	29.20	17.10	10.80	43.00	20.70	23.90
	(22.30)	(12.00)	(9.00)	(25.50)	(12.00)	(16.00)
#of intra sector links	678.00	204.00	262.00	1204.00	150.00	2524.00
	(334.00)	(136.00)	(184.00)	(536.00)	(88.00)	(1296.00)
#of inter sector links	530.00	340.00	382.00	558.00	126.00	2147.00
	(405.00)	(273.00)	(302.00)	(418.00)	(102.00)	(1663.00)
%of firms with inter sector links	0.59	0.81	0.63	0.81	0.65	0.69
	(0.59)	(0.81)	(0.63)	(0.81)	(0.65)	(0.69)

2007

	Finance	Consumer	Manuf	Hitech	Health	All
#of total links identified	1295	553	735	1691	372	4946
	(790)	(385)	(504)	(827)	(228)	(2958)
Average degree	14.89	6.74	5.93	19.89	10.94	10.89
	(9.08)	(4.69)	(4.06)	(9.73)	(6.71)	(6.52)
90th percentile of degree	28.80	14.00	13.00	47.60	23.40	23.70
	(18.20)	(10.90)	(9.00)	(22.60)	(15.00)	(15.00)
#of intra sector links	682.00	240.00	312.00	1202.00	210.00	2688.00
	(312.00)	(146.00)	(202.00)	(470.00)	(104.00)	(1266.00)
#of inter sector links	613.00	313.00	423.00	489.00	162.00	2258.00
	(478.00)	(239.00)	(302.00)	(357.00)	(124.00)	(1692.00)
%of firms with inter sector links	0.61	0.72	0.64	0.72	0.74	0.68
	(0.61)	(0.72)	(0.64)	0.72)	(0.74)	(0.68)

Table 12: Continued

2008

	Finance	Consumer	Manuf	Hitech	Health	All
#of total links identified	1324.00	611.00	714.00	1633.00	337.00	4876.00
	(857.00)	(432.00)	(504.00)	(890.00)	(205.00)	(3088.00)
Average degree	15.22	7.36	5.71	18.56	9.91	10.58
	(9.85)	(5.20)	(4.03)	(10.11)	(6.03)	(6.70)
90th percentile of degree	30.40	16.80	11.60	38.60	20.00	23.00
	(20.80)	(12.00)	(8.00)	(23.00)	(13.00)	(14.00)
#of intra sector links	774.00	236.00	312.00	1100.00	178.00	2630.00
	(408.00)	(154.00)	(194.00)	(476.00)	(88.00)	(1340.00)
#of inter sector links	550.00	375.00	402.00	533.00	159.00	2246.00
	(449.00)	(278.00)	(310.00)	(414.00)	(117.00)	(1748.00)
%of firms with inter sector links	0.61	0.82	0.64	0.75	0.76	0.71
	(0.61)	(0.82)	(0.64)	(0.75)	(0.76)	(0.71)

2009

	Finance	Consumer	Manuf	Hitech	Health	All
#of total links identified	1091.00	435.00	505.00	1317.00	295.00	3852.00
	(696.00)	(329.00)	(362.00)	(733.00)	(187.00)	(2472.00)
Average degree	12.54	5.12	4.01	14.16	8.68	8.21
	(8.00)	(3.87)	(2.87)	(7.88)	(5.50)	(5.27)
90th percentile of degree	33.80	9.60	10.00	27.80	21.70	18.00
	(17.80)	(8.00)	(7.50)	(18.60)	(14.70)	(12.00)
#of intra sector links	616.00	158.00	212.00	892.00	168.00	2078.00
	(302.00)	(110.00)	(140.00)	(398.00)	(86.00)	(1056.00)
#of inter sector links	475.00	277.00	293.00	425.00	127.00	1774.00
	(394.00)	(219.00)	(222.00)	(335.00)	(101.00)	(1416.00)
%of firms with inter sector links	0.60	0.66	0.53	0.65	0.68	0.62
	(0.60)	(0.66)	(0.53)	(0.65)	(0.68)	(0.62)

2010

	Finance	Consumer	Manuf	Hitech	Health	All
#of total links identified	1072.00	471.00	538.00	1234.00	286.00	3832.00
	(677.00)	(328.00)	(377.00)	(699.00)	(180.00)	(2442.00)
Average degree	12.04	5.48	4.45	13.41	8.67	8.24
	(7.61)	(3.81)	(3.12)	(7.60)	(5.45)	(5.25)
90th percentile of degree	27.80	12.00	11.00	28.00	21.00	17.60
	(15.80)	(9.00)	(8.00)	(16.90)	(11.80)	(11.60)
#of intra sector links	628.00	194.00	232.00	774.00	178.00	2050.00
	(320.00)	(126.00)	(148.00)	(372.00)	(94.00)	(1094.00)
#of inter sector links	444.00	277.00	306.00	460.00	108.00	1782.00
	(357.00)	(202.00)	(229.00)	(327.00)	(86.00)	(1348.00)
%of firms with inter sector links	0.63	0.63	0.62	0.64	0.70	0.64
	(0.63)	(0.63)	(0.62)	(0.64)	(0.70)	(0.64)

Table 12: Continued

2011

	Finance	Consumer	Manuf	Hitech	Health	All
#of total links identified	1330.00	453.00	549.00	1334.00	317.00	4282.00
	(815.00)	(324.00)	(416.00)	(767.00)	(204.00)	(2742.00)
Average degree	15.29	5.27	4.54	15.16	10.23	9.35
	(9.37)	(3.77)	(3.44)	(8.72)	(6.58)	(5.99)
90th percentile of degree	33.00	12.50	10.00	30.60	24.00	20.00
	(16.80)	(7.50)	(8.00)	(19.30)	(14.00)	(13.30)
#of intra sector links	846.00	196.00	214.00	852.00	192.00	2348.00
	(424.00)	(118.00)	(146.00)	(412.00)	(102.00)	(1236.00)
#of inter sector links	484.00	257.00	335.00	482.00	125.00	1934.00
	(391.00)	(206.00)	(270.00)	(355.00)	(102.00)	(1506.00)
%of firms with inter sector links	0.62	0.72	0.60	0.70	0.81	0.67
	(0.62)	(0.72)	(0.60)	(0.70)	(0.81)	(0.67)

2012

	Finance	Consumer	Manuf	Hitech	Health	All
#of total links identified	1262.00	483.00	574.00	1354.00	367.00	4338.00
	(768.00)	(366.00)	(407.00)	(794.00)	(213.00)	(2770.00)
Average degree	14.34	5.49	4.82	14.72	10.79	9.31
	(8.73)	(4.16)	(3.42)	(8.63)	(6.26)	(5.94)
90th percentile of degree	38.90	12.30	11.00	32.70	27.60	19.50
	(20.00)	(8.30)	(8.00)	(19.00)	(12.00)	(13.00)
#of intra sector links	814.00	200.00	272.00	870.00	244.00	2444.00
	(412.00)	(144.00)	(180.00)	(442.00)	(120.00)	(1330.00)
#of inter sector links	448.00	283.00	302.00	484.00	123.00	1894.00
	(356.00)	(222.00)	(227.00)	(352.00)	(93.00)	(1440.00)
%of firms with inter sector links	0.64	0.74	0.61	0.66	0.76	0.68
	(0.64)	(0.74)	(0.61)	(0.66)	(0.76)	(0.68)

2013

	Finance	Consumer	Manuf	Hitech	Health	All
#of total links identified	1158.00	477.00	518.00	1133.00	365.00	3990.00
	(752.00)	(359.00)	(388.00)	(695.00)	(230.00)	(2660.00)
Average degree	13.47	5.42	4.11	13.49	10.14	8.60
	(8.74)	(4.08)	(3.08)	(8.27)	(6.39)	(5.73)
90th percentile of degree	29.00	11.60	10.00	23.70	24.50	19.00
	(18.00)	(9.00)	(7.00)	(16.70)	(15.50)	(13.00)
#of intra sector links	736.00	180.00	206.00	678.00	226.00	2076.00
	(398.00)	(128.00)	(144.00)	(368.00)	(118.00)	(1190.00)
#of inter sector links	422.00	297.00	312.00	455.00	139.00	1914.00
	(354.00)	(231.00)	(244.00)	(327.00)	(112.00)	(1470.00)
%of firms with inter sector links	0.64	0.68	0.60	0.75	0.75	0.66
	(0.64)	(0.68)	(0.60)	(0.75)	(0.75)	(0.66)

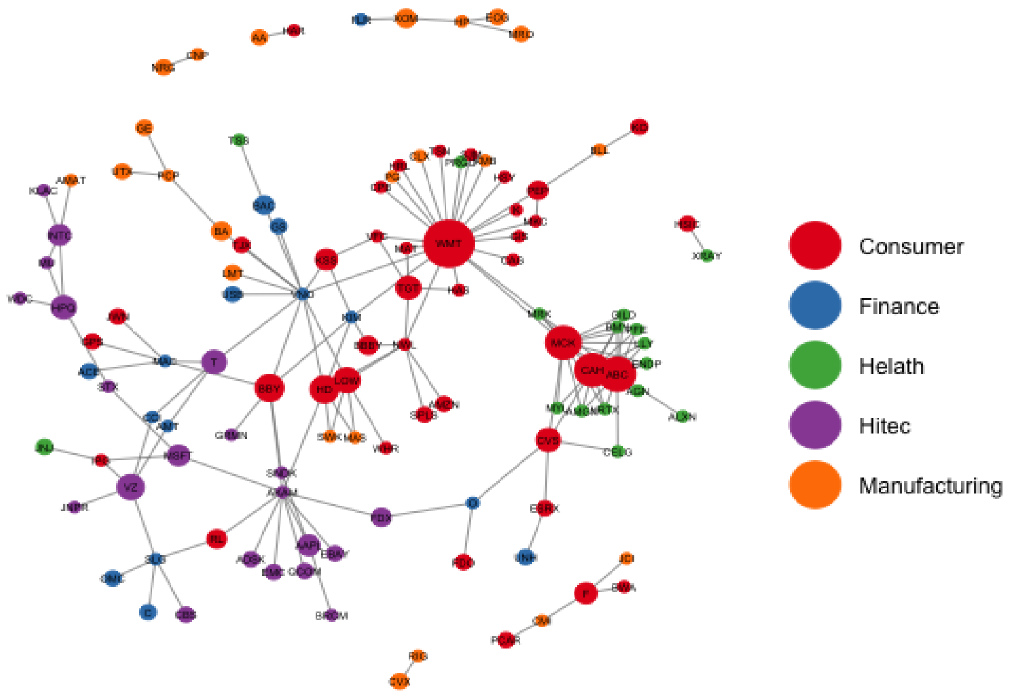
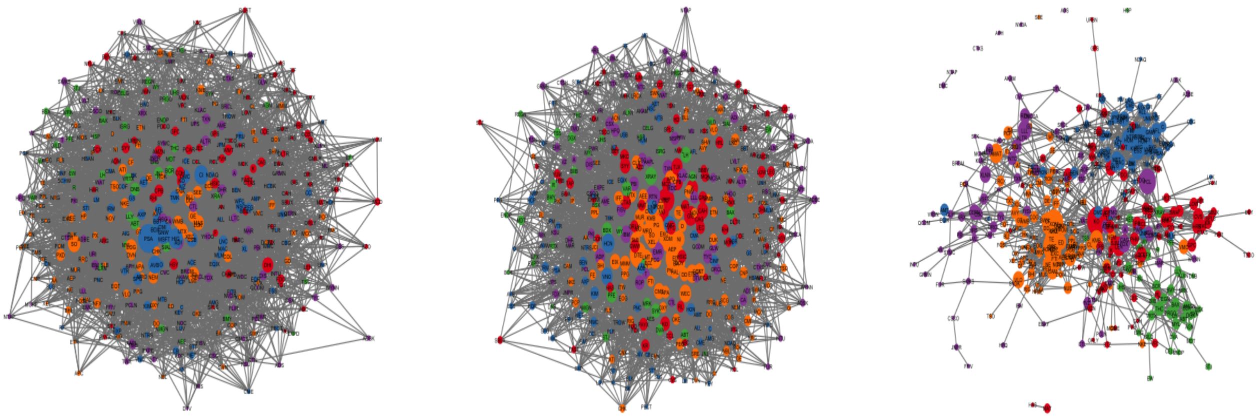
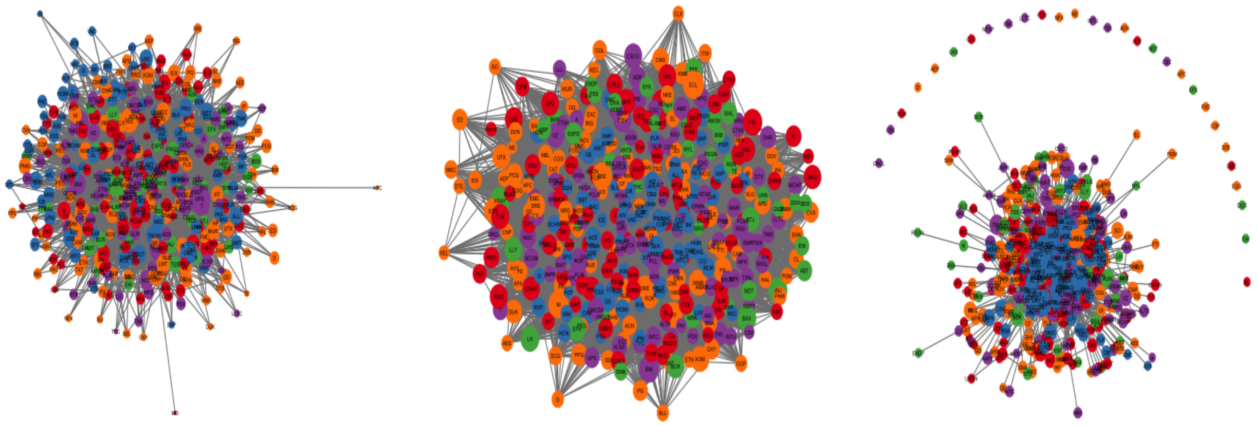


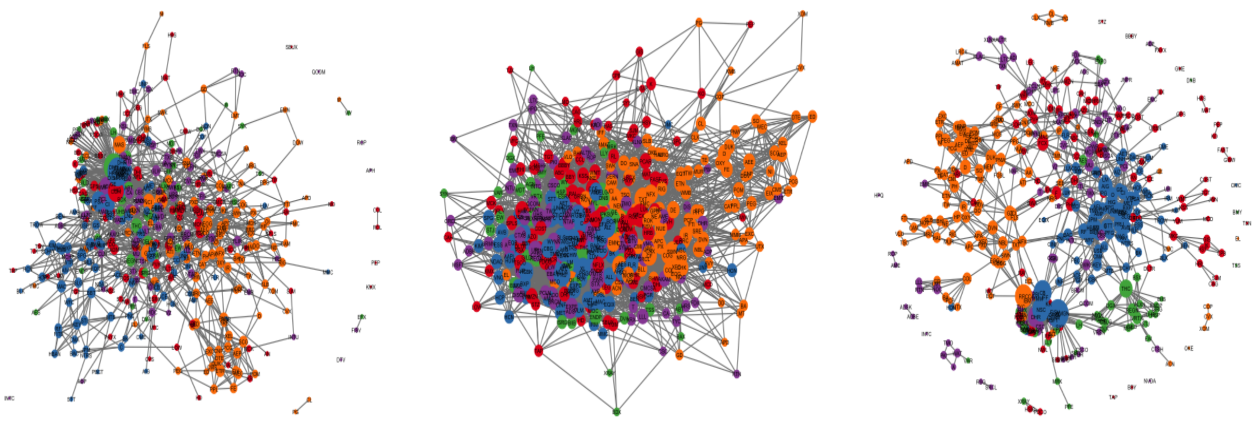
Figure 9: Customer-supplier links among $S\&P500$ companies $(2006-2013)$. Data source: Compustat segments files.



(a) Pre-crisis LVDN (2006-2007) (b) Crisis LVDN (2008-2009) (c) Full sample LVDN (2006-2013)



(d) Pre-crisis LGCN (2006-2007) (e) Crisis period LGCN (2008-2009) (f) Full sample LGCN (2006-2013)



(g) Pre-crisis PCN (2006-2007) (h) Crisis PCN (2008-2009) (i) Full sample PCN (2006-2013)

Figure 10: Long-run variance Decomposition network (LVDN), Long-run Granger causality network (LGCN) and Partial correlation network (PCN) applying the high-dimensional method from Barigozzi and Hallin (2017) on the de-factored returns from equation (1). Note: The LGCN and PCN are sparse given that the high-dimensional VAR and correlation matrix are regularized. LVDN, on the other hand, is dense (for the 3 samples, the link densities are all over 75%. Hard thresholding is applied and only links that contribute to more than 1% of the future variances are kept, thus the plotted LVDN is sparse. Link densities of LVDN applying different thresholds are presented in Table 8. Red, blue, green, purple and orange correspond to consumer, finance, health, hitech and manufacturing sector, respectively. Node size is proportional to out degree.

Threshold	Pre-crisis (2006-2007)	Crisis (2008-2009)	Full sample (2006-2013)
0	131935	162091	111165
1	3682	3835	1639
2	711	770	574
3	275	307	310
4	138	201	205
5	102	158	149

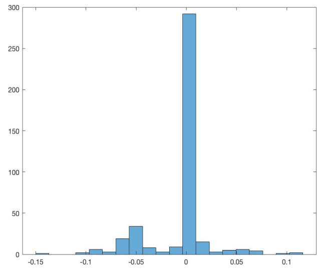
Table 13: Number of Long-run variance Decomposition network (LVDN) links after applying different hard thresholds (in percentage of future variance explained) for pre-crisis, crisis and full sample periods. Note: if threshold k is applied, a link from i to j is kept only if the shocks to i (the cumulative effect up to 10 lags) contribute to at least $k\%$ of j's variance.

	Pre-crisis (2006-2007)	Crisis (2008-2009)	Full sample (2006-2013)
LGCN	2005	9319	721
PCN	1614	3666	1486

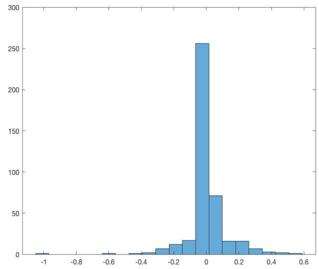
Table 14: Number of Long-run Granger causality network (LGCN) and the Partial correlation network (PCN) links identified from pre-crisis, crisis and full sample periods.

Threshold	Pre-crisis LVDN	Text-based networks
0	0.783	0.03
1	0.04	0.06
2	0.06	0.16
3	0.07	0.26
4	0.05	0.31
5	0.04	0.34

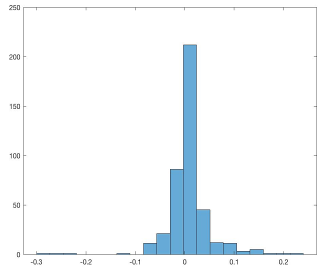
Table 15: Percentages of crisis period Long-run variance Decomposition network (LVDN) links that get identified using alternative pre-crisis network information. Note: Different hard thresholds are applied to the LVDN given the network implied by LVDN is very dense (the link densities for pre-crisis and crisis sample are 77.5% and 95.3%, respectively). We do not need to apply thresholding to text-based network since it is already very sparse (the link density of the full sample network is 4.5%, and for the short pre-crisis sample the density is even smaller.



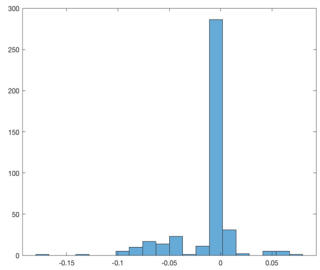
(a) own response ($h = 2$)



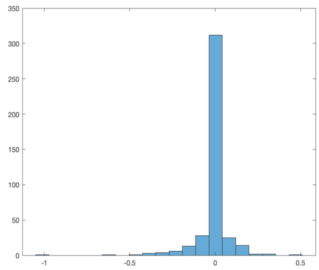
(b) in-degree ($h = 2$)



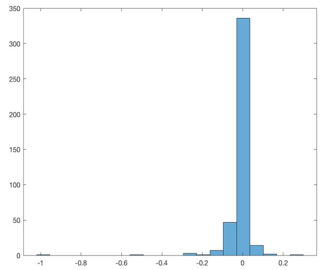
(c) out-degree ($h = 2$)



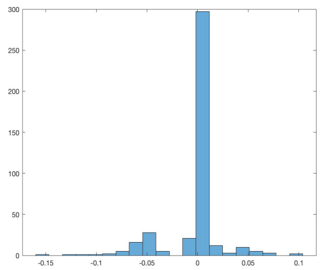
(d) own response ($h = 3$)



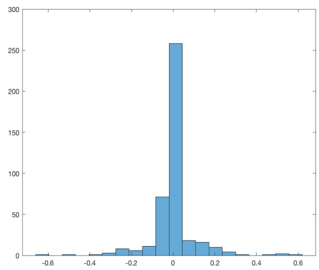
(e) in-degree ($h = 3$)



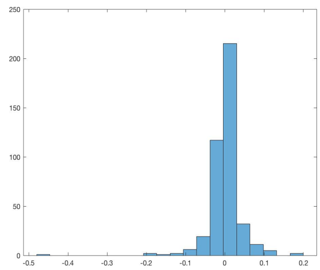
(f) out-degree ($h = 3$)



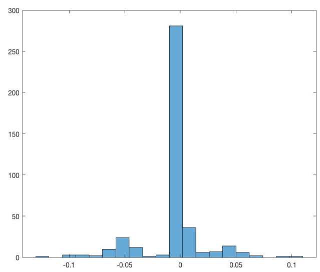
(g) own response ($h = 4$)



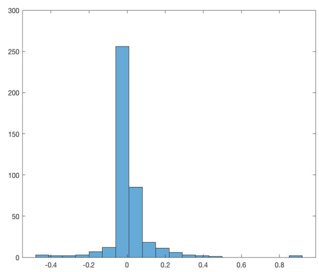
(h) in-degree ($h = 4$)



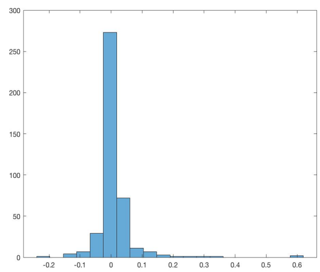
(i) out-degree ($h = 4$)



(j) own response ($h = 5$)



(k) in-degree ($h = 5$)



(l) out-degree ($h = 5$)

Figure 11: Histogram for own response, in-degree and out-degree at horizon $h = 2, 3, 4, 5$.

Table 16: In-sample and out-of-sample MSE (in basis point) of alternative models for rolling samples. For each sample, the training period is two years and testing period is one year. Note: for each panel, the best 3 (smallest MSE) cases are in bold.

Training period:03/01/2006-31/12/2007, Testing period :03/01/2008-31/12/2008

	Naive	BH-VAR	W_{empty}	W_{sector}	$W_{geographic}$	$W_{compustat}$	W_{news}
In Sample MSE							
(1)Heterogeneous coef	-	-	2.047	**2.012**	2.025	2.042	**1.992**
(2)Sectoral-heterogeneous coef	-	-	2.057	2.046	2.053	2.056	2.033
(3)Homogeneous coef	-	-	2.058	2.048	2.054	2.057	2.038
(4)	2.079	**1.272**	-	-	-	-	-
Out-of-Sample MSE							
(1)Heterogeneous coef	-	-	7.017	6.941	6.997	7.027	6.962
(2)Sectoral-heterogeneous coef	-	-	6.999	6.939	6.985	7.000	**6.911**
(3)Homogeneous coef	-	-	6.981	**6.930**	6.967	6.982	**6.894**
(4)	6.940	7.702	-	-	-	-	-

Training period:03/01/2007-31/12/2008, Testing period :03/01/2009-31/12/2009

	Naive	BH-VAR	W_{empty}	W_{sector}	$W_{geographic}$	$W_{compustat}$	W_{news}
In Sample MSE							
(1)Heterogeneous coef	-	-	4.413	**4.258**	4.340	4.401	**4.188**
(2)Sectoral-heterogeneous coef	-	-	4.514	4.473	4.506	4.514	4.437
(3)Homogeneous coef	-	-	4.502	4.467	4.492	4.500	4.430
(4)	4.532	**1.315**	-	-	-	-	-
Out-of-Sample MSE							
(1)Heterogeneous coef	-	-	5.298	5.271	5.294	5.308	5.270
(2)Sectoral-heterogeneous coef	-	-	5.188	**5.125**	5.182	5.187	5.158
(3)Homogeneous coef	-	-	5.199	**5.147**	5.187	5.197	**5.150**
(4)	5.253	6.106	-	-	-	-	-

Training period:03/01/2008-31/12/2009, Testing period :03/01/2010-31/12/2010

	Naive	BH-VAR	W_{empty}	W_{sector}	$W_{geographic}$	$W_{compustat}$	W_{news}
In Sample MSE							
(1)Heterogeneous coef	-	-	5.964	**5.753**	5.879	5.952	**5.679**
(2)Sectoral-heterogeneous coef	-	-	6.055	5.984	6.043	6.053	5.999
(3)Homogeneous coef	-	-	6.044	5.986	6.031	6.042	5.965
(4)	6.095	**1.440**	-	-	-	-	-
Out-of-Sample MSE							
(1)Heterogeneous coef	-	-	1.858	1.843	1.852	1.859	**1.822**
(2)Sectoral-heterogeneous coef	-	-	1.841	1.827	1.838	1.840	**1.824**
(3)Homogeneous coef	-	-	1.843	1.831	1.838	1.842	**1.820**
(4)	1.856	2.345	-	-	-	-	-

Training period:03/01/2009-31/12/2010, Testing period :03/01/2011-31/12/2011

	Naive	BH-VAR	W_{empty}	W_{sector}	$W_{geographic}$	$W_{compustat}$	W_{news}
In Sample MSE							
(1)Heterogeneous coef	-	-	3.421	**3.294**	3.369	3.413	**3.256**
(2)Sectoral-heterogeneous coef	-	-	3.471	3.428	3.464	3.470	3.422
(3)Homogeneous coef	-	-	3.469	3.436	3.461	3.468	3.425
(4)	3.555	**1.298**	-	-	-	-	-
Out-of-Sample MSE							
(1)Heterogeneous coef	-	-	2.009	1.997	2.009	2.010	1.989
(2)Sectoral-heterogeneous coef	-	-	1.993	**1.978**	1.990	1.992	**1.974**
(3)Homogeneous coef	-	-	1.991	1.978	1.989	1.991	**1.972**
(4)	1.991	2.349	-	-	-	-	-

Training period:03/01/2010-31/12/2011, Testing period :03/01/2012-31/12/2012

	Naive	BH-VAR	W_{empty}	W_{sector}	$W_{geographic}$	$W_{compustat}$	W_{news}
In Sample MSE							
(1)Heterogeneous coef	-	-	1.888	**1.844**	1.867	1.885	**1.841**
(2)Sectoral-heterogeneous coef	-	-	1.898	1.883	1.894	1.897	1.873
(3)Homogeneous coef	-	-	1.897	1.885	1.893	1.896	1.873
(4)	1.923	**1.239**	-	-	-	-	-
Out-of-Sample MSE							
(1)Heterogeneous coef	-	-	1.820	1.796	1.814	1.821	**1.796**
(2)Sectoral-heterogeneous coef	-	-	1.813	1.799	1.810	1.812	**1.792**
(3)Homogeneous coef	-	-	1.814	1.802	1.810	1.813	**1.794**
(4)	1.820	1.965	-	-	-	-	-

Training period:03/01/2011-31/12/2012, Testing period :03/01/2013-31/12/2013

	Naive	BH-VAR	W_{empty}	W_{sector}	$W_{geographic}$	$W_{compustat}$	W_{news}
In Sample MSE							
(1)Heterogeneous coef	-	-	1.879	1.833	1.857	1.875	**1.831**
(2)Sectoral-heterogeneous coef	-	-	1.887	1.872	1.884	1.887	1.871
(3)Homogeneous coef	-	-	1.887	1.875	1.883	1.886	**1.867**
(4)	1.906	**1.192**	-	-	-	-	-
Out-of-Sample MSE							
(1)Heterogeneous coef	-	-	1.472	1.458	1.471	1.472	**1.442**
(2)Sectoral-heterogeneous coef	-	-	1.467	1.457	1.464	1.467	**1.454**
(3)Homogeneous coef	-	-	1.468	1.460	1.465	1.468	**1.451**
(4)	1.468	1.602	-	-	-	-	-

Training period:03/01/2012-31/12/2013, Testing period :03/01/2014-31/12/2014

	Naive	BH-VAR	W_{empty}	W_{sector}	$W_{geographic}$	$W_{compustat}$	W_{news}
In Sample MSE							
(1)Heterogeneous coef	-	-	1.613	**1.574**	1.594	1.609	**1.562**
(2)Sectoral-heterogeneous coef	-	-	1.622	1.610	1.617	1.621	1.597
(3)Homogeneous coef	-	-	1.622	1.612	1.619	1.621	1.598
(4)	1.644	**1.089**	-	-	-	-	-
Out-of-Sample MSE							
(1)Heterogeneous coef	-	-	1.381	1.357	1.377	1.381	**1.344**
(2)Sectoral-heterogeneous coef	-	-	1.375	1.363	1.369	1.374	**1.351**
(3)Homogeneous coef	-	-	1.376	1.366	1.371	1.375	**1.354**
(4)	1.370	1.479	-	-	-	-	-

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Hierarchical factor model	0.144	0.420	0.521	0.522	0.617	0.787

Table 17: Summary statistics for corss-sectional regression R^2 for the hierarchical PCA

	(1)AR terms					(2) spatial-temporal terms						(3) σ
	λ_1	λ_2	λ_3	λ_4	λ_5	ψ_0	ψ_1	ψ_2	ψ_3	ψ_4	ψ_5	σ
Median	-0.026	-0.014	-0.014	-0.009	-0.002	0.221	0.021	0.003	0.005	0.005	0.003	1.395
MG Estimates	-0.026	-0.014	-0.016	-0.009	-0.005	0.270	0.032	0.004	0.002	0.008	0.008	1.491
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.019)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)	(0.025)
% Sig (at 5%)	40.9%	23.2%	22.0%	18.6%	19.6%	77.2%	24.5%	22.1%	15.7%	16.4%	14.5%	-
Non-zero coef.	413	413	413	413	413	408	408	408	408	408	408	413

Table 18: **QML estimation results of heterogeneous spatial temporal model (2) using** $\hat{\epsilon}_{it}^{pca}$

	(1)AR terms					(2) spatial-temporal terms						(3) σ
	λ_1	λ_2	λ_3	λ_4	λ_5	ψ_0	ψ_1	ψ_2	ψ_3	ψ_4	ψ_5	σ
Panel A: Consumer												
Median	-0.024	-0.017	-0.010	-0.014	-0.004	0.164	0.015	0.003	0.002	0.002	0.011	1.359
MG Estimates	-0.026	-0.019	-0.010	-0.011	-0.006	0.201	0.010	0.015	0.003	-0.001	0.008	1.431
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.039)	(0.011)	(0.009)	(0.008)	(0.010)	(0.010)	(0.051)
% Sig(at 5%)	32.5%	22.1%	19.5%	16.9%	19.5%	74.0%	15.6%	9.1%	11.7%	11.7%	6.5%	-
Non-zero coef.	77	77	77	77	77	77	77	77	77	77	77	77
Panel B: Finance												
Median	-0.032	-0.014	-0.024	-0.013	0.000	0.179	0.035	-0.015	0.015	0.000	0.020	1.574
MG Estimates	-0.035	-0.019	-0.028	-0.022	-0.002	0.257	0.074	-0.016	0.009	0.026	0.028	1.751
	(0.008)	(0.006)	(0.005)	(0.005)	(0.005)	(0.055)	(0.024)	(0.018)	(0.017)	(0.015)	(0.017)	(0.070)
% Sig(at 5%)	50.7%	41.3%	42.7%	33.3%	33.3%	78.7%	42.7%	36.0%	32.0%	28.0%	23.0%	-
Non-zero coef.	75	75	75	75	75	74	74	74	74	74	74	75
Panel C: Health												
Median	-0.009	-0.010	-0.002	-0.002	0.005	0.149	0.031	0.015	0.014	0.009	0.005	1.379
MG Estimates	-0.012	-0.005	-0.008	-0.006	0.007	0.203	0.032	0.016	0.013	0.013	0.024	1.459
	(0.006)	(0.005)	(0.005)	(0.005)	(0.004)	(0.068)	(0.015)	(0.014)	(0.010)	(0.017)	(0.015)	(0.087)
% Sig(at 5%)	22.9%	11.4%	14.3%	14.3%	17.1%	74.3%	20.0%	17.1%	8.6%	11.4%	17.6%	-
Non-zero coef.	35	35	35	35	35	34	34	34	34	34	34	35
Panel D: Hitech												
Median	-0.041	-0.020	-0.014	-0.009	-0.004	0.136	-0.004	0.002	0.020	0.005	-0.006	1.440
MG Estimates	-0.033	-0.019	-0.013	-0.008	-0.008	0.183	0.011	0.005	0.004	0.006	-0.008	1.553
	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)	(0.040)	(0.011)	(0.011)	(0.009)	(0.010)	(0.011)	(0.058)
% Sig(at 5%)	53.4%	17.8%	9.6%	13.7%	11.0%	65.8%	15.1%	17.8%	5.5%	12.3%	11.0%	-
Non-zero coef.	73	73	73	73	73	73	73	73	73	73	73	73
Panel E: Manufacturing												
Median	-0.013	-0.003	-0.016	-0.003	-0.005	0.407	0.010	0.012	-0.001	0.003	0.003	1.255
MG Estimates	-0.018	-0.006	-0.017	-0.003	-0.010	0.411	0.025	0.007	-0.003	0.005	0.002	1.288
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.033)	(0.009)	(0.008)	(0.007)	(0.007)	(0.007)	(0.039)
% Sig(at 5%)	33.6%	19.1%	19.1%	19.1%	20.9%	79.1%	18.2%	20.9%	11.8%	14.5%	16.7%	-
Non-zero coef.	110	110	110	110	110	108	108	108	108	108	108	110
Panel F: Other												
Median	-0.045	-0.018	-0.014	-0.008	0.001	0.208	0.044	-0.002	-0.003	0.011	0.008	1.471
MG Estimates	-0.032	-0.018	-0.019	-0.006	0.001	0.225	0.049	-0.005	-0.012	-0.004	0.003	1.590
	(0.006)	(0.004)	(0.005)	(0.003)	(0.004)	(0.065)	(0.018)	(0.017)	(0.014)	(0.016)	(0.015)	(0.078)
% Sig(at 5%)	51.2%	23.3%	25.6%	7.0%	9.3%	76.7%	30.2%	20.9%	14.0%	7.0%	11.9%	-
Non-zero coef.	43	43	43	43	43	42	42	42	42	42	42	43

Table 19: **QML estimation results of heterogeneous spatial temporal model (2) using** $\hat{\epsilon}_{it}^{pca}$**, parameters summarized by sector.**

	(1)AR terms					(2) spatial-temporal terms						(3) σ
	λ_1	λ_2	λ_3	λ_4	λ_5	ψ_0	ψ_1	ψ_2	ψ_3	ψ_4	ψ_5	σ
Panel A: $W_{baseline}$												
Median	-0.026	-0.013	-0.013	-0.009	-0.004	0.252	0.024	0.010	-0.009	0.008	0.008	1.470
MG Estimates	-0.028	-0.015	-0.016	-0.010	-0.006	0.292	0.036	0.010	-0.007	0.007	0.011	1.561
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.021)	(0.007)	(0.006)	(0.005)	(0.005)	(0.006)	(0.028)
% Sig(at 5%)	38.7%	22.5%	19.9%	17.2%	19.4%	77.0%	21.1%	20.6%	15.2%	16.4%	15.7%	-
Non-zero coef.	413	413	413	413	413	408	408	408	408	408	408	413
Panel B: $W_{unweighted}$												
Median	-0.024	-0.014	-0.014	-0.009	-0.003	0.267	0.022	0.018	0.001	0.012	0.008	1.478
MG Estimates	-0.026	-0.014	-0.016	-0.010	-0.005	0.286	0.035	0.017	-0.008	0.013	0.010	1.571
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.023)	(0.009)	(0.007)	(0.007)	(0.007)	(0.007)	(0.028)
% Sig(at 5%)	36.3%	22.0%	19.1%	17.2%	14.8%	74.8%	21.3%	17.4%	15.0%	17.4%	14.0%	-
Non-zero coef.	413	413	413	413	413	408	408	408	408	408	408	413
Panel C: $W_{weighted}$												
Median	-0.025	-0.013	-0.013	-0.009	-0.004	0.248	0.024	0.008	-0.008	0.005	0.006	1.475
MG Estimates	-0.028	-0.015	-0.016	-0.010	-0.006	0.281	0.035	0.007	-0.007	0.006	0.010	1.563
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.020)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)	(0.028)
% Sig(at 5%)	38.0%	22.8%	19.9%	16.2%	19.6%	77.5%	19.6%	20.3%	16.9%	16.2%	15.7%	-
Non-zero coef.	413	413	413	413	413	408	408	408	408	408	408	413
Panel D: $W_{noncompetitive}$												
Median	-0.026	-0.013	-0.013	-0.009	-0.003	0.253	0.020	0.009	-0.004	0.009	0.006	1.464
MG Estimates	-0.027	-0.015	-0.016	-0.010	-0.006	0.282	0.036	0.008	-0.005	0.008	0.008	1.563
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.020)	(0.007)	(0.006)	(0.006)	(0.005)	(0.006)	(0.028)
% Sig(at 5%)	39.5%	21.6%	20.6%	17.2%	18.6%	75.8%	20.5%	19.8%	16.1%	16.1%	14.9%	-
Non-zero coef.	413	413	413	413	413	409	409	409	409	409	409	413
Panel E: $W_{persistent}$												
Median	-0.025	-0.012	-0.014	-0.008	-0.006	0.163	0.017	0.008	-0.004	0.005	0.008	1.486
MG Estimates	-0.028	-0.015	-0.017	-0.009	-0.006	0.200	0.029	0.006	-0.005	0.006	0.011	1.567
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.018)	(0.006)	(0.005)	(0.004)	(0.004)	(0.005)	(0.028)
% Sig(at 5%)	38.5%	21.6%	22.3%	16.7%	17.2%	69.8%	22.8%	20.1%	17.4%	17.4%	14.1%	-
Non-zero coef.	413	413	413	413	413	368	368	368	368	368	368	413
Panel F: $W_{intrasector}$												
Median	-0.023	-0.012	-0.014	-0.009	-0.003	0.029	0.012	0.005	-0.001	0.009	0.001	1.516
MG Estimates	-0.025	-0.013	-0.016	-0.009	-0.005	0.060	0.018	0.008	-0.000	0.002	0.000	1.597
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.014)	(0.007)	(0.005)	(0.006)	(0.005)	(0.005)	(0.028)
% Sig(at 5%)	38.7%	20.6%	19.6%	14.8%	16.7%	46.2%	20.4%	16.3%	16.3%	16.6%	14.5%	-
Non-zero coef.	413	413	413	413	413	392	392	392	392	392	392	413
Panel G:$W_{intraindustry}$												
Median	-0.024	-0.013	-0.014	-0.009	-0.003	0.089	0.017	0.010	-0.011	0.003	0.009	1.499
MG Estimates	-0.026	-0.014	-0.016	-0.010	-0.005	0.143	0.020	0.012	-0.010	0.002	0.012	1.586
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.018)	(0.006)	(0.005)	(0.006)	(0.005)	(0.005)	(0.028)
% Sig(at 5%)	38.0%	21.6%	19.4%	16.7%	16.2%	62.7%	19.9%	18.2%	17.4%	17.4%	16.2%	-
Non-zero coef.	413	413	413	413	413	407	407	407	407	407	407	413

Table 20: **QML estimation results of heterogeneous spatial temporal model (2) alternative adjacency matrices W.**

	(1)inter-sector spatial-temporal terms						(2)intra-sector spatial-temporal terms					
	ψ_0^+	ψ_1^+	ψ_2^+	ψ_3^+	ψ_4^+	ψ_5^+	ψ_0^-	ψ_1^-	ψ_2^-	ψ_3^-	ψ_4^-	ψ_5^-
Median	0.032	0.010	0.004	0.002	0.007	0.001	0.324	0.016	0.020	0.002	0.001	-0.003
MG Estimates	0.063	0.015	0.005	-0.001	0.003	0.001	0.282	0.029	0.009	-0.005	-0.005	-0.007
	(0.013)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)	(0.027)	(0.010)	(0.009)	(0.010)	(0.009)	(0.008)
% Sig (at 5%)	45.7%	18.4%	16.1%	17.6%	15.8%	14.5%	75.5%	22.0%	17.7%	18.4%	14.3%	13.3%
Non-zero coef.	392	392	392	392	392	392	413	413	413	413	413	413

Table 21: **QML estimation results of two W heterogeneous spatial temporal model (18).** $\psi_0^+, \dots \psi_5^+$ correspond to the spatial-temporal parameters associated with inter-sector links. And $\psi_0^-, \dots \psi_5^-$ correspond to the spatial-temporal parameters associated with intra-sector links. We use Fama-French 6 sector definition described in section 4.1. Only estimated results for spatial-temporal terms are presented here. Full set of estimates are available upon request.

	(1)inter-sector spatial-temporal terms						(2)intra-sector spatial-temporal terms					
	ψ_0^+	ψ_1^+	ψ_2^+	ψ_3^+	ψ_4^+	ψ_5^+	ψ_0^-	ψ_1^-	ψ_2^-	ψ_3^-	ψ_4^-	ψ_5^-
Panel A: Consumer												
Median	0.062	0.011	0.003	-0.001	0.004	0.001	0.314	-0.001	0.001	-0.004	0.017	-0.012
MG Estimates	0.086	0.021	0.005	0.008	0.011	-0.002	0.216	0.019	0.002	-0.012	0.047	-0.014
	(0.025)	(0.010)	(0.010)	(0.009)	(0.010)	(0.009)	(0.073)	(0.020)	(0.021)	(0.020)	(0.018)	(0.017)
% Sig(at 5%)	48.7%	9.2%	6.6%	19.7%	17.1%	6.6%	84.4%	22.1%	16.9%	16.9%	6.5%	9.1%
Non-zero coef.	76	76	76	76	76	76	77	77	77	77	77	77
Panel B: Finance												
Median	-0.016	0.025	0.008	0.001	-0.007	0.017	0.437	-0.025	0.017	0.021	0.024	-0.003
MG Estimates	-0.093	0.029	0.002	0.031	-0.021	0.018	0.469	-0.003	0.004	0.040	0.042	-0.007
	(0.042)	(0.023)	(0.016)	(0.020)	(0.016)	(0.018)	(0.054)	(0.026)	(0.023)	(0.024)	(0.026)	(0.018)
% Sig(at 5%)	46.3%	23.9%	11.9%	23.9%	14.9%	17.9%	81.3%	33.3%	29.3%	33.3%	32.0%	22.7%
Non-zero coef.	67	67	67	67	67	67	75	75	75	75	75	75
Panel C: Health												
Median	0.048	0.010	-0.004	0.003	0.019	0.020	0.262	-0.006	0.012	0.001	0.025	-0.012
MG Estimates	0.040	0.003	0.005	0.003	0.031	0.030	0.075	-0.009	0.012	-0.015	0.002	0.018
	(0.022)	(0.011)	(0.014)	(0.010)	(0.018)	(0.009)	(0.093)	(0.023)	(0.021)	(0.026)	(0.027)	(0.026)
% Sig(at 5%)	39.4%	6.1%	9.1%	9.1%	12.1%	9.1%	80.0%	5.7%	14.3%	11.4%	11.4%	5.7%
Non-zero coef.	33	33	33	33	33	33	35	35	35	35	35	35
Panel D: Hitech												
Median	0.057	-0.012	-0.008	0.013	0.009	-0.010	0.219	0.003	0.041	-0.002	-0.033	0.003
MG Estimates	0.090	-0.011	-0.017	0.004	-0.002	-0.019	0.256	0.015	-0.003	0.009	-0.033	0.004
	(0.024)	(0.010)	(0.009)	(0.008)	(0.009)	(0.009)	(0.063)	(0.023)	(0.019)	(0.019)	(0.017)	(0.018)
% Sig(at 5%)	47.2%	12.5%	13.9%	6.9%	13.9%	12.5%	69.9%	26.0%	12.3%	8.2%	6.8%	9.6%
Non-zero coef.	72	72	72	72	72	72	73	73	73	73	73	73
Panel E: Manufacturing												
Median	0.033	0.012	0.014	-0.000	0.007	-0.005	0.381	0.073	0.013	0.002	0.008	-0.038
MG Estimates	0.137	0.014	0.021	-0.009	0.007	-0.000	0.313	0.060	-0.004	-0.040	-0.015	-0.015
	(0.025)	(0.008)	(0.008)	(0.008)	(0.007)	(0.006)	(0.048)	(0.024)	(0.023)	(0.026)	(0.018)	(0.020)
% Sig(at 5%)	46.1%	19.6%	24.5%	21.6%	12.7%	15.7%	74.5%	24.5%	17.3%	15.5%	13.6%	8.2%
Non-zero coef.	102	102	102	102	102	102	110	110	110	110	110	110
Panel F: Other												
Median	-0.011	0.023	0.004	-0.016	0.021	-0.003	0.207	0.049	0.084	0.016	-0.012	-0.066
MG Estimates	0.032	0.030	0.006	-0.038	0.012	-0.010	0.274	0.088	0.111	0.000	-0.021	-0.066
	(0.044)	(0.025)	(0.016)	(0.019)	(0.019)	(0.015)	(0.071)	(0.030)	(0.031)	(0.037)	(0.029)	(0.029)
% Sig(at 5%)	52.4%	31.0%	16.7%	14.3%	21.4%	11.9%	69.8%	20.9%	14.0%	18.6%	16.3%	18.6%
Non-zero coef.	42	42	42	42	42	42	43	43	43	43	43	43

Table 22: **QML estimation results of two W heterogeneous spatial temporal model (18), parameters summerised by sector.** $\psi_0^+, \dots \psi_5^+$ correspond to the spatial-temporal parameters associated with inter-sector links. And $\psi_0^-, \dots \psi_5^-$ correspond to the spatial-temporal parameters associated with intra-sector links. We use Fama-French 6 sector definition described in section 4.1. Only estimated results for spatial-temporal terms are presented here. Full set of estimates are available upon request.

	(1)inter-sector spatial-temporal terms						(2)intra-sector spatial-temporal terms					
	ψ_0^+	ψ_1^+	ψ_2^+	ψ_3^+	ψ_4^+	ψ_5^+	ψ_0^-	ψ_1^-	ψ_2^-	ψ_3^-	ψ_4^-	ψ_5^-
Median	0.077	0.016	0.010	-0.002	0.002	0.006	0.262	0.010	0.006	0.005	0.002	0.004
MG Estimates	0.132	0.019	0.012	-0.006	0.001	0.007	0.275	0.012	0.005	0.001	0.004	0.007
	(0.018)	(0.006)	(0.005)	(0.006)	(0.005)	(0.005)	(0.027)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)
% Sig (at 5%)	60.7%	21.6%	17.9%	16.0%	15.7%	17.9%	79.4%	26.2%	20.0%	17.8%	16.6%	16.9%
Non-zero coef.	407	407	407	407	407	407	320	320	320	320	320	320

Table 23: **QML estimation results of two W heterogeneous spatial temporal model (18).** $\psi_0^+, \ldots \psi_5^+$ correspond to the spatial-temporal parameters associated with inter-sector links. And $\psi_0^-, \ldots \psi_5^-$ correspond to the spatial-temporal parameters associated with intra-sector links. Stocks with same four-digits Standard Industrial Classification (SIC) code belong to the same industry. Only estimated results for spatial-temporal terms are presented here. Full set of estimates are available upon request.

	(1)inter-sector spatial-temporal terms						(2)intra-sector spatial-temporal terms					
	ψ_0^+	ψ_1^+	ψ_2^+	ψ_3^+	ψ_4^+	ψ_5^+	ψ_0^-	ψ_1^-	ψ_2^-	ψ_3^-	ψ_4^-	ψ_5^-
Panel A: Consumer												
Median	0.122	0.016	0.014	0.009	-0.001	-0.003	0.215	0.016	0.007	-0.005	-0.004	0.003
MG Estimates	0.169	0.020	0.025	0.014	-0.007	-0.012	0.134	0.018	0.006	-0.005	0.000	0.004
	(0.041)	(0.010)	(0.011)	(0.015)	(0.012)	(0.009)	(0.090)	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)
% Sig(at 5%)	66.2%	14.3%	9.1%	16.9%	14.3%	6.5%	72.9%	20.8%	10.4%	12.5%	14.6%	16.7%
Non-zero coef.	77	77	77	77	77	77	48	48	48	48	48	48
Panel B: Finance												
Median	0.019	0.007	0.013	-0.003	0.004	0.031	0.425	0.029	0.019	0.016	0.013	0.009
MG Estimates	0.095	0.004	0.025	0.001	0.000	0.046	0.414	0.024	0.001	0.008	0.019	0.014
	(0.044)	(0.018)	(0.016)	(0.021)	(0.014)	(0.014)	(0.059)	(0.015)	(0.012)	(0.010)	(0.010)	(0.010)
% Sig(at 5%)	59.5%	32.4%	29.7%	31.1%	21.6%	39.2%	91.5%	38.0%	50.7%	31.0%	33.8%	28.2%
Non-zero coef.	74	74	74	74	74	74	71	71	71	71	71	71
Panel C: Health												
Median	0.023	0.027	-0.007	0.012	0.007	0.029	0.088	-0.009	0.008	0.012	0.010	0.003
MG Estimates	-0.011	0.030	-0.000	0.007	0.012	0.038	0.079	-0.005	0.000	0.000	0.007	0.011
	(0.057)	(0.019)	(0.017)	(0.016)	(0.017)	(0.015)	(0.067)	(0.014)	(0.009)	(0.014)	(0.010)	(0.012)
% Sig(at 5%)	52.9%	23.5%	23.5%	8.8%	11.8%	14.7%	71.9%	25.0%	3.1%	9.4%	9.4%	6.2%
Non-zero coef.	34	34	34	34	34	34	32	32	32	32	32	32
Panel D: Hitech												
Median	0.053	0.012	-0.007	0.002	0.001	-0.002	0.169	0.019	0.006	0.004	-0.014	0.004
MG Estimates	0.050	0.010	-0.010	-0.006	0.006	-0.006	0.186	0.016	0.007	-0.003	-0.017	0.005
	(0.041)	(0.012)	(0.013)	(0.010)	(0.010)	(0.015)	(0.047)	(0.007)	(0.006)	(0.007)	(0.006)	(0.007)
% Sig(at 5%)	51.4%	18.1%	9.7%	2.8%	6.9%	11.1%	72.3%	20.0%	10.8%	12.3%	10.8%	13.8%
Non-zero coef.	72	72	72	72	72	72	65	65	65	65	65	65
Panel E: Manufacturing												
Median	0.155	0.025	0.017	-0.019	-0.002	0.003	0.470	-0.005	0.003	0.001	0.002	0.001
MG Estimates	0.233	0.026	0.012	-0.014	0.001	0.003	0.387	-0.003	0.007	-0.002	0.009	0.003
	(0.034)	(0.009)	(0.008)	(0.008)	(0.008)	(0.007)	(0.045)	(0.006)	(0.005)	(0.006)	(0.004)	(0.005)
% Sig(at 5%)	69.4%	18.5%	16.7%	18.5%	17.6%	16.7%	86.6%	22.0%	14.6%	18.3%	9.8%	13.4%
Non-zero coef.	108	108	108	108	108	108	82	82	82	82	82	82
Panel F: Other												
Median	0.041	0.021	-0.002	-0.026	0.020	-0.017	0.323	0.016	0.004	0.027	0.001	0.010
MG Estimates	0.123	0.031	0.010	-0.046	0.004	-0.022	0.258	0.029	0.005	0.014	0.004	0.013
	(0.063)	(0.025)	(0.020)	(0.019)	(0.020)	(0.018)	(0.106)	(0.012)	(0.011)	(0.009)	(0.009)	(0.009)
% Sig(at 5%)	52.4%	28.6%	26.2%	9.5%	21.4%	19.0%	59.1%	36.4%	13.6%	13.6%	18.2%	18.2%
Non-zero coef.	42	42	42	42	42	42	22	22	22	22	22	22

Table 24: **QML estimation results of two W heterogeneous spatial temporal model (18), parameters summarised by sector** $\psi_0^+, \ldots \psi_5^+$ correspond to the spatial-temporal parameters associated with inter-industry links. And $\psi_0^-, \ldots \psi_5^-$ correspond to the spatial-temporal parameters associated with intra-industry links. Stocks with same four-digits Standard Industrial Classification (SIC) code belong to the same industry. Only estimated results for spatial-temporal terms are presented here. Full set of estimates are available upon request.