What Economic Factors Underlie Connectedness in Corporate Credit Default Swaps: News vs. Macroeconomic Factors?*

Andrew Castro
Michigan State University

Neville Francis†
University of North Carolina - Chapel Hill

February 14, 2018

Abstract

We examine the economic factors that underlie return and price volatility networks for corporate credit default swaps. After examining company-level networks we then aggregate the data to study sector-level net connectedness. By introducing company news measures and macroeconomic covariates, we explore their effects on net connectedness. We find that no one factor explains connectedness across all sectors. Additionally, while returns networks are mainly influenced by news releases and macroeconomic factors are the main impetus behind price volatility networks, neither explain the general changes to *net connectedness* over time.

JEL classification: C32, G23.

Keywords: connectedness, credit default swap, news, Vector autoregression

*We thank Klara Peter for helpful commentary and feedback.
†nfrancis@unc.edu

1. Introduction

The recent financial crisis of 2007 has fueled greater and renewed interest in credit markets. Relatively new securities, such as the credit default swap (CDS)-invented in 1994, and the collateralized debt obligation (CDO)-invented in 1987, are void of empirical studies. A large contributor to the lack of empirical studies is difficulty obtaining data.[1] Contrary to equity instruments, such as stocks, CDSs and CDOs are fixed-income instruments, which have underlying names owing to a debt. Despite these issues, the credit market, which includes but is not limited to bonds, CDOs, and CDSs, has grown rapidly since the inception of these new instruments. For instance, the CDO market grew from $69 billion in 2000 to $500 billion in 2006, while the CDS market grew (in notional out-standings) from $919 billion in 2000 to $62 trillion in 2006.

Credit derivatives are of particular interest due to the recent credit crisis, a period with low issuance of credit. Usually, a credit crisis is a tightening of loans due to high failure in the credit market. Due to the difficulty in acquiring data on credit derivatives, CDS indices are commonly used as proxies for the market. The market share of these instruments varies over time, but consistently composes a majority of the credit market. As such, they act as a well-fitted representation in order to study the events which cause high failure in the underlying asset(s). By understanding the factors that are intrinsic to the underlying corporate credit default swaps of these indices, the groundwork is laid for developing models which aim to predict, and possibly prevent, credit crises.

As with many financial instruments, credit derivatives are part of a network, an intuitive representation of the connections that are formed by the actions and reactions that take place in the market, where investors are constantly acting and reacting to new information. Connectedness represents an investor's expectations, as part of the evaluation of individual asset risk and overall portfolio risk. That is, these expectations of risk are partially a function of connectedness. This approach does not aim to calculate an individual investor's risk, but instead the network formed by many such individuals. As such, networks are the result of the aggregate of investors' expectations. Take, for example, news that a financial firm with a credit default swap (CDS) may crash. When an investor hears this news, he/she takes into consideration the risk of his/her portfolio, including a consideration of how this crash may affect other assets. Right away, we can see this consideration has formed a network of the original firm's CDS with other assets.

Networks are not exclusive to financial markets or even economics, as they are applicable, but not limited to, game theory, evolutionary biology, and social media. However, their study

[1]Data on CDOs is more difficult to obtain than on CDSs

and development are fairly new to the field of economics, so difficulty arises in constructing networks, as they have no inherent measure that can be gathered by studying investors.

We contribute to the literature in several ways. First, we study connectedness of firm-level CDS's which, as stated earlier, are important instruments in the credit market. Second, we study how the inclusion of news measures, covering various topics, and macroeconomic covariates impact connectedness. For the latter contribution we use principal component analysis to collapse firm-level data to sector-level. This allows us to analyze network features in a compact manner. By adding in a large data set of events, the connectedness measure is further developed as a tool for understanding the network, despite its open-ended interpretation.

We summarize our results as follows: The significant covariates for returns are news items (save for a few exceptions), while the significant covariates for price volatilities are macroeconomic covariates. Furthermore, the inclusion of our covariates lessens the variance of *net connectedness* for returns, while it increases the variance of *net connectedness* for price volatilities. Lastly, there are some factors which are common across sectors, such as *Acquisitions and Mergers*, while some factors are unique to a sector, such as *War Conflict* for the *Transportation* sector.

The remainder of the paper is as follows: Section 2 provides more background on credit derivatives and available data. Next, Section 3 gives an overview of the literature, both on credit derivatives, network theory, and estimation of networks. Section 4 discusses the theoretical setup following a modified Modern Portfolio Theory that includes connectedness as an input for investors risk expectation. Section 5 starts by detailing the construction of the Diebold-Yilmaz index in a simple case, before transitioning to the general measure. Section 6 provides information on the data, including construction of our news measure and considerations for the macroeconomic variables. Section 7 presents the empirical model, using a VAR with inclusion of our covariates. Section 8 discusses the results, beginning with a static analysis, then transitioning to the dynamic spillover index, and finally concluding with analysis of *net connectedness* without and with our covariates. Section 9 discusses the sensitivity of our dynamic results to use of the Cholesky to calculate connectedness, changes in the window size and changes in the H-step ahead forecast error choice. Lastly, Section 11 concludes the paper.

2. Background

Before examining the role of CDS's in crises and the tools used to evaluate their risk, we must first understand their operation. A credit default swap is a contract, in which a buyer makes a payment to an insurer for some fixed amount of time. Usually, a credit a default swap is traded in basis points (BPS), referred to as the spread (1 basis point = 0.01%), where the spread is the payment made over some reoccurring period, commonly each quarter.[2] The remaining principal is paid to the buyer by the insurer if the debtor defaults. In this way, the underlying asset(s) is insured against default by a credit default swap. Furthermore, CDS's are commonly used as measurements of default probability, as investors hedge against the default of the underlying (in the case of our paper, a corporate bond). As the spread of a CDS increases, investors are further aiming to protect against a credit event.

Among other fixed-income instruments is the collateralized debt obligation (CDO). A CDO begins when a special purpose vehicle (SPV), a specialized legal entity, aggregates the names into a CDO.[3] Each tranche from this aggregation process is given a rating by common credit rating agencies. Cash flows of principal and interest payments are directed into the tranches in order of decreasing seniority, with the highest senior tranche having the least risk. As a result, coupon payments are inversely related to tranche seniority. Let us take, for example, a CDO sliced into six tranches. The highest tranche is known as a senior or super senior, whereas the lower tranches are junior or mezzanine. The lowest tranche of all is called an equity tranche, and it has the highest risk of loss. Each tranche has a percent spread of N-M percent, where N is known as the attachment point and M as the detachment point. For the purposes of this example, we will assign the tranches, in order of increasing seniority, 0-3, 3-7, 7-10, 10-15, 15-30, 30-100 percent. If a name defaults, the equity tranche will absorb all losses up to 3 percent of the total notional value (detachment point). This absorption will continue up the rank of seniority until the senior tranche stops receiving cash flows. When credit losses exceed the percent of the tranche, the entire amount for the investor is eliminated, and the investor loses his/her principal, and coupon payments will cease.

Overall, the instruments of the credit market are highly interdependent. A CDO can be composed of any form of a loan, such as a mortgage, corporate bond, or government bond. This is also true for a CDS, which in the latter case, it is referred to as a sovereign credit default swap (SCDS). A credit default swap can also be taken out on a CDO or tranches of a CDO, in which case the investor is hedging against default of all or parts of the CDO. Lastly, a CDO can be composed of credit default swaps, which is referred to as a synthetic CDO (SCDO). In this way, a credit event on a group of loans can cause rippling effects through many instruments in the credit market. This makes the credit market of particular interest for the study of networks, as we wish to empirically estimate this connectedness. However,

[2]The spread is not to be confused with a bid-ask spread, as with stocks.

[3]A name is any asset where an investor can take a position on its default

many data complications arise, which limit our ability to study the entire market.

With the financial instruments of this study well defined, we can now discuss the issues of data and use of indices. First, data on credit default swaps are quite hard to procure. As there is no direct market, pricing information is hidden in contracts and communications between investors and issuers. The underlying names, which are of particular interest, are so vast, that investors do not know their specifications. Rather, only the rating information is known to them. For these reasons, several indices have been created to monitor and study the market indirectly. In this study, we use the CDX North American Investment Grade Index and the CDX North American High Yield Index. Both indices feature the top 125 liquid companies on a rolling six-month basis. These two indices are also of particular interest as they have index CDO tranches issued, tie to the CDX index. As opposed to CDO's, these CDX.HY and CDX.IG indices function as synthetic collateralized debt obligations. Despite this, they function well as proxies for the credit market. Several studies have found that the ratio of risk premium to the credit risk component is constant (Pan and Singleton (2008), and Longstaff et al. (2011)). Given this analysis only takes into account the ordinal properties of the connectedness estimation, there are no issues with the use of spreads as proxies.

3. Literature Review

We begin by reviewing literature on CDO's and CDS's, which is strongly focused on pricing models. In addition to establishing groundwork for the model in this paper, this review aims to demonstrate the deficit in studies focusing on other aspects of systemic risk and the need for further data on the credit market. After, we then transition to literature on the Diebold-Yilmaz methodology, beginning with Diebold and Yilmaz (2009).

The pricing of CDO's began when Li (2000) introduced a pricing model using a copula coefficient, which grew rapidly in popularity among financial markets. CDS pricing follows similarly, constructing a credit curve that considers probability of default. However, the incorrect use of this model has been cited as a major contributor of the sub-prime mortgage crisis (Duffie, Saita, and Wang 2007). In terms of a network, the underlying connectedness of credit derivatives begins in the pricing model, which takes into account a copula coefficient.[4]

Many investment banks now focus heavily on developing better models for the credit market, which differ heavily from the equity market in terms of accuracy and depth of knowledge. For studying the effect of events, in an aim to predict prices, Longstaff and Rajan (2008) examined a three-factor model, each factor being a firm failure, failure of

[4]a copula is a multivariate probability distribution for which the marginal probability of each variable is one form

multiple firms, or wide scale failure. Although they found that multiple factors were needed to better account for pricing using a Poisson process, they also found the current market to be efficient in its pricing of overall portfolios.

The Diebold-Yilmaz Connectedness Measure began in Diebold and Yilmaz (2009), which introduced the methodology and its derivatives, such as the spillover index.[5] Further work was Diebold and Yilmaz (2011), which examined the connectedness of five countries' equity markets. They found volatility spillovers were sensitive to economic events by aligning the occurrence of these select events with changes in connectedness. The literature rapidly adopted this methodology, applying it to other markets and data. Klößner and Sekkel (2014) examined policy spillovers.

For credit markets, the literature has recently expanded beyond pricing. Bostanci and Yilmaz (2015) applied the Diebold-Yilmaz connectedness index to sovereign credit default swaps for 38 countries between 2009 and 2014. They found two clusters within the network, composed of developed and developing countries. Within this, they determined emerging market countries were the dominant factor in sovereign credit risk shocks, while developed countries displayed less risk. Additionally, Fonseca and Gottschalk (2012) applied the Diebold-Yilmaz measure to firm-level CDS data from Asia-Pacific Markets, analyzing the relation between stocks and CDS spreads. They determined a discrepancy exists between the implied volatility at firm-level and the realized volatility at market-level.

There are also several papers whose aim is to expand the methodology and provide new algorithms. Baruník, Kočenda, and Vácha (2016) extended to capture "good" and "bad" spillovers for the US stock market. In this framework, the networks are differentiated for spillovers that increase or decrease the variable of interest. For instance, we can have positive return spillovers and negative return spillovers. When using a Cholesky decomposition, there is also consideration of the ordering. Diebold and Yilmaz 2009 resolved this by examining random orders. In later work, the average spillover was taken across random permutations, but this was shown in Klößner and Wagner (2014) to underestimate the spillover index by up to a third. Klößner and Wagner (2014) introduced an algorithm to calculate the full range of the spillover index and individual spillovers.

We also discuss the factors which can affect the underlying default probability. Although credit ratings agencies attempted to undertake this task, they largely failed to include a national default probability. Most agencies believed the diversification by location lowered risk, as geographic isolation would contain crises to the sub-network. Connectedness is largely a result of investors expectations and risk behaviors (Bostanci and Yilmaz 2015). As a result, factors, both observed and unobserved, affect connectedness. Thus far, the literature

[5] Full details for the methodology are given in Section 5

has examined a few particular events, such as Bernanke's speech during the financial crisis (Bostanci and Yilmaz 2015), but it becomes difficult to examine a large number of events and provide quantitative meaning on their affects for a network.

4. Theoretical Setup

We follow Modern Portfolio Theory (Markowitz 1952). We begin with a single period model, where an investor considers available credit default swaps available before making his/her investment decision, beginning with an expectation of risk. As a credit default swap is a contract, the market is less volatile, meaning an investor cannot trade assets as easily. In this sense, this greatly reduces the need for a multi-period model, as investor need not re-evaluate his/her position often.

We now consider a multi-period model, where the investor is able to invest again, due to more cash flows. Now, the network expectation is subject to change the impact of an event on an investor. As these markets are traded at high volumes, most investors are professional, meaning they have constant access to information. When an event occurs, an investor compares his/her expectation of the network to the current realization, and then changes his/her position depending on the difference between his/her expectation and the realization.

Let us denote the expected network as $E(\mathcal{O})$. Let the expectation of risk be $\mathbf{E}(R_i) = f(E(\mathcal{O}), o_i)$, where o_i are other factors. An investor places weight w_i on each asset as a split of his/her portfolio, where $\sum_i^n w_i = 1$. We define the portfolio return variance as:

$$\psi_p^2 = \sum_i w_i^2 \psi_i^2 + \sum_i \sum_j w_i w_j \psi_i \psi_j \rho_{ij} \tag{1}$$

where ψ is the standard deviation for the periodic returns of the asset, and ρ_{ij} is the correlation coefficient between the returns of asset i and asset j. We denote Portfolio return volatility as

$$\psi_\rho = \sqrt{\psi_\rho^2} \tag{2}$$

Investors then choose a budget set $B(p, \mathsf{f})$, given available funds f as:

$$B(p, \mathsf{f}) = \{x \in \mathbb{R}_+^n : p \cdot x \leq \mathsf{f}\}, \tag{3}$$

where p is the price vector for the CDS's and x is an vector of volumes for the CDS's. Given limited funds f, an investor chooses to invest in CDS's which fit his/her expectation of return. An investor considers both return volatility and expected risk when making an investment decision. Our main conclusion of this model is investors examine the given information and form an expected value of risk. When an exogenous shock occurs, investors change his/her expectation of risk. As such, the connectedness between CDO's for each company changes, affecting the overall network structure.

5. Connectedness

5.1. Network Theory and Analysis

A network $\mathcal{N}$ is a collection of M nodes (or vertices) and K edges. Intuitively, a network represents relationships between these objects (nodes) as edges. A network can be directed, meaning each pair of nodes can have at most 2 edges between them, or undirected, meaning two nodes share a single edge. Mathematically, a network is an adjacency matrix $\mathcal{A} = [\dashv_{ij}]$, where $\dashv_{ij}$ represents the edge from node i to node j. In an unweighted network, $\dashv_{ij} = 0$ or 1, with 0 being no relationship and 1 being a relationship. These types of networks are common in sociology, such as when studying social networks. In many economic applications, we would like to use directional, weighted networks, as we believe a set of nodes (such as stocks) have a relationship we would like to represent in a directed, numerically ordered manner.

Network Theory is a relatively new area of study and tools to understand networks are still being developed. We begin with the connectedness table (Table 1), where the entries from x_1 to x_N are analogous to the transpose adjacent matrix.

In estimating networks using variance decompositions, as discussed in the next section, we form M^2 edge weights, where entry $\tilde{\theta}_{ij}(H)$ gives the connectedness from node j to node i. Networks are often large, dense datasets, which make gleaning useful information an impossible task. We can use several methodologies to gain useful information about a network, starting with *net connectedness*. For a pair of nodes i, j we define *pairwise directional connectedness* as

$$C^H_{i\leftarrow j} = \tilde{\theta}_{ij}(H)$$

As Diebold and Yilmaz note, it is not necessarily true that $C^H_{i\leftarrow j} = C^H_{j\leftarrow i}$, so there are $N^2 - N$ pairwise edge weights. We then define *net pairwise directional connectedness* as

$$C^H_{ij} = C^H_{j\leftarrow i} - C^H_{i\leftarrow j}$$

Calculating the pairwise connectedness for all pairs allows a great degree of combinations for analysis. In similar convention, we examine *to connectedness* and *from connectedness*. *To connectedness* is the effect of a CDS i to the system, which is high-yield and investment grade CDS's. *From connectedness* is the effect of the system to a CDS i of either index (we provide detailed calculations of connectedness in the following section). We can also calculate *total directional connectedness from all others to i* as

$$C^H_{i\leftarrow\bullet} \sum_{\substack{j=1\\ j\neq i}}^{N} \tilde{\theta}_{ij}(H) \tag{4}$$

Similarly, we can calculate *total directional connectedness to all others from j* as

$$C^H_{\bullet\leftarrow\, j} \sum_{\substack{i=1\\ j\neq i}}^{N} \tilde{\theta}_{ij}(H) \tag{5}$$

Differencing these two total measures gives us *net total directional connectedness*, which we will from here on refer to as just *net connectedness*, defined as:

$$C^H_i = C^H_{\bullet\leftarrow\, i} - C^H_{i\leftarrow\bullet} \tag{6}$$

Another aspect of interest in networks is communities, which are groups of nodes that are more densely connected among each other than other nodes. In this paper, we use the Louvain Method to detect communities (Blondel et al. 2008). This method has been used before with the Diebold-Yilmaz Connectedness Measure in Bostanci and Yilmaz (2015) - see citation for details. The Louvain Method is an iterative process that maximizes modularity, a metric which measures how much more dense the detected community connections are compared to a random network. After detecting communities, we can examine which nodes (in this paper, companies or sectors) belong to the communities and the underlying features which drive their modularity.

5.2. Diebold-Yilmaz Connectedness Measure

We begin with N-dimensional VAR(p) model,

$$\boldsymbol{Y}_t = \boldsymbol{\Phi}_1 \boldsymbol{Y}_{t-1} + \ldots + \boldsymbol{\Phi}_p \boldsymbol{Y}_{t-p} + \varepsilon_t \tag{7}$$

where ε is white noise and $\boldsymbol{\Phi}_1, \ldots, \boldsymbol{\Phi}_p$ are coefficient matrices. Assuming the VAR(p)

process is invertible, it's MA(∞) representation becomes

$$\boldsymbol{Y}_t = \varepsilon_t + \boldsymbol{A}_1\varepsilon_{t-1} + \boldsymbol{A}_2\varepsilon_{t-2} + \dots \tag{8}$$

with $N \times N$ matrices $\boldsymbol{A}_h$. Now for a H-step ahead forecast at time t, $P(Y_{t+h}|Y_t, Y_{t-1}, \dots)$ and for the corresponding forecast error

$$\boldsymbol{Y}_{t+H} - P(Y_{t+H}|Y_t, Y_{t-1}, \dots) = \varepsilon_{t+H} + \boldsymbol{A}_1\varepsilon_{t+H-1} + \boldsymbol{A}_2\varepsilon_{t+H-2} + \dots + \boldsymbol{A}_{H-1}\varepsilon_{t+1} \tag{9}$$

We denote $\boldsymbol{A}_0 := \boldsymbol{I}_N$, the forecast error's covariance matrix is $\boldsymbol{\Sigma}_\varepsilon^H = \sum_{h=0}^{H-1} \boldsymbol{A}_h \boldsymbol{\Sigma}_\varepsilon \boldsymbol{A}'_H$, with $\boldsymbol{\Sigma}_\varepsilon$ denoting the covariance matrix of ε.

5.2.1. Cholesky Decomposition

We make use of the lower-triangular Cholesky factor $\boldsymbol{L}$ of $\boldsymbol{\Sigma}_\varepsilon$, i.e. the lower-triangular matrix $\boldsymbol{L}$ with $\boldsymbol{L}\boldsymbol{L}' = \boldsymbol{\Sigma}_\varepsilon$ to decompose the forecast error variances. Using $\boldsymbol{L}$, $\boldsymbol{A}_h \boldsymbol{\Sigma}_\varepsilon \boldsymbol{A}'_h$ can be written as $(\boldsymbol{A}_h\boldsymbol{L})(\boldsymbol{A}_h\boldsymbol{L}')$, from which we have $(\boldsymbol{A}_h \boldsymbol{\Sigma}_\varepsilon \boldsymbol{A}'_h)_{ii} = \sum_{j=1}^{N} (\boldsymbol{A}_h\boldsymbol{L})^2_{ij}$ variable i's forecast error variance. Consequently,

$$\theta_{ij}(H) = \sum_{h=0}^{H-1} (\boldsymbol{A}_h\boldsymbol{L})^2_{ij} \tag{10}$$

may be considered as the contribution of (shocks to) variable j to variable i's forecast error variance. To form the spillover measure, we normalize $\theta_{ij}(H)$ as

$$\tilde{\theta}_{ij}(H) = \frac{\theta_{ij}(H)}{\sum_{j=1}^{N} \theta_{ij}(H)} \tag{11}$$

Where $\sum_{j=1}^{N} \tilde{\theta}_{ij}(H) = 1$, by construction. Finally, we have $\tilde{\theta_{ij}}(H)$ is the pairwise connectedness from variable j to variable i.

In the case where we cannot impose identifying assumptions on the ordering for the Cholesky decomposition, we may view the range of various spillover measures, such as *net connectedness*, by using a number of random permutation matrices. Consider a single permutation matrix $\boldsymbol{P}$, such that $\boldsymbol{P}_{ij} = 1$ if variable j appears at position i within the permuted matrix and $\boldsymbol{P}_{ij} = 0$ otherwise. We then apply the permutation matrix to the MA(∞) representation from Eq. (8) as

$$\boldsymbol{P}\boldsymbol{Y}_t = \varepsilon_{t,\boldsymbol{P}} + \boldsymbol{A}_{1,\boldsymbol{P}}\varepsilon_{t-1,\boldsymbol{P}} + \boldsymbol{A}_{2,\boldsymbol{P}}\varepsilon_{t-2,\boldsymbol{P}} + \dots \tag{12}$$

where $\varepsilon_{t,p} = \boldsymbol{P}\varepsilon_t$ and $\boldsymbol{A}_{h,\boldsymbol{P}} = \boldsymbol{P}\boldsymbol{A}_h\boldsymbol{P}'$ and with covariance matrix $\Sigma_{\varepsilon,\boldsymbol{P}} = \boldsymbol{P}\Sigma_\varepsilon\boldsymbol{P}'$. We then calculate the spillover as before. By using a number of these permutations randomly, we can view moments, such as the average, and the range for values such as individual spillover, *net connectedness*, and the spillover index.

5.2.2. Generalized Variance Decomposition

Cholesky decompositions are sensitive to ordering, so we may be reluctant to impose particular identifying assumptions; there are several ways to combat this sensitivity. One way, as in Diebold and Yilmaz (2009) is to take random permutations before the rotation and find the average, minimum, and maximum across these permutations. However, this was shown in Klößner and Wagner (2014) to underestimate the spillover index by up to a third.[6] Additionally, we can use generalized variance decompositions, which avoids using a rotation matrix and, thus, does not impose any identifying assumptions. Generalized variance decomposition (GVD) was introduced in Koop et al. (1996) and Pesaran and Shin (1998). To calculate it, we first measure pairwise directional connectedness, which measures the effect of company i on company j. This H-step-ahead generalized forecast error variance is given as:

$$\theta(H)^g_{ij} = \frac{\sigma_{jj}^{-1}\sum_{h=0}^{H-1}(\delta_i'\boldsymbol{A}_h\Sigma_\varepsilon\delta_j)^2}{\sum_{h=0}^{H-1}(\delta_i'\boldsymbol{A}_h\Sigma_\varepsilon\boldsymbol{A}_h'\delta_i)} \tag{13}$$

where σ_{jj}^{-1} is the inverse of the standard deviation of the error term for the j^{th} equation, and δ_i is a selection vector with one as the i^{th} element and zeros otherwise (Bostanci and Yilmaz 2015). Following Bostanci and Yilmaz (2015) we can normalize the measure as:

$$\tilde{\theta^g_{ij}}(H) = \frac{\theta^g_{ij}(H)}{\sum_{j=1}^{N}\theta^g_{ij}(H)} \tag{14}$$

As before, we have $\sum_{j=1}^{N}\tilde{\theta}^g_{ij}(H) = 1$, by construction. Similarly, $\tilde{\theta^g_{ij}}(H)$ is the pairwise connectedness from variable j to variable i under the GVD.

5.2.3. Spillover Index

The spillover index is a single measure which captures the total connectedness between all different pairs of assets (that is, except self-connectedness). We define the spillover index,

[6] Klößner and Wagner (2014) resolve this by using a permutation matrix in optimization problems, along with a "divide and conquer" strategy, that allows us to explore all VAR orderings of the Cholesky decomposition. Their algorithm is available for R in the CRAN repository under "fastSOM", which we utilize in calculating the spillover index and tabular values under a Cholesky decomposition, unless otherwise noted.

as in Diebold and Yilmaz (2009), as

$$SOI := 100 \times \frac{\sum_{\substack{i,j \\ i \neq j}}^{N} \tilde{\theta}_{ij}(H)}{N} \tag{15}$$

where N is the size of the matrix $\tilde{\theta}_{ij}(H)$. We can also form the spillover index for the GVD case, in which case we use $\tilde{\theta}^{g}_{ij}(H)$ instead of $\tilde{\theta}_{ij}(H)$.

5.2.4. Dynamic Analysis Methodology

The methodology discussed thus far is useful for giving a static view of connectedness, but a dynamic view is also of great interest. By employing a rolling window, we can view the connectedness up to the end of a window. Doing this over a series of several windows forms an overall series. Robustness checks are later done that analyze the effects of the window size on our results.

6. Data

6.1. Price Data

We use the CDX index, which is a daily price index for synthetic CDO's. The CDX index is distributed by Markit, a company which aggregates and parses CDS contracts into pricing data, and accessible at a number of financial data vendors. The series is rolled over every six months, where a dealer poll is used to drop illiquid firms. Markit makes these changes publicly available. Additionally, if a credit event occurs, the price index is stopped, re-evaluated, and started with a new combination of firms. A credit event is defined by the International Swaps and Derivatives Association (ISDA) as a bankruptcy, failure to pay, debt restructuring, or an obligation default/acceleration.

We aggregate data from 2003 to 2017 for all available companies. Since the market can be illiquid, we follow Bostanci and Yilmaz (2015) and impute missing days for up to 8 trading days. From here, we then calculate all available continuous series. Since the financial crisis is of particular interest, we then restrict the possible ranges such that they occur before 2006 and after 2010. After these restrictions, there are still several contiguous runs available, with a trade-off between the length of a run and the number of companies in a run. To make the choice of the run simple, we choose the run with the most companies. This run has 89 companies and 2,292 trading days (ranging from 2007-2013). In Fig. 2, we can see the distribution of companies in our sample among each sector.

In Table 2, we see the mean of the *Offer* is above the mean of the *Ask*, meaning investors who are attempting to sell the CDS value it above investors who are attempting to buy them. Additionally, the *Last* price of the spreads (given in basis points (BPS)) are well varied, so that when they are used to calculate returns, there is enough variance in the data for our regression.

6.1.1. Returns and Volatility

We use the daily difference of log spreads, as in Bostanci and Yilmaz (2015), which has several benefits. Log normal returns allow us to normalize to a common unit (returns) across different pricing units, as not all credit default swaps are traded in BPS. Given that the price of a CDS is normally distributed, then $\log(1 + r_i)$, where r_i is the daily return for given company i, is also normally distributed.

For estimating volatility, we use the method by Yang and Zhang (2000), which can handle overnight jumps and non-zero drift. We follow Yang and Zhang (2000) by keeping with the notation introduced by Garman and Klass (1980). Let $f \in [0, 1]$ be the fraction of the period that trading is closed, C_0 be the closing price of the previous period (at time 0), O_1 be the opening price of the current period (at time f), H_1 be the current period's high during the trading interval (between $[f, 1]$), L_1 be the current period's low during the trading interval (between $[f, 1]$), and C_1 be the closing price of the current period (at time 1). The variance (square of volatility) estimator is given by:

$$V = V_0 + kV_c + (1 - k)V_{RS} \tag{16}$$

where

$$\begin{aligned}
k &= \frac{0.34}{1.34 + \frac{n+1}{n-1}} \\
V_0 &= \frac{T}{n-1}\sum_{i=1}^{n}(o_i - \bar{o})^2 \\
V_c &= \frac{T}{n-1}\sum_{i=1}^{n}(c_i - \bar{c})^2 \\
V_{RS} &= \frac{T}{n}\sum_{i=1}^{n}[u_i(u_i - c_i) + d_i(d_i + c_i)]
\end{aligned}$$

where k is a constant chosen such that it minimizes the variance of the estimator V, T is the time period, n is the number of periods (in our case, $n = 191$, which results in $k = 0.144$),

$o = \ln O_1 - \ln C_0$, $c = \ln C_1 - \ln O_1$, $u = \ln H_1 - \ln O_1$, $d = \ln L_1 - \ln O_1$, $\bar{o} = \frac{1}{n}\sum_{i=1}^{n} o_i$, and $\bar{c} = \frac{1}{n}\sum_{i=1}^{n} c_i$. Finally, $T = 12$, which is the number of non-overlapping bins of size $n = 191$ in the sample.

6.2. *News Measure*

We capture news data on all firms available from accessible news sources to test whether these events have an impact on connectedness. News data is from RavenPack Analytics [Dow Jones Edition], a company which parses and categorizes news articles and press releases.[7] The database contains over 119 million observations of news releases from DowJones sources (Dow Jones Newswires, regional editions of the Wall Street Journal, Barron's and MarketWatch) at the milisecond level, along with the corresponding International Securities Identification Number (ISIN), a company identification code, for the topical company, sentiment analysis, and categorization into progressively finer topic levels. To transform this database we begin with the second coarsest level (Group) of categorization, which is composed of 51 possible different topics (not all topics appear in our sample, however). For the company-level analysis, we form a topical index for each company. Additionally, for sector-level analysis, we form a revenue weighted index for all available companies.

6.2.1. *Company Level*

We wish to construct an index for each company across the different news types. This index is denoted $F_{i,j,t}$, meaning the index for company i concerning topic j at time t. In order to construct an index for company-level events, we begin with $\mathbb{I}_{i,j,l,t}$ an indicator function for an event, defined as

$$\mathbb{I}_{i,j,l,t} = \begin{cases} 1 & \text{if the } l^{\text{th}} \text{ news release about topic } j \text{ for company } i \text{ occured at time } t \\ 0 & \text{otherwise} \end{cases} \tag{17}$$

To then aggregate these indicators to a single measure, we sum $\mathbb{I}_{i,j,t}$ across each day and divide by the maximum sum for the time period T, as:

$$F_{i,j,t} = \frac{\sum_{l=1}^{N} \mathbb{I}_{i,j,l,t}}{\max\limits_{\forall t}\{\sum_{l=1}^{N} \mathbb{I}_{i,j,l,t}\}} \tag{18}$$

where F is the total number of events for topic j and company i at time t. In sum, we

[7] http://www.ravenpack.com/

have constructed an index between 0 and 1 which indicates how much news activity there was for a topic on each day.

6.2.2. Sector Level

We construct a basic sector-level news measure by aggregating the sum of the company-level news indicator across each sector. Let $\mathcal{S} = \{\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_{m-1}, \mathcal{S}_{m-2}, \ldots, \mathcal{S}_q\}$ be the set of sectors, where $\mathcal{S}_m = \{\chi_1, \chi_2, \ldots, \chi_M\}$ for $M \leq N$, such that $\mathcal{S}_i \cap \mathcal{S}_j = \emptyset$ for all $i \neq j$, and where χ is the vector of companies. Thus, the basic sector level news measure $\mathbf{F}_m$ for sector m is given by

$$\mathbf{F}_{m,j,t} = \sum_{i \in M} F_{i,j,t} \tag{19}$$

For the sector level data, we include news data of companies outside the sample of those with credit default swaps. In this model, companies without credit default swaps can still affect those that do have CDS's.

6.3. Macroeconomic Data

Our remaining data includes macroeconomic covariates, denoted by the 9×1 vector W, which includes: 10-Year Treasury Constant Maturity Rate, Civilian Unemployment Rate, Consumer Price Index for All Urban Consumers, Effective Federal Funds Rate, Quarterly Seasonally Adjusted Gross Domestic Product, Real Median Household Income in the United States, Trade Weighted U.S. Dollar Index, University of Michigan: Consumer Sentiment, and University of Michigan: Inflation Expectation. All data are from Federal Reserve Economic Data (FRED) which are observed from 2007 to 2013. There are a few important aspects to note. First, except for both of the Michigan covariates (which are monthly), the macroeconomic covariates are quarterly, while our financial data is daily (trading days). Given several years of data, there is enough variation in the macroeconomic variables to make the regression meaningful. Second, the macroeconomic variables are the value for the prior quarter, so we use these to represent economic activity for each day. Additionally, the quarterly macroeconomic covariates act as controls for the economic climate, as well as other unseen factors not included in the news measure. Third, given the high frequency of the University of Michigan covariates (Consumer Sentiment and Inflation Expectation), they quite likely represent the expected value by consumers at the time of observation. We make the simlifying assumption that the investors for which we observe financial data are the same (of the same type) as the consumer forming expectations in these surveys. In this way, these two covariates do represent investor knowledge (expectation) at the time.

7. Empirical Model

7.1. Company Level

We follow Bostanci and Yilmaz (2015) and estimate a third order vector autoregression (VAR)

$$y_t = \sum_{h=1}^{3} B_h y_{t-h} + \varepsilon_t \tag{20}$$

where y_t is a 89×1 vector of returns or estimated volatilities for 89 companies at time t and A_h's are 89×89 coefficient matrices.

7.2. Sector Level

With 89 companies and a third order VAR, we would have 264 independent variables on the right-hand side. If we wished to include the company news measure F as a covariate, the regression becomes non-singular. We could impose identifying assumptions to reduce the number of covariates at the company level, but this still causes issues with non-singular matrix when estimating the coefficients. We resolve the issue of high-dimensionality by using principal component analysis (PCA) to view sector connectedness. Each credit default swap is grouped by its sector and PCA is then applied by group. For each of these PCAs, we use the first principal component when forming the new VAR, denoting the vector of first principal components as

$$z = [z_1, z_2, z_3, \ldots, z_{11}]^T, \tag{21}$$

where each z_i for $i = 1, 2, 3, \ldots, 11$ is a vector of the first principal component of returns or price volatilities for the 11 sectors.

The use of PCA follows intuitively, as it explains the greatest source of variance in the forecast error, which the index is meant to capture. Thus, we are able to gain a strong and accurate view of connectedness, without dealing with the issues of high-dimensional analysis.[8] PCA thus allows us an simpler network to interpret; instead of having to view 89 nodes and their $89 \times 2 = 178$ edge weights, we are limited to 10-20 nodes and their 20-40 edges, which is much more manageable. On the other hand, we lose information on

[8]We will discuss the specifics of the industries in the data and results section.

connectedness of specific companies, preventing us from viewing company-level spillovers and changes in its network over time.

In Table 3, we can see the amount of variance captured by the first principal component. The *Basic Industries* principal component captures the most variance for both returns and price volatilities, in columns 2 and 3, respectively. For the minimum variance, we have *Consumer Services* for returns in column 2 and *Finance* for price volatilities in column 3. Overall, we have an average of 51% and 86% of the variances of return and price volatility, respectively, captured by the first principal components.

We again estimate a VAR(3)

$$z_t = \sum_{h=1}^{3} B_h z_{t-h} + \Lambda \tilde{\mathbf{F}}_t + \Xi W_t + \varepsilon_t \tag{22}$$

where z_t is a 11×1 vector of returns or estimated volatilities for 11 sectors at time t, the A_h's are 11×11 coefficient matrices, the news measure is $\tilde{\mathbf{F}}$, and macroeconomic variables are W, with covariate matrices Λ and Ξ, respectively. Additionally, Λ is a $11 \times \dim(\tilde{F})$ coefficient matrix and D_t is a $\dim(\tilde{F}) \times 1$ matrix of $\dim(\tilde{F})$ vectors at time t, where $\dim(\tilde{F})$ is the total number of news covariates (the column size of $\tilde{\mathbf{F}}$. Lastly, Ξ is a $K \times \dim(W)$ coefficient matrix for the macroeconomic covariates W.

Each endogenous variable z_i, for $i = 1, \ldots, K$, has $\dim(\tilde{F})$ exogenous variables, meaning the $\dim(\tilde{F})$ variables are specific to sector i. We now discuss our identifying assumptions on the coefficient matrix Λ, imposing structure so that $\Lambda_{ij} = 0$ for $j \notin \{\dim(\tilde{\mathbf{F}}) \cdot (i-1) + 1, \dim(\tilde{\mathbf{F}}) \cdot (i-1) + 2, \ldots, \dim(\tilde{\mathbf{F}}) \cdot (i-1) + \dim(\tilde{\mathbf{F}})\}$, where $\tilde{\mathbf{F}}_m$ are the factors specific to sector m. However, we do not impose any structure on the A_h's or Ξ. In other words, factors for a sector only affect that specific sector contemporaneously, but macroeconomic variables and lags of the price are allowed to affect all sectors without restriction. This is not to say that factors of other sectors do not affect a particular sector. Instead, these factors first affect the price (and therefore the returns or volatilities) for the corresponding sector, which then affects all other sectors through their lags. In this way, there is a delayed response from investors reacting to shocks; investors spend time learning and evaluating the possible impacts of some shocks. Of course, the delay of the investors reaction may range widely amongst investors, so there is some importance in the unit t. A large t, such as a year, would imply a long, delayed reaction by the market in response to these shocks. As our unit is daily, we assume the market takes at least a day to react to shocks.

The inclusion of the news and macroeconomic covariates allows us to test for their significance in the VAR for sector returns and sector price volatilities After we determine which covariates are significant, we can then observe how they change the network. These results could be important for policymakers at all levels, as they allow us to distinguish, given the right conditions, how particular events underlie connectedness. For example, if we see a change in *net connectedness* for *Basic Industries* when including news for Regulation in the VAR, supposing that it is the only news at that time, then a policymaker could determine how further news on regulation would result in spillovers. This application is possible at all levels, from monetary policy at the Federal level to regulation at a state or even city level, depending on what level of connectedness we are examining.

8. Results

We begin by viewing static network plots for both company returns and price volatilities. As in Diebold and Yilmaz (2015), we use Gephi, a program for network plots and analysis, to graph the networks.[9] For all network graphs (including those later in this paper), we use ForceAtlas2, a node position algorithm which uses edge weights as a gravitational force.

Our company-level results show clear evidence of clustering which motivated us to aggregate to the sector-level. After analyzing the static sector-level networks for returns and price volatilities, we employ a rolling window to view changes in sector connectedness over time. As it is impractical to view the thousands of networks given by our rolling window analyses, we aggregate the sector connectedness measures to *net connectedness* of each sector. Lastly, we examine robustness checks to view the sensitivity of our results to the sizes of the rolling window and H-step ahead forecast error. We present our findings for the GVD measure of connectedness. Similar results were found using the Cholesky approach and are available upon request from the authors.

8.1. Static (Full-Sample, Unconditional) Company Level Results

Our analysis begins in Fig. 4, which is a network of company returns. The network is symmetrically connected for all companies, as it has a circular shape, indicating equal opposing and attracting forces. There are a few noticeable exceptions, such as the strong link between two blue nodes at the bottom of the graph, which are Campbell Food and General Mills, both in the *Consumer Non-Durables* sector. We note that these companies are in the same sector, which follows our intuition that companies in the same sector are more highly connected. Since shocks that affect a sector are propagated throughout all companies in the sector, the residuals are strongly correlated and, thus, their edge weight is stronger.

In Fig. 5, we plot a network for company price volatility. As opposed to company returns,

[9] Bastian, Heymann, and Jacomy (2009)

this network has more sparse regions along the outer edge. There still exists a strong centrality to the shape of the network, indicating equal opposing and attracting forces. Many of the companies in the same community, such as the orange in the bottom right, belong to the same sector. From the differing network shape, we infer that shocks affect volatility differently than returns or that the shocks themselves are different. In this way, price volatility has a greater range of spillovers, as there are asymmetric spillovers throughout the network, giving an asymmetric shape.

In Fig. 6, we see the distribution of communities amongst each sector. There are two clusters of bins, one for Returns and another for Price Volatility within each facet, with the number of bins in a cluster indicating the number of groups in a sector for each VAR. The bin width is irrelevant and simply fills the space. Each bin's height gives the number of companies in that sector for the relevant group. We note for many of the return networks, there is a strongly dominant community. While volatility also has a dominant community, there exist smaller communities in a sector. The return network is less sensitive to shocks that affect all sectors, while price volatility network is more sensitive to shocks that affect all sectors. Therefore, since companies belonging to the same sector are often more highly connected and thus form a group, we are motivated to reduce our data to a sector-level; this allows us to view sector spillovers and explore the factors which underlie the spillovers in the next subsection.

8.2. Static (Full-Sample, Unconditional) Sector Level Results

We begin this section with a static network of sector returns from our PCA analysis in Fig. 7. Now, we can distinguish interesting features of the network. For example, we observe that *Technology* is connected to many sectors, while not many sectors are connected to it. For example, in the center, we have a group of heavily connected sectors: *Public Utilities*, *Consumer Non-Durables*, *Consumer Services*, and *Capital Goods*. These central sectors all provide goods or services directly to households, while there are intermediary sectors between the outer edge networks. We expect the shocks to these sectors are common, such as changes in retail tax law (which would affect the aforementioned sectors), while not affecting other networks.

In Fig. 8, we plot the network of price volatilities. We see that *Consumer Durables*, *Consumer Non-Durables*, *Energy*, *Technology*, *Finance*, and *Consumer Services* are not as heavily connected as those in the center. Note the strong link between of *Energy* to *Public Utilities* and *Consumer Non-Durables*, which would follow as *Energy* is reliant on these sectors. *Consumer Services*, in particular, is far from many of the other sectors, indicating it is independent of shocks from all other sectors. Detecting communities as before, we note two communities, one of which contains *Consumer Non-Durables*, *Energy*, and *Finance*. This community has a common attribute of being fast-moving, meaning there is high quantity demanded and volatile changes are made to supply. The other community, consisting of *Consumer Services*, *Consumer Durables*, *Transportation*, *Capital Goods*, *Basic Industries*, *Health Care*, *Public Utilities*, and *Technology*, experiences a more consistent level of demand.

8.3. Dynamic (Rolling-Sample, Conditional) Sector Level Results

Due to dimensionality issues, we focus on sector level results for dynamic analysis, for which we use a window size of 500 and a 12-step ahead forecast error. We examine different network measures, starting from the broad view of the spillover index to the micro view of *net connectedness* by sector. In these analyses, we look at the changes before and after the addition of covariates and examine the significant factors.

8.3.1. Without Covariate

We begin with the spillover index in Fig. 9. For the index, we observe an increase in the average around 2008, at the beginning of the financial crisis. Although the index begins to decline thereafter, it does not reach it's previous trough. A sudden change occurs around 2010 that normalizes thereafter. This coincides with the standardization of credit default swaps, increasing their liquidity that, combined with a re-acceptance as a valuable finance instrument, renewed their growth. Throughout our analyses, we will see this trend in various measures. Hereafter, this event is referred to as the CDS big bang or big bang, as deemed in the previous literature.

In Fig. 10 we plot the spillover index for price volatility. Around 2008, we see a sharp increase in the spillover index, around the time of the financial crisis. We also see the advent of the big bang around mid-2010, with a return to normalcy thereafter.

8.3.2. With Covariate

We can see the effect of including covariates (news and macro variables) in the VAR for average sector Return *net connectedness* in Fig. 13 by comparing it to Fig. 12. Overall, many of the abrupt changes in *net connectedness* are lessened when covariates are added. For example, the sharp decrease for *Finance* in 2008 is lessened in magnitude. However, long-run trends, such as the slow fall of average *net connectedness* in *Technology* around 2010, have increased in magnitude.

For sector price volatility, we observe average sector *net connectedness* before the inclusion of covariates in Fig. 14. For many sectors, such as *Basic Industries*, the *net connectedness* measure is around 0, indicating equal *to connectedness* and *from connectedness*. There are notable exceptions, such as *Energy*, which experiences sharp movements in both directions (up and down). We expect, then, that energy shares covariates with other sectors the most compared to other sector covariates.

Introducing covariates to the volatility VAR in Fig. 15, we see the overall effect is an increase in variance of the *average net connectedness* measure by comparing it to Fig. 14. For example, the decrease in the average around 2008 for *Energy* is lower than the average for sector volatility *net connectedness* without the covariates. If *net connectedness* is becoming more varied, that means that *from connectedness* and *to connectedness* are also varied. In conclusion, including covariates for price volatility increases the variance of *net connectedness* for each sector.

We present the significant covariates in Table 4. For the Return VARs in the second column, we see that news items compose most significant covariates; macroeconomic covariates are not significant. We see that sectors that had strong connections share common covariates, such as news about Insider Trading among *Basic Industries*, *Capital Goods*, *Consumer Services*, and *Energy*. Additionally, we observe that there are factors unique to some sectors, such as Price Targets news for *Consumer Services* or War Conflict news for *Transportation*.

In the third column, we examine the significant covariates for sector volatilities. The most frequently occurring covariates are the macroeconomic ones. Thus, the volatility of the price resulting from changes in investor preferences is caused by the economic climate and macroeconomic forecasts. We do see some News Items, but these are less frequent. Interestingly, those News Items that are significant for volatility are not significant for returns, indicating different dynamics underlie volatilities and returns of CDS's.

9. Robustness

There are many considerations to take into account when testing the robustness of our results. In this section, we vary the window size from our rolling VAR for both returns and price volatility, as well as the H-step ahead forecast error and compare it to our baseline results; all baseline results are with a window size of 500 days and a 12-step ahead forecast error. We will view changes brought on by both decreases and increases for window size of 450, 500, and 550 and H-step ahead of 6, 12, and 18, giving us a total of 8 different results to compare to our baseline result of 500 window size and 12-step ahead error. Since this process entails the creation of many figures, we select certain sectors and results which we believe demonstrate the most dramatic change to our results, while maintaining their robustness. That is to say, even the most dramatic changes from varying window size and H-step do not significantly affect our results, thus our results are robust.

We begin with Fig. 16, where we plot the Spillover Index for Returns without Covariates across 3 different window sizes and 3 different H-step ahead forecast errors. We find, as expected, an increasing window size (as can be seen in all three rows going from left to right) decreases the variance of the spillover index. This to to be expected, as a larger window size means less variance when shift to the next iteration of the window (removing and adding a data point changes the data less for a larger window). Additionally, we see an increasing H-step, which can be seen in each column from top to bottom, does not significantly impact the spillover index. Thus, the spillover index is robust to changes in window size and H-step.

Next, we examine the robustness of average *Basic Industries* Returns *net connectedness* without covariates and in Fig. 17. We see, again, that a larger window size decreases the magnitude of the *net connectedness* measure. As the H-step ahead error size increases in each column from top to bottom, the magnitude of the *net connectedness* measure increases as well. Overall, the findings of our baseline remain the same among the various window sizes and H-step ahead errors. For our baseline window size (500) and H-step (12), we see, comparatively, there is a difference in the rise of *net connectedness* around 2011.

Lastly, we select and plot average *Energy* Price Volatility with covariates in Fig. 18. We see as window size increases, the magnitude decreases. Additionally, as H-step ahead error increases, the magnitude of *net connectedness* increases. Even with the smallest magnitude of *net connectedness* across our window size is 450 and 6-step ahead forecast errors, we still see a noticeable drop in *net connectedness* around 2008. Thus the trends we detect throughout our analyses are apparent despite changes in window size and H-step ahead error. Our choices find an appropriate middle ground that still covers sudden changes, such as the dramatic decrease around 2008, with the value not being as extreme as for a window size of 500 days and a 12-step ahead forecast error.

10. Extended News Measures

10.1. Revenue Weighted

Additionally, we expect some companies to play a more vital role in representing the stability (or instability) of a sector. The above sector-level news measures are the sum of the indicator measures for all companies in the sector, each company given an equal weight in the index. To represent the impact a company has on a sector, we use revenue data from

Institutional Brokers' Estimate System (I/B/E/S).[10] We take the sum of total revenue in the sector and use a company's share of the revenue as its weight in the aggregation function. Thus, we have

$$\tilde{\mathbf{F}}_{m,t} = \frac{\sum_{i \in M} |R_{i,t}| \mathbf{F}_{m,t}}{\sum_{i \in M} |R_{i,t}|} \tag{23}$$

where $R_{i,t}$ is the revenue for company i at time t.

To view some features of this measure, $\tilde{\mathbf{F}}$, we take the average of the measure across 11 sectors for each topic, which we plot in Fig. 3. We see some news items, such as *Investor Relations* and *Revenues*, have a quarterly frequency, as they are tied to quarterly company announcements. Other news items, such as *Analyst Ratings* and *Bankruptcy*, have a random distribution over time. Lastly, there are also news items which are rare, such as *Government* and *Crime*. The news measure also lends insight into the distribution of these items among companies. Take, for example, *Acquisitions/Mergers*, which has one of the larger ranges. This indicates that the revenues of the companies which belong to *Acquisitions/Mergers* are higher than other items, as the average of the measure is higher. We can also see news items which occur less, such as *Public Opinion*, with one of the smallest spreads.

10.2. Inclusion of Volatility

The news measures thus far have been some of a weighted mean. However, we lose information on high order moments. To control for this, to some extent, we include a volatility index along with a weighted average index.

10.3. Good News vs. Bad News

RavenPack Analytics features a sentiment index based on training data of an event being positive or negative. Thus, we may separate the news into good news (sentiment < 0.5) and bad news (sentiment > 0.5).[11]

11. Conclusion

We use the Diebold-Yilmaz connectedness measure to study the spillovers between 89 company credit default swaps over the period of 2007 to 2013. For company level-results,

[10] Although I/B/E/S is commonly used to evaluate analyst positions on measures of company performance, such as revenue, employment, etc., it also contains actual quarterly values for each company

[11] There may also be neutral news (exactly 0.50), but this is rare enough that it is useless to include

we find that the Return network is uniform in spillovers, while the Price Volatility network features more variance in spillovers. Additionally, many companies that were highly connected and formed communities in the network belonged to the same sector. We then sought to examine the effects of news items about all available firms from Dow Jones sources and general macroeconomic conditions on connectedness. Given parameter proliferation, we were forced to reduce the dimensionality in the VAR. To do so, we employ Principal Component Analysis by selecting the first principal component of the returns and price volatiles of the firms in each sector.

We find that inclusion of these covariates lessens the volatility of net sector return connectedness, while increasing volatility of net sector price volatility connectedness. For returns, news factors are frequently significant. However, for price volatility, macroeconomic factors are commonly significant. There do exist some news items which are significant for price volatility, but they differ from those that were significant for returns. Additionally, we see some common news items affect each sector, along with news items that only affect one sector. While we can never explain all factors that underlie connectedness, our approach allows to us to gain insight on connectedness by adding the ability to discuss significant factors, such as news items and macroeconomic conditions, among the companies of the network.

We conduct robustness exercises on the spillover index, sector-level *net connectedness* without covariates, and sector-level *net connectedness* with covariates by varying the window size and H-step ahead forecast error. We find a larger window size decreases the variance of all three results, while a larger forecast error increases sudden changes in *net connectedness*. Nevertheless, across different window sizes and H-step ahead forecast errors, we are able to distinguish our results and maintain they are robust.

Appendix A. Tables

Table 1: Connectedness Table

	x_1	x_2	$\cdots$	x_N	From Others
x_1	$\tilde{\theta}_{11}(H)$	$\tilde{\theta}_{12}(H)$	$\cdots$	$\tilde{\theta}_{1N}(H)$	$\sum_{\substack{j=1 \\ j\neq 1}}^{N} \tilde{\theta}_{1j}(H)$
x_2	$\tilde{\theta}_{11}(H)$	$\tilde{\theta}_{12}(H)$	$\cdots$	$\tilde{\theta}_{1N}(H)$	$\sum_{\substack{j=1 \\ j\neq 2}}^{N} \tilde{\theta}_{1j}(H)$
$\vdots$	$\vdots$	$\tilde{\theta}_{12}(H)$	$\ddots$	$\vdots$	$\vdots$
x_N	$\tilde{\theta}_{N1}(H)$	$\tilde{\theta}_{N2}(H)$	$\cdots$	$\tilde{\theta}_{NN}(H)$	$\sum_{\substack{j=1 \\ j\neq N}}^{N} \tilde{\theta}_{1j}(H)$
To Others	$\sum_{\substack{i=1 \\ i\neq 1}}^{N} \tilde{\theta}_{i1}(H)$	$\sum_{\substack{i=1 \\ i\neq 2}}^{N} \tilde{\theta}_{i2}(H)$	$\cdots$	$\sum_{\substack{i=1 \\ i\neq N}}^{N} \tilde{\theta}_{iN}(H)$	$\frac{1}{N} \sum_{\substack{i=1 \\ i\neq j}}^{N} \tilde{\theta}_{ij}(H)$

Connectedness table, as given in Diebold and Yılmaz (2014). Each entry $\tilde{\theta}_{ij}^{H}$ gives the connectedness from j to i, which is analgous to the transpose of the adjacency matrix.

Table 2: Summary Statistics

Variable	**Mean**	**Std. Dev.**
Last	112.725	173.773
Bid	109.146	169.769
Ask	116.209	177.675
Open	114.932	1085.593
Low	110.743	171.499
High	117.646	1143.145
$N = 89, T = 2,292$		

Average and standard deviation for each price category across all companies. *Last* is the last price of each trading day, *Open* is the opening price of each trading day, *Bid* is the last bididng price at which a trade was completed, *Ask* is the last asking price at which a trade was completed, *High* was highest price for a completed trade on each day, and *Low* is the lowest right for a completed trade on each day. All values are rounded to three decimal places.

Table 3: Principal Component Variance Share

Sector	**Returns**	**Volatilities**
Basic Industries	67.91	89.88
Capital Goods	45.45	89.81
Consumer Durables	51.14	79.49
Consumer Non-Durables	46.25	89.33
Consumer Services	33.42	78.77
Energy	52.90	87.28
Finance	49.63	76.29
Health Care	58.88	92.49
Public Utilities	51.99	83.31
Technology	45.26	83.97
Transportation	59.49	94.64
Average	51.12	85.93

Percentage of variance for returns and volatilities captured by the first principal component. All values are rounded to two decimal places.

Table 4: p-values and Coefficient Sign for Static VARs

Covariate	Return VARs	Volatility VARs
Basic Industries		
News Item		
Balance Of Payments	$.0000000000056^{+}$	.
Bankruptcy	0.0132^{+}	.
Credit Ratings	0.0861^{-}	.
Exploration	.	0.0672^{+}
Government	0.00158^{-}	.
Insider Trading	0.0433^{+}	.
Regulatory	0.0306^{+}	.
Revenues	0.0747^{-}	.
Macroeconomic		
Federal Funds	.	0.000000646^{-}
Inflation Expectation	.	0.0794^{+}
Consumer Sentiment	.	0.075^{-}
Unemployment Rate	.	0.000716^{-}
Capital Goods		
News Item		
Acquisitions Mergers	.	0.00601^{-}
Assets	0.0975^{+}	.
Credit	.	0.0195^{+}
Equity Actions	.	0.06^{+}
Indexes	0.00456^{+}	.
Insider Trading	0.0755^{+}	.
Marketing	.	0.00876^{-}
Products Services	0.0232^{-}	.
Macroeconomic		
CPI	.	0.0372^{+}
Federal Funds	.	0.00036^{-}
Inflation Expectation	.	0.0508^{+}
Dollar Index	.	0.0832^{-}
Consumer Sentiment	0.0187^{-}	0.00443^{-}
Consumer Durables		
News Item		
Bankruptcy	.	0.0121^{-}
Civil Unrest	.	0.000934^{+}
Corporate Responsibility	0.0952^{-}	.
Equity Actions	.	0.0404^{-}
Investor Relations	0.0298^{+}	.
Revenues	0.0347^{+}	.
Macroeconomic		

Continued on next page

Table 4 – continued from previous page

Covariate	Return VARs	Volatility VARs
Federal Funds	.	0.000000686^{-}
Inflation Expectation	.	0.0748^{+}
Unemployment Rate	.	0.0299^{-}
Consumer Non-Durables		
News Item		
Acquisitions Mergers	0.076^{+}	.
Corporate Responsibility	.	0.0769^{+}
Dividends	.	0.0237^{+}
Government	0.00549^{-}	.
Indexes	.	0.044^{-}
Order Imbalances	0.0782^{+}	.
Macroeconomic		
CPI	.	0.0134^{+}
Federal Funds	.	0.0000747^{-}
Inflation Expectation	.	0.0304^{+}
Consumer Sentiment	.	0.00387^{-}
Consumer Services		
News Item		
Acquisitions Mergers	0.0692^{+}	.
Civil Unrest	0.0467^{+}	.
Credit	0.0933^{-}	.
Insider Trading	.	0.00709^{-}
Labor Issues	.	.00000193^{+}
Marketing	.	0.00428^{+}
Partnerships	0.0152^{-}	.
Price Targets	0.0673^{-}	.
Macroeconomic		
CPI	0.00736^{+}	.
GDP	0.0615^{-}	.
Energy		
News Item		
Acquisitions Mergers	0.057^{-}	.
Dividends	.	0.00553^{-}
Earnings	.	0.031^{-}
Insider Trading	.	0.0169^{-}
Labor Issues	0.0428^{+}	.
Partnerships	.	0.0732^{+}
Price Targets	.0000549^{+}	.
Products Services	0.0894^{-}	.
Revenues	.	0.00934^{+}
Macroeconomic		

Continued on next page

Table 4 – continued from previous page

Covariate	Return VARs	Volatility VARs
CPI	.	0.0758^{+}
Dollar Index	.	0.0954^{-}
Finance		
News Item		
Acquisitions Mergers	.	0.06^{-}
Assets	0.0944^{-}	.
Balance Of Payments	0.0948^{+}	.
Bankruptcy	0.000313^{+}	.
Credit	0.0959^{-}	0.0148^{+}
Labor Issues	.	0.0155^{+}
Partnerships	.	0.0562^{-}
Regulatory	.	0.0246^{-}
Stock Prices	.	0.00646^{-}
Macroeconomic		
Treasury Rate	.	$.0000117^{+}$
GDP	0.072^{-}	.
Inflation Expectation	.	0.0976^{-}
Dollar Index	.	0.051^{-}
Consumer Sentiment	.	0.0178^{-}
Health Care		
News Item		
Credit Ratings	0.048^{+}	.
Equity Actions	0.0859^{+}	.
Indexes	0.0914^{+}	.
Legal	.	0.0368^{-}
Partnerships	.	0.0511^{+}
Macroeconomic		
CPI	.	0.0224^{+}
Treasury Rate	.	0.0629^{+}
Federal Funds	.	0.00000796^{-}
Inflation Expectation	.	0.00965^{+}
Consumer Sentiment	.	0.00405^{-}
Public Utilities		
News Item		
Assets	.	0.0258^{-}
Exploration	0.0233^{+}	.
Order Imbalances	.	0.018^{+}
Regulatory	.	0.00465^{-}
Revenues	.	0.0881^{-}
Macroeconomic		
CPI	.	0.015^{+}

Continued on next page

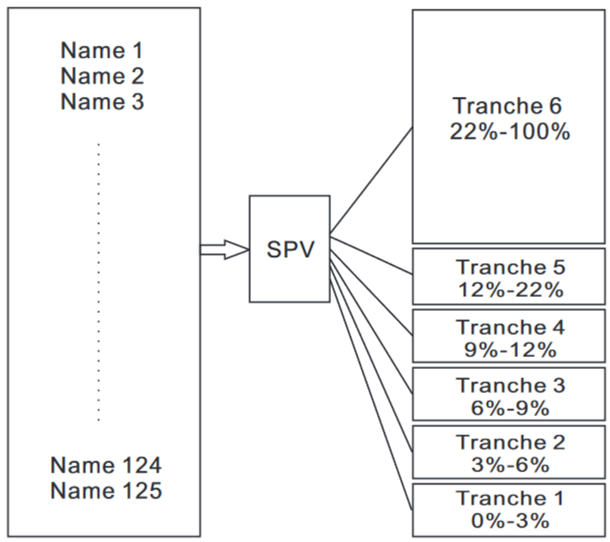
Table 4 – continued from previous page

Covariate	Return VARs	Volatility VARs
Federal Funds	.	0.0012^{-}
Inflation Expectation	.	0.0566^{+}
Consumer Sentiment	.	0.0962^{-}
Technology		
News Item		
Acquisitions Mergers	.	0.00174^{-}
Equity Actions	$.0000386^{+}$	.
Insider Trading	.	0.0894^{+}
Investor Relations	.	0.0571^{-}
Revenues	0.0653^{-}	.
Macroeconomic		
CPI	.	0.0671^{+}
Federal Funds	.	0.000395^{-}
Inflation Expectation	.	0.00273^{+}
Consumer Sentiment	.	0.0000754^{-}
Transportation		
News Item		
Assets	.	0.0266^{+}
Earnings	0.0398^{+}	.
Indexes	.	0.0573^{+}
Investor Relations	.	0.0452^{-}
Transportation	0.0000000247^{-}	.
War Conflict	0.00153^{-}	.
Macroeconomic		
CPI	.	0.0469^{+}
Federal Funds	.	0.000259^{-}
Inflation Expectation	.	0.0807^{+}
Dollar Index	.	0.0849^{-}
Consumer Sentiment	.	0.0189^{-}
Unemployment Rate	.	0.0406^{-}

p-values for covariates in each VAR that were of 10% significance or better. All values are given to three significant digits. A "." indicates the value was not significant in the estimated VAR. A minus sign ($^{-}$) indicates the sign of the coefficient in the VAR is negative, while a plus sign ($^{+}$) indicates the sign of the coefficient in the VAR is positive.

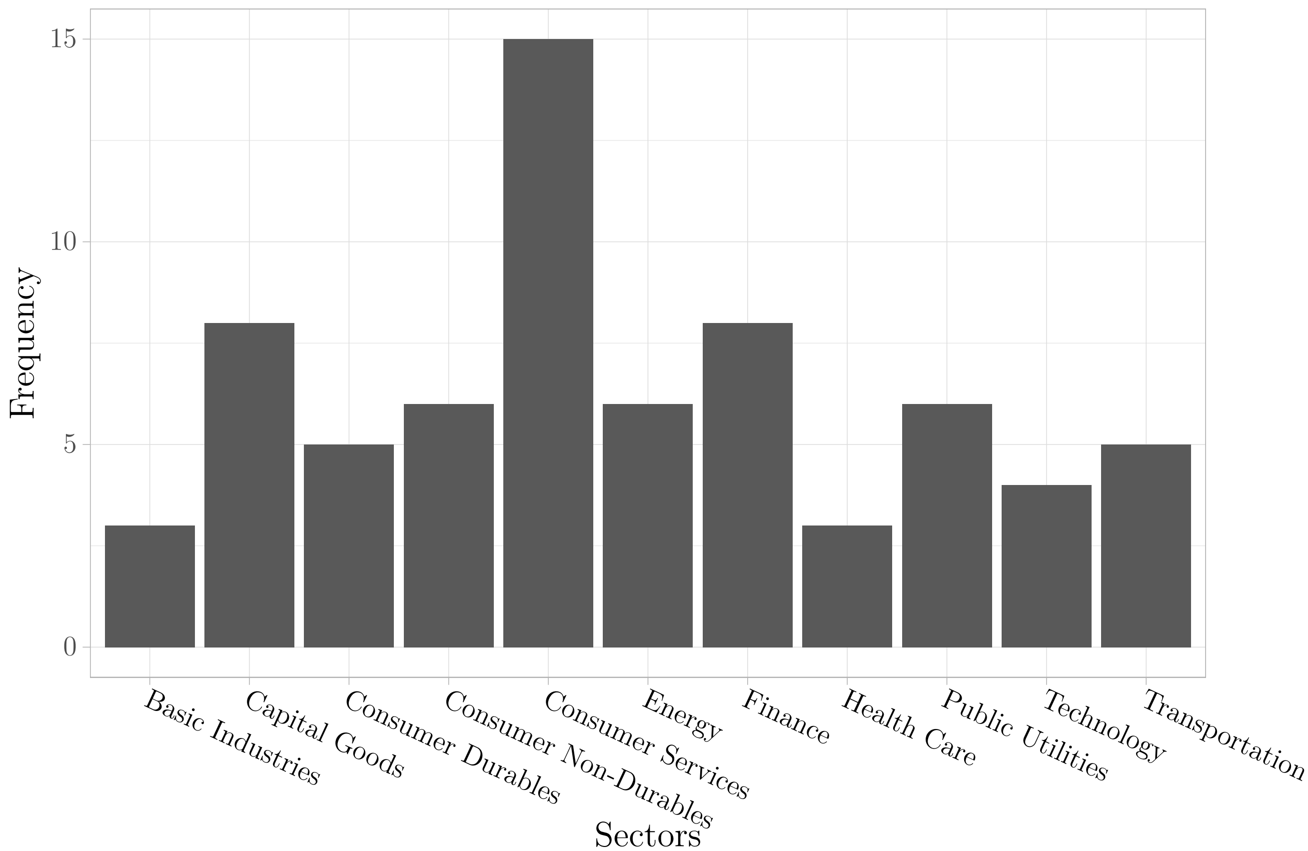
Appendix B. Figures

Fig. 1. CDO Structural Diagram



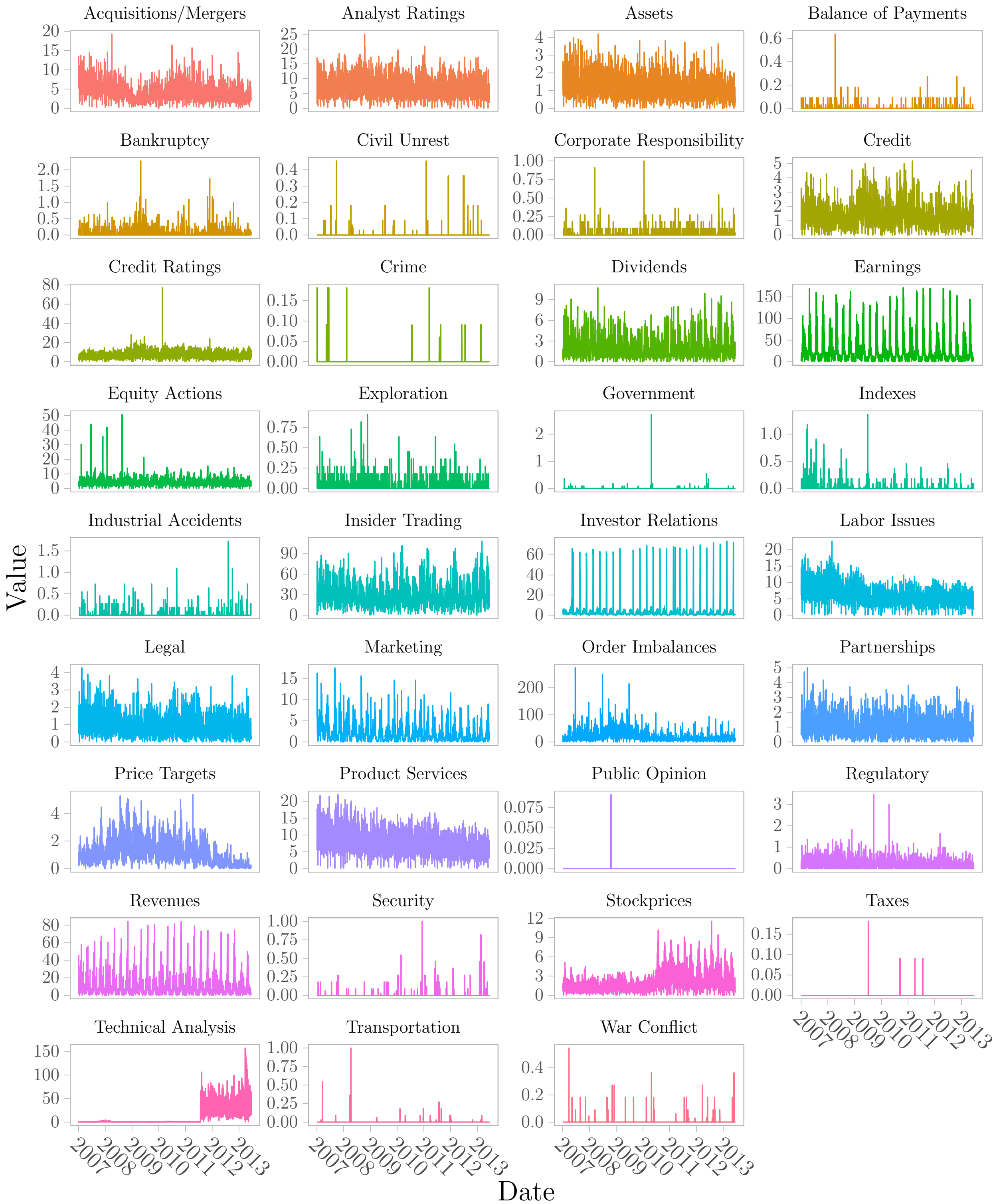
A typical CDO structure, including tranche percentages. A name would be any underlying asset, such as a bond, mortgage, etc. SPV is the special purpose vehicle, a corporate entity created specifically to create CDOs. Peng, Kou (2009)

Fig. 2. Distribution of Companies by Sector



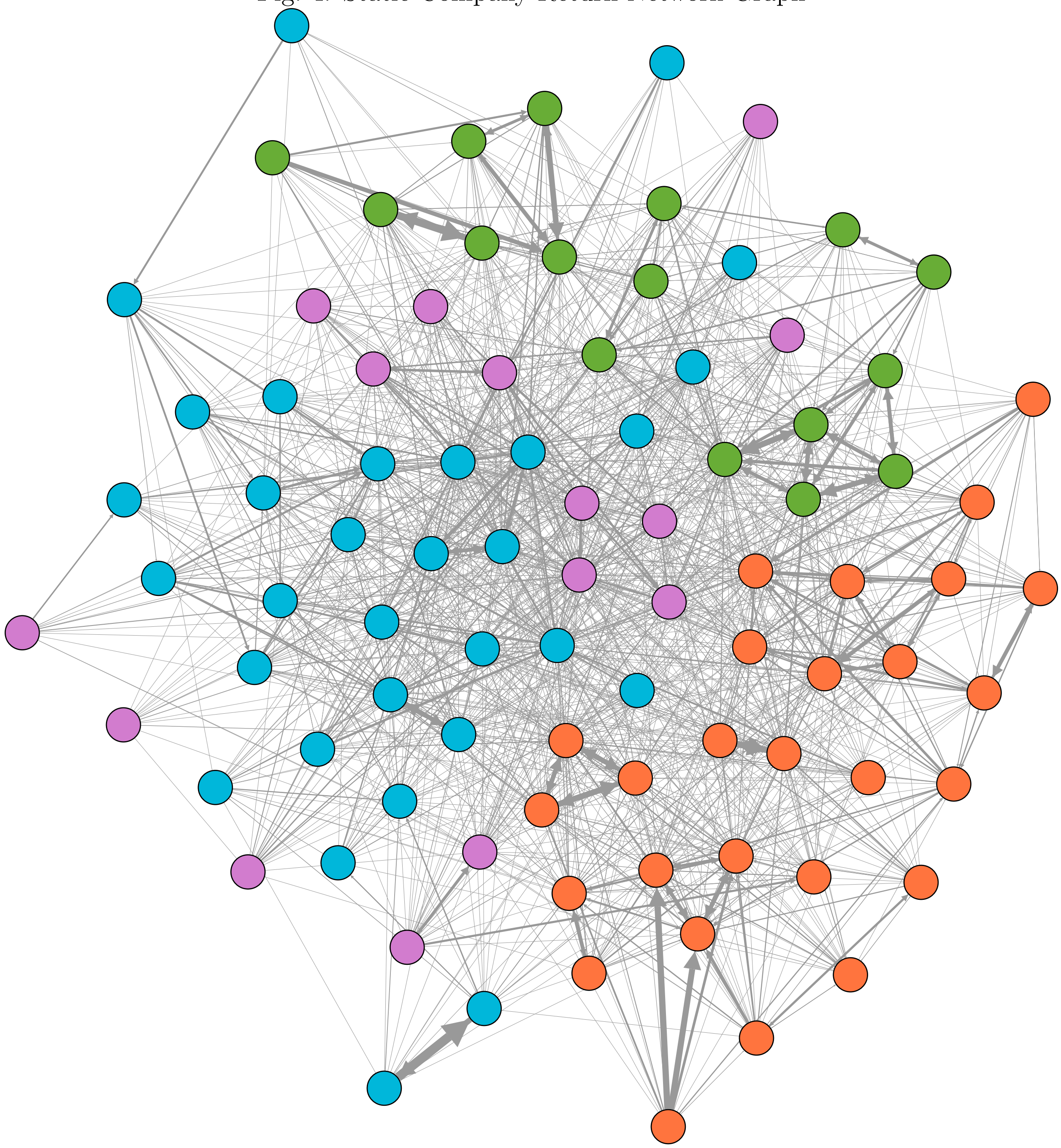
Histogram displaying number of companies in each sector.

Fig. 3. Average News Measure by Type



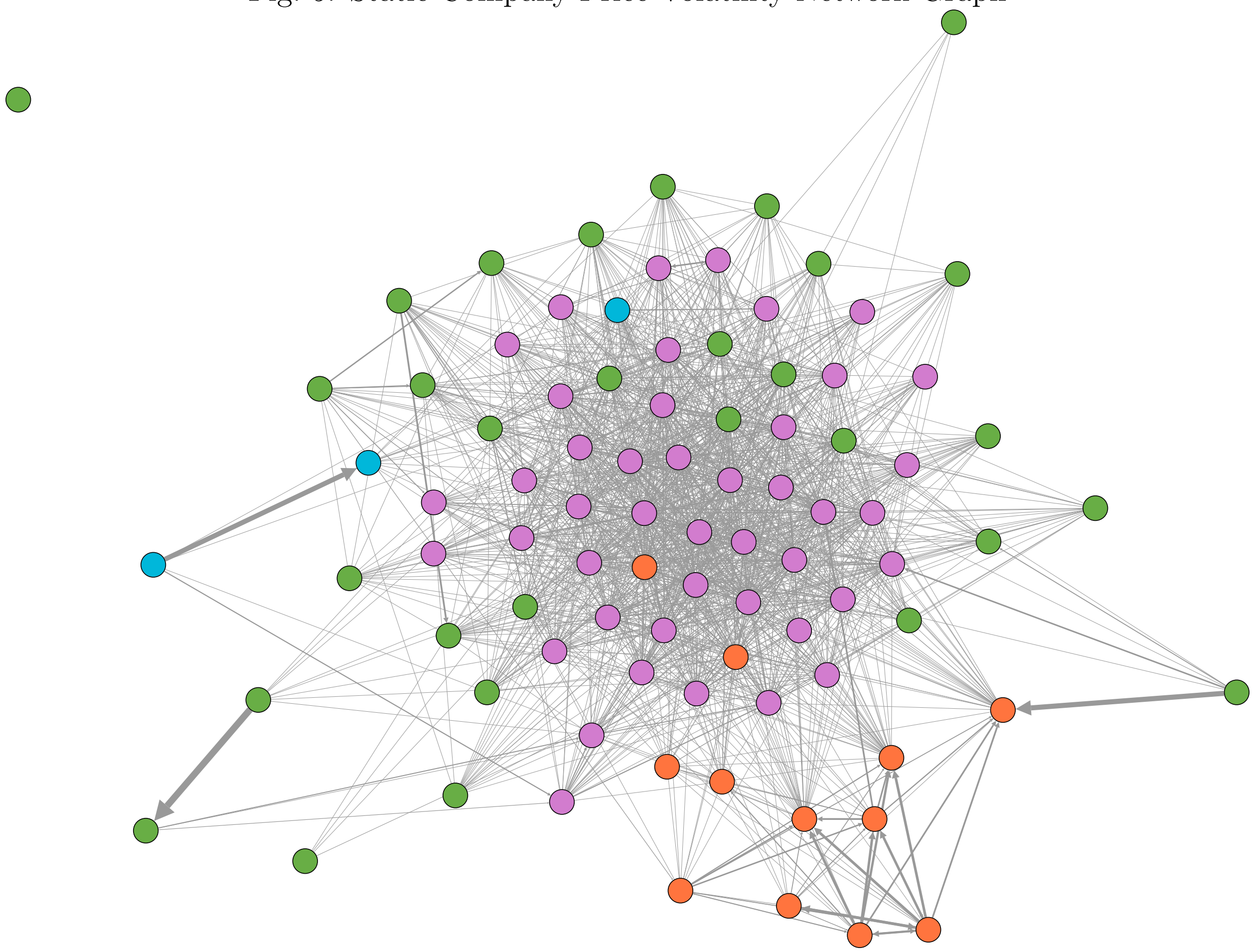
Average of the weighted sector news measure across all sectors, faceted by type.

Fig. 4. Static Company Return Network Graph



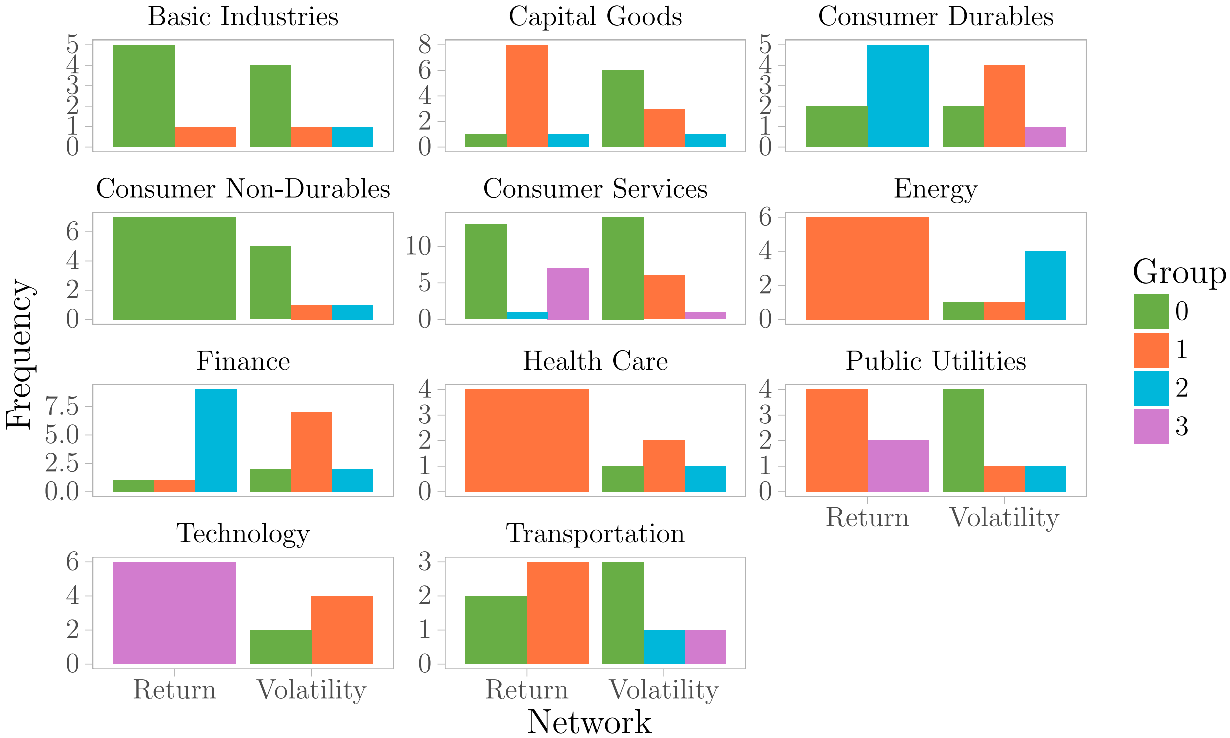
Network graph for returns of 89 companies CDS's. Communities are detected in Gephi using an algorithm by Blondel et al. (2008). Each color indicates a community (4 colors total). Size is indicative of the out-degree of connectedness for each company. The more a company affects all others, the larger the node. Edges are filtered so top 25% of edges by weight are visible. GVDs are used to decompose and H-step ahead error is 12.

Fig. 5. Static Company Price Volatility Network Graph



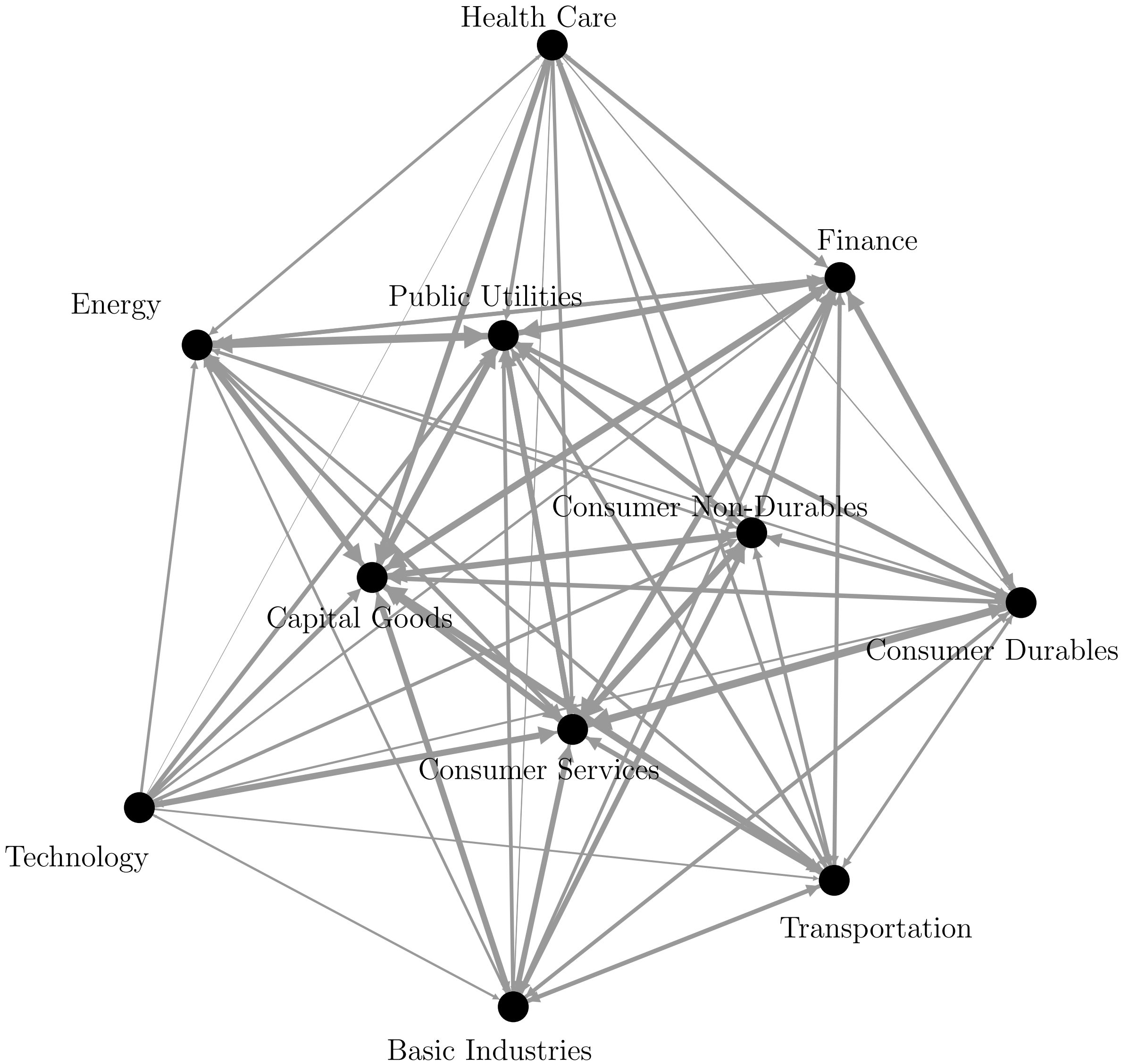
Network graph for price volatility of 89 company CDS's. Communities are detected in Gephi using an algorithm by Blondel et al. (2008). Each color indicates a community (4 colors total). Size is indicative of the out-degree of connectedness for each company. The more a company affects all others, the larger the node. Edges are filtered so top 25% of edges by weight are visible. GVDs are used to decompose and H-step ahead error is 12.

Fig. 6. Distribution of Companies Among Modularities



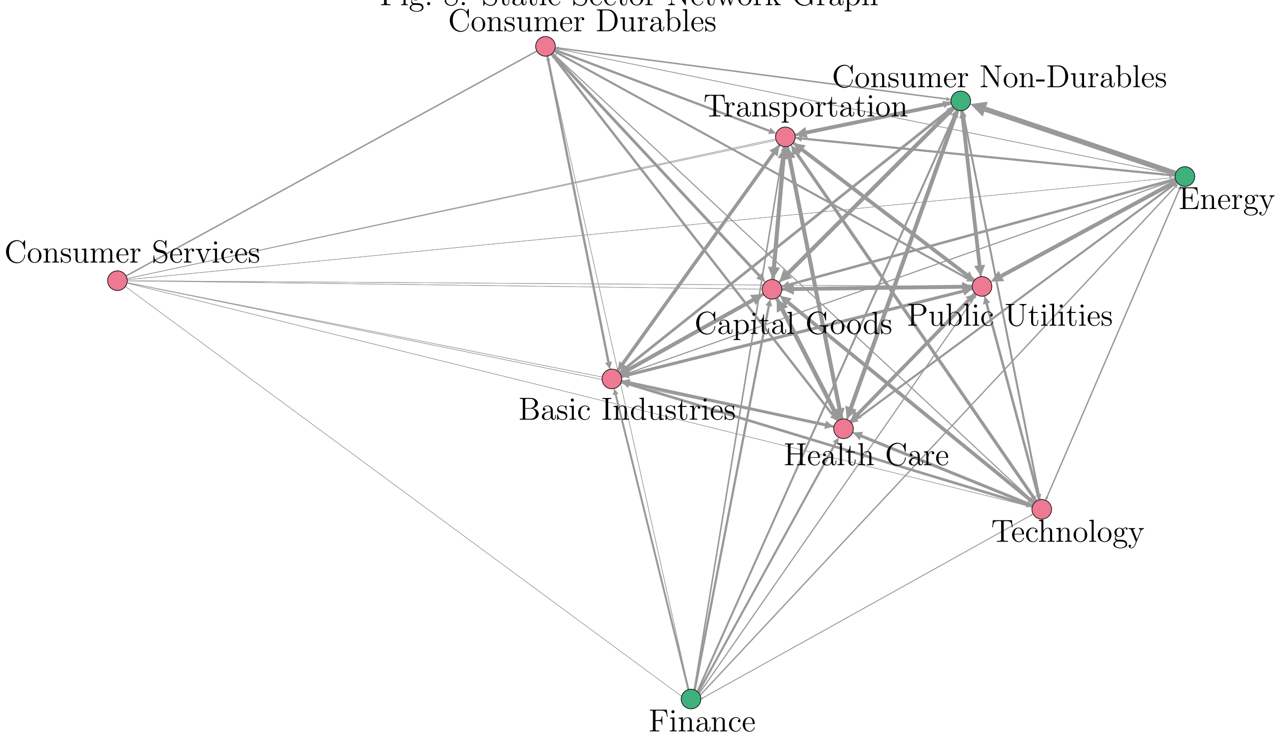
Communities (modularities) Blondel et al. (2008) in each static network for company returns and company price volatility. Each color for a group correspond to the node color in 4 and 5. The two groups are not necessarily the same, but represent detected groups in decreasing order of number of companies of group (0 being the highest, 3 being the lowest).

Fig. 7. Static Sector Network Graph



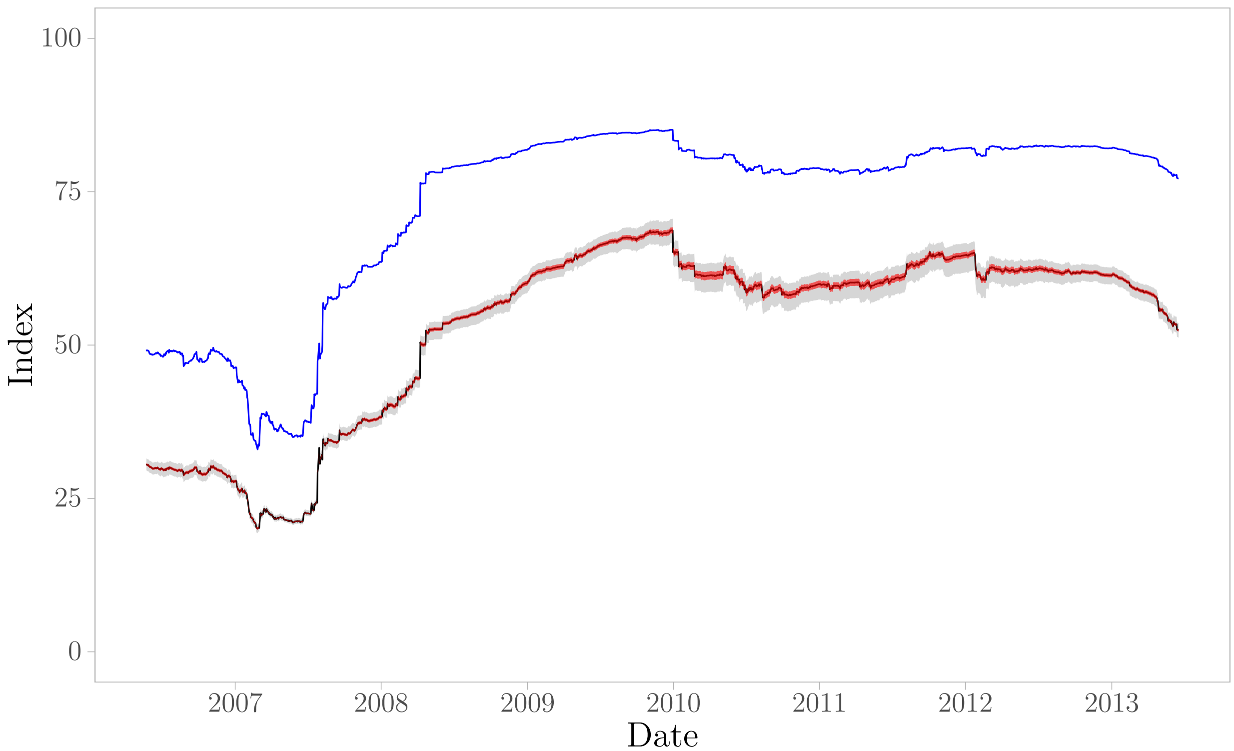
Network graph for First Principal Component of company returns by sector (11 sectors total). No community detection as in Figures 4, 5. All edges are visible. GVDs are used to decompose and H-step ahead error is 12.

Fig. 8. Static Sector Network Graph



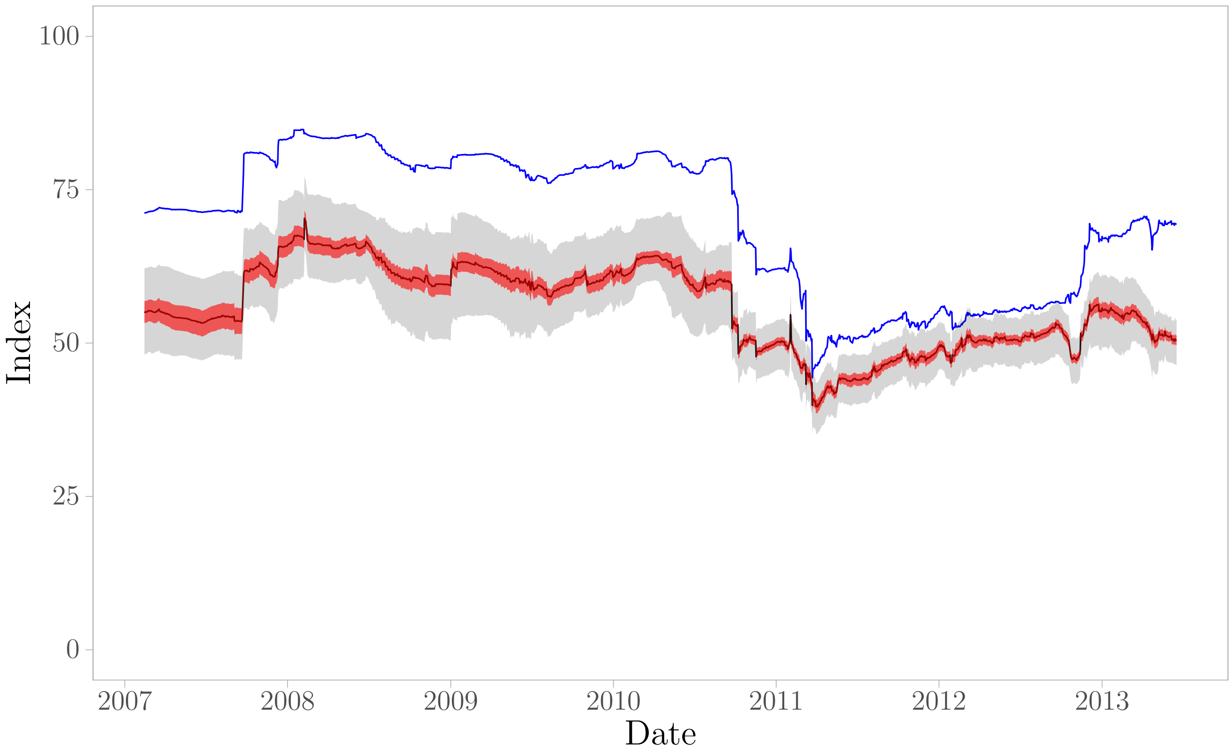
Network graph for First Principal Component of company price volatility by sector (11 sectors total). Communities are detected in Gephi using an algorithm by Blondel et al. (2008), with each color corresponding to community (2 colors total). All edges are visible. GVDs are used to decompose and H-step ahead error is 12.

Fig. 9. Spillover Index for Returns without Covariates



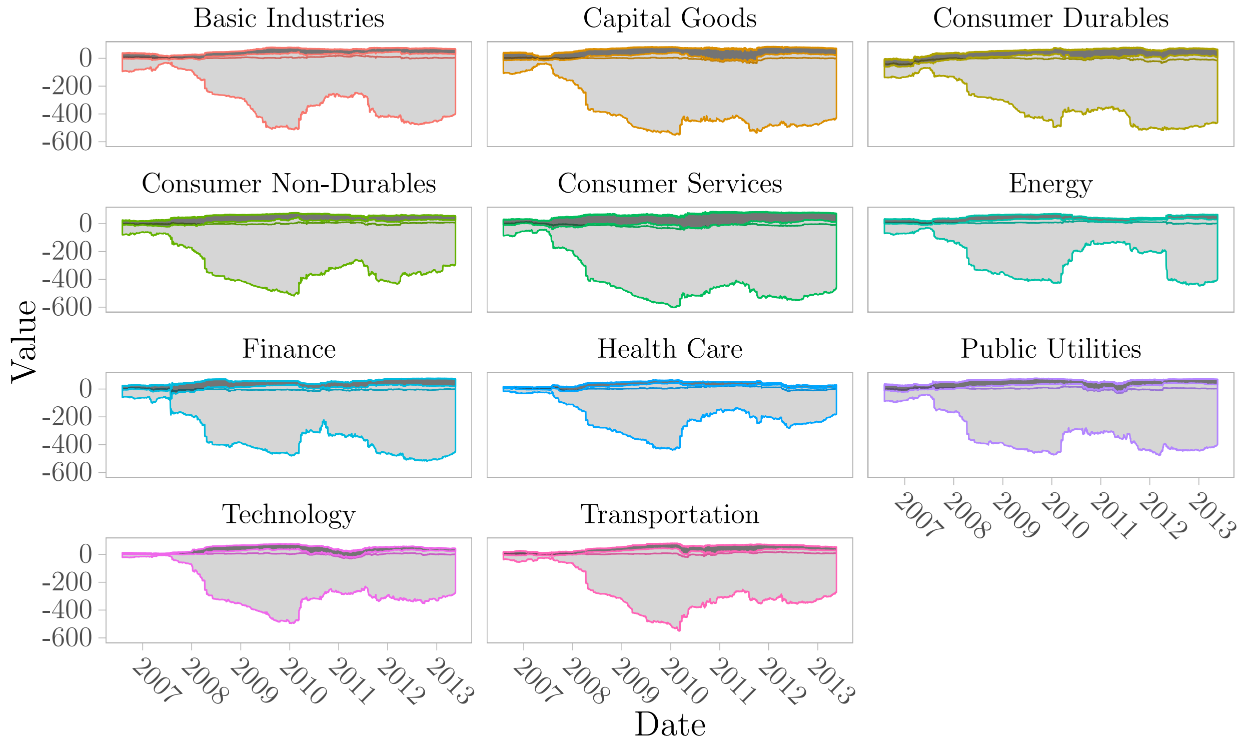
Diebold-Yilmaz spillover index for returns without sector-specific and macroeconomic covariates. Average value and maximum and minimum spillover index across all VAR orderings for Cholesky decomposition are given by the black line and gray ribbon, respectively. 25% and 75% bounds for the Cholesky Decomposition across 10,000 random permutations are given by the red ribbon. Index for generalized variance decomposition (GVD) is given by the blue line. Window is 500 days and H-step ahead error is 12.

Fig. 10. Spillover Index for Volatilities without Covariates



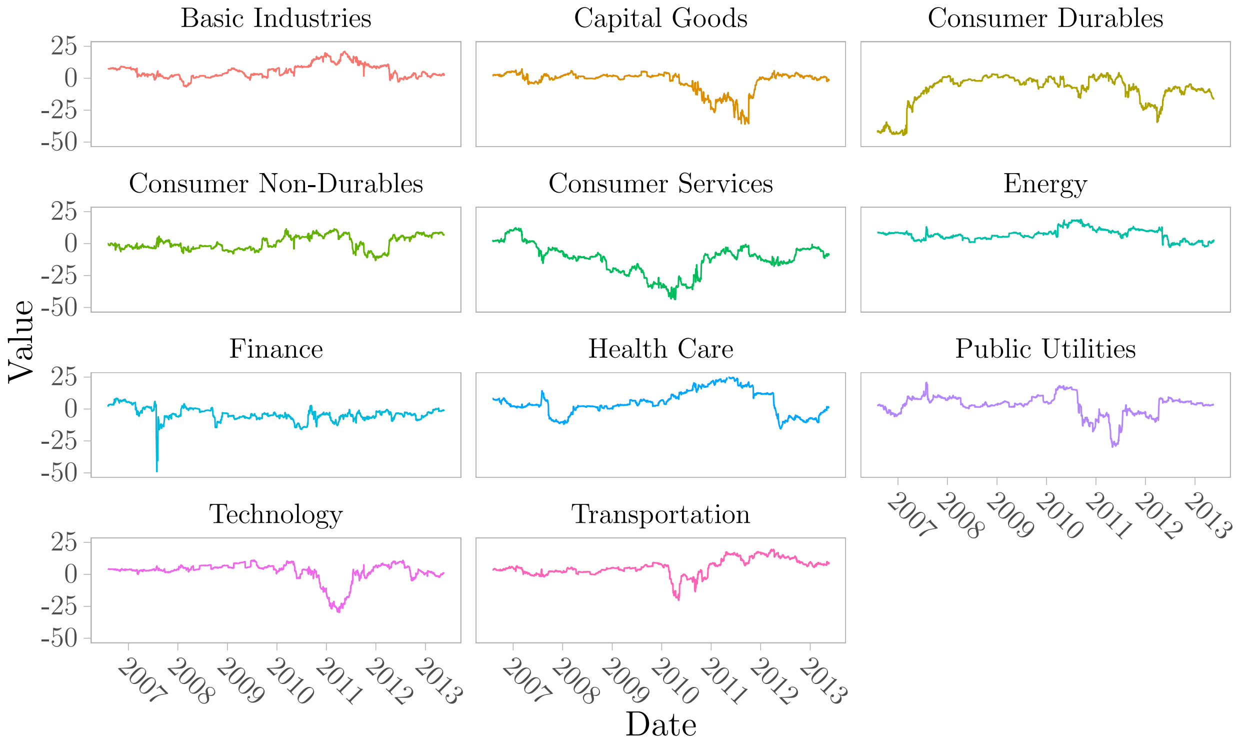
Diebold-Yilmaz spillover index for volatilities without sector-specific and macroeconomic covariates. Average value and maximum and minimum spillover index across all VAR orderings for Cholesky decomposition are given by the black line and gray ribbon, respectively. 25% and 75% bounds for the Cholesky Decomposition across 10,000 random permutations are given by the red ribbon. Index for generalized variance decomposition (GVD) is given by the blue line. Window is 500 days and H-step ahead error is 12.

Fig. 11. Range of Net Sector Return Connectedness without Covariates



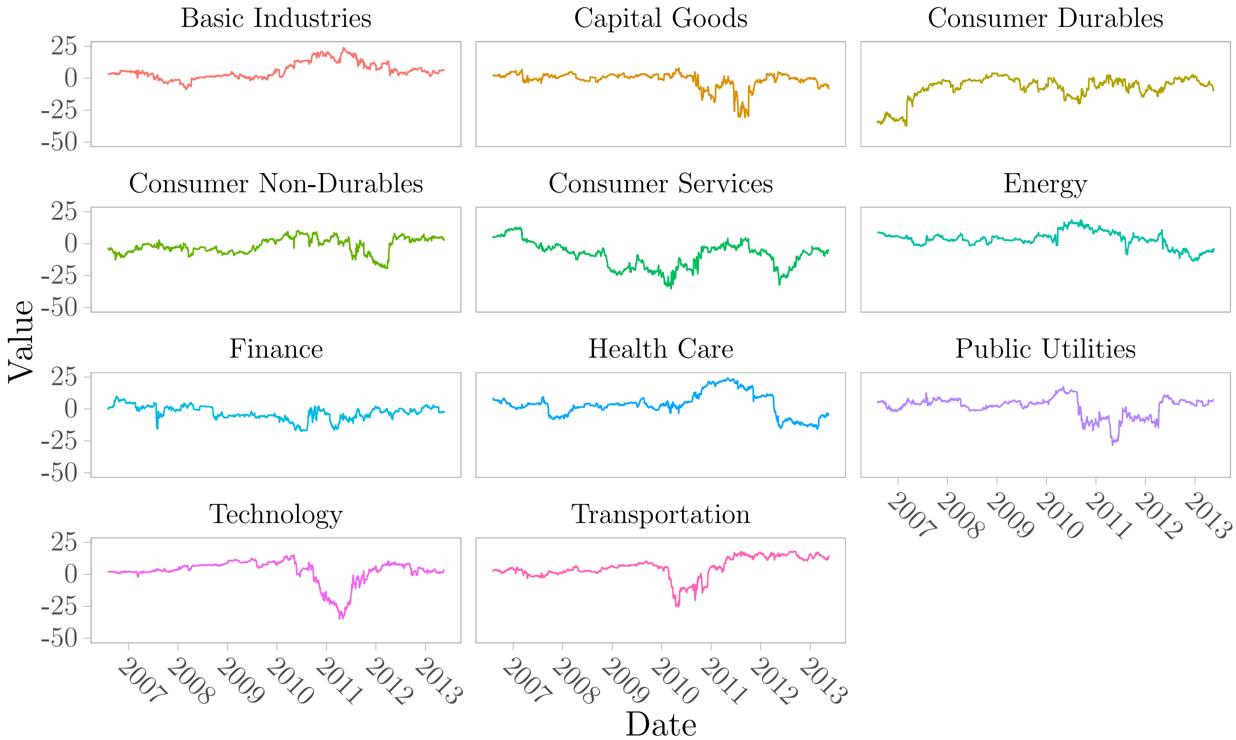
Minimum, maximum, and average *net connectedness* by sector for returns with sector-specific and macroeconomic covariates, calculated across 10,000 random permutations. Window is 500 days and H-step ahead error is 12.

Fig. 12. Net Sector Return Connectedness without Covariates



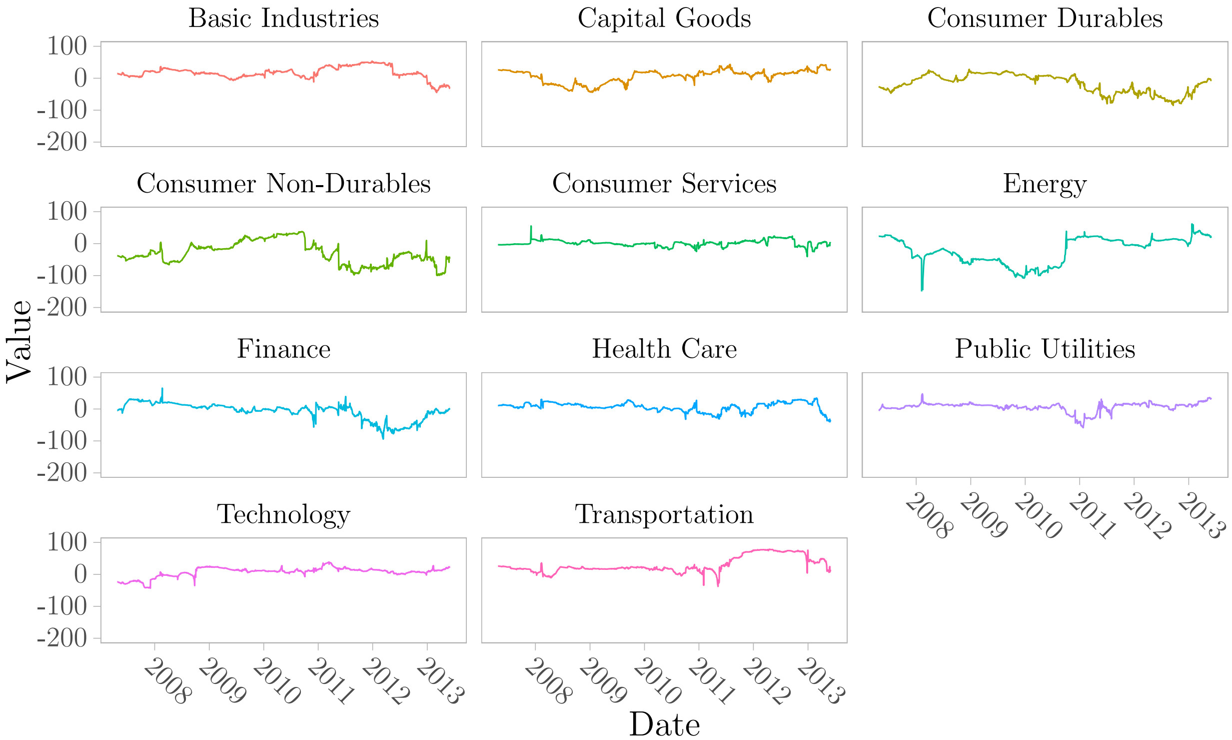
Average *net connectedness* by sector for returns without sector-specific and macroeconomic covariates, calculated across 10,000 random permutations. Window is 500 days and H-step ahead error is 12.

Fig. 13. Net Sector Return Connectedness with Covariates



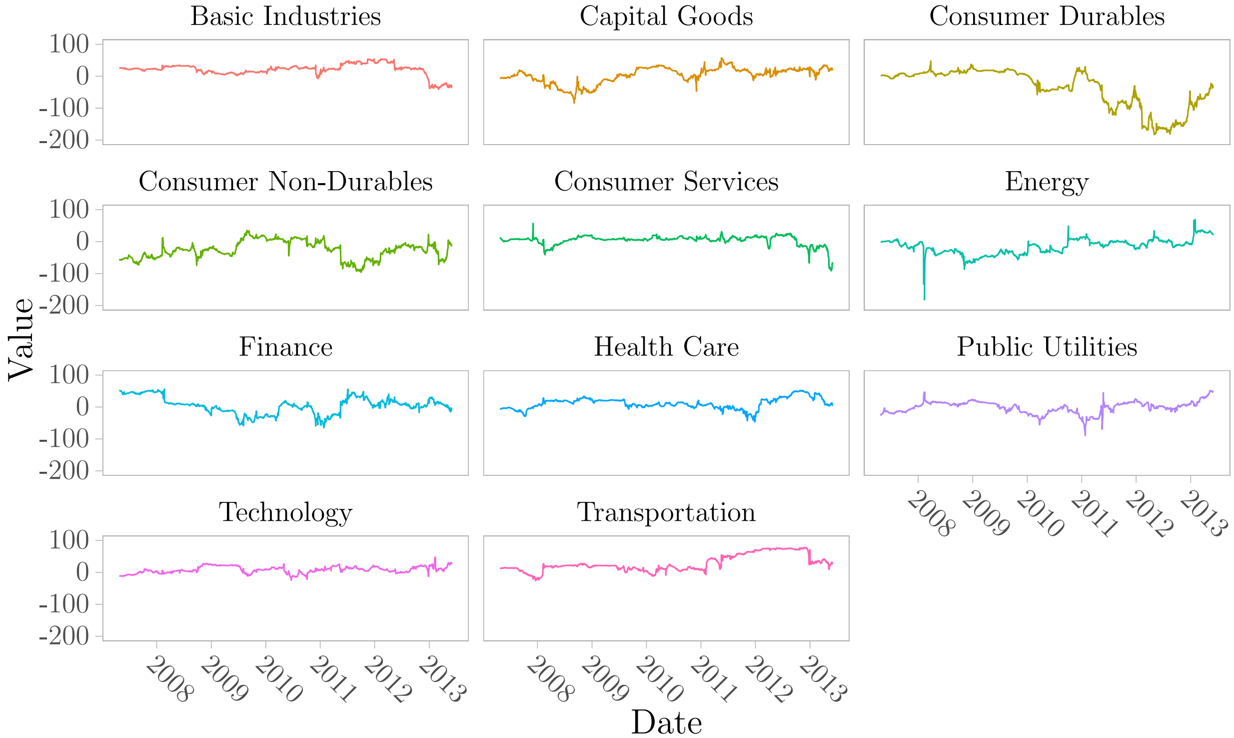
Average *net connectedness* by sector for returns with sector-specific and macroeconomic covariates, calculated across 10,000 random permutations. Window is 500 days and H-step ahead error is 12.

Fig. 14. Net Sector Volatility Connectedness without Covariates



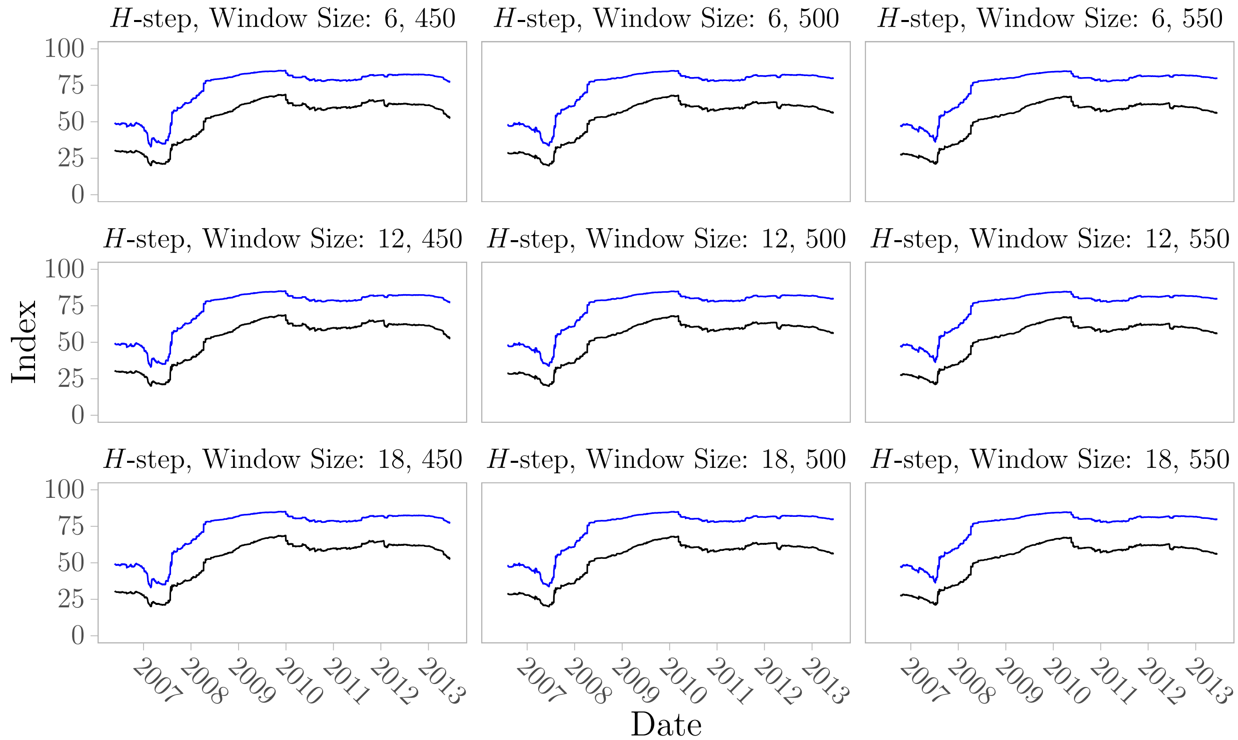
Average *net connectedness* by sector for volatilities with sector-specific and macroeconomic covariates, calculated across 10,000 random permutations. Window is 500 days and H-step ahead error is 12.

Fig. 15. Net Sector Volatility Connectedness with Covariates



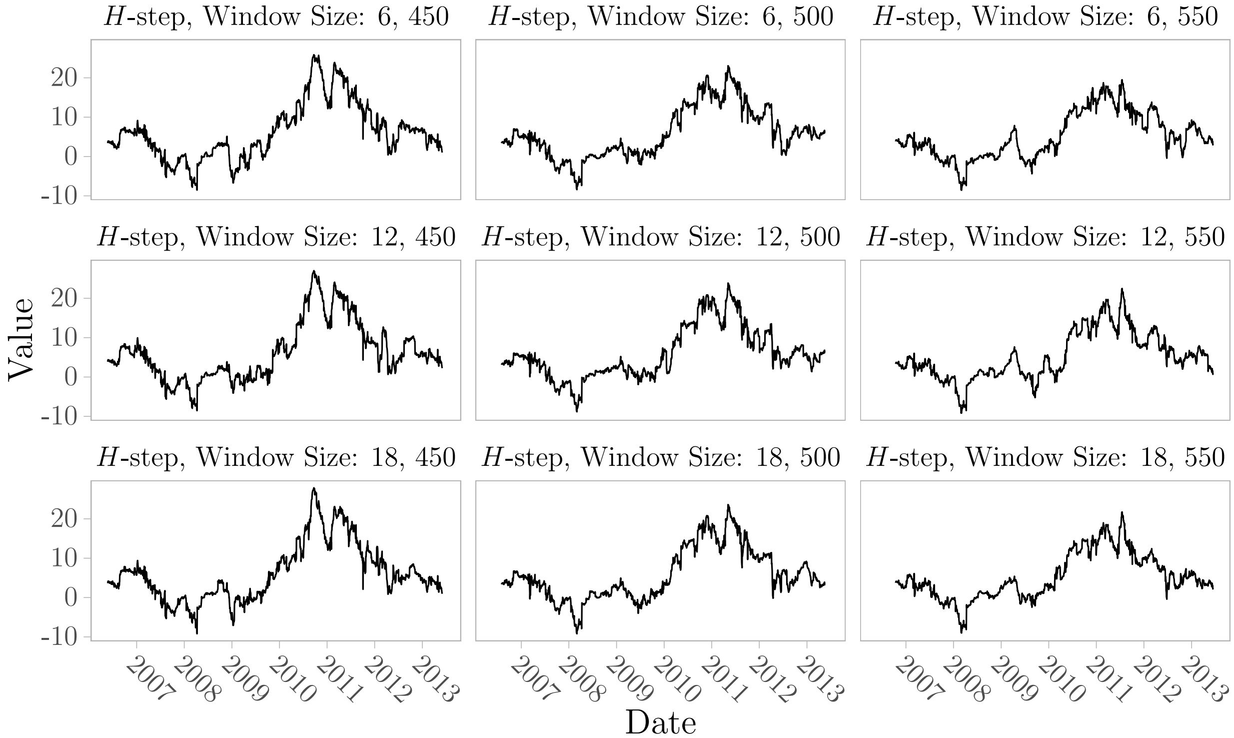
Average *net connectedness* by sector for volatilities with sector-specific and macroeconomic covariates, calculated across 10,000 random permutations. Window is 500 days and H-step ahead error is 12.

Fig. 16. Spillover Index for Returns without Covariates



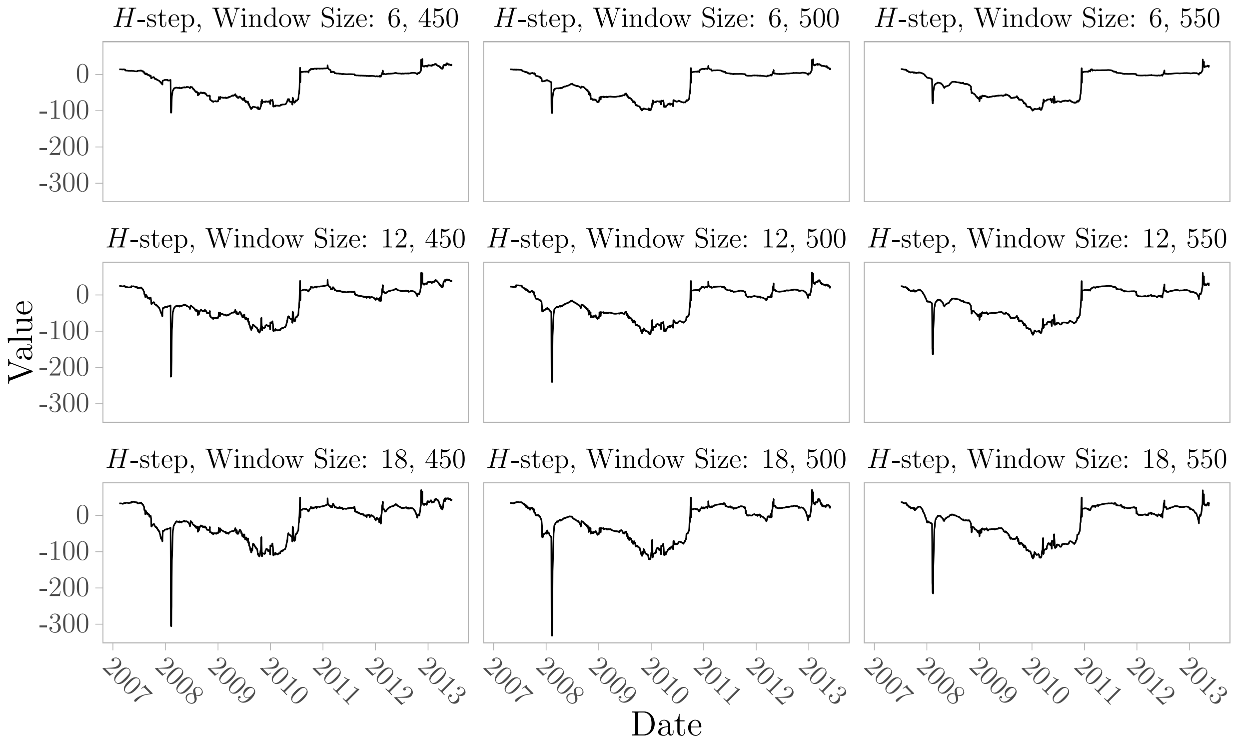
Diebold-Yilmaz spillover index for returns without sector-specific and macroeconomic covariates across various H-steps and window sizes. Average value across all VAR orderings for Cholesky decomposition are given by the black line and generalized variance decomposition (GVD) is given by the blue line

Fig. 17. Net Basic Industries Return Connectedness with Covariates



Average *net connectedness* for Basic Industries return with sector-specific and macroeconomic covariates, calculated across 10,000 random permutations and various H-steps and window sizes.

Fig. 18. Net Energy Volatility Connectedness without Covariates



Average *net connectedness* for Energy volatility without sector-specific and macroeconomic covariates, calculated across 10,000 random permutations and various H-steps and window sizes.

References

Baruník, Jozef, Evžen Kočenda, and Lukáš Vácha. 2016. "Asymmetric connectedness on the US stock market: Bad and good volatility spillovers". *Journal of Financial Markets* 27:55–78.

Bastian, Mathieu, Sebastien Heymann, and Mathieu Jacomy. 2009. *Gephi: An Open Source Software for Exploring and Manipulating Networks.* `http://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154`.

Blondel, Vincent D, et al. 2008. "Fast unfolding of communities in large networks". *Journal of statistical mechanics: theory and experiment* 2008 (10): P10008.

Bostanci, Gorkem, and Kamil Yilmaz. 2015. "How Connected is the Global Sovereign Credit Risk Network?" *Available at SSRN 2647251.*

Diebold, Francis X, and Kamil Yilmaz. 2015. *Financial and Macroeconomic Connectedness: A Network Approach to Measurement and Monitoring.* Oxford University Press, USA.

— . 2009. "Measuring financial asset return and volatility spillovers, with application to global equity markets". *The Economic Journal* 119 (534): 158–171.

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

Diebold, Francis X, Kamil Yilmaz, et al. 2011. "Equity market spillovers in the Americas". *Financial stability, monetary policy, and central banking* 15:199–214.

Duffie, Darrell, Leandro Saita, and Ke Wang. 2007. "Multi-period corporate default prediction with stochastic covariates". *Journal of Financial Economics* 83 (3): 635–665.

Fonseca, Jose De, and Katrin Gottschalk. 2012. *The co-movement of credit default swap spreads, stock market returns and volatilities: evidence from Asia-Pacific markets.* Tech. rep. working paper.

Klößner, Stefan, and Rodrigo Sekkel. 2014. "International spillovers of policy uncertainty". *Economics Letters* 124 (3): 508–512.

Klößner, Stefan, and Sven Wagner. 2014. "Exploring all VAR orderings for calculating spillovers? yes, we can!a note on Diebold and Yilmaz (2009)". *Journal of Applied Econometrics* 29 (1): 172–179.

Li, David X. 2000. "On default correlation: A copula function approach". *The Journal of Fixed Income* 9 (4): 43–54.

Longstaff, Francis A, and Arvind Rajan. 2008. “An empirical analysis of the pricing of collateralized debt obligations”. *The Journal of Finance* 63 (2): 529–563.

Longstaff, Francis A, et al. 2011. “How sovereign is sovereign credit risk?” *American Economic Journal: Macroeconomics* 3 (2): 75–103.

Markowitz, Harry. 1952. “Portfolio selection”. *The journal of finance* 7 (1): 77–91.

Pan, Jun, and Kenneth J Singleton. 2008. “Default and recovery implicit in the term structure of sovereign CDS spreads”. *The Journal of Finance* 63 (5): 2345–2384.

Pesaran, M Hashem, and Yongcheol Shin. 1998. “An autoregressive distributed-lag modelling approach to cointegration analysis”. *Econometric Society Monographs* 31:371–413.

Yang, Dennis, and Qiang Zhang. 2000. “Drift-independent volatility estimation based on high, low, open, and close prices”. *The Journal of Business* 73 (3): 477–492.