Interconnectedness in the Corporate Bond Market

CELSO BRUNETTI MATTHEW CARL JACOB GERSZTEN
CHIARA SCOTTI CHAEHEE SHIN*

February 2022

ABSTRACT

In this paper, we develop an alternative and complementary network structure—the investor similarity network—which mirrors the traditional notion of the portfolio similarity network. Leveraging on the richness of the eMAXX database, and matching its quarterly holding from to the security-level data on corporate bond trading volume, liquidity, and volatility derived from TRACE, we analyze the role of interconnectedness in the corporate bond market. We find that corporate bonds that are held across several portfolios are those that require a lower compensation for risk and that are more liquid. This relationship is stronger when a financial asset is under stress, i.e.spread and liquidity of an asset are in the upper tail of their conditional distributions.

Keywords: financial stability, institutional investors, big data, quantile regressions.

JEL Classification Codes: C13, C55, C58, G1

PRELIMINARY, DO NOT CIRCULATE

*Celso Brunetti, Jacob Gerszten, Chiara Scotti, and Chaehee Shin are with the Federal Reserve Board, 20th and C Streets NW, Washington, DC 20551, USA. Matthew Carl is a Ph.D. student at the University of Wisconsin-Madison but was a research assistant at the Board of Governors of the Federal Reserve System at the time work on this paper began. The authors can be reached via email at celso.brunetti@frb.gov, mcarl@wisc.edu, jacob.e.gerszten@frb.gov, chiara.scotti@frb.gov, and chaehee.shin@frb.gov. The views expressed in this article are those of the authors and not necessarily of the Federal Reserve System. We thank Gene Kang for excellent research assistance.

1. Introduction

In this paper, we develop an alternative and complementary network structure—the investor similarity network—which mirrors the traditional notion of portfolio similarity network. This new measure is derived at the asset level and it is based on the idea that assets are interconnected if they are held by the same firms. Its advantage is that it is specific to a given asset and it can be used to investigate the relationship between interconnectedness and the spread, liquidity or volatility of an asset, or group/portfolio of assets.

We focus on interconnectedness in the corporate bond market, leveraging on the rich information available in the Reuters eMAXX database, which draws from the quarterly N-CSR, N-CSRS, and N-Q filings with the Security and Exchange Commission (SEC) and contains holdings data at the institutional investor-subaccount-bond-year-quarter level. We complement this dataset with the TRACE database, which allows us to match the quarterly holding from eMAXX to the security-level data on corporate bond trading volume, liquidity, and volatility.

The new interconnectedness measure and the complexity of our data allow us to use a rich panel regression analysis to investigate the link between interconnectedness and spread, liquidity, and volatility of corporate bonds. We find that the higher the investor similarity of an asset—meaning that the asset is common to many investors' portfolios—the lower its spread and the higher its liquidity. This result highlights that, as expected, corporate bonds that are held across several portfolios are those that require a lower compensation for risk and that are more liquid. This relationship is, however, affected by market conditions. We explore the heterogeneous links/effects of interconnectedness throughout the conditional distribution of the response variables (spreads, liquidity, and volatility), while controlling for individual and time-specific bond characteristics, through a panel data quantile regression. We find that the relationship we have just highlighted is stronger when a financial asset is under stress, i.e.spread and liquidity of an asset are in the upper tail of their conditional distributions.

Our paper contributed to several strands of the literature. First, we contribute to the interconnectedness literature. Networks in finance have been mapped using three main techniques: (i) correlation networks, in which edges among financial institutions are based on estimates of the variance-covariance matrix of publicly available data, such as asset returns (see, Billio, Getmanski, Lo and Pellizzon, 2012; and Diebold and Yilmaz, 2014); (ii) physical networks, in which edges capture contractual agreements among counterparties, such as interbank transactions (see, Brunetti, Harris, Mankad and Michailidis, 2019); (iii) common holdings networks, in which firms are connected if they hold similar portfolios (see, Caccioli, Farmer, Foti and Rockmore, 2015; and Greenwood, 2015). In this paper, we propose a new approach of mapping financial networks which mirrors the notion of overlapping portfolios, and which we call *overalapping investors* or *investor similarity network*.

Second, we connect to the asset price literature, especially with regard to the emerging studies on institutional asset pricing and institutional price pressure. Haddad and Muir (2021) find support for the channel where risk-bearing capacity of intermediaries matter for asset prices and risk premiums. He, Kelly, and Manela (2017) show that capital risk factors of intermediaries explain expected returns in many different asset classes, including corporate bonds. Ben-David et al. (2021) show how the rising concentration of holdings by institutional investors affect stock volatility and price inefficiency. Coval and Stafford (2007) find that institutional investors have temporary but sizable impacts on equity prices, via fire sales. We cotribute to this literature by showing the importance of interconnectedness for corporate bond prices.

Third, we relate to the recent financial stability literature that tries to determine whether high interconnectedness is a vulnerability or a virtue of the financial system. Conflicting views exist in the literature, from Allen and Gale (2000) that finds interconnectedness as a virtue to more recent empirical works finding evidence for financial linkages and overlapping holdings of assets as a contagion or fire sales mechanism (Allen and Carletti (2012), Duarte and Eisenbach (2021), Falato et al. (2021), and Greenwood, Landier, and Thesmar (2015), among others). Somewhere in between these two conflicting views, many recent works study the non-monotonic tradeoff between contagion and risk-sharing, social optimality of interconnectedness and conditions for which one type of network is better than

another (Acemoglu et al. (2015), Cabrales, Gottardi, and Vega-Redondo (2017), Elliott, Golub and Jackson (2014), and Elliott, Hazell and Georg (2021), Gofman (2017), among others).

The paper is organized as follows. Section 2 describes our novel network approach, illustrating the building blocks of the investor similarity network. Section 3 summarizes the wealth of data that we use in the empirical analysis. Section 4 describes the resulting measures that we use in the analysis. Section 5 explains the regression analysis and its results, including those for the quantile regressions. Section 6 concludes.

2. Network Approach

Networks in finance have been mapped using three main techniques: (i) correlation networks, in which edges among financial institutions are based on estimates of the variance-covariance matrix of publicly available data, such as asset returns (see, Billio, Getmanski, Lo and Pellizzon, 2012; and Diebold and Yilmaz, 2014);[1] (ii) physical networks, in which edges capture contractual agreements among counterparties, such as interbank transactions (see, Brunetti, Harris, Mankad and Michailidis, 2019); (iii) common holdings networks, in which firms are connected if they hold similar portfolios (see, Caccioli, Farmer, Foti and Rockmore, 2015; and Greenwood, 2015). In this paper, we propose a new approach of mapping financial networks which mirrors the notion of overlapping portfolios, and which we call *overalapping investors* or *investor similarity network*.

The starting point for both network structures—the traditional overlapping portfolios and the new overlapping investors networks—is a bipartite network with two sets of nodes: firms (Fs) and financial assets (As). As shown in Figure 1, if a financial institution holds an asset in its portfolio, there is an edge between that asset and that financial institution. For example, because firm F_1 holds assets A_1 and A_3, there are edges between F_1 and A_1, and between F_1 and A_3. The traditional network of common asset holdings, or overlapping portfolios, implies that since A_1 is held also by F_2 and F_3, all three firms are interconnected

[1] A related approach adopts quantile regression analyses, see Ando, Li and Lu (forthcoming), and Härdle, Wang and Yu (2016).

through their common holdings of A_1. Similarly, because A_2 is held by F_2 and F_3, there is a link between these two firms (see, Barucca, Mahammod and Silvestri, 2021).

An alternative novel network structure—the investor similarity or overlapping investors network—can be derived at the assets level. In the bottom panel of Figure 2, A_1 is linked to A_3 because these assets are held in the same portfolio of firm F_1. Similarly, A_1 and A_2 are connected since they are held by firm F_2. The network of investor similarity is based on the idea that two assets, A_1 and A_2, are interconnected if they are held by the same firm. If A_1 and A_2 are held by all firms in the network, we consider these two assets highly connected: if F_1 is hit by a shock and forced to sell its assets, A_1 and A_3, there is a link between the two assets.[2]

In what follow, we will describe the investor similarity network structure in more details and we will highlight the network measures that we will use in the analysis.

2.1. Network of Financial Assets and Institutions

We denote the network of financial assets and financial institutions as $Q = (A, F, \mathbf{E})$, where $A = A_1, A_2, ..., A_S$ is the set of nodes corresponding to financial assets (corporate bonds only, in our case), $F = F_1, F_2, ..., F_N$ represents the set of financial institutions, and $\mathbf{E}$ is a $S \times N$ matrix representing the dollar amount, E_{ik}, invested by F_k in A_i:

[2]Alternatively, assume a shock hits A_1 (e.g., downgraded to junk), firms holding A_1 will be forced to rebalance their portfolio, for example, to raise more capital. This implies that holdings of A_3 will also change. If A_1 and A_3 are negatively correlated, the price of A_3 may increase, while if they are positively correlated, the price of A_3 will decrease. Even if the correlation is zero but A_1 is sold and is not part anymore of the different portfolios, portfolio weights will change and this will affect A_3.

$$
\mathbf{E} = \begin{array}{c|cccc|c}
 & F_1 & F_2 & \cdots & F_N & \\
\hline
A_1 & E_{11} & E_{12} & \cdots & E_{1N} & V_1^A \\
A_2 & E_{21} & E_{22} & \cdots & E_{2N} & V_2^A \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
A_S & E_{S1} & E_{S2} & \cdots & E_{SN} & V_S^A \\
\hline
 & V_1^F & V_2^F & \cdots & V_N^F &
\end{array} \tag{1}
$$

Summing across columns gives the total amount of security i held by the system (firms in our data), V^{A_i}, known as the *strength of the network*:

$$
V^{A_i} = \sum_{k=1}^{N} E_{ik} \tag{2}
$$

and summing across rows produces the total amount invested by firm k in all assets, V^{F_k}, which is the more traditional measure.

We define as $\overset{\circ}{\mathbf{E}}$ the corresponding adjacency matrix

$$
\overset{\circ}{\mathbf{E}} = \begin{array}{c|cccc|c}
 & F_1 & F_2 & \cdots & F_N & \\
\hline
A_1 & \overset{\circ}{E}_{11} & \overset{\circ}{E}_{12} & \cdots & \overset{\circ}{E}_{1N} & D_1^A \\
A_2 & \overset{\circ}{E}_{21} & \overset{\circ}{E}_{22} & \cdots & \overset{\circ}{E}_{2N} & D_2^A \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
A_S & \overset{\circ}{E}_{S1} & \overset{\circ}{E}_{S2} & \cdots & \overset{\circ}{E}_{SN} & D_S^A \\
\hline
 & D_1^F & D_2^F & \cdots & D_N^F &
\end{array} \tag{3}
$$

where the generic element $\overset{\circ}{E}_{ik} = 1$ if $E_{ik} > q$ and zero otherwise. The parameter q denotes a threshold and in traditional network analysis $q = 0$.[3]

[3]Given the richness of our data, we could also adopt $q > 0$ to select the strongest links among nodes.

Similarly to before, the sum across columns gives the total number of financial institutions holding security i, D^{A_i} , known as *network degree*:

$$D^{A_i} = \sum_{k=1}^{N} \overset{\circ}{E}_{ik}. \tag{4}$$

and the sum across rows generates the total number of assets firm k has invested in, D^{F_k}.

2.2. Network of Investor Similarity

From the network of financial assets and financial institutions Q, we can derive two additional networks: the network of investor similarity, which captures interconnectedness among assets as a function of belonging to the same portfolios.[4]

We define the network of asset similarity as $O^A = \left(A, \mathbf{P}^A\right)$, where $A = \{A_1, A_2, ..., A_S\}$ represents the set of assets, and $\mathbf{P}^A$ is the matrix measuring asset similarities. Several distance measures exist to quantify similarities (see Newman 2010, Delpini et al 2015, Barucca et al 2021, Brunetti et al 2021). In this work we adopt two definitions of similarities:

$$W^{A_i,1} = \frac{1}{N(S-1)} \sum_{j \in \{1,...,S\}: j \neq i} P_{ij}^{A,1} \tag{5}$$

$$W^{A_i,2} = \frac{1}{N(S-1)} \sum_{j \in \{1,...,S\}: j \neq i} P_{ij}^{A,2}. \tag{6}$$

These measures differ in the way $P^{A,1}$ and $P^{A,2}$ are computed. $P^{A,1}$ simply counts the number of portfolios two assets are part of and does not account for the size of the investment:

$$P^{A,1} = \overset{\circ}{E}\left(\overset{\circ}{E}\right)^{\top}. \tag{7}$$

[4]Similarly, from the network of financial assets and institutions, we can compute the network of portfolio similarity—a traditional measure in the network literature (e.g., Braverman and Minca, 2014)—which captures portfolio characteristics and is based on the idea that two firms are connected if they hold a similar portfolio. More information about the network of portfolio similarity can be found in the appendix.

$P_{ij}^{A,2}$ captures the distance between two non-zero vectors of an inner-product space and is referred to as *cosine* similarity (or distance):

$$P_{ij}^{A,2} = \frac{\sum_{k=1}^{N} \overset{\circ}{E}_{ik} \overset{\circ}{E}_{jk}}{\left\|D_i^A\right\| \left\|D_j^A\right\|} \tag{8}$$

with $\left\|D_i^A\right\|$ is the norm of the degree of the network, D_i^A, which represents the vector of firms holding asset i (see Getmansky et al 2016, Barucca et al 2021).[5]

We compute the asset-by-asset measures in equations (7) and (8) by summing across assets. To account for the fact that assets and firms change over time in our data, we normalize (5) and (6) by $(S-1) * N$, where S is the total number of assets and N is the total number of firms.

2.3. A Simple Example

An example may help to explain these concepts. Consider a network consisting of only three assets and three firms, where the entries in the matrix represent the (hypothetical) dollar amount of each asset held by each firm:

$$\mathbf{E}_{example} = \begin{array}{c|ccc} & F_1 & F_2 & F_3 \\ \hline A_1 & 6 & 5 & 4 \\ A_2 & 0 & 3 & 2 \\ A_3 & 1 & 0 & 0 \end{array}$$

which can be represented by the following adjacency matrix:

[5] A potential third overlapping investor measure we could adopt derives from the notion of Euclidean distance, namely, $P_{ij}^{A,3} = \frac{1}{2} \sum_{k=1}^{N} \left| \overset{\circ}{E}_{ik} - \overset{\circ}{E}_{jk} \right|$. However, due to the sparsity of our network, this measure has the inconvenient characteristic of positively correlating with the other measures of similarity described in equations (7) and (8) despite the intended function of capturing distance. Hence, we omit this measure from our analysis.

$$\mathring{\mathbf{E}}_{example} = \begin{array}{c|ccc} & F_1 & F_2 & F_3 \\ \hline A_1 & 1 & 1 & 1 \\ A_2 & 0 & 1 & 1 \\ A_3 & 1 & 0 & 0 \end{array},$$

where 1 in the top-left cell of the matrix $\mathring{\mathbf{E}}_{example}$ indicates that firm F_1 is invested in asset A_1, and a 0 in $cell(2,1)$ indicates that firm F_1 is not invested in asset A_2.

The strength, V^A, and degree, D^A, of the network are $\begin{bmatrix}15 & 5 & 1\end{bmatrix}'$ and $\begin{bmatrix}3 & 2 & 1\end{bmatrix}'$, respectively.

We can compute a simple measure of similarities across assets by counting the number of portfolios (firms) two assets are part of:

$$\mathbf{P}^{A,1}_{example} = \begin{array}{c|ccc} & A_1 & A_2 & A_3 \\ \hline A_1 & - & 2 & 1 \\ A_2 & 2 & - & 0 \\ A_3 & 1 & 0 & - \end{array},$$

where the numbers in the matrix represent the *edges*.[6]

Similarly, using the cosine similarity metric, we have

$$\mathbf{P}^{A,2}_{example} = \begin{array}{c|ccc} & A_1 & A_2 & A_3 \\ \hline A_1 & - & 0.82 & 0.58 \\ A_2 & 0.82 & - & 0.00 \\ A_3 & 0.58 & 0.00 & - \end{array}.$$

[6]Similarly, we could compute a simple measure of portfolio similarity between firms based on the number of common assets they hold: $\mathbf{P}^{F,1}_{example} = \begin{array}{c|ccc} & F_1 & F_2 & F_3 \\ \hline F_1 & - & 1 & 1 \\ F_2 & 1 & - & 2 \\ F_3 & 1 & 2 & - \end{array}$, where the numbers in the matrix are the *edges*. Here, the corresponding adjacency matrix using the cosine similarity measure yields: $\mathbf{P}^{F,2}_{example} = \begin{array}{c|ccc} & F_1 & F_2 & F_3 \\ \hline F_1 & - & 0.50 & 0.50 \\ F_2 & 0.50 & - & 1.00 \\ F_3 & 0.50 & 1.00 & - \end{array}$. The corresponding vectors of interconnectedness measures are: $W^{F,1} = \begin{bmatrix}0.33 & 0.50 & 0.50\end{bmatrix}'$ and $W^{F,2} = \begin{bmatrix}0.17 & 0.25 & 0.25\end{bmatrix}'$.

Accordingly, the vectors of interconnectedness measures corresponding to $\mathbf{P}^{A,1}_{example}$ and $\mathbf{P}^{A,2}_{example}$ are:

$$W^{A,1} = \begin{bmatrix} 0.50 & 0.33 & 0.17 \end{bmatrix}'$$
$$W^{A,2} = \begin{bmatrix} 0.23 & 0.14 & 0.097 \end{bmatrix}'$$

Both $W^{A,1}$ and $W^{A,2}$ are consistent with the intuition that asset 1, which is held by all three hypothetical firms, has the highest level of interconnectedness, whereas asset 3, which is only held by firm 1, has the lowest interconnectedness.

3. Data

Our analysis relies on data from different sources. Primarily, we rely on the Thomson Reuters eMAXX database and the Financial Industry Regulatory Authority's (FINRA) fixed income market Trade Reporting and Compliance Engine (TRACE) database. In addition, we use data from S&P Global and the Mergent Fixed Income Securities Database (FISD).

3.1. The eMAXX Database

We obtain information on institutional investors and their bond holdings from the Thomson Reuters eMAXX database, which draws from the quarterly N-CSR, N-CSRS, and N-Q filings with the Security and Exchange Commission (SEC). The eMAXX holdings data is at the individual institutional investor-subaccount-bond-year-quarter level, and runs from 1998:Q3 until 2020:Q4. In each quarter, we observe the full fixed-income portfolios of subaccounts belonging to an institutional investor. Our focus is in the bond holdings of the largest institutional investors, so we aggregate the bond holdings data to the institutional investor level by summing across an institutional investor's sub-accounts' bond holdings in each quarter.

We supplement the holdings data with further detail on the institutional investors, including the reported investor name, type, and headquarter location. All U.S. institutional

investors, with the exception of pension funds, are mandated to report their entire portfolio each quarter, a rule that has been in effect since May 2004.[7] Thus, we focus on the set of investors domiciled in the United States. Because the eMAXX data further distinguish between the subsidiaries of institutional investors, for example, JP Morgan Chase (New York) and JP Morgan Chase (Los Angeles), some investors belonging to a single parent company (i.e., JP Morgan Chase) are coded with different investor identifiers. This property of the data is inconvenient given our research objective of constructing networks that link assets together based on overlapping investors. We do not wish to differentiate between an institutional investor's subsidiaries' bond portfolios, and so we take great care to aggregate these subsidiaries' bond holdings into a single institutional investor portfolio.

Specifically, we utilize a string matching algorithm on the reported institutional investor names to match investors that plausibly belong to the same parent company, but which potentially receive separate investor identifiers in eMAXX. Following the string matching algorithm, we then conduct a manual audit on the matches to ensure their validity. Ultimately, we obtain a dictionary mapping parent companies to their subsidiaries, and use this dictionary to replace the subsidiaries' identifiers with a new investor identifier. Finally, we aggregate the bond holdings data to the institutional investor level.

Following the string matching algorithm, we identify 4,972 unique institutional investors. While eMAXX reports a type for each institutional investor (see Table 1), we uncovered several discrepancies between the true type of an investor and the type reported by eMAXX (for example, JP Morgan Chase is classified as a mutual fund). We therefore further audit the investor type in the final set of institutional investors, and categorize each investor as a bank, investment manager, insurance company, pension fund, or other investor type. Because we are interested in the largest institutional investors, and for the purpose of feasibility in the investor type classification, we further subset the list of institutional investors to those whose assets under management (AUM) fall within the top 50th percentile of the AUM distribution each quarter. We include corporate bonds as well as other assets such as government, municipal, and mortgage-backed securities to calculate total assets for each firm.

[7] https://www.sec.gov/rules/final/33-8393.htm.

Following this process, we obtain 190 banks, 703 investment managers, 604 insurance companies, 209 pension funds, and 36 other funds which uniquely appear across the panel. Our subsample of institutional investors represents the lion share of the total par amount held within the eMAXX universe. Specifically, the 1,742 unique parent firm identifiers whose AUM eclipses the 50th percentile of the cross-sectional AUM distribution hold 78.89% of the total par amount of corporate bonds held within the eMAXX universe.

eMAXX also provides detailed information on the underlying securities—identified at the 8-digit CUSIP level—that the institutional investors hold, including the market sector to which each security belongs. The eMAXX data distinguish between several different markets: asset-backed securities, including collateralized debt obligations and covered bonds; corporate bonds, including high-yield and investment grade; government bonds, including sovereign and government agency; mortgage-backed securities, including agency and private label pass through, collateralized mortgage obligations, collateralized mortgage-backed securities, and residential mortgage-backed securities; regional and municipal bonds, including U.S. muni and international cities, states, and provinces; private placements, including 144A and non-144A; and emerging markets.[8] Due to data limitations, the focus of this paper is on corporate bond holdings.

We also obtain information on corporate bond ratings, maturity, coupon rates, and spread from eMAXX, and supplement these variables with other security-specific information from TRACE and FISD (see below).

Table 2 shows the summary statistics of the data from eMAXX, that include information on outstanding amount, maturity, coupon rate, spread and rating. All variables are winsorized at top and bottom 1 percentiles.[9] Panel A shows summary statistics at the 8-digit CUSIP level, while Panel B shows the same statistics at the company level (6-digit CUSIP).

[8] https://www.thomsonreuters.com/content/dam/openweb/documents/pdf/tr-com-financial/fact-sheet/emaxx-bond-holders.pdf

[9] Note that Table 2 describes summary statistics related to issuance amount, maturity, coupon rate, spread, and rating collected from the eMAXX database but after this database is matched with the TRACE data, which results in only working with a subset of the initial eMAXX data. The lower number of observations for the outstanding amounts highlights the importance of complementing the eMAXX information related to outstanding amounts with that of FISD.

3.2. The TRACE Database

We obtain security-level data on corporate bond trading volume, liquidity, and volatility from the intraday trading information available from the TRACE database. We aggregate the intraday TRACE data at the quarterly frequency to match with the quarterly-level eMAXX dataset. Because trade frequencies may be extremely sparse for some bonds, we check robustness of our analyses by using alternate ways of quarterly aggregation, including mean, median, and the last quarterly observation of a variable.

For bond illiquidity, we use two proxies that are widely used in the literature: the Amihud measure and the interquartile range (IQR). Amihud (2002) price impact is defined as:

$$Amihud_{it} = \frac{1}{D_{it}} \sum_{k=1}^{D_{it}} \frac{r_{ikt}}{Q_{ikt}} \tag{9}$$

where D_{it} is the total number of trades on bond i at time (day) t, and r_{ikt} and Q_{ikt} each refer to the return and traded volume of the kth transaction of bond i on day t. IQR of traded prices is defined and calculated as the difference between the 75th and the 25th percentiles of daily prices.

Volatility of bond prices is measured as the quarterly standard deviation of traded prices of a bond. Table 2 shows the summary statistics of the data from the quarterly aggregated TRACE. All variables are winsorized at top and bottom 1 percentiles.

3.3. Other Data Sources

Additional information for each bond issuance comes from the Mergent Fixed Income Securities Database (FISD). Data include issuer-specific, issue-specific, and transaction information. In addition to basic bond characteristics such as maturity, issuer identity, etc., the database includes pricing at issue (but no pricing information thereafter), ratings, sinking fund and call information including an estimate of the amount outstanding at any given

time, covenants, defaults, and more. We obtain the total amount outstanding for each asset and take the mean amount outstanding for each quarter for each CUSIP. We then link this information to the eMAXX holdings data at the CUSIP-quarter level. Since eMAXX does not have complete coverage of bond ratings, we supplement the missing observations with data from FISD that cover ratings from S&P, Fitch and Moody's.

Whenever eMAXX and FISD were missing information on bond ratings, we supplemented it with ratings data from the S&P Global database. In the end, about 3 million observations on rating have been filled in accordingly. The ratings have been transformed into a numerical scale between 1 (lowest) and 21 (highest). Whenever multiple observations on a bond rating are available from S&P, Fitch and Moody's, we take the average of ratings.

4. Interconnectedness in the Corporate Bond Market

Table 3 reports summary statistics for the company-level interconnectedness measures. Panel A contains three main results: (i) overall interconnectedness is low, the average level of overlapping investors is below 0.3 percent indicating that only a small portion of bonds of the same corporation are held by multiple institutional investors; (ii) in the pre-crisis period, 2002-2008, interconnectedness is even lower; and (iii) it increases after the crisis, from 2009 to 2020.

Panel B reports descriptive statistics for the cross-section of assets. Specifically, for each asset, we take the average of the variables across our sample period. Average degree is 38, indicating that, on average, a bond from a given corporation is held by 38 institutional investors. There is great heterogeneity in our sample, though; degree varies between 1 and 414 investment firms—see Figure C3 in the appendix. Bonds from a given corporation are in the sample, on average, for 15 quarters. Importantly, some bonds only appear for a quarter, confirming that our network is sparse, as shown by the histograms in Figure C1 in the appendix.

Overall, we find the investor similarity network changes over time. Its dynamics suggest that interconnectedness increased after the GFC. This is an interesting result. In physical networks, e.g. interbank market, we usually observe a rump up in connections leading to the crisis, followed by a rapid decrease as a consequence of the increased uncertainty. Our results indicate that the crisis generated higher assets interconnectedness through overlapping institutional investors. This implies that more corporate bonds are held in the same portfolio. In other words, after the GFC, institutional investors tried to diversify their holdings and, following traditional finance theory, attempted to mitigate risk. However, while doing so, corporate bonds became more interconnected. The effect of interconnectedness on market quality and on the volatility of corporate bonds is the subject of the next section.

5. Market Quality and Interconnectedness

The advantage of our measures of interconnectedness is that they are specific to a given asset. Given the wealth of information contained in the eMAXX database about corporate bond holdings, we can compute interconnectedness measures at the CUSIP/company level, and we can use those measures in a panel data setting to analyze, for example, the relationship between interconnectedness of a certain asset and its spread and liquidity, or look at how this relationship is affected during periods of stress.

5.1. Interconnectedness, Spread, and Liquidity

The first question we can ask is how interconnectedness affects the first moments of an asset, such as spread and liquidity, above and beyond the usual characteristics of the asset.

$$Spread_{it} = \alpha + \beta_1 IC_{it} + \gamma X_{it} + FE_i + FE_t + \epsilon_{it} \tag{10}$$

The dependent variable is spread of bond i at time t, measured as the difference between the average yield for all trades on the bond on a day and the comparable Treasury or interpolated maturity-matched swap rate on the same day, and aggregated at the quarterly level. The main variable of interest is our measure of investor similarity, or interconnectedness, of asset i at time t, IC_{it}, which we measure in two different ways according to equations (7) and (8).

X_{it} is a matrix of time-varying bond characteristics that include liquidity (Amihud, interquartile range of traded prices), trade volume, outstanding issuance size, coupon rate, credit rating, and time to maturity. FE_i refers to the issuer and the issue date (year-quarter) fixed effects. FE_t controls for time fixed effects (current year-quarter).

Table 4 shows the preliminary results of estimating equation (10) at the CUSIP-level. Both measures of interconnectedness find that the higher the interconnectedness of an asset, the lower its spread. In terms of economic magnitude, Cosine similarity has a larger coefficient than the number of overlapping investors, with a standard deviation increase in Cosine similarity being associated with a quarter of a standard deviation increase of the spread. This economic significance is comparable to that of the interquartile range of trade prices.

Table 5 shows the preliminary results of estimating equation (10) aggregated at the company-level, where the results are similar.

We also explore the same specification regressing $Liquidity_{it}$ on the same set of variables as in equation (10), omitting the liquidity measures from X_{it}.

$$Illiquidity_{it} = \alpha + \beta_1 IC_{it} + \gamma X_{it} + FE_i + FE_t + \epsilon_{it} \tag{11}$$

Table 6 shows the results of the estimating equation (11) at the CUSIP-level. Panel A and B each run the same regression using the Amihud illiquidity measure and the interquartile range (IQR) of trade prices. In Panel A, coefficient on the number of overlapping investors is not significant but that on the Cosine similarity is statistically significant at 1% level. In

Panel B, both measures of interconnectedness show statistically significant magnitudes of coefficients at 1% levels. Economic magnitudes are slightly lower, as a standard deviation increase in either of the interconnectedness measures is associated with less than one-tenth of a standard deviation decrease in illiquidity.

Table 7 shows the results of the estimating equation (11) aggregated at the company-level, where the overall results are similar and consistent with the previous findings that high interconnectedness is associated with lower spreads.

5.2. Interconnectedness and Volatility

After examining the effects of interconnectedness on the asset's first moments, the next question we can ask is if interconnectedness can explain the second moments, for instance, financial market volatility.

The estimating equation for asset volatility is:

$$Volatility_{it} = \alpha + \beta IC_{it} + \gamma X_{it} + FE_i + FE_t + \epsilon_{it} \tag{12}$$

where volatility is measured as the quarterly standard deviation of trade prices. Table 8 shows the results at the CUSIP-level. while the number of overlapping investors does not show any significance, interconnectedness measured in Cosine similarity show negative coefficients on asset volatility. A one standard deviation increase in Cosine similarity leads up to one-tenth of a standard deviation of asset volatility. This finding is consistent with the previous finds that high interconnectedness is associated with lower spreads and higher liquidity.

Table 9 shows the results at the company-level. The results are similar to those at the CUSIP-level : the higher the interconnectedness, the lower the volatility of asset price.

6. Conclusion

In this paper, we develop an alternative and complementary network structure derived at the asset level and based on the idea that assets are interconnected if they are held by the same firms. We focus on interconnectedness in the corporate bond market, to investigate the link between interconnectedness and spread, liquidity, and volatility of corporate bonds. We find that the higher the investor similarity of an asset—meaning that the asset is common to many investors' portfolios—the lower its spread and the higher its liquidity. This result highlights that, as expected, corporate bonds that are held across several portfolios are those that require a lower compensation for risk and that are more liquid. This relationship is, however, affected by market conditions. We explore the heterogeneous links/effects of interconnectedness throughout the conditional distribution of the response variables (spreads, liquidity, and volatility), while controlling for individual and time-specific bond characteristics, through a panel data quantile regression. We find that the relationship we have just highlighted is stronger when a financial asset is under stress, i.e.spread and liquidity of an asset are in the upper tail of their conditional distributions.

Table 1

Institutional Investor and Subaccount Types

Notes: This table reports the institutional investor and subaccount types given in the eMAXX data. We organize the institutional investor types into four categories: banks, investment managers, insurance companies, and other investors. While subaccounts are not the focus of this paper, for the sake of illustration, we also organize the subaccount types into four categories, including insurance investment accounts, mutual funds, pension funds, and other funds.

Institutional Investor Type	Subaccount Type
Banks	*Insurance Investment Accounts*
Bank-Management Division	Insurance Co-Diversified
Bank-Portfolio	Insurance Co-Life/Health
Bank-Savings/Bldg Society	Insurance Co-Prop & Cas
Bank-Trust	*Mutual Funds*
Broker/Dealer-Fund Mgr	Mut Fd-O/E/Unit Tr/SICAV
Broker/Management Sub	Mut Fd-C/E/Invst Tr
Investment Managers	Mutual Fund-Equity
Investment Manager	Mutual Fund-Fund of Funds
Mutual Fund Manager	*Pension Funds*
Insurance Companies	Pension Fund-Corporate
Insurance Co-Diversified	Pension Fund-Government
Insurance Co-Life	Pension Fund-Union
Insurance Co-Mgmt Div	*Other Funds*
Insurance Co-Prop & Cas	401K
Reinsurance Company	Annuity/Variable Annuity
Other Investors	Bank-Portfolio
Pension Fund-Government	Bank-Trust
Pension Fund-Union	Church/Religious Org
Corporation	Corporation
Credit Union	Credit Union
Equity Manager	Finance Company
Finance/Credit Company	Fonds Commun de Placement
Foundation/Endowment	Foundation/Endowment
Government	Health Care Systems
Health Care Systems	Hedge Fund
Hedge Fund	Hospital
Nuclear De-Comm Trust	Investment Manager
Other-General	Nuclear De-Comm Trust
Pension Fund-Corporate	Other
Trust Company	Small Business Invst Co
Unit Investment Trust	Unit Investment Trust

Table 2
Basic Statistics on Bond Characteristics in the Network

This table presents summary statistics of bond-level characteristics in our network, at each CUSIP-level in Panel A and aggregated at each company-level in Panel B. Spread is calculated as yield minus the Treasury rate of comparable maturity. Rating is calculated as the average rating of three rating agencies, S&P, Fitch, and Moody's, where the categorical ratings are transformed into a numerical scale between 1 (lowest rating) and 21 (highest rating). Volatility is calculated as the standard deviation of traded price during each quarter. We use two measures of bond illiquidity, the Amihud (2002) price impact measure and the interquartile range ("IQR") of traded prices. All variables are winsorized at top and bottom 1 percentiles. Source: eMAXX, TRACE.

	Panel A: CUSIP-level				
Variable	Obs	Mean	Std. Dev.	Min	Max
Outstanding issue amount ($bil)	281,998	0.63	0.57	0.00	3.00
Remaining maturity (quarter)	325,611	35.02	30.87	2.00	121.00
Coupon rate	322,280	5.59	1.97	1.25	11.00
Spread (quarterly median)	327,553	1.26	2.82	(4.84)	35.04
Spread (last quarterly observation)	327,553	1.22	2.95	(5.26)	38.31
Rating	318,815	13.46	3.60	4.00	21.00
Trade volume (quarterly mean; $bil)	327,553	89.80	127.00	0.20	1,590.00
Trade volume (quarterly median; $bil)	327,553	32.30	69.60	0.10	1,010.00
Volatility	327,553	1.61	1.50	0.02	11.14
Illiquidity: Amihud (quarterly mean)	327,553	3.17E-09	9.02E-09	1.82E-13	2.30E-07
Illiquidity: Amihud (quarterly median)	327,553	1.40E-09	2.62E-09	1.11E-13	3.57E-08
Illiquidity: Amihud (last quarterly observation)	327,553	2.68E-09	6.43E-09	1.59E-13	8.33E-08
Illiquidity: IQR (quarterly mean)	327,553	0.57	0.55	0.01	5.03
Illiquidity: IQR (quarterly median)	327,553	0.43	0.50	0.00	4.12
Illiquidity: IQR (last quarterly observation)	327,553	0.65	0.84	0.00	6.89
	Panel B: Company-level				
Variable	Obs	Mean	Std. Dev.	Min	Max
Outstanding issue amount ($bil)	109,648	1.42	2.27	0.02	17.10
Remaining maturity (quarter)	109,558	32.31	23.67	3.00	118.00
Coupon rate	109,055	6.14	2.00	1.50	11.75
Spread (quarterly median)	109,648	3.25	3.52	-4.65	35.89
Spread (last quarterly observation)	109,648	3.24	3.65	-4.91	39.30
Rating	106,301	12.13	3.77	3.76	20.67
Trade volume (quarterly mean; $bil)	109,648	202.00	276.00	0.20	6820.00
Trade volume (quarterly median; $bil)	109,648	72.80	122.00	0.11	2330.00
Volatility	109,648	1.71	1.51	0.02	11.13
Illiquidity: Amihud (quarterly mean)	109,648	2.86E-09	9.11E-09	1.84E-13	2.30E-07
Illiquidity: Amihud (quarterly median)	109,648	1.09E-09	2.01E-09	1.14E-13	3.07E-08
Illiquidity: Amihud (last quarterly observation)	109,648	2.16E-09	4.75E-09	1.61E-13	8.27E-08
Illiquidity: Interquartile range ("IQR"; quarterly mean)	109,648	0.58	0.51	0.01	5.01
Illiquidity: Interquartile range ("IQR"; quarterly median)	109,648	0.41	0.44	0.00	4.12
Illiquidity: Interquartile range ("IQR"; last quarterly observation)	109,648	0.64	0.74	0.00	6.85

Table 3
Company-Level Interconnectedness Measures

This table presents summary statistics for the corporation-level interconnectedness measures utilized in this paper. Panel A reports the descriptive statistics in the CUSIP-quarter-level panel dataset. Panel B reports descriptive statistics for the cross-section of assets. Specifically, for each asset, we take the arithmetic average of the variables across the time period in which that asset appears in the sample. Degree is defined as in equation (4), while Strength is defined as in equation (2). Overlapping Investors and Cosine Similarity are defined as in equations (5) and (6), respectively. The Number of Quarters variable captures how many quarters a particular asset is in the sample. Source: eMAXX.

Panel A: Interconnectedness Measures in CUSIP-quarter Panel

Variables	*N*	*Mean*	*Median*	*Std. Dev.*	*Minimum*	*Maximum*	*Skew.*	*Kurt.*
				2002:Q3–2020:Q4				
Overlapping Investors	142,357	2.70E-02	2.09E-02	2.30E-02	0.00	0.13	1.08	3.72
Cosine Similarity	142,357	3.71E-04	3.78E-04	1.71E-04	0.00	7.74E-04	-0.16	2.09
Degree	142,357	63.54	44.00	61.96	1.00	503.00	1.97	8.09
Strength	142,357	696289.00	265959.00	1459221.25	1.00	29810172.00	7.07	79.59
				2002:Q3–2008:Q4				
Overlapping Investors	40,450	1.24E-02	9.04E-03	1.13E-02	0.00	0.07	1.45	5.02
Cosine Similarity	40,450	2.37E-04	2.40E-04	1.05E-04	0.00	4.83E-04	-0.16	2.21
Degree	40,450	46.10	29.00	53.14	1.00	503.00	2.84	14.15
Strength	40,450	415410.38	168016.00	845318.56	1.00	14240261.00	6.26	59.09
				2009:Q1–2020:Q4				
Overlapping Investors	101,907	3.27E-02	2.82E-02	2.39E-02	0.00	0.13	0.80	3.13
Cosine Similarity	101,907	4.25E-04	4.55E-04	1.62E-04	0.00	7.74E-04	-0.62	2.65
Degree	101,907	70.46	52.00	63.82	1.00	486.00	1.77	7.08
Strength	101,907	807778.31	319559.00	1627008.50	1.00	29810172.00	6.62	68.61

Panel B: Interconnectedness Measures in CUSIP Cross-section

Variables	*N*	*Mean*	*Median*	*Std. Dev.*	*Minimum*	*Maximum*	*Skew.*	*Kurt.*
Overlapping Investors	9,147	1.80E-02	1.27E-02	1.70E-02	0.000	9.84E-02	1.331	4.558
Cosine Similarity	9,147	3.08E-04	3.01E-04	1.58E-04	0.000	6.78E-04	1.24E-01	2.080
Degree	9,147	38.29	26.80	39.36	1.00	414.10	2.48	12.81
Strength	9,147	350841.16	157165.13	748009.56	3.00	15275439.00	8.22	107.88
Number of Quarters	9,147	15.563	9.000	17.394	1.000	73.000	1.620	4.926

Table 4
Analysis of Interconnectedness and Spread CUSIP-level

This table presents results from the analysis of interconnectedness and spread using equation (10) at CUSIP-level. The dependent variable is spread of bond i at time t, measured as the average yield for all trades for bond i over comparable Treasury or interpolated maturity-matched swap rate on the same day, and aggregated at the quarterly-level. Main variables of interest are the two measures of interconnectedness, "Overlapping investors" and "Cosine similarity" based on equations (7) and (8). We use two widely used proxies for illiquidity, Amihud (2002) illiquidity measures and interquartile range (IQR) of trade prices. Credit rating has been converted from the average of the three rating agencies, S&P, Fitch and Moody's to a numerical scale between 1 (lowest) and 21 (highest). Time to maturity is in quarters. Outstanding issuance amount and trade volume ($thous) are both in logarithm.

Spread	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Overlapping investors	-21.81***	-37.61***	-8.782*	-33.51***				
	(7.272)	(2.603)	(4.434)	(2.379)				
Cosine similarity					-2,432***	-3,850***	-784.7*	-3,230***
					(732.3)	(268.6)	(445.3)	(239.0)
Amihud illiquidity	1.875e+07	2.065e+07			1.817e+07	1.930e+07		
	(1.462e+07)	(1.550e+07)			(1.414e+07)	(1.455e+07)		
IQR			1.662***	1.406***			1.652***	1.377***
			(0.143)	(0.104)			(0.142)	(0.103)
Rating	-0.483***	-0.332***	-0.440***	-0.285***	-0.488***	-0.338***	-0.442***	-0.291***
	(0.0332)	(0.0191)	(0.0255)	(0.0174)	(0.0337)	(0.0189)	(0.0259)	(0.0173)
Coupon rate	0.206***	0.440***	0.241***	0.443***	0.211***	0.428***	0.241***	0.433***
	(0.0284)	(0.0276)	(0.0260)	(0.0256)	(0.0281)	(0.0274)	(0.0258)	(0.0255)
Time to maturity	0.00357***	-0.00230**	-0.00566***	-0.00985***	0.00383***	-0.00160	-0.00545***	-0.00910***
	(0.00117)	(0.00110)	(0.00162)	(0.00107)	(0.00114)	(0.00112)	(0.00159)	(0.00109)
Outstanding issue amount (log)	0.0975*	0.152***	-0.0507	0.0940***	0.118**	0.205***	-0.0552	0.126***
	(0.0561)	(0.0349)	(0.0371)	(0.0314)	(0.0565)	(0.0343)	(0.0359)	(0.0304)
Trade volume (log)	-0.132***	-0.0967***	-0.0193*	0.0561***	-0.138***	-0.0631***	-0.0202*	0.0832***
	(0.0229)	(0.0238)	(0.0115)	(0.0136)	(0.0227)	(0.0223)	(0.0109)	(0.0141)
FE	Issuer	Time	Issuer	Time	Issuer	Time	Issuer	Time
Observations	265,976	266,707	265,976	266,707	265,976	266,707	265,976	266,707
R-squared	0.597	0.552	0.660	0.592	0.599	0.554	0.660	0.593

Table 5
Analysis of Interconnectedness and Spread: Company-level

This table presents results from the analysis of interconnectedness and spread using equation (10), aggregated at company-level. The dependent variable is spread of a company i's average bond at time t, measured as the average yield for all trades for each of company i's bond over comparable Treasury or interpolated maturity-matched swap rate on the same day, and aggregated at the quarterly-level. Main variables of interest are the two measures of interconnectedness, "Overlapping investors" and "Cosine similarity" based on equations (7) and (8). We use two widely used proxies for illiquidity, Amihud (2002) illiquidity measures and interquartile range (IQR) of trade prices. Credit rating has been converted from the average of the three rating agencies, S&P, Fitch and Moody's to a numerical scale between 1 (lowest) and 21 (highest). Time to maturity is in quarters. Outstanding issuance amount and trade volume ($thous) are both in logarithm.

Spread	(1)	(2)	(3)	(4)
Overlapping investors	-12.11***	-9.251***		
	(2.897)	(2.875)		
Cosine similarity			-1,881***	-1,380***
			(376.1)	(340.2)
Amihud illiquidity	1.438e+07		1.413e+07	
	(1.100e+07)		(1.098e+07)	
IQR		1.585***		1.581***
		(0.117)		(0.117)
Rating	-0.539***	-0.472***	-0.537***	-0.471***
	(0.0338)	(0.0299)	(0.0332)	(0.0292)
Coupon rate	0.186***	0.222***	0.187***	0.223***
	(0.0207)	(0.0198)	(0.0214)	(0.0205)
Time to maturity	0.000512	-0.00703***	0.000627	-0.00693***
	(0.00118)	(0.00120)	(0.00120)	(0.00122)
Outstanding issue amount (log)	0.256***	0.117**	0.236***	0.0985**
	(0.0529)	(0.0493)	(0.0440)	(0.0373)
Trade volume (log)	-0.133***	-0.0146	-0.131***	-0.0131
	(0.0145)	(0.00984)	(0.0143)	(0.00980)
FE	Issuer, time	Issuer, time	Issuer, time	Issuer, time
Observations	104,879	104,879	104,879	104,879
R-squared	0.726	0.751	0.726	0.751

Table 6
Analysis of Interconnectedness and Liquidity: CUSIP-level

This table presents results from the analysis of interconnectedness and liquidity using equation (11) at CUSIP-level. The dependent variable is liquidity of bond i at time t, measured following Amihud (2002) in Panel A and interquartile range (IQR) of trade prices in Panel B. Main variables of interest are the two measures of interconnectedness, "Overlapping investors" and "Cosine similarity" based on equations (7) and (8). Credit rating has been converted from the average of the three rating agencies, S&P, Fitch and Moody's to a numerical scale between 1 (lowest) and 21 (highest). Time to maturity is in quarters. Outstanding issuance amount and trade volume ($thous) are both in logarithm.

	Panel A: Amihud illiquidity				Panel B: Interquartile range (IQR) of trade prices			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Overlapping investors	-2.36e-08	9.59e-09			-8.103***	-2.775***		
	(1.64e-08)	(1.53e-08)			(1.882)	(0.424)		
Cosine similarity			-4.45e-06***	-1.94e-06			-1,046***	-477.3***
			(1.32e-06)	(1.34e-06)			(170.5)	(43.34)
Rating	-9.42e-11***	-1.85e-10***	-1.01e-10***	-1.82e-10***	-0.0271***	-0.0359***	-0.0291***	-0.0362***
	(0)	(0)	(0)	(0)	(0.00577)	(0.00227)	(0.00566)	(0.00218)
Coupon rate	-3.57e-10***	-1.95e-10***	-3.35e-10***	-2.04e-10***	-0.0251***	-0.00501	-0.0218***	-0.00666**
	(5.36e-11)	(5.46e-11)	(5.22e-11)	(5.11e-11)	(0.00360)	(0.00315)	(0.00350)	(0.00300)
Time to maturity	0***	0***	0***	0***	0.00592***	0.00580***	0.00597***	0.00587***
	(0)	(0)	(0)	(0)	(0.000252)	(0.000248)	(0.000254)	(0.000245)
Outstanding issue amount (log)	9.45e-10***	5.49e-10***	1.09e-09***	7.55e-10***	0.0999***	0.0491***	0.117***	0.0675***
	(1.24e-10)	(7.09e-11)	(1.04e-10)	(6.90e-11)	(0.0190)	(0.00804)	(0.0176)	(0.00839)
Trade volume (log)	-1.05e-09***	-1.47e-09***	-1.08e-09***	-1.47e-09***	-0.0794***	-0.130***	-0.0831***	-0.127***
	(1.08e-10)	(1.27e-10)	(1.08e-10)	(1.38e-10)	(0.00743)	(0.00656)	(0.00742)	(0.00644)
FE	Issuer	Time	Issuer	Time	Issuer	Time	Issuer	Time
Observations	265,976	266,707	265,976	266,707	265,976	266,707	265,976	266,707
R-squared	0.103	0.354	0.104	0.354	0.400	0.444	0.411	0.448

Table 7
Analysis of Interconnectedness and Liquidity: Company-level

This table presents results from the analysis of interconnectedness and liquidity using equation (11) at company-level. The dependent variable is liquidity of bond i at time t, measured following Amihud (2002) in Panel A and interquartile range (IQR) of trade prices in Panel B. Main variables of interest are the two measures of interconnectedness, "Overlapping investors" and "Cosine similarity" based on equations (7) and (8). Credit rating has been converted from the average of the three rating agencies, S&P, Fitch and Moody's to a numerical scale between 1 (lowest) and 21 (highest). Time to maturity is in quarters. Outstanding issuance amount and trade volume ($thous) are both in logarithm.

	Panel A: Amihud illiquidity		Panel B: Interquartile range (IQR) of trade prices	
	(1)	(2)	(3)	(4)
Overlapping investors	1.46e-08		-1.672***	
	(4.33e-08)		(0.453)	
Cosine similarity		-1.13e-06		-327.4***
		(3.45e-06)		(51.53)
Rating	-2.47e-10***	-2.29e-10***	-0.0445***	-0.0439***
	(0)	(0)	(0.00360)	(0.00359)
Coupon rate	-2.39e-10***	-2.56e-10***	-0.0247***	-0.0249***
	(0)	(5.05e-11)	(0.00327)	(0.00316)
Time to maturity	0***	0***	0.00493***	0.00494***
	(0)	(0)	(0.000204)	(0.000205)
Outstanding issue amount (log)	4.82e-10	7.28e-10**	0.0918***	0.0935***
	(6.27e-10)	(3.46e-10)	(0.00959)	(0.00789)
Trade volume (log)	-9.58e-10***	-9.52e-10***	-0.0833***	-0.0828***
	(1.24e-10)	(1.24e-10)	(0.00402)	(0.00399)
FE	Issuer, time	Issuer, time	Issuer, time	Issuer, time
Observations	104,879	104,879	104,879	104,879
R-squared	0.410	0.410	0.535	0.536

Table 8
Analysis of Interconnectedness and Volatility: CUSIP-level

This table presents results from the analysis of interconnectedness and volatility using equation (12) at CUSIP-level. The dependent variable is volatility of bond i at time (quarter) t, measured as standard deviation of trade prices of bond i during each quarter. Main variables of interest are the two measures of interconnectedness, "Overlapping investors" and "Cosine similarity" based on e quations (7) and (8). We use two widely used proxies for illiquidity, Amihud (2002) illiquidity measures and interquartile range (IQR) of trade prices. Credit rating has been converted from the average of the three rating agencies, S&P, Fitch and Moody's to a numerical scale between 1 (lowest) and 21 (highest). Time to maturity is in quarters. Outstanding issuance amount and trade volume ($thous) are both in logarithm.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Overlapping investors	-3.422	-2.493	7.283	0.623				
	(6.416)	(1.671)	(5.114)	(1.512)				
Cosine similarity					-968.6*	-531.9***	395.7	-56.07
					(528.4)	(142.2)	(403.0)	(125.6)
Amihud illiquidity	1.370e+07	1.657e+07			1.341e+07	1.641e+07		
	(1.181e+07)	(1.108e+07)			(1.153e+07)	(1.094e+07)		
IQR			1.361***	1.065***			1.361***	1.064***
			(0.0986)	(0.0558)			(0.0996)	(0.0558)
Rating	-0.126***	-0.0942***	-0.0906***	-0.0590***	-0.128***	-0.0944***	-0.0894***	-0.0589***
	(0.0164)	(0.00808)	(0.0107)	(0.00626)	(0.0163)	(0.00796)	(0.0106)	(0.00622)
Coupon rate	0.0147	0.0362***	0.0439**	0.0383***	0.0209	0.0342***	0.0461**	0.0379***
	(0.0181)	(0.0129)	(0.0173)	(0.0117)	(0.0187)	(0.0128)	(0.0178)	(0.0118)
Time to maturity	0.0187***	0.0182***	0.0111***	0.0126***	0.0186***	0.0183***	0.0109***	0.0126***
	(0.000931)	(0.00101)	(0.000914)	(0.000957)	(0.000906)	(0.00100)	(0.000913)	(0.000943)
Outstanding issue amount (log)	0.0537	0.0598***	-0.0692*	0.0166	0.0958**	0.0841***	-0.0483	0.0248*
	(0.0515)	(0.0165)	(0.0414)	(0.0138)	(0.0454)	(0.0167)	(0.0349)	(0.0131)
Trade volume (log)	-0.0711***	-0.0650***	0.0225**	0.0494***	-0.0787***	-0.0617***	0.0200**	0.0491***
	(0.0208)	(0.0160)	(0.00974)	(0.00781)	(0.0215)	(0.0155)	(0.00977)	(0.00739)
FE	Issuer	Time	Issuer	Time	Issuer	Time	Issuer	Time
Observations	265,976	266,707	265,976	266,707	265,976	266,707	265,976	266,707
R-squared	0.300	0.427	0.440	0.504	0.302	0.428	0.440	0.504

Table 9
Analysis of Interconnectedness and Volatility: Company-level

This table presents results from the analysis of interconnectedness and volatility using equation (12) at company-level. The dependent variable is volatility of bond i at time (quarter) t, measured as standard deviation of trade prices of bond i during each quarter. Main variables of interest are the two measures of interconnectedness, "Overlapping investors" and "Cosine similarity" based on e quations (7) and (8). We use two widely used proxies for illiquidity, Amihud (2002) illiquidity measures and interquartile range (IQR) of trade prices. Credit rating has been converted from the average of the three rating agencies, S&P, Fitch and Moody's to a numerical scale between 1 (lowest) and 21 (highest). Time to maturity is in quarters. Outstanding issuance amount and trade volume ($thous) are both in logarithm.

	(1)	(2)	(3)	(4)
Overlapping investors	-2.863**	-0.870		
	(1.256)	(1.050)		
Cosine similarity			-561.4***	-211.5
			(155.3)	(133.6)
Amihud illiquidity	1.117e+07		1.111e+07	
	(7.468e+06)		(7.447e+06)	
IQR		1.099***		1.098***
		(0.0517)		(0.0517)
Rating	-0.137***	-0.0904***	-0.136***	-0.0899***
	(0.0127)	(0.00967)	(0.0126)	(0.00957)
Coupon rate	-0.0209**	0.00354	-0.0212**	0.00325
	(0.0104)	(0.00908)	(0.0103)	(0.00910)
Time to maturity	0.0177***	0.0124***	0.0177***	0.0124***
	(0.000763)	(0.000760)	(0.000763)	(0.000759)
Outstanding issue amount (log)	0.0767***	-0.0192	0.0794***	-0.0157
	(0.0255)	(0.0198)	(0.0222)	(0.0169)
Trade volume (log)	-0.0687***	0.0124*	-0.0679***	0.0127**
	(0.00953)	(0.00631)	(0.00949)	(0.00629)
FE	Issuer, time	Issuer, time	Issuer, time	Issuer, time
Observations	104,899	104,899	104,899	104,899
R-squared	0.501	0.565	0.501	0.565

Notes: This figure illustrates a network of firms constructed via their overlapping portfolios. In this example, Firm 1 holds positive amounts of Assets 1 and 3, and Firms 2 and 3 hold positive amounts of Assets 1 and 2. The resulting network of overlapping portfolios draws links between all firms, since all firms are exposed to asset 1. Notice that Firms 2 and 3 are connected not only through their holding of Asset 1, but also through Asset 2.

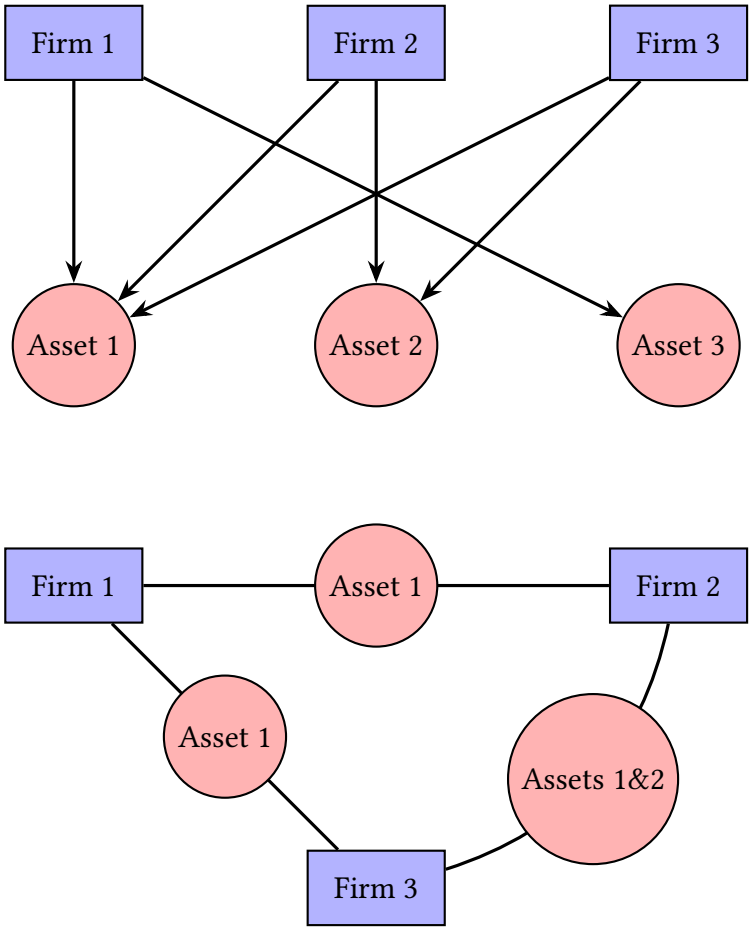


Figure 1. Network of Overlapping Portfolios

Notes: This figure depicts a network of assets constructed via the overlapping investors (firms) which hold them. In this example, Asset 1 is held by Firms 1 and 3, Asset 2 is held by Firms 2 and 3, and Asset 3 is held only by Firm 1. In the network of overlapping investors, Assets 1 and 2 are connected via their common exposure to firms 2 and 3, Assets 1 and 3 are connected through Firm 1, and Assets 2 and 3 are not connected.

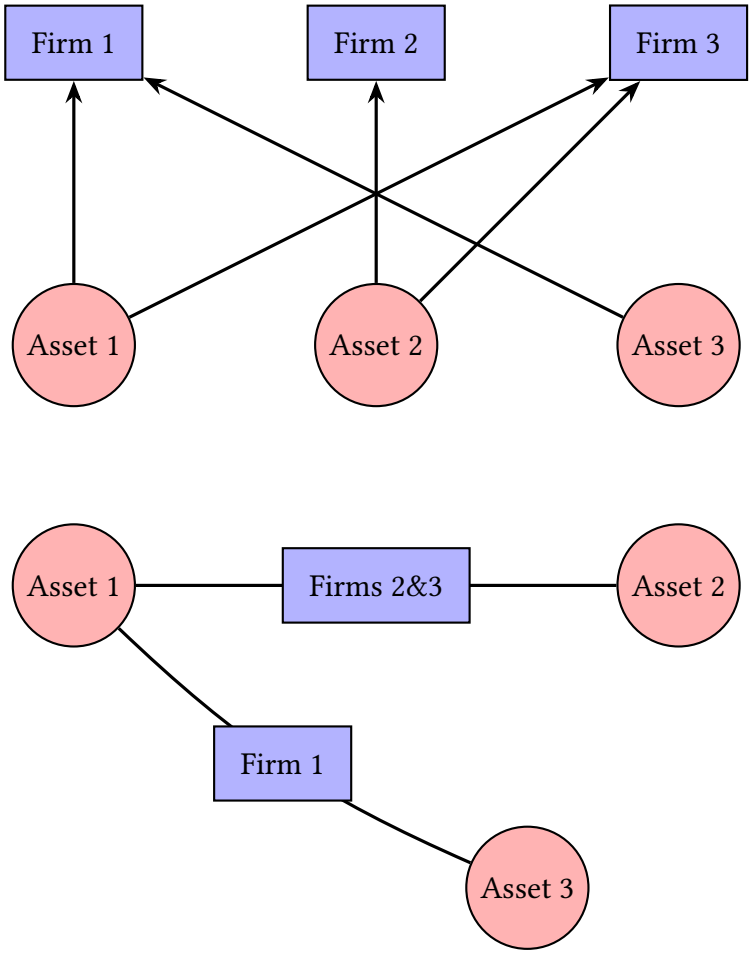
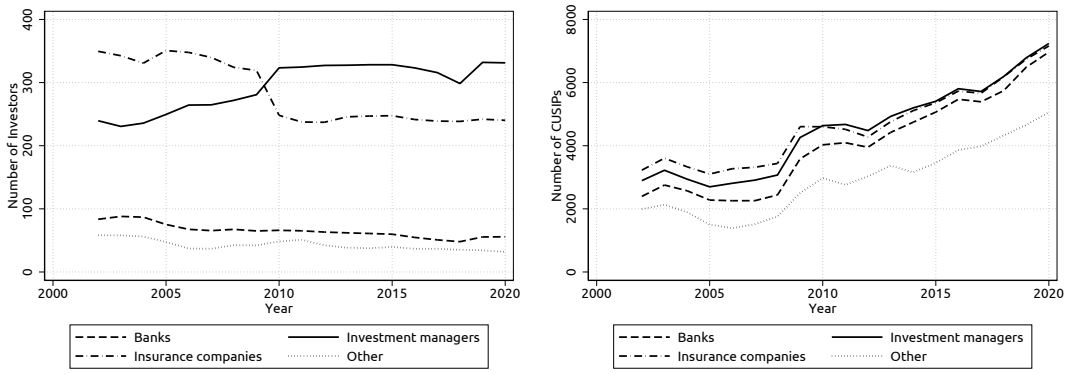


Figure 2. Network of Overlapping Investors

Notes: This figure reports simple counts in our final sample of bond holding data by firm type. Firms' types were carefully verified and assigned via a manual auditing process. Specifically, in subfigure (a), we plot the number of unique firms in the network of financial institutions and assets. In subfigure (b), we plot the number of unique corporate bonds held by these firms over time. Quarterly figures are averaged within a year. Sources: eMAXX.



(a) Number of Unique Investors (b) Number of Unique Corporate Bonds

Figure 3. Basic Statistics on the Bond Holding Data

Notes: This figure depicts how much each firm type holds out of the total outstanding amount of bonds in our final sample of bond holding data. That is, each point represents the sum of bond holdings by the firm type—as shown in eMAXX—divided by the sum of outstanding amount of the bonds based on FISD. Bonds in eMAXX and FISD are matched based on CUSIPs. Quarterly statistics are averaged within each year. Sources: eMAXX and FISD.

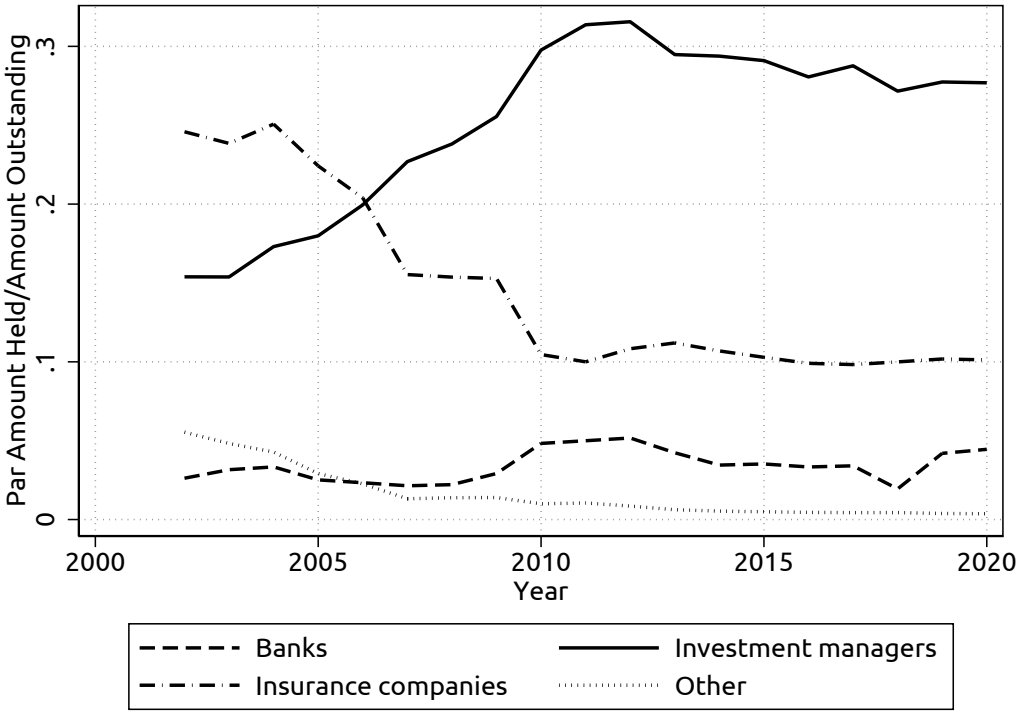


Figure 4. Bond Holdings as a Share of Total Outstanding on FISD

Appendix A Data

A.1 Cleaning eMAXX

- We drop observations for which the external manager is not disclosed (firmid =0).
- Because we focus on institutional investors, we drop observations relating to the holdings of co-managed subaccounts.
- We focus on firms that are domestically domiciled (firm_domicile = "USA")
- We initially sort institutional investors into 4 categories based on the firm_code variable.
    1. Banks: BKM, BKT, BMS, BFM, BKP
    2. Investment managers: INM
    3. Insurance companies: ILF, IMD, IND, IPC, REI
    4. Pension/other firms: GPE, UPE, CPE; EQM, FEN, GVT, HGE, CRP, CRU, FCC, HLC, OTG, SVG, TRT, UIT
- There are some instances in which the market sector of a CUSIP changes over time. To enforce consistency of this variable over time, we collapse the market sector variable to its modal value for each CUSIP.

Appendix B Network of overlapping portfolios

The network of overlapping portfolios is based on the following idea: assume two firms, F_1 and F_2, hold a similar portfolio and F_1 is hit by a shock; F_1 is forced to liquidate (part of) its portfolio which, in turn, will affect the price of the asset holdings of F_2. Even if the shock only affected F_1, the interconnectedness due to the overlapping portfolios creates a contagion effect.

We define the network of overlapping portfolios as $O^F = \left(F, \mathbf{P}^F\right)$, where $F = \{F_1, F_2, ..., F_N\}$ represents the set of financial institutions, and $\mathbf{P}^F$ is the matrix measuring portfolio overlapping.

The first definition simply counts the common assets across two portfolios,

$$P^{F,1} = \left(\overset{\circ}{E}\right)^T \overset{\circ}{E}.$$

The second measure is the *cosine* distance,

$$P_{ij}^{F,2} = \frac{\sum_{k=1}^{N} \overset{\circ}{E}_{ki} \overset{\circ}{E}_{kj}}{\left\|V^{F_i}\right\| \left\|V^{F_j}\right\|}$$

where $\left\|V^{F_i}\right\|$ is the norm of holdings of firm i.

The last overlapping portfolio measure we adopt derives from the notion of Euclidean distance,

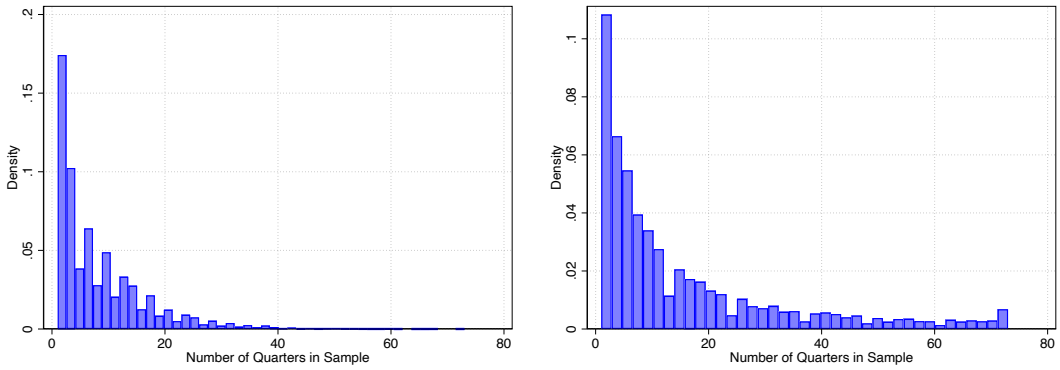
$$P_{ij}^{F,3} = \frac{1}{2} \sum_{k=1}^{N} \left| \overset{\circ}{E}_{ki} - \overset{\circ}{E}_{kj} \right|.$$

Appendix C Interconnectedness Measures

Table C1
Interconnectedness Measures

This table presents summary statistics for the interconnectedness measures utilized in this paper. Panel A reports the descriptive statistics in the CUSIP-quarter-level panel dataset. Panel B reports descriptive statistics for the cross-section of assets. Specifically, for each asset, we take the average of the variables across the time period in which that asset appears in the sample. Degree is defined as in equation (4), while Strength is defined as in equation (2). Overlapping Investors and Cosine Similarity are defined as in equations (5) and (6), respectively. The Number of Quarters variable captures how many quarters a particular asset is in the sample.

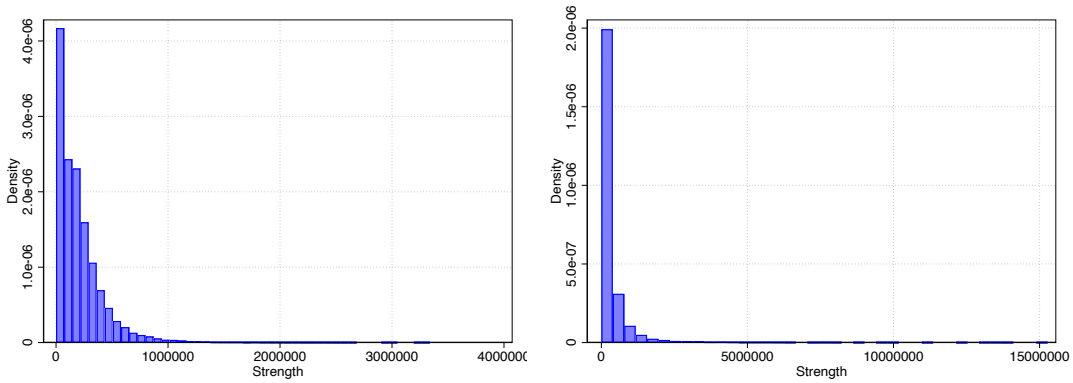
Panel A: Interconnectedness Measures in CUSIP-quarter Panel								
Variables	*N*	*Mean*	*Median*	*Std. Dev.*	*Min.*	*Max.*	*Skew.*	*Kurt.*
				2002:Q3–2020:Q4				
Overlapping Investors	361,292	1.81e-02	1.68e-02	1.37e-02	0.00	7.87e-02	0.46	2.36
Cosine Similarity	361,292	3.45e-04	3.68e-04	1.80e-04	0.00	7.36e-04	-0.29	1.96
Degree	361,292	43.85	41.00	31.24	1.00	237.00	0.74	3.51
Strength	361,292	2.66E05	2.03E05	2.64E05	1.00	7.50E06	2.51	16.75
				2002:Q3–2008:Q4				
Overlapping Investors	89,660	6.49e-03	5.36e-03	5.63e-03	0.00	3.26e-02	0.90	3.36
Cosine Similarity	89,660	1.73e-04	1.84e-04	9.52e-05	0.00	3.95e-04	-0.22	1.98
Degree	89,660	29.55	24.00	26.78	1.00	208.00	1.43	5.80
Strength	89,660	1.83E05	1.21E05	2.21E05	1.00	5.77E06	2.90	19.26
				2009:Q1–2020:Q4				
Overlapping Investors	271,632	2.20e-02	2.24e-02	1.33e-02	0.00	7.87e-02	0.12	2.38
Cosine Similarity	271,632	4.01e-04	4.45e-04	1.66e-04	0.00	7.36e-04	-0.85	2.91
Degree	271,632	48.58	46.00	31.17	1.00	237.00	0.59	3.40
Strength	271,632	2.94E05	2.31E05	2.71E05	1.00	7.50E06	2.47	16.65
Panel B: Interconnectedness Measures in Cross-Section of Assets								
Variables	*N*	*Mean*	*Median*	*Std. Dev.*	*Min.*	*Max.*	*Skew.*	*Kurt.*
Overlapping Investors	43,508	1.48e-02	1.23e-02	1.26e-02	0.000	6.59e-02	0.649	2.512
Cosine Similarity	43,508	3.06e-04	3.09e-04	1.82e-04	0.000	6.80e-04	-8.74e-02	1.750
Degree	43,508	34.42	30.71	27.32	1.00	206.90	0.86	3.66
Strength	43,508	2.04E05	1.53E05	2.19E05	1.00	3.34E06	2.65	16.66
Number of Quarters	43,508	8.304	6.000	8.089	1.000	73.000	1.799	7.172



(a) 8-Digit CUSIP Level

(b) Company Level

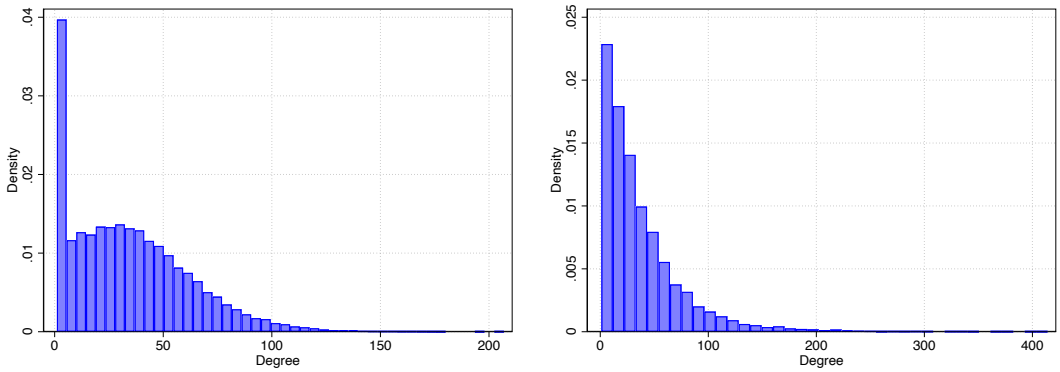
Figure C1. Number of Quarters in Sample



(a) 8-Digit CUSIP Level

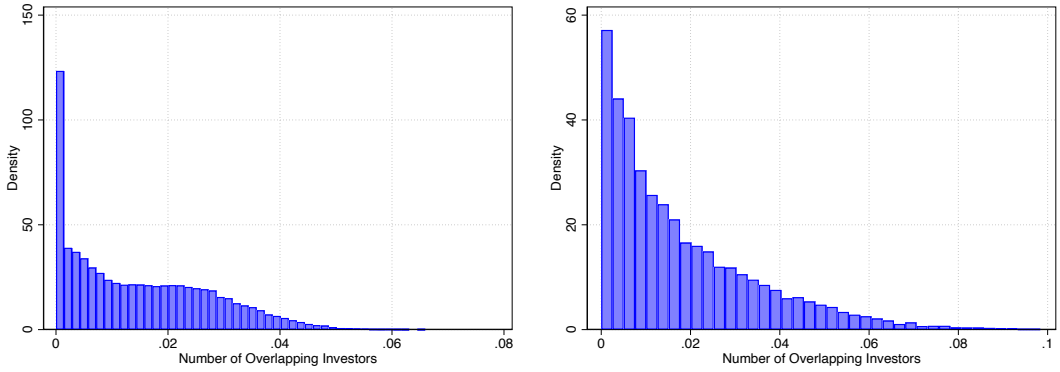
(b) Company Level

Figure C2. Strength



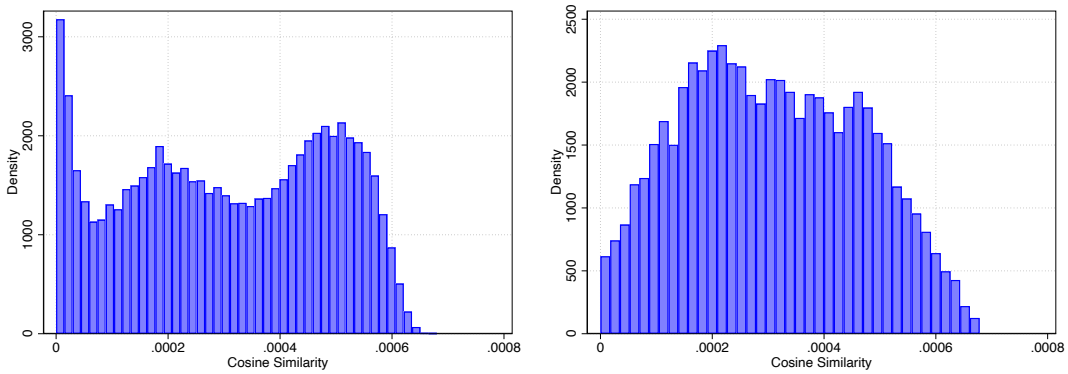
(a) 8-Digit CUSIP Level (b) Company Level

Figure C3. Degree



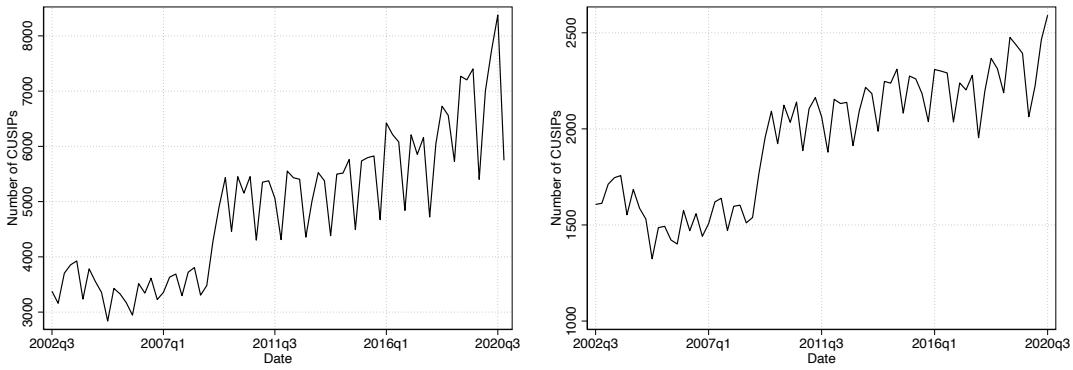
(a) 8-Digit CUSIP Level (b) Company Level

Figure C4. Number of Overlapping Investors



(a) 8-Digit CUSIP Level (b) Company Level

Figure C5. Cosine Similarity



(a) 8-Digit CUSIP Level (b) Company Level

Figure C6. Number of CUSIPs

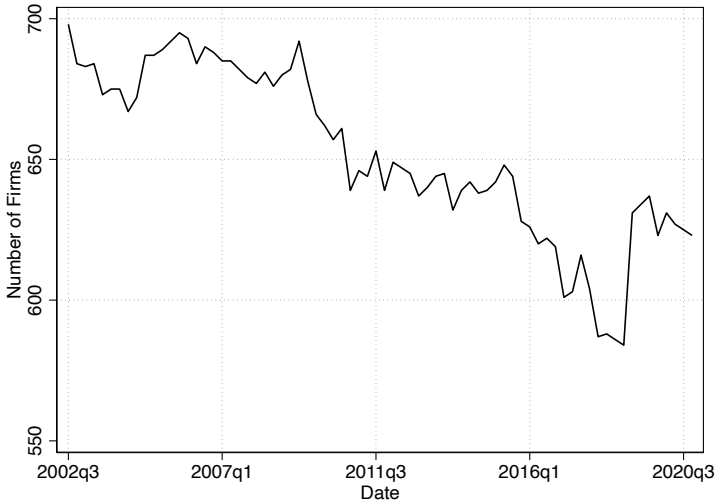
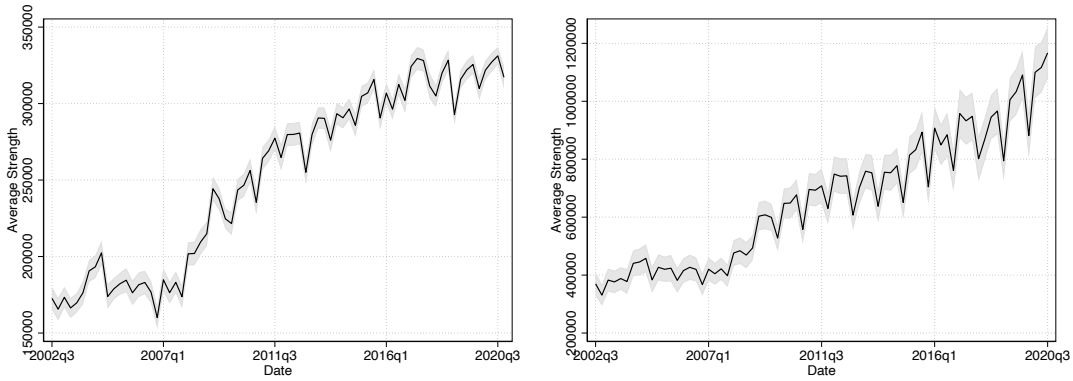


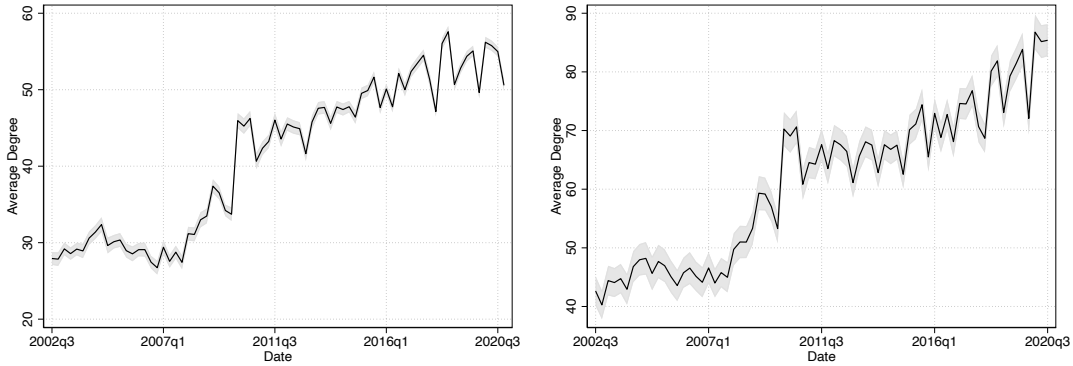
Figure C7. Number of Firms



(a) 8-Digit CUSIP Level

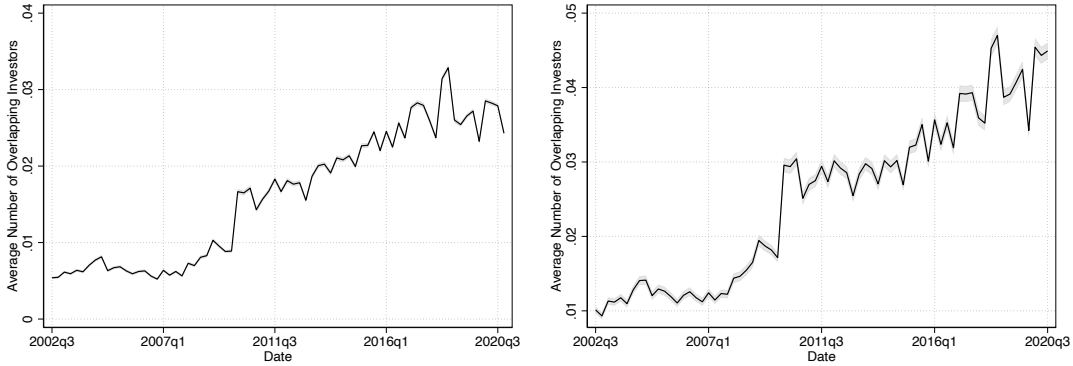
(b) Company Level

Figure C8. Strength



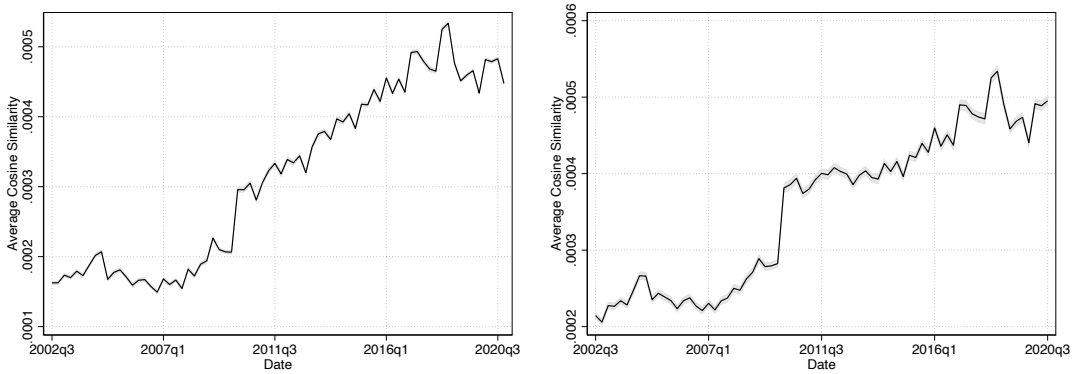
(a) 8-Digit CUSIP Level (b) Company Level

Figure C9. Degree



(a) 8-Digit CUSIP Level (b) Company Level

Figure C10. Number of Overlapping Investors



(a) 8-Digit CUSIP Level

(b) Company Level

Figure C11. Cosine Similarity