Do market-based networks reflect true exposures between banks?

Ben Craig [*] Madina Karamysheva [†] Dilyara Salakhova [‡§]

October 6, 2023

Abstract

We compare networks constructed using five commonly used methods and publicly available daily market data to networks based on reported exposures along several dimensions of the balance sheet, i.e., loans, bonds, equity. Our findings suggest that while the global network structure remains stable, individual exposures are more dynamic. The main message from the regression analysis is that the market-based networks do their job relatively well, however, various market-based networks capture different types of exposures. All the measures reflect common portfolios of bonds and loans. Equity-based measures match better direct and indirect equity, while credit-risk measures capture direct bonds. None of the measures robustly identify direct interbank lending.

Keywords: banking regulation, financial networks, interconnections, market-based networks, true-exposure networks

JEL codes : G20, L14, D85, C63

[*]Ben Craig: Federal Reserve Bank of Cleveland, ben.r.craig@clev.frb.org

[†]Madina Karamysheva: NRU Higher School of Economics, mkaramysheva@hse.ru

[‡]Dilyara Salakhova: Banque de France, IESEG School of Management, d.salakhova@ieseg.fr

[§]Data work was done when Dilyara Salakhova was affiliated with the ECB. We are grateful to Mert Demirer, Laura Liu, Martín Saldías, Julia Schaumburg, Jin-Chuan Duan, Christian Brownlees, Christina Hans, for providing us with the codes. We thank anonymous ECB referees, Matthiew Jackson, Christian Brownlees, Anatoly Peresetsky, Mert Demirer, Laura Liu, and all seminar participants at CEBA webinar series at St.Petersburg State University, ECB DGMF webinar series, CIER Online Research Seminar (MGIMO), Nova SBE seminar series, conference participants at iCEBA conference 2021, CFE Conference 2021, FMND 2022 conference, ICEE at PennState-Altoona 2023. The views expressed in the paper are those of the authors and do not necessarily reflect the views of the European Central Bank or the Eurosystem, Banque de France, the Federal Reserve Bank of Cleveland, or the Board of Governors of the Federal Reserve System. Any errors and omissions are the sole responsibility of the authors.

Contents

Non-technical summary

Since 2008-2009 financial crisis, network analysis focusing on systemic risk and contagion propagation has been used to access stability of the financial system.[1] Regulatory agencies have invested heavily in collecting granular data on exposures across different financial institutions, often focusing on banks. However, these granular data are available to only a small circle of researchers and regulators and only for a subset of banks within a subset of banking systems. To overcome this problem, multiple methods were proposed to reconstruct networks using publicly available data.

In this paper, we focus on how well and what market prices reflect regarding interbank exposures observed in the confidential balance sheet data. First, we construct several benchmark networks based on balance sheet data. Four of these networks reflect direct exposures between banks in loans, cross-holdings of bonds and equity, and cross-holdings of non-traded equity or participation/ownership. The remaining four networks capture indirect exposures, or overlapping portfolios, across the same set of assets. Second, we construct five networks using market data based on estimation techniques widely used in the literature. These approaches use either data on simple partial correlations of equity returns Craig, Saldías (2016) (CS), estimates constructed from both equity returns and volatility (Diebold, Yılmaz (2014) (DDLY), Demirer et al. (2018), tail-risk probabilities Hautsch et al. (2014), Hautsch et al. (2015) (HSS)), default probabilities (Duan et al. (2012) (PD), Chan-Lau et al. (2016)) or credit default swaps (CDS) price data (Brownlees et al. (2020)). Finally, we compare these market-based networks to the balance-sheet based networks to uncover which particular exposures they reflect.

These comparisons go in several directions. First, we look at graphical representation of market-based networks (MBN). Next, we compare their global network characteristics to those of the balance-sheet networks (BSN). We find that most of the networks exhibit properties common to all financial networks such as small-world structure with all nodes being connected in one graph despite a relatively low density of links. Second, we compare MBN to BSN edge by edge using similarity measures such as simple correlations, Jaccard, Hamming and accuracy measures for adjacency (0-1) matrices, as well as cosine similarity

[1] See details about systemic risk in financial networks in survey by Jackson, Pernoud (2021).

for weighted links. We document that various measures highlight different aspects with no single measure being ideal for assessing the similarity between two networks.

Finally, the core of our study is a regression analysis. We run four types of regressions that look at a relationship between an increase in intensity of balance-sheet exposures by one standard deviation and the probability/intensity of observing a link in a market-based network. We study the extensive margins using logit regressions while the intensive margins are analysed using Poisson specification that addresses multiple zeros in our database due to networks sparsity. We also run these regression in two directions. Indeed, as we are exploring associations between the MBN and BSN rather than causal models, we first look how MBN can correctly guess balance-sheet exposures. In other words, each link of a single balance sheet item is regressed on the corresponding links of all of the market-based networks. Then we examine the associations in the other direction by analyzing which BSN exposures are important in explaining a single MBN set of exposures.

Our findings suggest that the various market-based networks reflect different information in the balance sheet exposures, but, overall, all market-based networks reflect better exposures related to common portfolio holdings. On the other hand, credit-risk market-based networks such as the default probability and CDS networks identify well balance sheet variables related reflecting credit risk, i.e., indirect loan and bond portfolios as well as direct bonds. Measures based on returns and return volatility capture mostly common loan and bond portfolios that are sources of banks' variability in returns but also direct and indirect equity exposures.

Granular data on interbank interconnections are not available to academics and most regulatory institutions, but systemic risk and contagion analysis are widely studied using simulated networks. Our policy implication suggests that one should be judicious in the choice of method to construct a network from market data. For example, to investigate propagation of a shock via bilateral links, one should use a network that captures credit risk. In this case, a network reconstructed using an approach based on equity prices may be less relevant than those constructed via techniques that focus, for example, on probability defaults.

1 Introduction

Network analysis is used as an important tool to assess stability of financial systems. Following the 2008-2009 financial crisis, many central banks have invested heavily in collecting granular data on exposures across different financial institutions, particularly focusing on banks. However, even in those banking systems where the data are collected, granular data are available to only a small circle of researchers and regulators. To overcome this problem, multiple methods have been proposed to reconstruct networks using publicly available data, either partial balance sheet data or market data.

The reconstructed networks are often used to run contagion analysis, a recognised tool for stress testing a financial system. For example, the International Monetary Fund (IMF) applies the approach by Diebold, Yılmaz (2014) to analyze stability of a banking system in the Financial Sector Assessment Program[2]. However, a question arises how well reconstructed networks reflect the reality. As for the reconstruction methods applied to partial balance sheet data, Anand et al. (2018) provide a comprehensive analysis of those, concluding that the choice of the best methods depends on the choice of a specific network property one seeks to preserve during the reconstruction.

In our paper, we are interested in focusing on networks reconstructed using publicly available market data. We aim to investigate what these networks represent in terms of granular banks' balance sheet exposures. Do they reflect exposure-based networks well, and if so which exposures: direct cross-holdings of claims on each other or indirect common portfolios? Which market-based approach reflects a particular balance-sheet network (BSN) most closely, and which aspects of the balance sheets do a single market-based network (MBN) represent? Finally, is there a combination of market-based approaches that achieves a closer representation of a balance sheet network?

For the market-based networks, key is a question about information flows from assets held by a bank to market participants who evaluate bank's profitability (as reflected in equity prices) or bank's credit risk (as reflected in default probabilities or CDS spreads). These differences in available information are reflected in two types of networks: indirect

[2] For instance, IMF.FSAP.Norway (2015), IMF.FSAP.Germany (2016), IMF.FSAP.Spain (2017).

networks that shed light on what participants believe about assets that might be held by banks, and how the two banks are related via the assets they hold. Often the market is quite knowledgeable about banks' assets or at least asset categories. The second type of networks is direct networks which reflect information about bilateral relationships among banks, notably, interbank lending or security holdings. This information is potentially hidden from the market.[3]

In the analysis, we use rich granular proprietary data on interbank exposures and portfolio holdings of a set of major European banks. These are quarterly data from 2013q3 to 2019q4. Large Exposures data set from COREP is employed to construct direct networks of bilateral interbank loans and indirect networks of common loan portfolios. Securities Holdings Statistics (SHS-G) database is used to build direct networks of bilateral securities holdings and indirect networks of common securities portfolios. Indirect exposures are computed as a cosine similarity of portfolios of two banks.

To construct market-based networks, we take five different methods that are well-known in the literature. Those methods cover different aspects of the market data and are based on co-movement of volatilities as well as co-movement in returns (Diebold, Yılmaz (2014), Demirer et al. (2018)) - DDLY further on; partial correlation of equity returns (Craig, Saldías (2016)) - CS; partial correlation of idiosyncratic default intensities using the CDS contracts (Brownlees et al. (2020)) - CDS; partial correlation of forward-looking probabilities of default (Chan-Lau et al. (2016)) - PD, and tail-risk network (Hautsch et al. (2015)) approach - HSS. From a methodological point of view, these approaches also cover three main types of network construction, namely, linear and contemporaneous (CS and CDS), linear and dynamic (DDLY and PD), and non-linear and contemporaneous (HSS).

We compare market-based and balance-sheet-based networks to identify which of the MBN methods can better guess presence and size of links in BSN, and which of the balance-sheet variables reflect the existence and intensity of links in MBN. We also invert our analysis to see which MBN methods should be used in order to reconstruct different types of the BSN networks. In our exercise we do not talk about causal relationship as all the links are formed

[3] Some information flows also involve a bank learning about its asset, as in bank monitoring or in the private information implicit in relationship lending. For a survey see Ongena et al. (2000)

endogenously, and balance-sheet and market variables all affect each other. Nevertheless, we believe that our analysis is important as it explores the connections between the market-based networks and balance-sheet exposures. And to the best of our knowledge, it is the first exercise of such a scale.

Our analysis consists of three steps. We start with a graphical analysis that illustrates differences between the market-based networks visually.[4] We observe significant differences among networks in their density, link distribution, cross- and within-country connections, etc. We further describe the networks using global characteristics. We find that while networks evolve over time as we can see on the charts, the global network structure remains stable for both the market- and balance-sheet based networks. We also document that most networks exhibit small world properties characterising financial networks (Bech, Atalay (2008), Cont et al. (2010)), such as low density, negative assortativity and low average shortest paths.[5] This suggests that these networks have a hub-type structure with some well-connected banks connecting all the other banks in one graph. Exceptions include the DDLY network that is complete by construction, the almost complete CDS network due to very low number of nodes (17 nodes in the CDS network vs 55 in all other networks), and the indirect balance-sheet networks with a density higher that normally observed for financial networks. The main reason for the last observation is that the indirect connections arise due to common portfolios and in some sense appear more easily than direct connections which require costly monitoring efforts (Craig, Ma (2018)). For the same reason, these networks also exhibit high clustering where two banks connected with a third one also connected with each other. The CS network attracts additional attention by being somewhat different from other networks. It is the most sparse network among the market-based network, and it has mostly positive assortativity.

In the second step, we run bivariate comparison of network edges among various pairs of networks. In particular, we use simple correlation, three measures comparing only presence or absence of links (Jaccard, Hamming and Accuracy), and the Cosine similarity compar-

[4]We do not plot the balance-sheet networks due to confidentiality constraints

[5]Density corresponds to a share of existing links over all possible links. Assortativity is a correlation between densities. Negative assortativity means that well-connected nodes are connected with less-connected nodes. Average shortest path refers to a number of links needed to connect to randomly chosen banks.

ing also link size. We document that various measures highlight different aspects with no single measure being ideal for assessing the similarity between two networks. Despite this complexity, we note several consistent patterns across these measures. Most market-based networks effectively mirror indirect bonds, indirect loans, and direct bonds, while showing less agreement with indirect and direct equity, and direct loans. These results serve us the basis for further econometric analysis performed in the next step.

We run four types of regressions that look at a relationship between an increase in intensity of balance-sheet exposures by one standard deviation and the probability/intensity of observing a link in a market-based network. We study the extensive margins using logit regressions while the intensive margins are analysed using Poisson specification. The latter allows us to address multiple zeros in our database as most of the networks are very sparse. We also run these regression in two directions. Indeed, as we are exploring associations between the MBN and BSN rather than causal models, we first look if MBN can correctly guess balance-sheet exposures. In other words, each link of a single balance sheet item is regressed on the corresponding links of all of the market-based networks. Then we examine the associations in the other direction by analyzing which BSN exposures are important in explaining a single MBN set of exposures.

Our main findings are the following. The market-based networks capture largely indirect exposures reflecting essentially a common business model among banks as such information is more easily available to public investors. Indirect bonds and indirect loans, constituting the largest part of banks' balance sheets, are the links most often represented by links in the market-based networks. To a lesser extent, direct bonds are also represented by some of the networks, particularly by those that capture credit risk such as the default-probability network and CDS. On the other hand, networks based on equity prices (volatility and returns) such as the DDLY and CS reflect direct and indirect equity exposures. Finally, direct interbank loans that often serve as an input to the interbank contagion analysis cannot be robustly estimated by any of the market-based networks. This is potentially due to the fact that this information is proprietary and not available to the market, but also probably because direct interbank loans constitute a relatively tiny share of banks' exposures.

Our policy implication suggests that one should be judicious in the choice of method to construct a network from market data. Whether the use of a specific MBN or a particular BSN is appropriate for contagion analysis depends on the precise mechanism of financial contagion that is used.

While we find that market-based networks reflect balance-sheet information, our results are somewhat different from the results of Abbassi et al. (2017), whose paper is the closest to ours. The authors provide a comparison of a market-based network constructed using CDS prices to German balance-sheet-networks covering three types of exposures: direct interbank lending, common lending portfolios, common securities exposure to core and peripheral European countries. The sample consists of 13 publicly traded German banks. Abbassi et al. (2017) show that CDS network reflects well true exposures: interbank lending, similarity in lending practice and asset holdings towards troubled European countries. In our case, we find that the CDS measure[6] captures only direct and indirect bond exposures, and this result is robust across all specifications. Unlike Abbassi et al. (2017), we focus on a large set of European banks and compare several methods in terms of their performance to reflect balance-sheet information. This provides a wider insight of what various statistical dependencies in market data represent for observed data on balance sheet connections.

Our contribution to the literature is twofold. First, we contribute to the literature that aims to reconstruct an interbank network from market data or from partial balance sheet data. As granular data on interbank interconnections are not available to academics and most regulatory institutions, but systemic risk and contagion analysis are widely studied using simulated networks, multiple approaches exist for network reconstruction. For example, Diebold, Yılmaz (2014), Brownlees et al. (2020), Hautsch et al. (2015) build networks from market data, while Anand et al. (2018) run a horse race of methods to reconstruct networks from partial balance sheet data. We compare different approaches aiming to build networks using publicly available market data and evaluate their performance relative to the true balance-sheet data.

Second, we contribute to a larger network literature by providing information on the structure of European interbank networks across different assets. Thus we contribute to

[6] We are using exactly the same approach by Brownlees et al. (2020) to construct the CDS network.

the empirical literature describing network characteristics over a relatively long period of time (Boss et al. (2004), Bech, Atalay (2008), Cont et al. (2010)) and confirm that financial networks, particularly, of direct exposures are very sparse and exhibit a small-world property. While individual links may vary over time, the structure and aggregate network properties remain stable over time. As expected, networks reflecting indirect exposures are denser than those of direct exposures with many more banks being connected to each other, but they are far from complete. They still have a hub structure with smaller banks tending to connect to big banks seen in their negative assortativity.

We further contribute to the multiple theoretical studies that focus their analysis on interbank exposures (Freixas et al. (2000), Allen, Gale (2000), Allen et al. (2012)) by providing information on the network structure and differences in networks based on various asset classes.[7] We provide statistics and characterisation of the European banking networks that can be helpful to this theoretical literature to design their models. For example, using our data, we provide further evidence to the result of Elliott et al. (2021) and Jackson, Pernoud (2019) who show that banks are exposed to the same counterparties via different assets.

The rest of the paper is structured as follows. Section 2 describes the data used to construct balance sheet data. Section 3 reports all the five methodologies to construct market-based networks. Section 4 compares market-based and balance-sheet based networks using graphical analysis, global characteristics and bivariate edge comparison via similarity measures. Section 5 discusses results of the econometric analysis. Section 6 provides results of the robustness tests. Finally, Section 7 concludes.

[7] Allen, Babus (2009) is a great survey to look at for further details.

2 Balance sheet networks: data and construction

In this section, first, we describe how we define a sample of banks to construct both balance-sheet and market-based networks for our analysis. Second, we focus on balance-sheet networks and explain the data and the methodology used. We finish the section by providing descriptive statistics of true exposures among banks in our sample.

We start from a list of systemically important banks supervised by the ECB for which the ECB collects granular data on balance sheet exposures with quarterly frequency. We then retain banks that have either ISIN or LEI in databases of securities holdings statistics by banking group (SHS-G) and large exposures (LE COREP) that span the period from 2013q3 to 2019q4.[8]

We use Large Exposures data set from COREP to construct direct networks of bilateral interbank loans and indirect networks of common loan portfolios. Banks report quarterly all their exposures to individual counterparties by instrument type, loans, debt securities and equity, that are higher than EUR 300 million or 10% of their capital.

SHS-G data is used to build direct networks of bilateral securities holdings and indirect networks of common securities portfolios. Banks report quarterly holdings of individual debt and equity securities at a security (ISIN) level.

While banks report data at the consolidated group level, counterparties are reported at the entity level. To construct a network, the counterparty data are aggregated to the group level (see Adam et al. (2019) for more details on the construction of the database). We complement the database with data on total assets and CET1 capital at the group level obtained from Bloomberg.

We then proceed with construction of true exposure networks by making the following modelling choices. We scale each exposure by total assets TA_{it}. Then we build directed networks based on scaled exposures, E_{ijt}/TA_{it}, and undirected networks by taking an average of the weighted links going in both directions $(E_{ijt}/TA_{it} + E_{jit}/TA_{jt})/2$.

Finally, we construct networks based on portfolio, or indirect, exposures using cosine similarity measure.[9] Cosine similarity measures the cosine of the angle between two vectors

[8] In this way we cover about 70% of total assets of the European banking system and time period from 2013q3 - 2019q4.

[9] Words "undirected" and "indirect" refer to different notions. We use undirected networks to refer to

projected in a multi-dimensional space, and it does not depend on size. In our case, a portfolio of assets represents such a vector in a multi-dimensional space equal to the number of assets in the portfolio.

$$\cos\theta = \frac{\sum_{i=1}^{n} a_i b_i}{\sqrt{\sum_{i=1}^{n} a_i^2}\sqrt{\sum_{i=1}^{n} b_i^2}} \tag{1}$$

Our data allow us to define portfolio at different granularity: by security or by loan, by counterparty/borrower, by country and sector. We choose a borrower level to compute cosine similarity for loans and securities. The original cosine similarity leads to an undirected network. But as it makes a significant difference if a bank is similar to another bank by 10% or 90% of its portfolio, we prefer to scale links by common exposures, thus also making similarity links directed. To make a link from bank i to bank j, ij, scaled and directed, we scale the cosine value by bank i' common exposures with bank j weighted by bank i' total assets, TA_i. We do a similar operation for the link ji where we scale the cosine value by bank j' common exposures with bank i weighted by bank j' total assets, TA_j.

To sum up, we construct eight true-exposure networks: direct (loans, securities (equity, bonds, ownership (unlisted equity holdings)) and indirect, reflecting common portfolios, (loans, securities (equity, bonds, ownership (unlisted equity holdings)) for each reported date from 2013q4 to 2019q3.

Table 1 reports the descriptive statistics of true exposure variables we use further in regressions. All true exposure variables are positive. As market-based links can be both directed and undirected depending on the method, we construct BSNs links in both versions as well, directed and undirected. Direct exposures are transformed into undirected networks by taking an average of the direct exposures scaled by respective total assets. Mean indirect exposures are larger than direct exposures by an order of magnitude. This is largely expected since directs links are more expensive to maintain while indirect links are also under a lesser control of banks.

a graph characteristic meaning that a link between nodes A and B is equal to the link between B and A. Indirect exposures refer to common portfolios of nodes A and B. By definition, indirect exposures are represented by undirected networks. In our analysis, we modify indirect and direct exposures to become directed and undirected graphs, respectively.

Table 1: Descriptive statistics of balance-sheet exposure variables

Statistics:	Min	Mean	Median	Max	SD	Observations
Variables						
Panel A: weighted directed networks						
wDL directed	0	0.0025	0.0006	0.434	0.01	5181
wIL directed	0	0.02	0.0008	0.807	0.07	31236
wDB directed	0	0.0005	0.0001	0.01	0.001	11894
wIB directed	0	0.05	0.023	0.98	0.32	27936
wDE directed	0	0.00007	0.000006	0.0057	0.0003	7620
wIE directed	0	0.0006	0.000016	0.041	0.0021	19207
wDP directed	0	0.0003	0.00003	0.006	0.0008	621
wIP directed	0	0.0002	0.000004	0.012	0.0009	1720
Variables						
Panel B: weighted undirected networks						
wDL undirected	0	0.0015	0.0005	0.217	0.006	5181
wIL undirected	0	0.022	0.0009	0.57	0.065	31236
wDB undirected	0	0.0003	0.0001	0.0068	0.0006	11894
wIB undirected	0	0.059	0.0247	0.99	0.228	27936
wDE undirected	0	0.00005	0.000004	0.0029	0.0002	7620
wIE undirected	0	0.0006	0.000035	0.029	0.0018	19207
wDP undirected	0	0.0002	0.000014	0.003	0.0004	621
wIP undirected	0	0.00025	0.000013	0.0075	0.0007	1720

Notes: Exposures are measured as a fraction of total assets. In the name of each variable w stands for weighted by total assets. D - direct exposure, I - indirect or common portfolio exposure, L - loans, B - bonds, E - equity, P - equities non securities or fund shares. Panel A: direct exposures are directed by construction, indirect exposures are made directed by weighing them by the amount of common exposures between the two banks scaled by total assets. Panel B: indirect exposures are undirected networks by construction; direct exposures are transformed into undirected networks by taking an average of the direct exposures scales by respective total assets. Number of observations varies as descriptive statistics is for non-zero links only.

3 Market-based networks

In this section, we describe the market data and methods we use to construct market-based networks.

We obtain daily market data for the sample of banks used to construct true-exposure networks.[10] The data include daily equity prices and CDS contracts (1, 2, 3, 5, 7, and 10-year) denominated in Euros. We remove banks with zero or constant prices and banks with zero intraday volatility (no intraday changes in prices). We also remove subsidiaries that belong to the same banking group since our true exposure data (LE COREP, SHS-G) are consolidated at the group level.[11]

It is worth noting that only large publicly traded banks have CDS instruments, thus significantly limiting the sample size in the networks based on CDS data.

The market-based networks rely on daily data and have to be computed on a window of observations. On the other hand, the balance networks are snapshots at quarter ends. To reconcile the two, we choose to calculate a market-based network using daily observations over the six months preceding a reporting date of balance-sheet exposures. For example, we observe a balance-sheet network of direct loan exposures as of December 31, 2018. Then we take daily data over the previous two quarters, i.e., from July 1, 2018 to December 31, 2018,

[10]Daily equity prices are obtained from Bloomberg. In cases where data is missing, we use Eikon for imputation.

[11]We keep a subsidiary if two conditions are met: first, it represents significant part of total assets of the holding company, and second, if it is the only traded (public) part of the holding company.

to compute a market-based network.[12] We use a 6-month rolling window to have enough observations to estimate market-based networks.

Among the various methods available, we select five distinct approaches from the literature, which encompass the majority of techniques commonly used to construct market-based networks. Each approach relies on various data sets and covers different types of risk. Diebold, Yılmaz (2014) and Craig, Saldías (2016) capture volatility and return market co-movement using equity returns. Hautsch et al. (2015) also uses equity prices but focuses on tail-risk and correlation in extreme events. Finally, Brownlees et al. (2020) and Chan-Lau et al. (2016) capture credit risk using CDS prices and forward-looking default probabilities, respectively.

The approaches also cover main types of network construction, namely, linear and contemporaneous (Craig, Saldías (2016), Brownlees et al. (2020)), linear and dynamic (Diebold, Yılmaz (2014), Chan-Lau et al. (2016)), and non-linear and contemporaneous (Hautsch et al. (2015)). Linear methods capture average effects while non-linear methods focus on extreme events or distribution effects. Dynamic networks enrich the assessment with information from time series unlike contemporaneous networks that use only information available at time t. This is important as often market players and policy-makers are interested not in observing current links but how links will change under certain circumstances like spillover of a market shock.

We now discuss each method in detail.

3.1 Diebold-Yilmaz network. DDLY

The first method, proposed by Diebold, Yılmaz (2014) and Demirer et al. (2018) (DDLY further on),[13] uses a forecast error variance decomposition (FEVD) as a measure of interconnectedness. This is one of the most used methods, particularly, among policy-makers. For example, the International Monetary Fund (IMF) applies it to analyze stability of the financial system in the Financial Sector Assessment Program (IMF.FSAP.Norway (2015), IMF.FSAP.Germany (2016), IMF.FSAP.Spain (2017)).

[12] The exact dates may change depending on the trading/weekend days.

[13] We thank Laura Liu and Mert Demirer for sharing the code with us and useful suggestions regarding the code.

Following Demirer et al. (2018), we rely on equity prices to construct daily range-based realized stock return volatility.

$$\tilde{\sigma}_{it}^2 = 0.511(H_{it} - L_{it})^2 - 0.019[(C_{it} - O_{it})(H_{it} + L_{it} - 2O_{it}) - \\ -2(H_{it} - O_{it})(L_{it} - O_{it})] - 0.383(C_{it} - O_{it})^2 \tag{2}$$

where H_{it}, L_{it}, C_{it}, O_{it} - are logs of daily high, low, opening and closing prices for bank stock i at day t. Their connectedness measure answers the following question: how much of entity $i's$ future uncertainty (at horizon H) is due to shocks arising from entity j? Bank j contribution to bank i H - step-ahead generalized forecast error variance :

$$\theta_{ij}^g(H) = \frac{\sigma_{jj}^{-1} \sum_{h=0}^{H-1} (e_i' A_h \Sigma e_j)^2}{\sum_{h=0}^{H-1} (e_i' A_h \Sigma A_h' e_j)}, H = 1, 2, ... \tag{3}$$

where Σ - is covariance matrix of disturbance vector ϵ, σ_{jj} - is standard deviation of disturbance of the equation j , e_i - is selection vector with one at position i and zeros otherwise. While different pairwise directional connectedness from j to i at different horizons is:

$$\tilde{\theta}_{ij}^g(H) = \frac{\theta_{ij}^g(H)}{\sum_{j=1}^{N} \theta_{ij}^g(H)} \tag{4}$$

Following the authors, we use Adaptive Elastic Net to estimate the large Vector Autoregression model. The DDLY network has several important features. It allows for pseudo-dynamic networks, which estimate different static networks for different periods, as well as dynamic networks, which can be obtained by rolling window estimation. The resulting network is directed and complete with positive weights.

We take the original DDLY approach (Demirer et al. (2018)) as a baseline measure. We further modify it in several ways to create alternative measures. Similar to the approach suggested in Hale, Lopez (2019) and Barigozzi, Brownlees (2019), we remove common factors by using defactored version of the initial model. We construct a network of residuals after regressing bank's return volatility on four common factors - European stock market index, European banking sector index, option-based implied volatility European market index, and option-based implied volatility European banking sector index.[14] By doing so we capture

[14] We choose a maturity of 6 months as the most liquid one.

only idiosyncratic movements across banks. For robustness we also construct a network on demeaned volatility using cross-sectional mean instead of factors. Finally, using Demirer et al. (2018) we build a network, based on log returns.

3.2 Craig-Saldias network. CS

The second approach, proposed by Craig, Saldías (2016), uses an interconnection measure based on correlations and assessed in a two-step procedure.[15] This method is careful about removing strong common factors, which may often be the reason behind identified links. Following Craig, Saldías (2016) (hereafter referred to as CS), we remove strong factors using principal component analysis (PCA), then conduct cross-sectional dependence (CD) and compute the correlation matrix. In the second step, we apply regularization or thresholding to the correlation matrix. This step selects the correlation coefficients, $\hat{\rho}_{ij}$, among weakly-dependent residuals that are statistically different from zero at a given significance level (five percent) from all possible $N(N-1)/2$ elements of the correlation matrix using the Holm-Bonferroni method). The resulting network, using the Craig, Saldías (2016) approach, is a partial correlation network of bank equity returns. It is undirected and sparse with weights in a range between -1 and 1.

As a baseline measure we take one-factor and a five percent threshold (CS further on). As a sensitivity check, we also consider the following combinations: one-factor and a 10% threshold (CS_TH10_F1 further on), one-factor and a 25% threshold (CS_TH25_F1), and three-factors with a 25% threshold (CS_TH25_F3).

3.3 Tail-risk network. HSS

The third approach, developed in Hautsch et al. (2014) and Hautsch et al. (2015), proposes to measure interconnections between banks at extreme events.[16] A tail risk network is based on the conditional Value-at-Risk (VaR) of equity returns. It shows how much market capitalization of a bank will drop in one day (with a given probability 10%) following a distress or a sharp decline in other banks' equity prices. The firm-specific VaRs are estimated as func-

[15] We thank Craig, Saldías (2016) for providing us with the code.

[16] We thank Julia Schaumburg for giving helpful suggestions regarding the code.

tions of firm-specific characteristics, macroeconomic fundamentals, and tail-risk spillovers from other banks, which are captured by loss exceedances. Loss exceedance is a common tail-risk metric, it equals to return values when returns are below the empirical 10% quantile, and zeros otherwise.

The complete set of risk-drivers is given by: $R_t^i = (Z_{t-1}^i, M_{t-1}, N_t^{-i})$, where Z_{t-1}^i - balance sheet characteristics of bank i, M_{t-1} - general market conditions, N_t^{-i} - "network impacts" from all other banks (loss exceedances).[17] To estimate VaR, we use a linear quantile regression with Lasso-regularization:

$$Return^{i,t} = \xi_0^i + \xi_1^{i,t} Z_{t-1}^i, + \xi_2^i M_{t-1} + \xi_k^i N_t^{-i} + e_0^{i,t} \tag{5}$$

$$V\hat{a}R_t^i = \hat{\xi_0^i} + \hat{\xi^{i,t}} * R_t^i \tag{6}$$

where, $Return^{i,t}$ is a daily equity return of bank i at day t. The use of loss exceedance assumes that bank k affects the VaR of bank i only if bank k is under stress. The idea is to estimate how a relatively large decline in daily equity returns of bank k, (N_t^k), feeds through bank's i returns (more precisely, to its VaR), other things being equal. The vector of parameters of interest is ξ_k^i. Each element of the vector shows how much Value-at-Risk of bank i would change if the loss exceedance of bank k rose by one percentage point. These parameters, which can be both positive and negative, serve as our weights in the network.

Following Hautsch et al. (2015), we address the high-dimensionality issue (potentially there are a lot of unimportant risk-drivers in R_t^i, but we are interested in choosing only a subset of relevant factors), using a Least Absolute Shrinkage and Selection Operator (LASSO) technique (see Belloni, Chernozhukov (2011)) that allows for identification of relevant tail risk drivers for each bank in a data-driven way. Two types of cross-validation (CV) are used: K-fold (out-of-sample) and BIC (in-sample). After finding an optimal model (or, alternatively, an optimal level of network scarcity) via cross-validation, we discard those factors in R_t^i which are smaller than the threshold (0.0001, see Hautsch et al. (2014)). We then re-estimate the model without shrinkage using the final subset of regressors left after the

[17] Balance sheet characteristics include: leverage, that corresponds to total assets divided by total equity; maturity mismatch measured as the ratio of short-term debt to total debt; and size, proxied by the logarithm of total assets. General market conditions are proxied by variables that are described in Table 19 in Appendix A.1.

shrinkage and thresholding process.

The analysis is repeated for all banks in the system, and tail-risk interconnections can then be depicted and summarized in a network graph. The resulting tail-risk network is directed and sparse with weights being positive or negative.

3.4 Default probability network. PD

Chan-Lau et al. (2016) approach focuses on a forward-looking probability of default (PD). We choose this approach for two reasons: first, default probabilities measure credit risk and financial distress. Second, forward-looking measure dynamically responds to the state of the economy.

The data for this approach is provided directly by the authors.[18] The Probability of Default (PD), defined by the Credit Research Initiative (CRI), measures the likelihood that an individual bank will be unable to fulfill its financial obligations (Duan et al. (2012)). The CRI PD features term structures ranging from 1 to 60 months. More specifically, to construct default correlations Duan, Miao (2016) proceed as follows. First, the authors identify a set of predetermined credit risk factors, estimate the factor model, and produce the factor model residuals. In the second step, they estimate the time series dynamics of the predetermined credit risk factors and individual factor model residuals. Next, the authors construct a sparse correlation matrix for the factor model residuals after taking out their individual time series effect. Finally, they further calibrate the model to the term structure of PDs at the time of application to take advantage of the information embedded in longer-term PDs.

The full correlation matrix is then transformed into a partial correlation matrix that by definition reflects the residual correlation after subtracting any indirect impact from other parties in the system, or direct default risk linkages among institutions. The sparsity of the network is achieved through the use of the regularisation algorithm CONCORD (Khare et al. (2015)). The resulting partial correlation network, representing forward-looking default probabilities is sparse and undirected with the weights ranging from -1 to 1.

[18]Thanks to Chan-Lau et al. (2016) and Duan et al. (2012) for providing us directly with the network for our sample of banks.

3.5 CDS network. CDS

This approach is based on a credit risk model proposed by Brownlees et al. (2020).[19] We choose this method as it allows us to observe interconnections reflecting credit risk. CDS prices are particularly convenient as they explicitly price credit risk, unlike share prices.

The authors start with a reduced-form credit risk model where the dependency on defaults among financial entities originates through three channels: a global factor, a country-specific factor, and a banking network channel. Default intensity interdependence is induced both by exposure to common systematic components as well as by dependence between idiosyncratic shocks. To bootstrap risk-neutral idiosyncratic default intensities from CDS data the authors apply standard pricing formulas for single-name Credit Default Swap (CDS) contracts derived in Ang, Longstaff (2013). Next, the credit-risk network is constructed based on a partial correlation of idiosyncratic components of default intensities using a LASSO procedure, Adaptive Graphical LASSO.[20] The resulting network is sparse and undirected with weights ranging from -1 and 1.

3.6 Descriptive statistics

Table 2 displays the descriptive statistics of market-based variables.

Panel A of table 2 shows that all the DDLY networks are complete with positive weights as expected. The original DDLY network as well as modified networks based on intraday price volatility and return-based network look rather similar with the distributions being mostly of the same order of magnitude. The only network that stands out is the one with eliminated common factors but without scaling. Eliminating common factors increases the maximum value of weights, but it decreases the mean and median. Original paper Demirer et al. (2018) does not eliminate common factors, thus no need to apply z-score normalization. However, elimination of common factors leads to weights explosion due to potentially increased heterogeneity across banks. Indeed, common factors are a big component in prices and volatilities that make banks look similar to each other. Once

[19] We refer the reader for details both on the model and methodology to Brownlees et al. (2020). We thank Brownlees et al. (2020) for kindly providing us with the code.

[20] For details on Adaptive GLASSO, please see Yuan, Lin (2007), Friedman et al. (2008), Fan et al. (2009).

these factors are suppressed, banks become extremely heterogeneous, particularly, in their variance. Heteroskedasticity might cause a problem in the estimation of the adaptive elastic net, as the penalty term would be very different for different banks.

Panel B of table 2 displays descriptive statistics of Craig, Saldías (2016) network. First, this is a partial correlation network, and weights can be both positive and negative. A positive correlation is interpreted as a co-movement between bank returns, while negative correlations can be interpreted as a diversification benefit. We report statistics for positive and negative weights, as well as all weights combined. Panel B suggests that CS network is very sparse, as the number of non-zero links is less than 2%. This is partly due to the fact that CS approach filters out common factors using principal component analysis.

Panel F shows statistics for variations of CS approach with the alternative cut-off threshold for correlation and a number of factors. All the CS networks show very similar statistics. The only difference worth noting is a drastic reduction in the number of positive links in the network with three factors.

Panels C, D, and E display the descriptive statistics for the tail-risk network by Hautsch et al. (2015), for the default probabilities network by Chan-Lau et al. (2016), and for the credit risk network by Brownlees et al. (2020) respectively. For HSS, PD, and CDS networks the weights can be also positive and negative. Number of observations is much smaller than in the DDLY networks but larger than in the CS networks. All the three types of networks have correlations much lower than in the CS networks with the HSS method producing the lowest weights.

Table 2: Descriptive statistics of Market-based variables

Statistics:	Min	Mean	Median	Max	SD	Observations
Variables						
Panel A						
DDLY_vol	0.0000	1.4290	1.0500	25.4000	1.3646	55058
DDLY_vol_4factors_scaled	0.0000	1.3064	0.6730	49.3000	1.7489	55058
DDLY_vol_demeaned_scaled	0.0000	1.1670	0.5990	35.7000	1.5882	55058
DDLY_vol_demeaned	0.0000	1.5134	0.4410	77.5000	4.9823	55058
DDLY_return_4factors_scaled	0.0000	1.3381	0.8890	25.5000	1.4300	55058
Panel B						
CS_positive	0.3539	0.5285	0.4819	0.8935	0.1452	820
CS_negative	-0.4679	-0.3917	-0.3840	-0.3553	0.0301	100
CS	-0.4679	0.0072	0.0000	0.8935	0.0686	55058
Panel C						
HSS_positive	0.0000	0.0027	0.0018	0.0444	0.0030	6887
HSS_negative	-0.0225	-0.0017	-0.0011	0.0000	0.0019	2122
HSS	-0.0225	0.0003	0.0000	0.0444	0.0015	51296
Panel D						
PD_positive	0.0001	0.1094	0.0665	0.8604	0.1207	7188
PD_negative	-0.3066	-0.0571	-0.0417	-0.0002	0.0512	2626
PD	-0.3066	0.0153	0.0000	0.8604	0.0687	41588
Panel E						
CDS_positive	0.0001	0.1267	0.0789	0.9401	0.1415	2682
CDS_negative	-0.5777	-0.0685	-0.0501	-0.0001	0.0636	1644
CDS	-0.5777	0.0439	0.0053	0.9401	0.1398	5174
Panel F						
CS_TH1_F1_positive	0.3416	0.5108	0.4545	0.8935	0.1478	910
CS_TH1_F1_negative	-0.4679	-0.3809	-0.3734	-0.3425	0.0325	132
CS_TH1_F1	-0.4679	0.0075	0.0000	0.8935	0.0705	55058
CS_TH25_F1_positive	0.3235	0.4889	0.4342	0.8935	0.1502	1038
CS_TH25_F1_negative	-0.4679	-0.3662	-0.3579	-0.3236	0.0355	188
CS_TH25_F1	-0.4679	0.0080	0.0000	0.8935	0.0730	55058
CS_TH25_F3_positive	0.3237	0.4103	0.3821	0.6919	0.0835	328
CS_TH25_F3_negative	-0.4701	-0.3605	-0.3537	-0.3229	0.0334	198
CS_TH25_F3	-0.4701	0.0011	0.0000	0.6919	0.0389	55058

*Notes:*In the name of each variable DDLY stands for Demirer et al. (2018) approach, CS - Craig, Saldías (2016) approach, HSS - Hautsch et al. (2015) approach, PD - Chan-Lau et al. (2016) approach, CDS - Brownlees et al. (2020) approach, vol - volatility based network, return - return based network, words "positive", "negative" refer to non-zero correlations with the respective sign.

4 Network comparison

In this section we aim to compare market-based and balance-sheet networks along various dimensions. We start by looking at a graphical representation of market-based networks over the full period and change in network structure in different periods.[21] Second, we compare market-based and balance-sheet networks along global network characteristics but also edge-by-edge. The latter provides us with a first glance into how well links in market-based networks match links in balance-sheet networks.

4.1 Graphical Analysis

In this subsection, we show how market-based networks differ visually. Networks are computed on daily trading data of 62 European banks from 2013q3 to 2019q4. It is also possible to observe the evolution of networks over time by drawing individual networks for each quarter. We report the graphs with network dynamics in Appendix A.3.

Figures 1a - 2d demonstrate a graphical representation of volatility network (DDLY), defactored volatility network (DDLY), defactored return network (DDLY), partial correla-

[21] We do not draw balance-sheet networks due to data confidentiality.

tion network of equity returns (CS), tail-risk network (HSS), default probability network (PD), credit-risk CDS network (CDS). In all graphs, a node corresponds to a bank with a node name being a bank ticker symbol from Bloomberg.[22] Node size reflects a weighted degree, and node color defines a country. A node location within a graph is defined by a Fruchterman-Reingold and a circle pack layout that allows us to draw illustrative charts with larger and more connected nodes being closer to the center and all banks of the same country being clustered together. Edge size and color reflect pairwise directional connectedness (to and from) for the directed and partial correlation for the undirected approach. The edge color goes from light purple (the lowest weight and the weakest connection) to dark purple (the highest weight and the strongest connection).

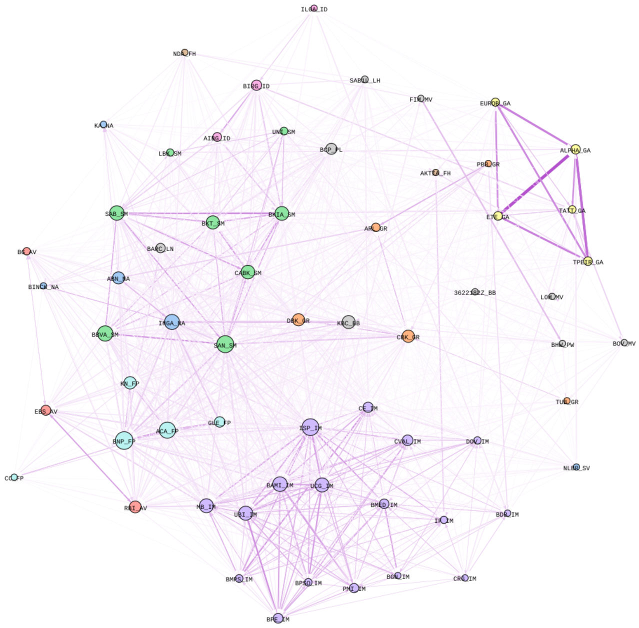
Figures 1a, 1b and 1c show networks based on forecast error variance decomposition with each link reflecting a share of forecast error variance in bank i due to the shock to bank j, thus weights are non-negative and can be higher than one. All the three networks are complete networks, with a density equal to 100. These graphs are directed, and the edge from node i to node j is not necessarily equal to the edge from node j to node i. The three networks based on volatility, defactored volatility, and defactored return over the full period of time visually show little difference, however, charts with the evolution of the networks over time presented in the Appendix A.3 show variability across construction methods.

Figures 2a-2d represent partial correlation networks, thus weights can be both positive and negative (not seen on the charts) and belong to the interval [-1,1]. Higher positive weights suggest higher correlation and higher risk of spreading a shock, while negative correlation implies diversification benefits. The CS network is the most sparse. The CDS network is almost complete as it consists of only 18 banks versus 62 banks in the full sample. Already visually, we can observe high heterogeneity across links with few connections being very strong and the majority of the links being rather weak.

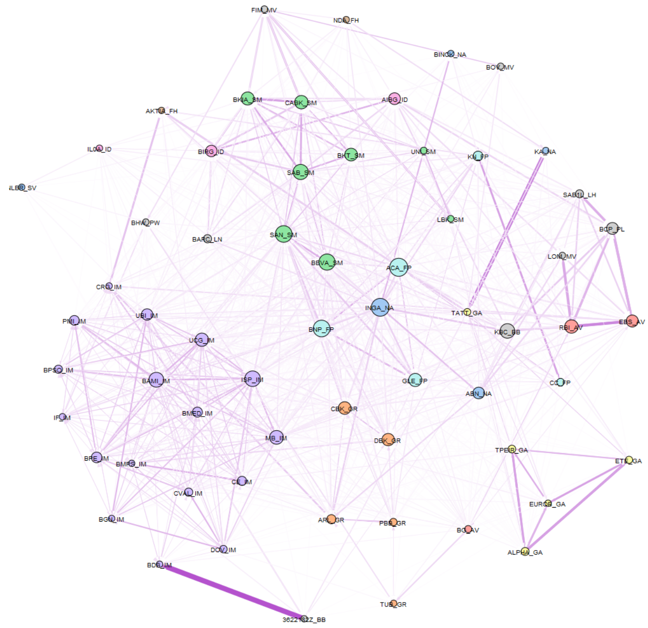
Various approaches capture differently within- and cross-country links. DDLY and CS methods demonstrate stronger links within-country (more intense connections for the DDLY networks and higher clustering for the CS networks) but also capture connections between countries (Figures 1a-1c and 2a). The networks of tail risk, CDS, and default probabilities

[22]The detailed list of banks is in the appendix A.5.

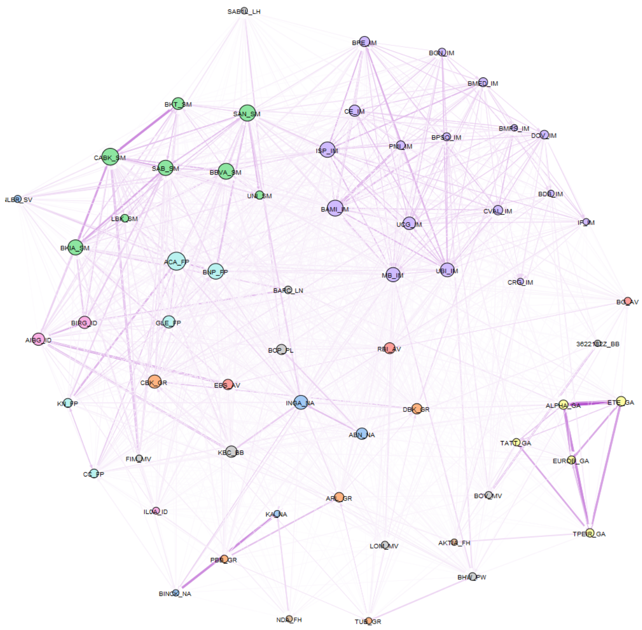
(HSS, CDS, and PD) tend to identify stronger cross-country than domestic links (Figures 2b-2d). In particular, we observe relatively low intensity and low clustering of banks of the same country.



(a) DDLY volatility network

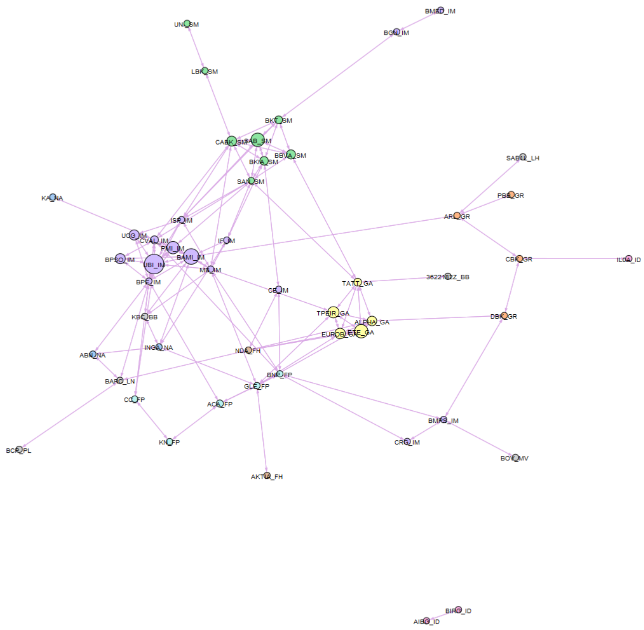


(b) DDLY volatility network defactored with four factors

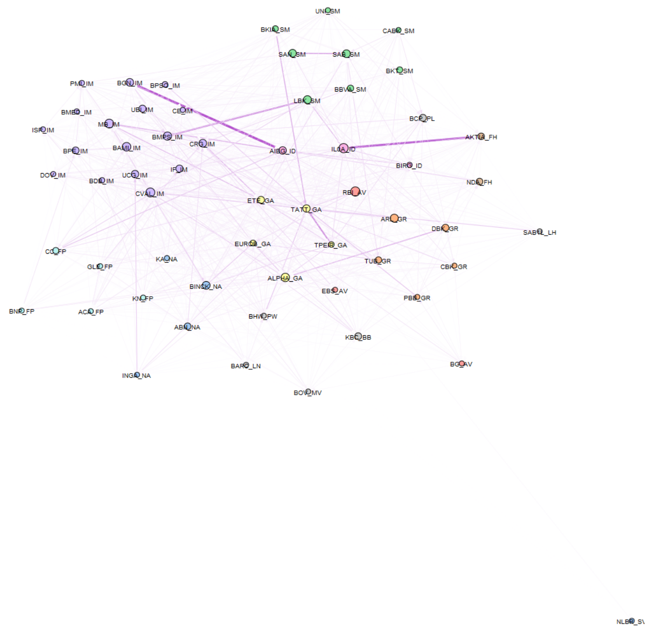


(c) DDLY return network

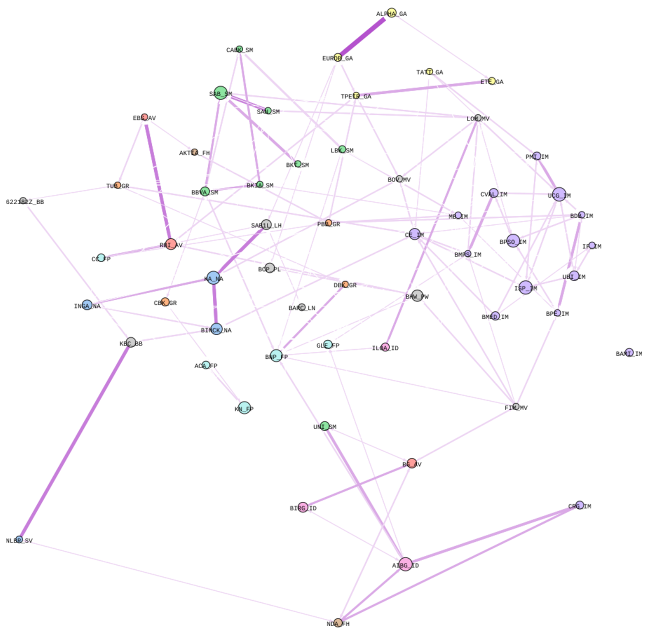
Figure 1: DDLY volatility network, DDLY volatility network defactored with four factors, DDLY return networks, sample 62 banks, 2013q3-2019q4, Node indicates bank, node size - weighted degree, node colour - country, node location - fruchterman reingold + circle pack, edges size and colour - pairwise directional connectedness (to and from)



(a) CS network



(b) HSS network



(c) PD network

(d) CDS network

Figure 2: CS, HSS, PD networks, sample 62 banks, CDS network, sample 18 banks 2013q3-2019q4. Node indicates bank, node size - weighted degree, node colour - country, node location - fruchterman reingold + circle pack, edges size and colour - pairwise directional connectedness (to and from)

4.2 Network Characteristics

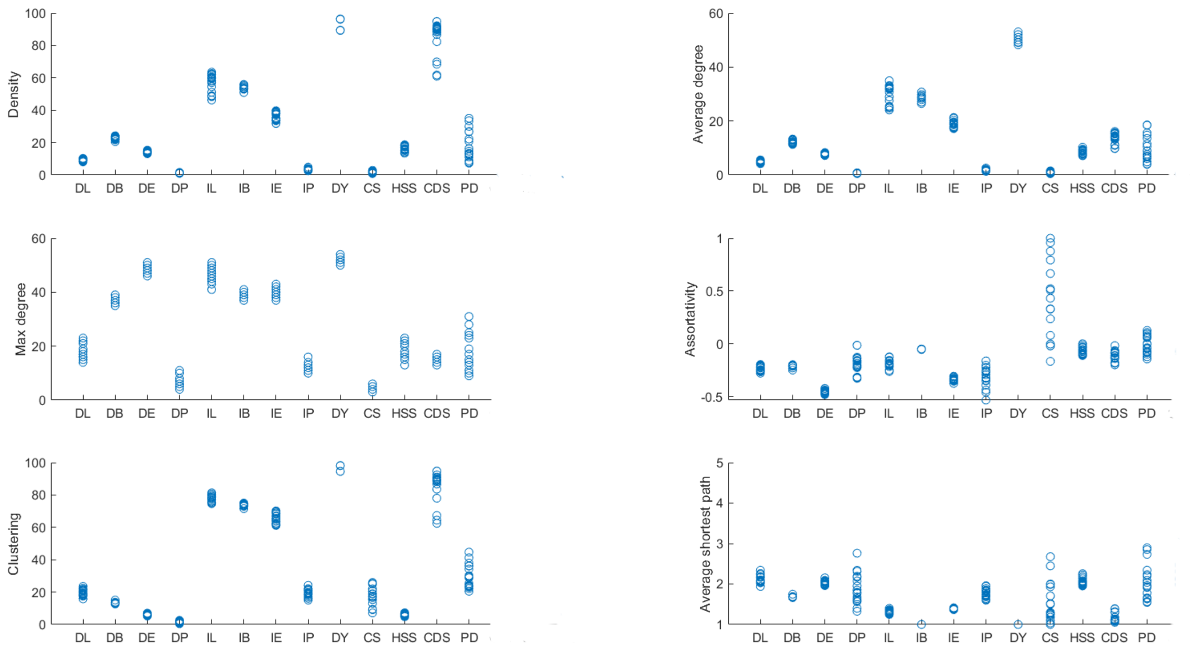
4.2.1 Aggregate network characteristics

In this subsection, we look at aggregate network characteristics of market-based and balance-sheet networks. We choose a list of measures following Anand et al. (2018) who run a horse race of methods to reconstruct financial networks using partial balance sheet data. The results presented in Figure 3 plot values of all measures for all considered networks for all time periods. The results Table 3 show average, median, and standard deviation of measures across years and networks. Below, we discuss network characteristics based on Figure 3, while the reader can check exact numbers in Table 3. The details on network characteristics measures can be found in Appendix A.4

For all types of networks, we document high stability of aggregate network characteristics as the dispersion of circles within each column is very small. Networks based on direct bilateral exposures (DL, DB, DE, DP) have the properties of well-documented financial networks (see, i.e., Anand et al. (2018), namely, low density, highly skewed density distribution (low average degree and high maximum degree), negative assortativity, low clustering and short average shortest path. All these characteristics indicate that direct networks exhibit a small-world property and a hub structure where few banks are extremely well-connected to the rest of the system, and the majority of the banks have very few connections. One of the reasons for such a structure is that creating and maintaining links is costly, particularly in lending (see Craig, Ma (2018) for more discussion). Consistent with this hypothesis, we can see that the degree of direct bond and equity holdings is significantly higher as low cost is associated with buying these securities on the market.

When we look at the networks of indirect exposures, through portfolio commonality (IL, IB, IE, IP), we see that networks are denser and both average and maximum degree are higher. This result is overall expected as it is easier for two banks being connected through exposure to the same portfolio of assets. Clustering is significantly higher while average shortest path is lower for indirect networks as there are more links. Assortativity is much less negative suggesting that less-connected banks are connected not only to more-connected banks but also to each other. IB network is somewhat special, it consists of a fully complete

Figure 3: Network Characteristics



*Notes:*On the x-axis there are different methods. True exposures: DL, DB, DE, DP - are direct loans, bonds, equity, and fund shares, IL, IB, IE, IP - are indirect loans, bonds, equity, and fund shares, market-based networks: DY - DDLY volatility network, CS - partial correlation network based on equity return, HSS - tail risk network, CDS - credit risk network, PD - default probability network. Each circle represents characteristics for each time period, March 2014 - September 2019.

subnetwork and a number of isolated nodes. For this reason, we see, for example, that its average shortest path is strictly equal to one meaning that any bank can be reached from any other bank using only one link. This structure of the IB network can be explained by two things: first, smaller banks may not invest in bonds, and thus they are disconnected from the IB network. Second, a relatively limited number of firms issuing bonds leads to a high probability that any two banks investing in bonds are exposed to the same set of firms. This is in turn reflected in connection of all banks to each via common bond portfolios.

Finally, regards the market-based networks, DDLY and CS networks as expected are the two extremes: DDLY is a complete network with all banks being connected to each other, while CS is an extremely sparse network. Interestingly, despite such low density, CS network is still quite well connected within its giant connected component (a subset of network where isolated nodes are excluded) with average shortest path below 2.5 for most of the periods. A significant difference of CS relatively to the other network is that its assortativity is positive suggesting that well-connected banks are connected among each other. CS characteristics exhibit also the highest volatility across years.

Table 3: Network Characteristics

		DL	DB	DE	DP	IL	IB	IE	IP	DY	CS	HSS	CDS	PD
	mean	54.3	54.3	54.3	54.3	54.3	54.3	54.3	54.3	54.3	54.3	54.3	17.0	54.3
Nodes	median	55	55	55	55	55	55	55	55	55	55	55	17	55
	sd	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4	0.6	1.4
	mean	273	661	416	35	1644	1569	1039	96	2740	48	473	231	528
Links	median	275	648	409	36	1712	1598	1040	94	2704	46	484	238	451
	sd	25	45	24	5	204	108	84	17	132	15	55	35	248
	mean	9.4	22.8	14.4	1.2	56.7	54.2	35.9	3.3	94.8	1.7	16.3	84.6	18.3
Densities	median	9.4	23.0	14.6	1.2	58.8	54.4	35.1	3.2	96.2	1.6	16.3	89.9	15.4
	sd	0.6	0.9	0.6	0.2	5.5	1.5	2.5	0.6	2.9	0.5	1.5	10.8	8.6
	mean	5.0	12.2	7.7	0.6	30.2	28.9	19.1	1.8	50.5	0.9	8.7	13.5	9.7
Mean degree	median	5.0	12.0	7.5	0.7	31.9	29.1	19.0	1.7	50.0	0.8	8.8	14.0	8.3
	sd	0.4	0.6	0.3	0.1	3.3	1.3	1.4	0.3	1.6	0.3	0.9	1.8	4.5
	mean	3.9	11.8	2.3	0.0	34.7	39.1	20.0	0.0	51.8	0.1	8.6	13.5	10.9
Median degree	median	4.0	12.0	2.0	0.0	35.5	39.5	20.5	0.0	51.5	0.0	8.8	14.0	9.3
	sd	0.4	1.2	0.4	0.0	3.5	1.4	2.0	0.0	1.3	0.3	1.0	1.9	5.3
	mean	17.2	36.6	48.6	7.0	46.4	39.1	39.6	12.8	51.8	4.7	17.7	15.6	17.8
Max degree	median	16.0	36.0	49.0	6.5	47.0	39.5	40.0	13.0	51.5	5.0	17.5	16.0	16.0
	sd	2.8	1.3	1.7	1.9	2.8	1.4	1.9	1.5	1.3	0.9	3.0	1.1	6.3
	mean	-0.2	-0.2	-0.5	-0.2	-0.2	-0.1	-0.3	-0.3	NA	0.4	-0.1	-0.1	0.0
Assortativity	median	-0.2	-0.2	-0.5	-0.2	-0.2	-0.1	-0.3	-0.3	NA	0.5	-0.1	-0.1	0.0
	sd	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.1	NA	0.4	0.0	0.0	0.1
	mean	19.8	13.2	6.1	1.5	78.2	73.8	65.6	19.6	97.4	17.3	6.1	85.1	29.4
Clustering	median	19.4	13.0	6.1	1.5	78.5	74.0	65.5	19.7	98.1	17.4	6.2	89.1	27.1
	sd	1.9	0.6	0.6	0.6	2.1	1.0	2.8	2.2	1.5	5.9	0.6	9.8	7.1
	mean	18.0	30.0	21.5	2.0	51.5	50.0	37.5	5.5	52.0	6.0	8.5	12.0	11.0
Core size	median	18.0	30.0	21.5	2.0	52.0	50.0	37.5	5.5	52.0	6.0	8.5	12.0	11.0
	sd	2.5	3.5	1.5	1.0	3.5	2.0	2.5	1.5	3.0	1.0	1.5	2.0	2.5
	mean	25.1	15.1	17.5	59.7	4.7	0.0	1.4	14.7	0.0	55.0	62.8	8.5	53.8
Core error	median	24.0	14.8	17.6	59.3	4.9	0.0	1.3	14.4	0.0	57.9	62.8	6.5	57.1
	sd	2.4	0.8	1.1	8.2	1.0	0.0	0.5	3.6	0.0	12.7	2.6	5.8	11.9
	mean	2.1	1.7	2.0	1.9	1.3	1.0	1.4	1.8	1.0	1.5	2.1	1.2	2.0
Average shortest path	median	2.1	1.7	2.0	1.8	1.3	1.0	1.4	1.8	1.0	1.3	2.0	1.1	1.9
	sd	0.1	0.0	0.0	0.4	0.0	0.0	0.0	0.1	0.0	0.5	0.1	0.1	0.4
	mean	4.9	3.2	4.4	4.2	2.7	1.9	2.7	3.2	1.0	3.2	4.3	2.0	4.0
Diameter	median	5.0	3.0	4.0	4.0	3.0	2.0	3.0	3.0	1.0	3.0	4.0	2.0	4.0
	sd	0.8	0.4	0.5	1.1	0.4	0.3	0.5	0.5	0.0	1.4	0.6	0.0	1.1

Notes: D - direct exposure, I - indirect or common portfolio exposure, L - loans, B - bonds, E - equity, P - equities non securities or fund shares. DDLY stands for Demirer et al. (2018) approach, CS stands for Craig, Saldías (2016) approach, HSS stands for Hautsch et al. (2014), Hautsch et al. (2015) approach, CDS stands for Brownlees et al. (2020) approach, PD stands for Duan et al. (2012) approach.

Interestingly, tail risk (HSS) and default probability (PD) networks have density, average and maximum degrees, and average shortest path very similar to direct balance sheet networks. Both network types have slightly negative and in certain periods even positive assortativity which resemble them more to indirect balance sheet networks. Default probability network exhibits quite strong variability of characteristics over time.

CDS networks are similar to other balance-sheet networks with the difference that these networks are almost complete. This is largely because these networks consist of only 18 banks (vs 62 for other networks). For this reason, clustering parameter is very large while average shortest path is very low.

The characteristics we present are computed for a sub-sample of the banking system as we do not have data either balance sheet or market data for all banks in the system. This may have an impact on aggregate network characteristics, though this is a general issue in the empirical network literature. Crain (2018) provides a summary of some of the philosophical problems in inference about network characteristics in a population from a sample of network nodes.

Table 4: Correlations between edges of Balance Sheet Networks

	IL	IB	IE	IP	DL	DB	DE	DP
IL	1.00							
IB	0.19	1.00						
IE	0.11	0.08	1.00					
IP	0.06	0.10	0.02	1.00				
DL	0.05	0.19	0.08	0.00	1.00			
DB	0.07	0.35	0.00	0.03	0.13	1.00		
DE	0.02	0.04	0.02	-0.01	0.08	0.09	1.00	
DP	0.03	0.12	0.03	0.02	0.56	0.06	0.01	1.00

Notes: D - direct exposure, I - indirect or common portfolio exposure, L - loans, B - bonds, E - equity, P - equities non-securities or fund shares.

4.2.2 Edge Comparison

The graphical diagrams of the market-based networks and aggregate network statistics for all networks presented above are somewhat indicative of structural similarities. However, to understand how well market-based networks match individual edges of balance-sheet networks, a more formal analysis is needed. In this section we start with the descriptive evidence, that we will further investigate using econometric analysis in the next section.

Edge comparison refers to understanding if two compared networks have the same present and the same absent links. To do this analysis, network literature proposes several approaches. We start with a simple correlation of weighted networks and then proceed with computing four main similarity measures, namely Jaccard, Hamming, Accuracy and Cosine measures.

We start by looking at the correlations between the types of balance sheet networks, between the types of market-based networks and between the balance-sheet and market-based networks. Table 4 shows that the balance sheet exposures are often correlated across types. Particularly, direct loan portfolios are highly correlated with direct ownership, and this makes sense as banks can lend more easily to companies they (partially) own. Direct loans are also correlated with direct indirect bond portfolios. And overall all correlations are positive. While this makes little sense in terms of portfolio diversification, it makes a lot of sense in terms of monitoring efforts of counterparties that can be used across balance sheet items. Indeed, as explained in Acharya et al. (2006) and explored most recently in Bednarek et al. (2022), monitoring efforts of companies in one asset class can be used to

Table 5: Correlations between MBN Edges

	CS	CS1	CS2	CS3	DDLY	DDLY1	DDLY2	DDLY3	HSS	CDS	PD
CS	1.00										
CS1	0.97	1.00									
CS2	0.94	0.97	1.00								
CS3	0.74	0.74	0.72	1.00							
DDLY	0.32	0.35	0.36	0.25	1.00						
DDLY1	0.30	0.34	0.35	0.24	0.64	1.00					
DDLY2	0.35	0.38	0.39	0.27	0.71	0.76	1.00				
DDLY3	0.32	0.35	0.36	0.25	1.00	0.64	0.71	1.00			
HSS	0.17	0.18	0.18	0.14	0.14	0.12	0.15	0.14	1.00		
CDS	0.30	0.31	0.34	0.25	0.28	0.26	0.26	0.28	0.14	1.00	
PD	0.19	0.21	0.22	0.20	0.23	0.20	0.22	0.23	0.13	0.31	1.00

Notes: CS stands for Craig, Saldías (2016) approach, CS1 - CS_TH1_F1, CS2 - CS_TH25_F1, CS3 - CS_TH25_F3 are alternative thresholds and factors CS networks. DDLY is the baseline Demirer et al. (2018) approach, DDLY1 is DDLY_return_4factors_scaled the 4 factor-scaled return version of Demirer et al. (2018) in the text, DDLY2 is DDLY_vol_4factors_scaled the 4 factor-scaled volatility version of Demirer et al. (2018) in the text, DDLY3 is DDLY_thr that is original DDLY with a threshold of Demirer et al. (2018), HSS is the Hautsch et al. (2014), Hautsch et al. (2015) approach, CDS is the network from the Brownlees et al. (2020) approach, PD is the Duan et al. (2012) approach.

increase the exposure in other assets.

Table 5 shows edge correlations between various types of market-based networks. We notice several patterns. First, all variations of DDLY and CS are correlated with each other. This is an indication that in regressions, we should use only one specification of the measure. Second, correlations between other measures are relatively high but not too much to pose a collinearity problem in the estimation. Third, we observe that equity and credit-based measures tend to correlate more within the classes. The only exception is HSS which has relatively low correlation with all other networks.

In Table 6, we observe the following patterns of edge correlations between balance-sheet and market-based networks. First, all market-based networks have the largest correlations with the network based on common bond holdings, or IB. These correlations are of the order of 0.23-0.32 except for the HSS network which has a correlation of 0.11 but it also has much smaller correlations overall. Such importance of common bond holdings makes sense as banks have a large share of bonds in their portfolio, and information on bond holdings is publicly available and thus can be assessed by market participants. CDS network stands out as it has equally high correlations with both common bond portfolios and cross-holdings of bonds. This is not surprising as CDS contracts often use non-repayment of bonds as a trigger of a credit event.

Table 6: Simple Edge Correlations

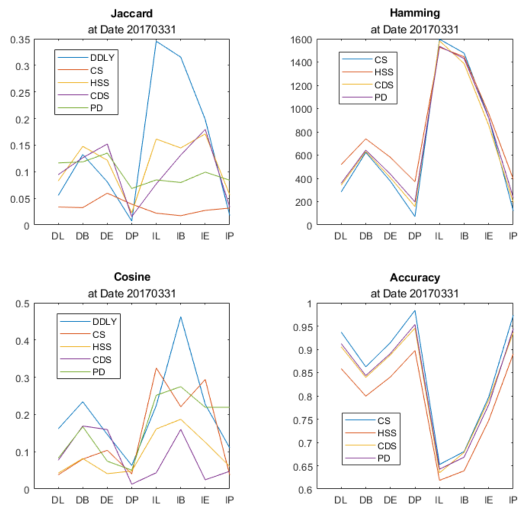
	CS	DDLY	HSS	CDS	PD
IL	0.06	0.14	0.07	0.18	0.13
IB	0.25	0.32	0.11	0.28	0.23
IE	0.09	0.15	0.03	0.03	0.07
IP	0.02	0.01	0.01	0.03	0.05
DL	-0.04	0.04	0.01	0.03	0.02
DB	0.07	0.15	0.09	0.28	0.13
DE	0.02	0.07	-0.01	0.10	0.03
DP	-0.02	0.00	0.01	0.01	0.03

Notes: Balance Sheets: D - direct exposure, I - indirect or common portfolio exposure, L - loans, B - bonds, E - equity, P - equities non-securities or fund shares. Market Base Measures: CS stands for Craig, Saldías (2016) network, DDLY is the baseline Demirer et al. (2018) network, HSS stands for Hautsch et al. (2014), Hautsch et al. (2015) network, CDS is the network from the Brownlees et al. (2020) network, PD is the Duan et al. (2012) network.

Direct bond holdings is the next category of assets largely correlated with all market-based networks. Again holdings of bonds are relatively well-known to the market. One may expect that banks' direct exposures should be more important for the markets since if a bank defaults its counterparty will be affected directly. On the other hand, banks' direct exposures represent a significantly smaller share of banks' total assets. Indirect and direct bond portfolio represent about 12.5% and 0.11% of total assets of a median bank with maximum share reaching 31% and 4.8% respectively. However, we can see how importance of an asset in banks' portfolio is affected by market knowledge. Indirect loans represents a lion share of banks' portfolios: 17.5% for a median bank with the maximum being up to 59%. But this information is largely proprietary, and thus markets may have only imperfect guess of these exposures. This is reflected in common loan portfolio being captured by market-based networks only as good as direct bond holdings despite its overall size.

Common bond and loan portfolios as well as direct bond holdings are the most correlated with networks reflecting credit risk, i.e., HSS, CDS, and PD. While debt assets are also captured by return-based networks, CS and DDLY are the only market-based measures that are strongly associated with common equity portfolio. The fact that return-based networks are correlated with debt assets is expected as market equity prices potentially reflect overall return of banks' business model or portfolio which consists to a large extent of loans and bonds. However, it is interesting to see that return-based networks also capture exposure in indirect equity holdings as they represent a relatively small share of banks' portfolio, 0.26% for a median bank and 6.2% maximum.

Figure 4: Adjacency Matrix Comparisons



*Notes:*Balance Sheets: D - direct exposure, I - indirect or common portfolio exposure, L - loans, B - bonds, E - equity, P - equities non-securities or fund shares. Market Base Measures: DDLY stands for Demirer et al. (2018) approach, CS stands for Craig, Saldías (2016) approach, HSS stands for Hautsch et al. (2014), Hautsch et al. (2015) approach, CDS stands for Brownlees et al. (2020) approach, PD stands for Duan et al. (2012) approach.

The remaining assets, direct trading and non-trading equity holdings as well as common portfolio of non-trading equity, have rather low correlation with market-based networks. This is probably due to low share of these assets on banks' balance sheets and little information available to the markets.

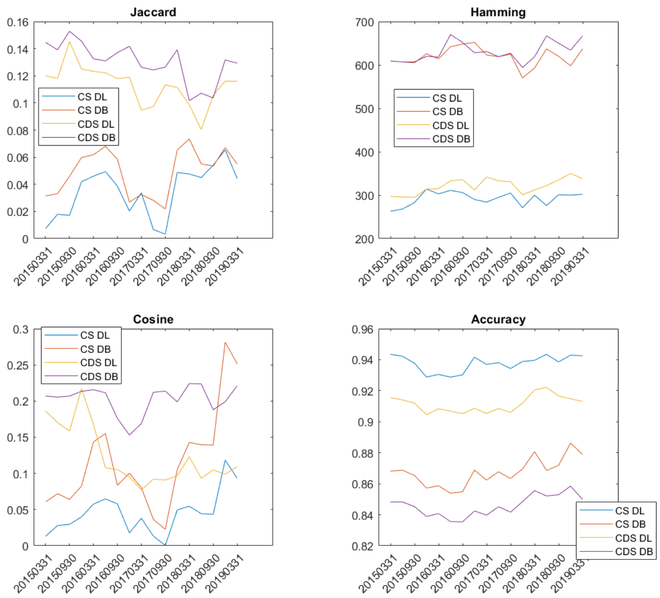
Next, we compare networks using similarity measures, namely Jaccard, Hamming, Accuracy and Cosine. The first three measures account only for presence or absence of links and not link weight. The Jaccard measure of similarity is defined as an intersection of common links divided by the union of the links present in both networks. The Hamming similarity sums up all links that are different in the two compared networks. The Accuracy similarity computes percentage of true-positive and true-negative links in the market-based network relative to the balance-sheet network. Finally, cosine similarity computes an angle between two vectors thus also taking into account weights of edges. It is a close analog to a correlation measure.

Figure 4 presents results of this comparison. Each sub-figure depicts each combination of the market-based versus balance-sheet-based networks. Hamming and Accuracy measures are only for adjacency matrices (0 - 1 networks), thus the complete network of DDLY is irrelevant.

Several patterns emerge from Figure 4. First, the Hamming and Accuracy measures of similarity are essentially the inverse of each other. This is indeed true, as Hamming measures the number of links which differ in two compared networks while Accuracy measures the share of correctly identified links over a total number of links in one of the networks. All market-based networks perform similarly with respect to these two measures: the accuracy is relatively high when comparing market-based networks with direct exposure networks and indirect participation networks. The accuracy turns out to be somewhat lower for direct bonds as well as indirect loan and bond networks. We know that balance-sheet networks are very sparse with direct networks being extremely sparse. Thus, these two charts confirm that Hamming and Accuracy measures perform better for more sparse networks. In this sense, they are not very useful as we are more interested in seeing market-based networks correctly matching non-zero connections.

The Jaccard and Cosine similarity measures give a different set of results. Interestingly, both measures show results somewhat similar to each other and similar to the correlations in Table 6. For both measures, almost all market-based networks reflect stronger exposures to a combination of direct bonds and equity, as well indirect bonds, equity and loans. The cosine measure emphasizes the market-based networks' similarity to the IB, IL and IE networks, with a secondary emphasis on the DB and DE networks. The DDLY network does better than the other in the indirect bonds, direct bonds and direct loans networks. The network based on default probabilities captures pretty well connections in all indirect networks and to a lesser extend in direct bonds networks. According to the Jaccard similarity, the HSS network does well in most of the categories of balance-sheet networks, in contrast to the simple edge correlations where it performed universally more poorly than the other market-based networks. CDS picks up the same categories according to both measures but it performs better according to the Jaccard measure than the Cosine. Interestingly, the CS network captures well the IL and IE networks when considering weighted network in the cosine measure but its performance is very poor for the Jaccard measure. Another observation concerns direct loans network: while market-based networks were poorly correlated with direct loans in Table 6, according to similarity measures they show much better performance, particularly, for DDLY using Cosine measure and almost all networks using Jaccard measure.

Figure 5: Adjacency Matrix Comparisons. Most Volatile over Time



*Notes:*Balance Sheets: D - direct exposure, I - indirect or common portfolio exposure, L - loans, B - bonds, E - equity, P - equities non-securities or fund shares. Market Base Measures: DDLY stands for Demirer et al. (2018) approach, CS stands for Craig, Saldías (2016) approach, HSS stands for Hautsch et al. (2014), Hautsch et al. (2015) approach, CDS stands for Brownlees et al. (2020) approach, PD stands for Duan et al. (2012) approach.

Figure 4 presents comparison for one randomly chosen date. In Figure 5, we plot the evolution of similarity measures for a subset of pairs of market-based and balance sheet networks. For better visibility, we focus on four pairs with the greatest variation over time instead of plotting forty pairs (five market-based networks times eight balance-sheet networks). Hamming and Accuracy measures show significant stability both in the performance ranking and absolute values since these measures are explained by a large amount of zeros in the matrices. Jaccard measure shows higher variation over time but ordering remains stable. These three relatively stable measures capture adjacency matrix, i.e. presence or absence of links. When we look at the Cosine measure that aims to assess weights in two compared networks, we observe significantly higher variation. The two pairs CDS-DL and CS-DB demonstrate particularly extreme volatility over time with performance switching from the best to the worst. For example, CDS network captures relatively well DL network at the beginning of the period and not that well at the end of the period. This variation potentially indicates that CDS and CS networks will not explain well DL and DB networks respectively in the regressions.

In summary, no single measure is ideal for assessing the similarity between two networks.

Various measures highlight different aspects, leading to contrasting results. Some market-based networks perform better according to some measures yet fall short when judged by others. Despite this, we note several consistent patterns across these measures. Most market-based networks effectively mirror indirect bonds, indirect loans, and direct bonds, while showing less agreement with indirect and direct equity, and direct loans. These observations serve as a basis for our forthcoming econometric analysis, the results of which are detailed in the next section.

5 Results of econometrics analysis

In our analysis we are interested in answering the following two questions. How an increase in intensity of balance-sheet exposures by one standard deviation affects the probability of observing a link in a market-based network? And how an increase in intensity of balance-sheet exposures by one standard deviation affects the intensity of an observed link in a market-based network?

The first question refers to an existence of a link, or an extensive margin, and we assess the relationship using a logit model in Section 5.1. The second question refers to a size of a link, or an intensive margin, and the relationship is estimated using a Poisson regression in Section 5.2.

Both questions we can also ask the other way around. For example, how an increase in market perception of a connection between two banks is related to: a probability of observing a link or a size of the observed link in the balance sheet exposures? In all cases, we talk about correlation and not causality, but intuitively we can think about these questions as follows. Let's look at the probability of observing a link in the direction from balance sheet exposure to market perception. First, we expect that when a balance sheet connection via direct asset holdings or exposure to common assets increases or intensifies, markets react by re-evaluating the available information and either create a new link or re-assess the existing one. This is a direct relationship in the direction of causality, though we do not claim causality. Second, to know if markets can predict a link in a balance-sheet network we run the opposite regression. We expect that higher intensity of a connection in market-based networks is related to an observation of a link in a balance-sheet network. In a linear model with one variable the coefficients would be the same in both regressions, while in logit and Poisson regressions with multiple variables this is not the case. Thus, in subsections on extensive and intensive margins we run regressions in both directions.

Our baseline specification is a dyadic panel data model, which is widely used in trade and social network analysis. We estimate each BSN as a function of the MBN as follows

$$BSN_{ijt} = \alpha_i + \alpha_j + \alpha_t + \alpha_{it} + \alpha_{jt} + \sum_{m=1}^{5} \beta_m MBN_{ijt}^m + controls_{ijt} + \epsilon_{ijt} \tag{7}$$

and a corresponding equation of each MBN as a function of the BSN edges is estimated as follows:

$$MBN_{ijt} = \alpha_i + \alpha_j + \alpha_t + \alpha_{it} + \alpha_{jt} + \sum_{k=1}^{4} \psi_k D_{ijt}^k + \sum_{k=1}^{4} \eta_k I_{ijt}^k + controls_{ijt} + \epsilon_{ijt} \quad (8)$$

where a unit of observation MBN_{ijt} is a network edge from bank i to bank j at time t calculated using a specific market-based approach, and BSN_{ijt} is an edge from bank i to bank j at time t in a specific balance-sheet network. An edge is either directed or undirected depending on the network. Depending on the regression model used - logit or Poisson - an edge is either a binary variable, or continuous variable reflecting link weight. The parameters α_i, α_j, α_t are source bank, target bank, time fixed effects respectively. α_{it} is a source bank time fixed effect that captures the change in source bank i's overall exposure in quarter t. Finally, α_{jt} is a similar fixed effect for target bank j.[23]

In the first equation, our variable of interest is β_m where m varies from 1 to 5 reflecting the five market-based networks. In the second equation we separate the BSN_{ijt} components into direct D_{ijt}^k and indirect I_{ijt}^k true exposures. k varies from 1 to 4, and covers four types of exposure: direct and indirect loans, bonds, equity, and equity non-security or fund shares. The coefficients of our interest are ψ_k and η_k. We expect a positive sign for all coefficients of interest as higher intensity of a link in one network should be reflected by a higher intensity of the same link in the other network.

We cluster standard errors at the dyad level. Consider two pairs of edges, for example, BSN_{ij} and BSN_{ik} at time t. Both edges share a common bank i and so co-vary. Failing to account for this type of dependence will result in too small standard errors. Thus, we cluster standard errors along both dimensions of the pair to take into account the correlation pattern of the dyadic regression.

We consider the following control variables. Log total assets (TA) captures the size effect. Common equity tier ratio (CET1 ratio) is a leverage ratio to take into account

[23] We consider quarterly data, and the sample period for our regressions is 2013q3 - 2019q4, so we are limited in the number of observations in terms of the time dimension. Taking into account this limitation we are not able to include bank pair fixed effects, α_{ij} and α_{ji}, which control for the effect of time-invariant bank-pair unobserved characteristics on the link between bank i and bank j. On the other hand, some studies have shown that banks tend to have long-lasting relationships, and we expect certain stability in bank pair relationships.

the health of the bank. Risk-weighted assets intensity (RWAI) captures riskiness, while Return on average equity (ROAE) reflects banks profitability. We include in the regression the product of source and target controls, in other words pair-wise controls. All control variables are lagged by one period.

In the next section we discuss results of the analysis at the extensive margin, explaining the presence of a link in the network with the logit model. And then, we explore the results of the intensive margin, explaining the intensity of the link in the dependent variable using a Poisson model.

5.1 Extensive margin results (Logit regression)

In this section, we discuss results of the regression of an extensive margin analysis using a logit model. We start with Equation 7 which estimates how well market-based networks can guess balance-sheet networks.

The dependent variable BSN^{01}_{ijt} is a dummy, it is equal to one if there is a link, and zero otherwise. The right-hand side variables are market-based networks MBN_{ijt}: DDLY, HSS, CS and PD.

Three comments are worth making. First, we exclude the CDS variable as CDS networks contain only 18 banks. Including CDS edges in the regression severely limits the sample size and thus the predictive power of the explanatory variables.

Second, DDLY and HSS networks are directed, while CS and PD networks are undirected. We choose to use an undirected version of the BSN to be conservative in our estimations. The difference between undirected and directed networks is an increased number of links in the network as for all existent links from node A to B we create links from B to A. As CS and PD variables already include links both from A to B and from B to A, the estimation of these two approaches is not affected in the case of the undirected dependent variable. On the other hand, DDLY and HSS are punished by the use of the undirected dependent variable as now they have to "predict" more links than otherwise. This contrasts to the case of the directed dependent variable, when the estimation of the directed DDLY and HSS is not affected, but undirected CS and PD benefit as they can guess either of the existing link, from A to B or from B to A.

Third, the IB network consists of a complete subnetwork and some standalone nodes. Thus, we cannot estimate the IB regression with fixed effects, as fixed effects will explain all the variation. To address this issue, we take a cumulative sum of ranked weights in the network, retain only weights up to the 95th percentile, and replace all values below the threshold of the 95th percentile with zeros.

Next, we run a logit using Equation 8 where the dependent variable, MBN^{01}_{ijt}, is a dummy equal to 1 for an existing link with positive weight in a market-based network, and 0 otherwise (no link or a link with negative weight). As in Equation 7, the MBN are used in their original version, namely, directed (DDLY and HSS) and undirected (CS and PD), while the right-hand side BSN are adjusted to match the direction of the MBN on the left-hand side. Regarding the DDLY, we cannot estimate the regression as it is a complete network. Thus, we use the same approach as for the IB network and replace all values below the threshold of 95th percentile with zeros.

The results are presented in Tables 7 and 8 below. One can observe that essentially all the coefficients in Table 7 are positive and significant. The positive sign reflects our expectations that an increase in a link weight in the MBN is associated with higher probability of having a link in the BSN. The significance of the coefficients suggests that the market-based networks do a good job in "predicting" existence of a link in various balance-sheet networks.

Table 8 reports the coefficients for the opposite exercise and shows how a change in the intensity of a link in a balance-sheet network affects the probability of observing a link in a market-based network. Most of significant coefficients have the expected positive sign. The two balance-sheet variables captured by almost all market-based networks are common portfolio exposures of loans and bonds. Equity price-based networks such as CS and DDLY reflect also direct and indirect equity exposures, while credit-risk-based networks (HSS and PD) are the only ones associated with direct bonds exposures. Not surprisingly, the coefficients in front of the direct loans either have low significance or have an unexpected negative sign as the market fails to reflect highly proprietary data on interbank lending.

One can wonder why all the coefficients in Table 7 are significant while the coefficients in the opposite regression in Table 8 are not. When we aim to explain the market-based measures by the balance-sheet variables, we observe a differential effect, namely, the sensitivity of the market-based measures vary depending on the balance-sheet networks. However, regressing a balance-sheet variable on the market-based networks shows all the market-based networks equally important in explaining the balance-sheet variable. The main reason behind this effect is quite important correlation between various balance sheet networks. For example, as shown in Table 4, direct loans are correlated with common bond portfolios at 19%, with direct bonds at 13%, direct and indirect equity at 8%. When we regress only direct loans on market-based networks, the latter may explain not links specific to the direct loans but rather similarity of direct loan network to other balance-sheet networks. The fact that banks are exposed to the same counterparties via different assets is also demonstrated by Elliott et al. (2021). The authors show, for example, that German commercial banks that lend to each other bilaterally also tend to have similar portfolios of loans to non-financial firms. Furthermore, the authors argue that a network structure with banks being interconnected in different networks arises in the equilibrium as banks benefit together from the same assets in good times while protected by limited liabilities and possibility of a bail-out if failing together in bad times. Jackson, Pernoud (2019) also find an incentive of banks to correlate their assets, but for various reason, that is 'risk matching'. Their more genral result can be extended to allow to relax limited liability.

In a linear regression model, the reported coefficient is a marginal effect, however, this is not the case in the logit model. For this reason, we compute marginal effects (ME) of an independent variable on the probability of observing a link.[24] This will allow us to compare sensitivities of various variables, to provide an economic interpretation to the results, and it may also shed light on the differential effects of market-based networks on balance-sheet variables in Table 7 with all significant coefficients.

The ME allows the following interpretation: one standard deviation increase in an inde-

[24] We compute a marginal effect in a logit regression as follows. The derivative with respect to any independent variable can be solved using the chain and quotient rules. Since the logit regressions are non-linear, the effect of an increase in an independent variable is modified because the marginal effect is now a function of the values of the x's themselves $X\beta$. We can thus find the marginal effect of an independent variable on the probability of having a link in the network defined by the dependent variable as follows: $\frac{\partial P(link)}{\partial x_k} = \frac{\exp(-X\beta)}{(1+\exp(-X\beta))^2}\beta_k$. To interpret the results, we compute the marginal effect at density (ME further on), as it allows accounting for difference in sparsity of the dependent variable networks and making ME comparable across specifications. Thus, we choose λ so that the $P(link)$ is set at the density of the network on the left-hand side in the regression model, that is $\lambda = -\ln\left(\frac{1}{d_n} - 1\right)$. Then the marginal effect at density is just $\beta_i(1-d_n)d_n$, where, d_n is the density of the network defined by the dependent variable.

pendent variable is associated with ME percent increase in the probability of having a link in the network defined by the dependent variable. Tables 9 and 10 demonstrate the marginal effects for the results of logit specifications presented in Tables 7 and 8 respectively. ME are comparable across all independent variables, as all the variables are z-score normalized. Unlike in Table 7 with all coefficients being significant and non-comparable, Table 9 allows us to differentiate between the effects of various regressors and regressions. One can observe that indeed marginal effects vary from about 1% to 16%, and certain market-based networks perform better for some balance-sheet networks but not for others. Each ME can be read as follows: for example, when the intensity of links in the PD network increases by one standard deviation the probability of the link in DL network increases by around 4.7%. Thanks to the normalization of the variables and computation of marginal effects at density, we can compare numbers both across rows and columns. When we look at ME within the same column, we compare which market-based network has the largest effect on a specific balance-sheet network. In this case, we can observe that PD performs best in guessing links in the common bond holdings network (IB) with the estimated marginal effect of 18.15% and IP at around 1%, CS - in the IL, DB, and IE networks with the ME being equal to 10.64%, 12.13% and 20.34% respectively, while DDLY outperforms other market measures in the DE and DL networks with the ME at 11.89% and 5.65% respectively, and DP at around 2%. As for HSS, it has relatively small ME in all regressions with the largest ME in the IB regression. On the other hand, when comparing the performance of the same market measures in different regressions, we notice that all market measures have the second or third-largest marginal effects for the DB network. HSS and PD have the largest ME for the IB network, while DDLY and CS - in the DE and IE regressions respectively.

In Table 10 marginal effects calculated for Table 8 vary from 0.2% to 29%. As we have seen in Table 8, all The largest marginal effects across all specifications (except for PD) has the IB network. In PD regression the largest marginal effect has an IL network. Consistently with our previous logic this current table has similar message. If we do not take into account the marginal effect of IB, the second largest marginal effect for CS network is DE 2.38%, while for DDLY is IE 5.52%. While second largest effect for CDS is DB and equal to 18.98% and for HSS it is IL equal to 1.45%. These results suggest that market-based networks

Table 7: Logit model. The probability of an existence of a link in the balance sheet networks given a change in intensity of a link in the market-based network

Dependent Variable:	DL	IL	DB	IB	DE	IE	DP	IP
Variables								
DDLY	0.4343***	0.1673**	0.5017***	0.2686***	0.6741***	0.5101***	0.6445***	0.2457
	(0.0977)	(0.0756)	(0.1108)	(0.0772)	(0.0908)	(0.1326)	(0.1110)	(0.1526)
CS	0.1151**	0.4312***	0.5575***	0.3404***	0.2116***	0.8845***	-0.0574	0.0893**
	(0.0551)	(0.0649)	(0.1132)	(0.0854)	(0.0577)	(0.1270)	(0.0651)	(0.0361)
HSS	0.0383	0.0989*	0.1491***	0.2157**	0.1649***	0.1323**	0.1972***	0.2111**
	(0.0290)	(0.0507)	(0.0416)	(0.0840)	(0.0372)	(0.0596)	(0.0481)	(0.1076)
PD	0.3381***	0.3146***	0.3772***	0.8240***	0.3348***	0.2362***	0.3337***	0.2941***
	(0.0432)	(0.0643)	(0.0631)	(0.2805)	(0.0633)	(0.0781)	(0.0853)	(0.0979)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed-effects								
source	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
target	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
source-date	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
target-date	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
date	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics								
Observations	18,038	18,316	17,085	13,058	21,108	15,302	4,929	3,218
Squared Correlation	0.45726	0.63960	0.60983	0.60205	0.64848	0.77223	0.26825	0.53750
Pseudo R^2	0.40133	0.60516	0.54333	0.55962	0.59059	0.73711	0.27712	0.47408
BIC	25,153.9	21,793.8	23,486.6	17,719.2	25,297.4	17,087.3	9,206.4	6,934.8

Notes: Clustered (source & target) standard-errors in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1. In the names of dependent variables: D - direct exposure, I - indirect or common portfolio exposure, L - loans, B - bonds, E - equity, P - equities non-securities or fund shares. The IB network consists of a complete subnetwork and some isolated nodes, and we cannot estimate the IB regression with fixed effects, as fixed effects will explain all the variation. For this reason we apply a 95th percentile threshold to the IB network, namely, we make the 5% smallest links equal to zero. Independent variables: DDLY - volatility-based network, CS - return-based partial correlation network, HSS - tail-risk based network, PD - forward-looking default probability-based network. Pair-wise controls: TA - total assets, CET1 - common equity tier ratio, RWAI - risk-weighted assets intensity, ROAE - return on average equity.

constructed using equity prices (returns and volatility) both reflect and predict better the direct equity and indirect equity balance-sheet networks. While the opposite holds for the market-based networks that capture credit risk, tail risk, and default probabilities: that is they do a better job in reflecting and predicting direct and indirect bond holdings.

Table 8: Logit model. The probability of an existence of a link in the market-based networks given a change in intensity of a link in the balance sheet network

Dependent Variable:	DDLY_thr	CS	HSS	PD	CDS
Variables					
DL	0.0113	-0.9803***	-0.0287	0.0590*	-0.0720*
	(0.0141)	(0.3762)	(0.0209)	(0.0333)	(0.0411)
IL	0.1233***	0.9271***	0.1320***	0.3475***	-0.0049
	(0.0401)	(0.2864)	(0.0249)	(0.0805)	(0.1285)
DB	-0.0150	-0.2175	0.0425	0.0900**	0.7609***
	(0.0326)	(0.4280)	(0.0309)	(0.0391)	(0.2246)
IB	0.3159***	7.174***	0.3193***	0.0835	1.165***
	(0.0505)	(1.690)	(0.0622)	(0.1473)	(0.2523)
DE	0.0623***	1.607***	0.0180	0.0123	-0.0196
	(0.0209)	(0.4767)	(0.0188)	(0.0173)	(0.0240)
IE	0.3481***	0.2770**	0.0314	0.2198***	-0.0233
	(0.1105)	(0.1313)	(0.0265)	(0.0652)	(0.0544)
DP	0.0682	0.4876***	0.0134	0.0156	0.0645***
	(0.0651)	(0.1041)	(0.0092)	(0.0118)	(0.0097)
IP	0.3116	-0.0215	0.0208**	-0.0317	0.0148
	(0.1978)	(0.1062)	(0.0082)	(0.0202)	(0.0214)
Controls	Yes	Yes	Yes	Yes	Yes
Fixed-effects					
source	Yes	Yes	Yes	Yes	Yes
target	Yes	Yes	Yes	Yes	Yes
source-date	Yes	Yes	Yes	Yes	Yes
target-date	Yes	Yes	Yes	Yes	Yes
date	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	26,268	2,586	26,136	22,378	3,076
Squared Correlation	0.55252	0.77184	0.12396	0.13968	0.20476
Pseudo R^2	0.50049	0.72212	0.12979	0.14002	0.16268
BIC	32,258.0	5,419.9	36,168.0	33,746.1	7,942.6

Notes: Clustered (source & target) standard-errors in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1. In the names of independent variables: D - direct exposure, I - indirect or common portfolio exposure, L - loans, B - bonds, E - equity, P - equities non-securities or fund shares. Dependent variables: CS - return-based partial correlation network, HSS - tail-risk based network, DDLY - original DDLY network with a threshold, PD - forward-looking default probability-based network, CDS - credit risk based network. The DDLY network is complete, and we cannot estimate the regression with all 1 in the dependent variable. For this reason we apply a 95th percentile threshold to the DDLY network, namely, we make the 5% smallest links equal to zero. Pair-wise controls: TA - total assets, CET1 - common equity tier ratio, RWAI - risk-weighted assets intensity, ROAE - return on average equity.

Table 9: Marginal effects for Table 7

	DL	IL	DB	IB	DE	IE	DP	IP
DDLY	5.65%***	4.13%**	10.92%***	5.92%***	11.89%***	11.73%***	1.49%***	0.80%
CS	1.50%**	10.64%***	12.13%***	7.50%***	3.73%***	20.34%***	-0.13%	0.29%**
HSS	0.50%	2.44%*	3.25%***	4.75%**	2.91%***	3.04%**	0.46%***	0.69%**
PD	4.40%***	7.76%***	8.21%***	18.15%***	5.91%***	5.43%***	0.77%***	0.96%***

Notes: Marginal effects at density are calculated using the following formula: $\beta_i(1 - d_n)d_n$, where, β_i is the coefficient in front of the independent variable, d_n is a density of the dependent variable network specified at the top of each column. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1. In the names of dependent variables: D - direct exposure, I - indirect or common portfolio exposure, L - loans, B - bonds, E - equity, P - equities non-securities or fund shares. Independent variables: DDLY - volatility-based network, CS - return-based partial correlation network, HSS - tail-risk based network, PD - forward-looking default probability-based network.

Table 10: Marginal Effects for Table 8

	DDLY	CS	HSS	PD	CDS
DL	0.24%	-1.45%***	-0.31%	0.68%*	-1.80%*
IL	2.66%***	1.38%***	1.45%***	4.01%***	-0.12%
DB	-0.32%	-0.32%	0.47%	1.04%**	18.98%***
IB	6.83%***	10.64%***	3.50%***	0.96%	29.05%***
DE	1.35%***	2.38%***	0.20%	0.14%	-0.49%
IE	7.52%***	0.41%**	0.34%	2.54%***	-0.58%
DP	1.47%	0.72%***	0.15%	0.18%	1.61%***
IP	6.73%	-0.03%	0.23%**	-0.37%	0.37%

Notes: Marginal effects at density are calculated using the following formula: $\beta_i(1 - d_n)d_n$, where, β_i is the coefficient in front of the independent variable, d_n is the density of dependent variable network specified at the top of each column. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1. In the names of independent variables: D - direct exposure, I - indirect or common portfolio exposure, L - loans, B - bonds, E - equity, P - equities non-securities or fund shares. Dependent variables: CS - return-based partial correlation network, HSS - tail-risk based network, DDLY - original DDLY network with threshold, PD - forward-looking default probability-based network, CDS - credit risk-based network.

5.2 Intensive margin results (Poisson regression)

In this section, we estimate equations 7 and 8 using the Poisson regression model. This time, our dependent variables are weights of edges in either balance-sheet or market-based networks. As most of the considered networks except DDLY and CDS are sparse, our data contains many zeros reflecting absence of connections, thus we choose Poisson specification to address this issue and we estimate our model using a Poisson likelihood following Silva, Tenreyro (2006), Silva, Tenreyro (2011).[25] [26]

Tables 11 and 12 report the coefficients β_m, ψ_k and η_k from equations 7 and 8.

We start with the regression of the balance-sheet links on the main market-based measures. Due to high correlation among variations of the CS and DDLY, we do not include

[25]We are grateful to Yoto Yotov and Laurent Berge for suggesting this. We use R package "fixest" for estimation purposes. For the details on theory used to obtain the fixed-effects estimation see Bergé, others (2018).

[26]The results, using other estimation methods, such as Negative Binomial (to take into account over-dispersion) and OLS are available upon request.

them in the regression. As in the case of the logit regression, we omit CDS network due to the low number of nodes in the network.[27]

Table 11 shows results for the prediction of the size of balance-sheet links by market-based measures. Unlike in the logit regression (Table 7), we see variation in significance of predictions. Credit networks (IL, DB and IB) are significantly correlated with all market-based links. DDLY and, to a lesser extent CS, additionally guess well equity networks (DE, IE and DP, IP). HSS and PD are the only market measures that also capture direct loan exposures.

Now we turn to the results of the regression of market-based measures on balance-sheet exposures in Table 12. This table contains two measures of DDLY: standard DDLY and DDLY threshold. The coefficients of both measures are almost identical in size and significance, indicating robustness of our results to such variation. Again, we use DDLY threshold to be able to compute marginal effects, as the original DDLY measure has density equal to 1. Overall, main messages of the estimation in Table 12 are qualitatively similar both to the results in the logit regression (Table 8) and the results of the other Poisson regression (Table 11). Indeed, loan portfolios (IL) are captured by all market-based networks. Direct bonds are correlated with CS and PD, while indirect bonds with DDLY and HSS. As previously, DDLY and CS additionally correlate with equity exposures. However, DL is no more captured by any market measure. This is consistent with our previous discussion that direct loans are mostly proprietary information and cannot be well-identified by banks. As for CDS, it mostly confirms the results in the logit regression, capturing well DB and IB.

Similar to the logit model, Poisson is a non-linear model, and reported coefficients have a different interpretation than a linear regression model. Thus, we compute marginal effects for the estimators of the Poisson regression to compare coefficients among each other but also to provide some economic interpretation to the estimators.[28]

In terms of the marginal effects, in Table 13, we see that DDLY has the largest effects

[27]Regressions on the reduced data set are reported in Table 22 in Appendix A.6

[28]It can be shown that the marginal effect on the probability of having a link between the dyadic pair is $\frac{\partial P(link)}{\partial x_k} = \exp(-\lambda)\gamma_k$ in case of Poisson regression. We assume for simplicity that $\lambda = X\gamma$. We choose λ so that the $P(link)$ is set at the density of the dependent network. This would set $\bar{\lambda} = -\ln(1 - d_n)$, where d_n is the density of the network on the left-hand side. This would mean that the marginal effect of x_k evaluated at $\bar{\lambda}$ is $(1 - d_n)\gamma_k$.

on equity exposures at 9.78% and 6.77%, and relatively large effects on all other significant estimators. For CS, HSS and PD, the top effects are on DB and IL, and interestingly, PD has also a strong effect on IE.

Regards marginal effects for the regression of market-based networks on balance-sheet exposures (Table 14), IB dominates all other balance-sheet variables for all market measures, except PD. For PD, it is the IL which has the strongest effect of 8.89%. CDS stands out with particularly large effects of 29% and 8.9% for IB and DB respectively.

When interpreting marginal effects at density of the network in Poisson model, one should keep in mind, that the effect is stronger for a sparser network, because estimated coefficients are multiplied by $(1 - d_n)$. Indeed, if a network is sparse, even a small change in an independent variable could have a large effect on the probability of a link existing. This results in larger coefficients and marginal effects in the regression with the sparse network. This is not necessarily true in logit model as coefficients are multiplied by $d_n(1 - d_n)$.

To conclude this section, we draw the following messages that are consistent for all four types of regressions: logit and Poisson, direct and inverse. First, all market measures are more or less associated with common holdings of loans and bonds. Furthermore, equity market measures are more correlated with equity holdings, direct or indirect; while credit-risk measures capture better direct credit risk exposures via loans and bonds. Second, overall, results for the direct loans are not very robust and reflect the difficulty of markets to price such information. Third, regarding the economic significance of our results, the original measure of DDLY is mostly closely related to direct equity holdings, consistently across both logit and Poisson "prediction" regression[29]. According to the the marginal effect for the logit regression, CS, HSS, and PD measures are good at capturing the existence of a link in the indirect equity, indirect bonds, and direct bonds respectively. However, they do not overall do a good job in predicting the magnitude or intensity of a link in balance sheet networks. As for the economic relevance of our results in "reflection" regression, regardless of whether we use logit or Poisson specifications, CDS is the most associated with direct and indirect bond holdings, while PD best reflects indirect bonds. As for DDLY, CS and HSS,

[29]Although we remind the reader that original measure of DDLY is a volatility-based network, which might partially reflect common factors and co-movement. We further address this issue in the robustness section 6.1.

they all reflect well both the existence and the magnitude of a link in common portfolios of bonds and loans, although with relatively smaller marginal effects.

Table 11: Extensive margin: Market Networks as Predictors of Balance Sheet exposure

Dependent Variable:	DL	IL	DB	IB	DE	IE	DP	IP
Variables								
DDLY	0.0143	0.0935***	0.0802**	0.0271	0.1268***	0.1056***	0.0372**	0.0432***
	(0.0116)	(0.0165)	(0.0386)	(0.0309)	(0.0407)	(0.0237)	(0.0171)	(0.0150)
CS	-0.0123	0.0389***	0.0341***	0.0348***	0.0175*	0.0011	-0.0066	0.0132*
	(0.0166)	(0.0111)	(0.0057)	(0.0095)	(0.0097)	(0.0145)	(0.0166)	(0.0077)
HSS	0.0084*	0.0256***	0.0323***	0.0112***	0.0019	0.0058	0.0111	0.0066
	(0.0044)	(0.0058)	(0.0082)	(0.0025)	(0.0065)	(0.0066)	(0.0087)	(0.0075)
PD	0.0156**	0.0748***	0.0314***	0.0352***	0.0002	0.0642***	0.0181***	0.0135
	(0.0062)	(0.0143)	(0.0106)	(0.0063)	(0.0060)	(0.0118)	(0.0061)	(0.0149)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed-effects								
source	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
target	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
date	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
source-date	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
target-date	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics								
Observations	28,233	28,233	28,233	28,233	28,233	28,233	28,233	28,233
Squared Correlation	0.16890	0.40851	0.23506	0.79889	0.15585	0.54192	0.15485	0.18489
Pseudo R^2	0.04364	0.08224	0.08414	0.16047	0.05222	0.12429	0.04221	0.04823
BIC	77,336.4	74,931.5	79,504.8	72,833.2	79,399.6	76,657.2	77,887.5	77,476.4

Notes: Two-way (source & target) standard-errors in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1. Fixed-effects: source,target,date,source-date,target-date for all regressions. In the names of dependent variables: D - direct exposure, I - indirect or common portfolio exposure, L - loans, B - bonds, E - equity, P - equities non-securities or fund shares. Independent variables: DDLY - original DDLY, volatility-based network, CS -return-based partial correlation network, HSS - tail-risk based network, PD - forward-looking default probability-based network. Pair-wise controls: TA - total assets, CET1 - common equity tier ratio, RWAI - risk-weighted assets intensity, ROAE - return on average equity.

Table 12: Poisson: Variables Comparison + controls

Dependent Variable:	DDLY	DDLY_thr	CS	HSS	PD	CDS
Variables						
DL	0.0009	0.0009	-0.0037	-0.0040	0.0064	0.0011
	(0.0035)	(0.0033)	(0.0047)	(0.0031)	(0.0086)	(0.0192)
IL	0.0324***	0.0316***	0.0402***	0.0325***	0.1026***	0.0563
	(0.0092)	(0.0092)	(0.0137)	(0.0050)	(0.0210)	(0.0358)
DB	0.0059	0.0055	0.0279***	0.0107	0.0266**	0.1874***
	(0.0073)	(0.0074)	(0.0101)	(0.0074)	(0.0116)	(0.0216)
IB	0.0932***	0.0905***	0.0364	0.0665***	0.0399	0.6124***
	(0.0143)	(0.0139)	(0.0396)	(0.0121)	(0.0410)	(0.1170)
DE	0.0090**	0.0087**	0.0186**	0.0045	0.0041	0.0110
	(0.0038)	(0.0037)	(0.0091)	(0.0037)	(0.0028)	(0.0092)
IE	0.0091	0.0087	0.0276*	0.0084	0.0433***	-0.0226***
	(0.0063)	(0.0061)	(0.0156)	(0.0058)	(0.0166)	(0.0069)
DP	-0.0015*	-0.0014*	0.0068	0.0042**	0.0008	0.0079**
	(0.0009)	(0.0008)	(0.0056)	(0.0019)	(0.0042)	(0.0040)
IP	-0.0002	-0.0001	0.0021	0.0053***	-0.0021	0.0130**
	(0.0005)	(0.0004)	(0.0026)	(0.0010)	(0.0063)	(0.0064)
Controls:	Yes	Yes	Yes	Yes	Yes	Yes
Fixed-effects						
source	Yes	Yes	Yes	Yes	Yes	Yes
target	Yes	Yes	Yes	Yes	Yes	Yes
date	Yes	Yes	Yes	Yes	Yes	Yes
source-date	Yes	Yes	Yes	Yes	Yes	Yes
target-date	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Observations	28,271	28,271	28,225	27,511	22,952	3,118
Squared Correlation	0.67118	0.67273	0.33737	0.19484	0.13110	0.52788
Pseudo R^2	0.13949	0.14027	0.03778	0.02423	0.02636	0.09915
BIC	87,875.3	87,958.6	87,446.1	87,443.8	76,501.9	12,609.2

Notes: Clustered (source & target) standard-errors in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1. In the names of independent variables: D - direct exposure, I - indirect or common portfolio exposure, L - loans, B - bonds, E - equity, P - equities non-securities or fund shares. Dependent variables: DDLY - original volatility-based network, DDLY_thr - original DDLY network with a threshold, CS - return-based partial correlation network, HSS - tail-risk based network, PD - forward-looking default probability-based network, CDS - credit risk based network. The DDLY network is complete, and we cannot estimate the regression with all 1 in the dependent variable. For this reason we apply a 95th percentile threshold to the DDLY network, namely, we make the 5% smallest links equal to zero. Pair-wise controls: TA - total assets, CET1 - common equity tier ratio, RWAI - risk-weighted assets intensity, ROAE - return on average equity.

Table 13: Marginal effects for Table 11

	DL	IL	DB	IB	DE	IE	DP	IP
DDLY	1.21%	4.14%***	5.45%**	1.28%	9.78%***	6.77%***	3.63%**	4.18%***
CS	-1.04%	1.72%***	2.32%***	1.65%***	1.35%*	0.07%	-0.65%	1.28%*
HSS	0.71%*	1.13%***	2.19%***	0.53%***	0.15%	0.37%	1.09%	0.63%
PD	1.32%**	3.31%***	2.13%***	1.67%***	0.01%	4.12%***	1.77%***	1.31%

Notes: Marginal effects at density are calculated using the following formula: $\beta_i(1 - d_n)$, where, β_i is the coefficient in front of the independent variable, d_n is a density of the dependent variable network specified at the top of each column. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1. In the names of dependent variables: D - direct exposure, I - indirect or common portfolio exposure, L - loans, B - bonds, E - equity, P - equities non-securities or fund shares. Independent variables: DDLY - volatility-based network, CS - return-based partial correlation network, HSS - tail-risk based network, PD - forward-looking default probability-based network.

Table 14: Marginal Effects for Table 12

	DDLY_thr	CS	HSS	PD	CDS
DL	0.03%	-0.37%	-0.35%	0.55%	0.05%
IL	1.00%***	3.96%***	2.85%***	8.89%***	2.68%
DB	0.17%	2.75%***	0.94%	2.31%**	8.91%***
IB	2.86%***	3.59%	5.81%***	3.46%	29.13%***
DE	0.28%**	1.83%**	0.39%	0.36%	0.53%
IE	0.27%	2.71%*	0.73%	3.76%***	-1.07%***
DP	-0.05%*	0.67%	0.37%**	0.07%	0.38%**
IP	0.00%	0.21%	0.46%***	-0.18%	0.62%**

Notes: Marginal effects at density are calculated using the following formula: $\beta_i(1 - d_n)$, where, β_i is the coefficient in front of the independent variable, d_n is the density of dependent variable network specified at the top of each column. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1. In the names of independent variables: D - direct exposure, I - indirect or common portfolio exposure, L - loans, B - bonds, E - equity, P - equities non-securities or fund shares. Dependent variables: DDLY_thr - DDLY network with threshold, CS - return-based partial correlation network, HSS - tail-risk based network, PD - forward-looking default probability-based network, CDS - credit risk-based network.

6 Robustness

In this section, we test if our baseline results hold under different specifications. First, we propose alternative measures of DDLY and see how they perform relative to the original DDLY. Second, we test sensitivity of CS network towards change in an estimation parameter and assess its performance in capturing balance-sheet exposures. Finally, we dig deeper into the problem of multiple zeros in our networks and assess if our results are biased by capturing mostly zero-zero links.

6.1 Alternative DDLY measures.

We start by looking at the estimation of alternative DDLY measures presented in section 3.1. Table 15 reports the results for the specification of equation 8, where the dependent variable is a DDLY network. DDLY variable is a directed link, so all explanatory variables are also taken as directed. The first column is an original DDLY, volatility based network, not scaled; the second is DDLY that is defactored using demeaning in a volatility based network; the third is DDLY defactored using demeaning, volatility based network, but not scaled; the fourth is DDLY defactored using 1 EU stock market factor, volatility based network; the fifth and the sixth is respectively return-based or volatility-based DDLY network that is defactored using 4 factors (stock market factor, banking sector factor, option-based volatilities both for stock market and banking sector); finally, the last, seventh, column is exactly like the first column but with a threshold.

Before moving to the interpretation of the results, let us note that columns (1) and (7) have almost identical results both in terms of size and significance. Thus from now on, we do not comment on column (7).

One big result emerges immediately after looking at Table 15. All DDLY measures are strongly significantly and positively associated with common portfolios of loans and bonds. In other words, our results suggest that two banks having similar loan or bond portfolios are perceived by the markets as more connected both in terms of higher volatility connectedness and higher return connectedness. This is expected as loan and bond portfolio makes a major share of banks' balance sheet and to a large extent defines banks' profitability which in its turn affects banks' equity returns and return volatility.

The only DDLY network for which this result does not hold is the non-scaled defactored DDLY in Column (3). Here we expect the results to be less robust since as explained in 3.6, z-score normalization of variables has a significant effect. Indeed, two banks may look similar due to strong common factors, but if one eliminates these factors, either by taking cross-sectional mean or by defactoring, banks become very different. To take this into account one needs to scale or normalize defactored variables.

Now looking at the results in more detail, we can note that the correlation in indirect loans and bonds cannot be explained by common factors, as comparing column (1) with column (6) of Table 15 one can see that coefficients are similar in terms of the magnitude and stay significant, even when the dependent variable is constructed using a market factor, banking sector factor, and option-based volatility both for market and for the banking sector. However, the same cannot be said about direct equity exposures as once co-movement with common factors is eliminated direct equity becomes insignificant.

Direct loans are not significant for columns (1)-(4), and become significant once we eliminate common factors (columns (5) and (6)). This result is similar to Abbassi et al. (2017), who find positive and significant association between direct interbank lending and a CDS-based market network for the German sample of banks.

Finally, it is worth mentioning that once controlling for common factors, DDLY networks based on equity return - column (5) - and volatility - column (6) - provide very similar results. Most of the coefficients have similar magnitude and significance in columns (5) and (6).

Table 15: Poisson: Baseline DDLY + controls with clustered errors at dyad level

Dependent Variable:	DDLY						
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variables							
DL	0.0009	-0.0031	-0.0020	0.0039	0.0074*	0.0068**	0.0009
	(0.0034)	(0.0051)	(0.0031)	(0.0055)	(0.0040)	(0.0031)	(0.0033)
IL	0.0324***	0.0241***	0.0025	0.0322***	0.0233***	0.0262***	0.0316***
	(0.0092)	(0.0082)	(0.0036)	(0.0074)	(0.0080)	(0.0083)	(0.0092)
DB	0.0059	0.0005	-0.0019	0.0015	0.0032	0.0027	0.0055
	(0.0073)	(0.0086)	(0.0037)	(0.0071)	(0.0090)	(0.0062)	(0.0074)
IB	0.0932***	0.1424***	0.0261***	0.1316***	0.1081***	0.1112***	0.0905***
	(0.0143)	(0.0271)	(0.0083)	(0.0221)	(0.0190)	(0.0185)	(0.0139)
DE	0.0090**	0.0097	-0.0002	0.0081	0.0068	0.0070	0.0087**
	(0.0038)	(0.0073)	(0.0015)	(0.0051)	(0.0053)	(0.0044)	(0.0037)
IE	0.0091	0.0123	0.0059*	0.0080	0.0082	0.0066	0.0087
	(0.0063)	(0.0118)	(0.0034)	(0.0115)	(0.0088)	(0.0086)	(0.0061)
DP	-0.0015*	-0.0057*	-0.0002	-0.0031	-0.0043*	-0.0035**	-0.0014*
	(0.0009)	(0.0035)	(0.0009)	(0.0022)	(0.0022)	(0.0016)	(0.0008)
IP	-0.0002	-0.0029***	-0.0055***	-1.9×10^{-5}	0.0018***	0.0025***	-0.0001
	(0.0005)	(0.0005)	(0.0011)	(0.0008)	(0.0005)	(0.0009)	(0.0004)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed-effects							
source	Yes	Yes	Yes	Yes	Yes	Yes	Yes
target	Yes	Yes	Yes	Yes	Yes	Yes	Yes
date	Yes	Yes	Yes	Yes	Yes	Yes	Yes
source-date	Yes	Yes	Yes	Yes	Yes	Yes	Yes
target-date	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics							
Observations	28,271	28,271	28,271	28,271	28,271	28,271	28,271
Squared Correlation	0.67118	0.34832	0.30411	0.41879	0.45751	0.52335	0.67273
Pseudo R^2	0.13949	0.07525	0.05043	0.08138	0.09440	0.10086	0.14027
BIC	87,875.3	90,829.4	87,801.4	90,545.5	90,409.1	89,227.4	87,958.6

Notes: Two-way (source & target) standard-errors in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1. Fixed-effects: source, target, date, source-date,target-date for all regressions. In the names of independent variables: D - direct exposure, I - indirect or common portfolio exposure, L - loans, B - bonds, E - equity, P - equities non-securities or fund shares. Dependent variables: Model (1) - original DDLY, volatility-based network, not scaled; Model (2) - DDLY defactored using demeaning, volatility-based network; Model (3) - DDLY defactored using demeaning, volatility-based network, not scaled; Model (4) - DDLY defactored using 1 EU stock market factor, volatility based network; Model (5) - DDLY defactored using 4 factors, return based network; Model (6) - DDLY defactored using 4 factors, volatility based network; Model (7) - original DDLY volatility-based network with the threshold. Pair-wise controls: TA - total assets, CET1 - common equity tier ratio, RWAI - risk-weighted assets intensity, ROAE - return on average equity.

6.2 Alternative CS measures

In this section, we look how alternative specifications of the CS measure affect our baseline results. As explained in Section 3.2, we compute CS baseline measure by removing one common factor and using a five percent threshold in Holm-Bonferroni thresholding method. As a sensitivity check, we consider the following combinations: one-factor and a 10% threshold (CS_TH10_F1 of $CS1$), one-factor and a 25% threshold (CS_TH25_F1 or $CS2$), and three-factors with a 25% threshold (CS_TH25_F3 or $CS3$). Table 16 reports the results for all CS measures. The key messages are the following: CS measures capture particularly well common loan portfolio and direct bond exposures as well direct and indirect equity connections. The results are robust across all specifications with positive and significant coefficients of a similar magnitude. The only exception is DB that loses its significance in CS3 - Column (4).

Table 16: Poisson: CS sensitivity to threshold

Dependent Variable:	CS	CS1	CS2	CS3
Variables				
DL	-0.0037	-0.0030	-0.0033	-0.0050
	(0.0047)	(0.0050)	(0.0051)	(0.0050)
IL	0.0402***	0.0469***	0.0494***	0.0232***
	(0.0137)	(0.0117)	(0.0110)	(0.0039)
DB	0.0279***	0.0297***	0.0345***	0.0153
	(0.0101)	(0.0108)	(0.0119)	(0.0094)
IB	0.0364	0.0402	0.0415	0.0341
	(0.0396)	(0.0438)	(0.0467)	(0.0347)
DE	0.0186**	0.0193**	0.0206**	0.0273**
	(0.0091)	(0.0096)	(0.0095)	(0.0123)
IE	0.0276*	0.0281*	0.0273*	0.0298**
	(0.0156)	(0.0162)	(0.0161)	(0.0151)
DP	0.0068	0.0062	0.0057	0.0035
	(0.0056)	(0.0056)	(0.0057)	(0.0081)
IP	0.0021	0.0017	0.0014	0.0018
	(0.0026)	(0.0027)	(0.0027)	(0.0031)
Controls	Yes	Yes	Yes	Yes
Fixed-effects				
source	Yes	Yes	Yes	Yes
target	Yes	Yes	Yes	Yes
date	Yes	Yes	Yes	Yes
source-date	Yes	Yes	Yes	Yes
target-date	Yes	Yes	Yes	Yes
Fit statistics				
Observations	28,225	28,225	28,225	28,225
Squared Correlation	0.33737	0.33516	0.33545	0.13831
Pseudo R^2	0.03778	0.03896	0.04077	0.02532
BIC	87,446.1	87,641.7	87,881.5	89,714.2

Notes: Two-way (source & target) standard-errors in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1. Fixed-effects: source, target, date, source-date,target-date for all regressions. In the names of independent variables: D - direct exposure, I - indirect or common portfolio exposure, L - loans, B - bonds, E - equity, P - equities non-securities or fund shares. Dependent variables: CS - original CS measure, CS1 - CS TH1 F1 is 10% threshold with 1 factor, CS2 - CS TH25 F1 is 25% threshold with 1 factor, CS3 - CS TH25 F3 is 25% threshold with 3 factors. Pair-wise controls: TA - total assets, CET1 - common equity tier ratio, RWAI - risk-weighted assets intensity, ROAE - return on average equity.

6.3 Selection 00 and Selection 11

The density of both market-based and balance-based networks is very low, except for DDLY and CDS, thus our data contains many zeros both for dependent and independent variables. Poisson specification is meant to deal with this issue. However, as a robustness check, we conduct an alternative specification to ascertain that significance of our results is not driven by cases when we have zeros on both sides of the equation. To address the issue, we first remove all NA values in the variables. Next, we create dummy $cs00$, that takes into account so-called zero-zero cases, that is we set $cs00$ to zero when a dependent variable is zero and all independent variables (the links in true exposure networks) are zero as well. We set $cs00$ equal to one otherwise. That being said we would like to estimate a model, when either dependent variable or at least one of the independent variables are non-zeros. But this might generate a potential selection problem. In order to address this issue we apply Heckman correction model. First, we run a probit model where we regress dummy $cs00$ on the pairwise characteristics which we used as controls in our baseline models: total assets, capital ratio CET1, risk-weighted assets and return on equity. Second, we calculate Inverse Mills ratio (IMR). Third, we run standard Poisson regression using the R package "fixest", but we augment our baseline set of regressors with an additional variable IMR. We use Poisson in the second stage because even though we eliminate zero-zero cases, still we do have a lot of zeros either on the left-hand side or on the right-hand side of the equation. This problem arises in all the market-based networks except for DDLY network, as it is a complete network.

Table 17 reports the results for the second stage of selection model for CS, HSS, PD, and CDS networks. As can be seen from the table, the results remain robust and consistent with the previous findings.

Next, we conduct even stricter robustness check. We construct a dummy $cs11$ setting it to one only when dependent and at least one of the independent variables is non-zero, and we set it to zero otherwise. This exercise potentially might create even more severe selection problem, that is why we use a Heckman correction model again. Results are reported in Table 18. As this exercise is overly restrictive and with significantly lower number of observations the results should be taken with the grain of salt. Nevertheless, the results for the HSS and CDS networks are confirmed. The PD retains only one positive and significant coefficient, IL. Interestingly, this corresponds to the variable that also exhibited the largest marginal effects in both baseline specifications, logit and Poisson. As for the CS network, the restriction reduces even further the number of observations (just 2% of observations from the baseline model) , this can potentially explain several unexpected negative signs in front of IL, IB and IE. However, the robust result for direct equity - DE - further underlines its importance.

Overall, we conclude that our results are not driven by zero connections. If one looks at interconnections both in terms of direct holdings and common portfolios as a source of credit risk, then one could expect the best approaches based on credit risk information. This is indeed confirmed byIL surviving the strictest robustness test for the CDS, PD, and HSS measures. DB and IB remain strongly robust for the CDS network as well. On the other hand, for measure based on returns such as CS, the only surviving variable with different robustness checks is DE - direct equity.

Table 17: Selection 00: CS, HSS, PD, CDS

Dependent Variable:	CS	HSS	PD	CDS
Variables				
DL	-273.4***	-14.29	6.917	10.32
	(91.73)	(10.02)	(18.58)	(38.31)
IL	8.545**	1.508**	6.183***	6.693**
	(4.019)	(0.6770)	(2.008)	(3.239)
DB	167.8	89.35**	258.5**	1,172.9***
	(401.2)	(41.53)	(129.8)	(240.9)
IB	25.58***	7.738***	3.139	7.077***
	(7.441)	(1.716)	(3.913)	(2.474)
DE	5,031.8***	181.3	237.5	234.4
	(619.0)	(113.9)	(188.8)	(233.1)
IE	65.92	44.56**	138.0***	-53.87
	(68.85)	(19.40)	(42.70)	(34.37)
DP	1,273.4***	266.1***	287.3	322.8**
	(417.2)	(52.67)	(238.7)	(135.9)
IP	570.0***	122.4***	95.65	327.9*
	(101.5)	(35.65)	(174.2)	(183.4)
IMR1	12.48***	1.501***	1.598***	118.2***
	(3.057)	(0.3573)	(0.5719)	(31.07)
Fixed-effects				
source	Yes	Yes	Yes	Yes
target	Yes	Yes	Yes	Yes
source-date	Yes	Yes	Yes	Yes
target-date	Yes	Yes	Yes	Yes
date	Yes	Yes	Yes	Yes
Fit statistics				
Observations	4,856	29,429	25,202	4,156
Squared Correlation	0.66165	0.25032	0.13493	0.44957
Pseudo R^2	0.39205	0.06692	0.10416	0.12784
BIC	7,468.5	18,588.6	21,360.3	6,891.2

Notes: Two-way (source & target) standard-errors in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1. Fixed-effects: source, target, date, source-date,target-date for all regressions. In the names of independent variables: D - direct exposure, I - indirect or common portfolio exposure, L - loans, B - bonds, E - equity, P - equities non-securities or fund shares, IMR - inverse Mills ratio. Dependent variables: CS - original CS measure, HSS - tail-risk based network, PD - forward-looking default probability-based network, CDS - credit risk based network. Pair-wise controls: TA - total assets, CET1 - common equity tier ratio, RWAI - risk-weighted assets intensity, ROAE - return on average equity.

Table 18: Selection 11: CS, HSS, PD, CDS

Dependent Variable:	CS	HSS	PD	CDS
Variables				
DL	8.532	0.1800	-27.48***	-34.23
	(11.25)	(7.959)	(9.597)	(29.22)
IL	-1.374***	0.9459***	3.079***	7.217***
	(0.3165)	(0.2238)	(1.181)	(1.695)
DB	-19.81	5.243	91.69	897.6***
	(34.17)	(16.55)	(60.60)	(132.9)
IB	-2.964***	1.499***	1.584	5.167***
	(0.4498)	(0.4060)	(1.283)	(1.764)
DE	521.5***	112.4**	86.51	443.1**
	(25.88)	(52.35)	(237.4)	(212.2)
IE	-1.355	11.14**	31.08	-57.53***
	(8.591)	(4.449)	(19.23)	(1.103)
DP	130.1	146.9***	25.59	180.8
	(91.08)	(15.03)	(154.9)	(154.6)
IP	157.5***	13.11	161.6*	146.6
	(25.90)	(25.74)	(96.26)	(90.46)
IMR1	-0.5664	6.251	-1.164	7.079***
	(0.3985)	(4.327)	(1.937)	(1.480)
Fixed-effects				
source	Yes	Yes	Yes	Yes
target	Yes	Yes	Yes	Yes
source-date	Yes	Yes	Yes	Yes
target-date	Yes	Yes	Yes	Yes
date	Yes	Yes	Yes	Yes
Fit statistics				
Observations	562	4,585	4,635	2,178
Squared Correlation	0.98576	0.70488	0.39325	0.59877
Pseudo R^2	0.02156	0.04054	0.06784	0.09534
BIC	5,323.6	14,340.6	15,946.1	6,201.0

Notes: Two-way (source & target) standard-errors in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1. Fixed-effects: source, target, date, source-date,target-date for all regressions. In the names of independent variables: D - direct exposure, I - indirect or common portfolio exposure, L - loans, B - bonds, E - equity, P - equities non-securities or fund shares, IMR - inverse Mills ratio. Dependent variables: CS - original CS measure, HSS - tail-risk based network, PD - forward-looking default probability-based network, CDS - credit risk based network. Pair-wise controls: TA - total assets, CET1 - common equity tier ratio, RWAI - risk-weighted assets intensity, ROAE - return on average equity.

7 Conclusion

Networks based on market data have several advantages. First, they are publicly available. This means that new academic theories of contagion can be tested using networks simulated from these techniques, where networks based on balance sheet data may not be available. This lack of availability can be due to the fact that data collection is costly and time consuming, and the data are sensitive and so subject to strict confidentiality standards.

Second, market-based reconstructed networks can provide necessary information more timely as they can be computed on a daily basis. This is a great advantage particularly during crisis times as regulators can get information about a change in a network following, for example, a bank failure with a one day delay instead of three to six months after a time-consuming process of collecting and cleaning balance sheet data.

Market-based networks however come with some severe disadvantages. First, they can be constructed only for publicly traded banks as price information is needed for the analysis. However, important nodes in the interbank network can be privately held so that there is no equity prices to input into the market-based network estimation. As a result, a network may be estimated using a non-representative sample. Further, CDS swaps are only estimated for larger banks that make such trades worthwhile to market makers. So market-based networks are estimated from a very special and censored sample. Second, market prices tend to be very volatile during times of crisis. It is not always clear that these prices reflect the same information during times of financial crisis as they do in normal times.

The third problem which is also the subject of this paper is lack of clarity regarding what exactly market-based networks represent. Just as each of the balance-sheet networks represent different types of exposures, with differing characteristics, the five market-based networks represent different aspects of available public knowledge regarding the connections between banks.

One of the networks is built upon those rare, but important banks for which there is a CDS market. Others have links that are built from equity prices data, some networks weigh tail events more heavily, some are based on default probabilities. Not surprisingly, each may have their advantages.

To investigate this question, we use a very complete balance sheet data set to construct a large network of various exposures between European banks: four networks reflect direct exposures between banks in loans, bonds and equity as well as non-traded equity or participation/ownership; the other four networks reflect exposures to similar types of assets or common portfolios in loans, bonds, equity and non-traded equity. Then we compare these baseline networks to five networks estimated from market data. These five approaches use either data on equity returns and volatility (Diebold, Yılmaz (2014), Demirer et al. (2018), Craig, Saldías (2016), Hautsch et al. (2014), Hautsch et al. (2015)), default probabilities (Duan et al. (2012), Chan-Lau et al. (2016)) or CDS (Brownlees et al. (2020)). In spite of the diversity of MBN tested in this paper, some patterns emerge for all of them. The balance sheet links most embodied in the links of the market-based networks are indirect exposures, and thus the MBN reflect essentially a common business model among banks more easily available to public investors. Indirect bonds and indirect loans are the links most often represented by links in the market-based networks. To a lesser extent, direct bonds are also represented by some of the networks, particularly by those that capture credit risk. On the other hand, networks based on equity prices (volatility and returns) reflect direct and indirect equity exposures. Finally, direct interbank loans that often serve as an input to the interbank contagion analysis cannot be robustly estimated by any of the market-based networks. This is potentially due to the fact that this information is proprietary and not available to the market, but also probably because direct interbank loans constitute a relatively tiny share of banks' exposures.

When we look at various market-based networks individually, they do reflect different information in the balance sheet exposures. If one is interested at interconnections in terms of direct holdings and common portfolios as a source of credit risk, then one might expect the best approaches to be based on credit risk information. This is indeed confirmed that tail risk measure and default probability measure are most associated with both direct and common exposures of loans and bonds. They are particularly good at capturing existence of a link, however, less good at matching link size. CDS-based networks capture mostly direct and indirect bond holdings which is indeed expected as CDS contracts are directly linked to bond performance. On the other hand, measures based on returns and return volatility such as DDLY and CS capture particularly well exposures in direct and indirect equity but also common factors and overlapping portfolios in loans and bonds which are sources of banks' variability in returns.

Granular data on interbank interconnections are not available to academics and most regulatory institutions, but systemic risk and contagion analysis are widely studied using simulated networks. Our policy implication suggests that one should be judicious in the choice of method to construct a network from market data. Whether the use of a particular MBN or a particular BSN is appropriate for contagion analysis depends on the particular mechanism of financial contagion that is used. Here the youth of structural contagion analysis hinders a final statement of which MBN or even whether any MBN would add to the analysis. Fairly crude models of cascading defaults in a sparse network, such as proposed by Allen, Gale (2000) or Eisenberg, Noe (2001) seem to depend on the structure of direct loan interbank networks, where no MBN matches the edges well enough, although the sparsity and other network quality measures can be matched. A more reduced form model of the propagation of a common shock via firesale contagion could possibly be captured in the credit risk models constructed via techniques that focus on tail-risk, probability defaults or CDS prices. A much more sophisticated model of network formation leading to financial stability problems such as proposed in Craig, Ma (2018) would seem to require a lot of information about direct lending exposure networks, which is not embedded in any of the MBN that we studied. As formal modeling of financial stability gets more sophisticated and more relevant to policy counterfactual analysis, the relevance of any of the networks, MBN or BSN, will come into a sharper focus. As is usually the case, richer data on the links between financial institutions, both in actual exposures and in exposures inferred from market data, will be needed to inform both the development and the interpretation of these formal models.

Bibliography

Abbassi Puriya, Brownlees Christian, Hans Christina, Podlich Natalia. Credit risk interconnectedness: What does the market really know? // Journal of Financial Stability. 2017. 29. 1–12.

Acharya Viral V, Hasan Iftekhar, Saunders Anthony. Should banks be diversified? Evidence from individual bank loan portfolios // The Journal of Business. 2006. 79, 3. 1355–1412.

Adam Michał, Bochmann Paul, Grodzicki Maciej, Mingarelli Luca, Montagna Mattia, Rodriguez d'Acri Costanza, Spaggiari Martina, others . Assessing the systemic footprint of euro area banks // Financial Stability Review. 2019. 2.

Allen Franklin, Babus Ana. Networks in finance // The network challenge: strategy, profit, and risk in an interlinked world. 2009. 367.

Allen Franklin, Babus Ana, Carletti Elena. Asset commonality, debt maturity and systemic risk // Journal of Financial Economics. 2012. 104, 3. 519–534.

Allen Franklin, Gale Douglas. Financial contagion // Journal of political economy. 2000. 108, 1. 1–33.

Anand Kartik, Lelyveld Iman van, Banai Ádám, Friedrich Soeren, Garratt Rodney, Hałaj Grzegorz, Fique Jose, Hansen Ib, Jaramillo Serafín Martínez, Lee Hwayun, others . The missing links: A global study on uncovering financial network structures from partial data // Journal of Financial Stability. 2018. 35. 107–119.

Ang Andrew, Longstaff Francis A. Systemic sovereign credit risk: Lessons from the US and Europe // Journal of Monetary Economics. 2013. 60, 5. 493–510.

Barigozzi Matteo, Brownlees Christian. Nets: Network estimation for time series // Journal of Applied Econometrics. 2019. 34, 3. 347–364.

Bech M., Atalay E. The topology of the federal funds market. // Federal Reserve Bank of New York Staff Reports. 2008.

Bednarek Peter, Dinger Valeriya, Schultz Alison, Westernhagen Natalja von. Banks of a Feather: The Informational Advantage of Being Alike // Deutsche Bundesbank Working Paper Series. 2022.

Belloni Alexandre, Chernozhukov Victor. l1-penalized quantile regression in high-dimensional sparse models // The Annals of Statistics. 2011. 39, 1. 82–130.

Bergé Laurent, others . Efficient estimation of maximum likelihood models with multiple fixed-effects: the R package FENmlm. 2018.

Boss Michael, Elsinger Helmut, Summer Martin, Thurner 4 Stefan. Network topology of the interbank market // Quantitative finance. 2004. 4, 6. 677–684.

Brownlees Christian, Hans Christina, Nualart Eulalia. Bank Credit Risk Networks: Evidence from the Eurozone // Journal of Monetary Economics. 2020.

Chan-Lau Jorge A, Chuang Chienmin, Duan Jin-Chuan, Sun Wei. Banking network and systemic risk via forward-looking partial default correlations // research paper of the National University of Singapore Risk Management Institute. 2016.

Cont R., Moussa A., Santos E. Network structure and systemic risk in banking systems // mimeo. 2010.

Craig Ben, Saldías Martín. Spatial dependence and data-driven networks of international banks. 2016.

Craig Ben R, Ma Yiming. Intermediation in the interbank lending market // Columbia Business School Research Paper Forthcoming. 2018.

Crain Harry. Probabilistic Foundations of Statistical Network Analysis. 2018.

Demirer Mert, Diebold Francis X, Liu Laura, Yilmaz Kamil. Estimating global bank network connectedness // Journal of Applied Econometrics. 2018. 33, 1. 1–15.

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

Duan Jin-Chuan, Miao Weimin. Default correlations and large-portfolio credit analysis // Journal of Business & Economic Statistics. 2016. 34, 4. 536–546.

Duan Jin-Chuan, Sun Jie, Wang Tao. Multiperiod corporate default prediction—A forward intensity approach // Journal of Econometrics. 2012. 170, 1. 191–209.

Eisenberg Larry, Noe Thomas H. Systemic risk in financial systems // Management Science. 2001. 47, 2. 236–249.

Elliott Matthew, Georg Co-Pierre, Hazell Jonathon. Systemic risk shifting in financial networks // Journal of Economic Theory. 2021. 191. 105157.

Fan Jianqing, Feng Yang, Wu Yichao. Network exploration via the adaptive LASSO and SCAD penalties // The annals of applied statistics. 2009. 3, 2. 521.

Freixas Xavier, Parigi Bruno M, Rochet Jean-Charles. Systemic risk, interbank relations, and liquidity provision by the central bank // Journal of money, credit and banking. 2000. 611–638.

Friedman Jerome, Hastie Trevor, Tibshirani Robert. Sparse inverse covariance estimation with the graphical lasso // Biostatistics. 2008. 9, 3. 432–441.

Hale Galina, Lopez Jose A. Monitoring banking system connectedness with big data // Journal of Econometrics. 2019. 212, 1. 203–220.

Hautsch Nikolaus, Schaumburg Julia, Schienle Melanie. Forecasting systemic impact in financial networks // International Journal of Forecasting. 2014. 30, 3. 781–794.

Hautsch Nikolaus, Schaumburg Julia, Schienle Melanie. Financial network systemic risk contributions // Review of Finance. 2015. 19, 2. 685–738.

IMF.FSAP.Germany . Stress testing the banking and insurance sectors — technical note // IMF Country Report No. 16/191. 2016.

IMF.FSAP.Norway . Technical note — linkages and interconnectedness in the Norwegian financial system // IMF Country Report No. 15/256. 2015.

IMF.FSAP.Spain . Technical note—interconnectedness and spillover analysis in Spain's financial system // IMF Country Report No. 17/344. 2017.

Jackson Matthew O, Pernoud Agathe. Optimal Regulation and Investment Incentives in Financial Networks // Available at SSRN 3311839. 2019.

Jackson Matthew O, Pernoud Agathe. Systemic risk in financial networks: A survey // Annual Review of Economics. 2021. 13. 171–202.

Khare Kshitij, Oh Sang-Yun, Rajaratnam Bala. A convex pseudolikelihood framework for high dimensional partial correlation estimation with convergence guarantees // Journal of the Royal Statistical Society: Series B: Statistical Methodology. 2015. 803–825.

Ongena Steven, Smith David C, others . Bank relationships: a review // Performance of financial institutions: Efficiency, innovation, regulation. 2000. 221.

Silva JMC Santos, Tenreyro Silvana. The log of gravity // The Review of Economics and statistics. 2006. 88, 4. 641–658.

Silva JMC Santos, Tenreyro Silvana. Further simulation evidence on the performance of the Poisson pseudo-maximum likelihood estimator // Economics Letters. 2011. 112, 2. 220–222.

Yuan Ming, Lin Yi. Model selection and estimation in the Gaussian graphical model // Biometrika. 2007. 94, 1. 19–35.