

# Bank Regulation, Network Topology, and Systemic Risk: Evidence from the Great Depression

Sanjiv R. Das  
Department of Finance  
Santa Clara University

Kris James Mitchener  
Department of Economics  
Santa Clara University  
CAGE, CESifo, CEPR & NBER

Angela Vossmeier  
Robert Day School of Economics and Finance  
Claremont McKenna College\*

July 3, 2021

## Abstract

We study how bank regulation interacts with network topology to influence systemic stability. Employing unique hand-collected data on the correspondent network for all U.S. banks on the eve of the Great Depression and a methodology that captures bank credit risk and network position, we explore how the pyramid-shaped network topology was inherently fragile and systemically risky. We measure its contribution to banking distress in the early 1930s, and show that a bank's network position as well as the risk of its network neighbors are strong predictors of bank survivorship. Institutional alternatives, such as branch banking, and alternative topologies appear to deliver networks that are more stable than the network that existed in 1929.

*JEL classification:* L1, E42, E44, G01, G18, G21, N12, N22.

*Keywords:* Systemic risk; banking networks; Great Depression; peer effects; branch banking; model comparison.

---

\*We are grateful to Charles Calomiris, Mark Carlson, Co-Pierre Georg, Gary Gorton, Taylor Jaworski, Ivan Jeliazkov, Christoffer Koch, Lorian Pelizzon, and conference and seminar participants at the NBER Summer Institute, Federal Reserve Bank of Dallas, Federal Reserve Bank of Atlanta, Bank of England, Federal Reserve Board, Mountain West Economic History Conference, Chicago Financial Institutions Conference, and Western Finance Association Annual Conference for helpful comments and suggestions. We thank Jack Brown, Kendall Greenberg, Angie Wang, and Peter Welch of the Lowe Institute of Political Economy, and Bhavika Booragadda, Lorraine Zhao, Tanisha Seth, Krystal Sung, Laira Aggarwal, and Hunter Olsen of the Financial Economics Institute for excellent research assistance. The authors may be reached at [srdas@scu.edu](mailto:srdas@scu.edu); [kmitchener@scu.edu](mailto:kmitchener@scu.edu); and [angela.vossmeier@cmc.edu](mailto:angela.vossmeier@cmc.edu).

# 1 Introduction

The Global Financial Crisis (GFC) focused attention on the importance of understanding the connections between financial institutions, which have been seen as a mechanism for propagating distress (Allen and Gale, 2000; Acharya et al., 2017; Elliott et al., 2014; Acemoglu et al., 2015; Allen et al., 2010; Freixas et al., 2000; Dasgupta, 2004; Caballero and Simsek, 2013). Researchers and policymakers have sought a better understanding of the characteristics that predispose financial systems to risk, but are often hindered by difficulties measuring connections between financial institutions, which can be opaque. In this paper, we analyze a banking system where the entire interbank network can be observed and precisely mapped in order to examine the importance of network topology in predisposing financial systems to risk.

The structure and shape of networks (hereafter “network topology”) are known to affect the speed and scale of the transmission of payments, information, people, and goods, such as commercial aviation travel routes or packets of information on the internet. For example, in the context of banking, scale-free networks can be both a blessing and a curse: they can permit efficient transmission of payments in good times, but can potentially propagate risk in periods of distress. Researchers often model extant financial network topologies as the outcome of responses of firms in the network. However, what is less well understood is how financial regulation can alter the shape of financial networks and how that, in turn, predisposes financial institutions to risk and influences their ability to weather financial crises. Our research seeks to understand this question. Focusing on network topology also allows us to examine the extent to which a financial institution’s ability to survive a crisis is dependent, not only on its ex ante balance-sheet risk and network position, but on its network-neighbors’ riskiness.

To shed light on these questions, we study the commercial bank correspondent network in 1929 – just prior to the largest American financial crisis of the 20<sup>th</sup> century. Using bank-level data, we demonstrate that the pyramid-like structure of the commercial banking system, which was a legacy of 19<sup>th</sup> century regulations on bank reserves, had direct consequences for systemwide risk. We do so by first constructing novel measures of systemic risk at

the individual bank level to examine how the correspondent network’s topology influenced the geographical distribution and concentration of risk prior to the Great Depression, and how “fragile” the 1929 network was relative to other possible topologies. Our systemic risk measure accounts for risk arising from a bank’s balance sheet and its position in the network, and allows us to quantify risk arising from linkages with “network neighbors” or “peers.” Next, we use Bayesian model comparison techniques and data on surviving banks in 1934 (after the severe banking crisis of the early 1930s had ended) to examine whether the inclusion of our systemic risk measures in the prediction of bank survival is supported by the data and, more broadly, to provide a validation test for policymakers interested in applying these measures. Finally, we analyze the extent to which a bank’s predicted probability of survival can be attributed to balance-sheet risk and network position as well as the risk due to linkages to network neighbors and spatial neighbors. Our findings not only shed light on how the correspondent-banking network predisposed banks to subsequent failure but, more generally, allow us to examine the biases that arise from the omission of network characteristics and peer effects when studying financial stability.

To do so, we hand collected data on every commercial bank in the U.S. in 1929 and each bank’s correspondent relationship with every other bank. These data and our historical setting prove crucial for studying the impact of network topology on systemic stability. *First*, in contrast to approaches that estimate financial linkages from financial flow data or price and return co-movements, our historical measure of connectedness uses stated, contractual bank correspondent relationships based on bank records. Because network connections between commercial banks are observed in our data and are not inferred from correlations, we are able to construct a full adjacency matrix of connections. The full matrix allows for a deeper examination of peer effects, and affords the consideration of counterfactual adjacency matrices that can be used to assess whether the observed 1929 network was riskier than alternative topologies. *Second*, in this earlier era, regulations on branch banking varied across state borders. This geographical variation in regulation allows us to consider how banking policy interacts with financial networks to shape risk and to examine spatial dimensions of systemic risk. *Third*, financial regulation affects the size of the financial system. Because of restrictions on branch banking at the state and national level, small banks flourished ([White](#),

1983). As a result, there were more than five times as many banks in 1929 than today. This feature of the data allows us to examine topologies that look different from those characterizing the American banking system in recent crises, such as the GFC. As we show, this has important implications for where risk can reside in a financial system. *Fourth*, and related to the previous point, because we collect data on all commercial banks (roughly 26,000), we have the full distribution of banks in terms of size – not just the upper tail. Thus, we are able to demonstrate that some financial crises, like the Great Depression, can start, unsuspectingly, at the bottom of the distribution, but can then reach the entire network of banks through the vast number of links and network-neighbor balance-sheet risk at the top nodes – the implications of which may be important for systemic-risk regulators.

In understanding how network topology influences systemic risk, we find that the pyramid-shape of the pre-existing network topology made the system more fragile and riskier than counterfactual random graph and alternative scale-free topologies. During panic periods, the pyramid shape of the network left positionally-important banks (e.g., reserve and central-reserve city banks) exposed to large clusters of banks making simultaneous withdrawals through the correspondent network in order to fend off local runs – a feature that magnified systemwide risk. In decomposing the predicted probability of bank survival, we show that, if the 1929 correspondent network instead had a random network topology, nearly 1,100 more banks (roughly 5% of all banks or 12% of those that closed) would have survived the Great Depression, making it more stable and less crisis prone. This result supports recent theoretical work showing that interconnectedness can have a negative effect on systemic stability when networks are dense (Acemoglu et al., 2015) and draws attention to the importance of network topology in understanding where systemic risk emanates.

We show that regulation made the network “fragile by design,” as it appears to have heightened systemic risk on the eve of the Depression. An alternative to the correspondent network existed in the 1920s, mainly branch banking, but it was prohibited nationally and restricted or prohibited in many states. In states where branch banking was limited or prohibited, correspondent banks had an oversized importance in the network on the eve of the Depression. We thus consider how systemic stability would have been affected under an

alternative scenario – if the pyramid had been “flattened” by allowing branch banking to flourish at the state level, as it had in states like California. We estimate that if regulators had permitted the correspondent network to be reshaped by allowing widespread branch banking (at the state level), the predicted probability of bank survival would have been 10 percentage points higher and that roughly 26% of those that closed would have survived the banking crises of the 1930s. Here, the beneficial effects of branching banking accrue from reduced reliance on the correspondent network during periods of distress, and could be complementary to previously identified benefits in the literature, such as portfolio diversification or industry consolidation.

We also analyze how much systemic risk in the banking system emanated from the balance sheets of banks before the banking crisis of the early 1930s began. Decreasing a bank’s risk by one standard deviation increases the predicted probability of bank survival by 8.35 percentage points, resulting in more than 2,000 additional banks surviving the Great Depression. However, the topology of the network also matters when considering a bank’s balance-sheet risk since risk can arise through its linkages via the correspondent network. When we account for this additional source of balance-sheet risk, approximately 250 more banks would survive the Depression if we lowered “network neighbor” risk by one standard deviation. Taken together, these balance sheet effects capture approximately 24% of banks that closed.

Our model comparison analysis demonstrates that network position and network peer effects should not be overlooked by researchers studying financial crises. Although peer effects have received considerable attention in applied microeconomics and randomized control trials in development economics, applications to financial networks and systemic risk have received little attention. Our findings demonstrate that a model with systemic risk and network-neighbor measures has a posterior model probability of nearly 1, relative to models that omit these variables. Further, we show that failing to account for network position and peer effects can materially bias estimates of other variables in models of bank distress.

The findings in this paper reveal new aspects of financial networks that have implications for policy, methodology, and history. Our work relates to a growing literature on systemic-

risk measurement and the evaluation of crises (e.g., [Acharya et al. \(2017\)](#)). (For a survey of methodologies, see [Bisias et al. \(2012\)](#)). To date, studies have largely focused on the period around the GFC. This is somewhat surprising, given the well-documented challenges of measuring connections between the “shadow banking system” and the regulated sector of the 2000s, a need to understand the out of sample properties of these tools, and the benefits from comparing the GFC to earlier crises. By exploring systemic-risk measures from 1929, our model selection exercise (which shows that the data strongly prefer models that include network and systemic risk measures) sheds light on their external validity. And by employing a rich, micro level data set on individual banks, we are able to demonstrate how they can be employed to pinpoint where ex ante risk resides in a financial network and what “components” of systemic risk might be of particular concern to policymakers. Early-generation systemic-risk measures estimated network linkages using financial-flow data. For example, [Billio et al. \(2012\)](#) construct networks of financial institutions using bivariate Granger causality regressions based on inferences from asset returns. Our analysis draws attention to physical connectedness, which has been explored more recently by [Burdick et al. \(2011\)](#) and [Brunetti et al. \(2019\)](#) for U.S. and European banks, respectively, during the 2007-08 crisis.

Topology and the concentration of connectedness are longstanding topics of interest in network theory ([Lasszlo-Barabasi and Albert, 1999](#); [Barabasi, 2002](#); [Barabasi and Bonabeau, 2003](#)), but detailed analyses and applications to financial networks and systemic risk are less well developed ([Iyer and Puri, 2012](#); [Markose et al., 2012](#)). We contribute to and augment these streams of research by considering the role that regulation plays in shaping network topology and, in turn, predicting subsequent distress.

Our paper also contributes to the fast-moving, complementary literature evaluating financial networks in earlier periods, which includes research on the national banking era ([Anderson et al., 2019](#); [Calomiris and Carlson, 2017](#); [Brownlees et al., 2020](#); [Dupont, 2017](#)), the founding of the Federal Reserve ([Carlson and Wheelock, 2016, 2018](#); [Anderson et al., 2018](#)), and the 1930s ([Richardson, 2007](#); [Heitfield et al., 2017](#); [Mitchener and Richardson, 2019](#)). One strand of the literature explores the structure of the correspondent network, examining how its pyramid-shape emerged in the 19<sup>th</sup> century ([James and Weiman, 2010](#)), and

how it was altered by state and national legislation. For example, [Anderson et al. \(2019\)](#), analyze banks in Pennsylvania, arguing that the National Banking Acts changed the network structure of banks, leading to the creation of “systemically-important banks” in this state. [Anderson et al. \(2018\)](#), [Anderson et al. \(2020\)](#) and [Carlson and Wheelock \(2018\)](#) study how the establishment of the Federal Reserve influenced the network structure, using samples of New York banks, Virginia banks, and aggregate time series data, respectively. [Jaremski and Wheelock \(2020\)](#) document longer-run changes in the correspondent banking network by calculating measures of network centrality between 1900 and 1940, without incorporating balance sheet characteristics or risk. Our research takes this interest in understanding the pyramid-shaped correspondent network and relates it explicitly to systemic risk measurement so that we can investigate whether the shape or topology of the network, as measured on the eve of the Great Depression, mattered for systemic stability and relative to alternative network topologies.<sup>1</sup>

Our research also relates to the literature on banking distress of the 1930s (e.g., [Friedman and Schwartz \(1963\)](#) and [Wicker \(1996\)](#)) and the role of bank fundamentals in explaining subsequent distress ([Calomiris and Mason, 2003](#); [White, 1983](#)) during the Great Depression. Some recent studies in this area have further focused on the role of correspondents in propagating distress during panic periods. For example, [Heitfield et al. \(2017\)](#) build a hazard model to explore suspensions during the first banking panic of the Great Depression, the Caldwell Panic, and examine how that panic spread through correspondents located in Tennessee, Mississippi, and Alabama.<sup>2</sup> And, using an analysis at the Federal Reserve District level, [Mitchener and Richardson \(2019\)](#) assess how the pyramid-shaped network affected bank lending during the local and regional panics of the early 1930s. [Calomiris et al. \(2019\)](#) restrict their attention to national banks during the Depression (a small portion of the overall network and of banks that failed in the 1930s) to explore how correspondents of

---

<sup>1</sup>[Brownlees et al. \(2020\)](#) relates to our interest in systemic risk measurement using historical data; however, their approach (measuring the correlation of stock returns for New York clearinghouse banks) and focus (“backtesting” standard systemic risk measurement models, such as CoVAR and SRisk) differ from our study.

<sup>2</sup>This line of inquiry has also been pursued in early historical panics. For example, [Calomiris and Carlson \(2017\)](#) examine roughly 200 national banks in western and southern cities to assess the role of liquidity risk and bank closure; they find that keeping more assets with correspondents led to a higher rate of failure during the Panic of 1893.

national banks impacted their failure.<sup>3</sup> We break new ground by mapping the entire correspondent network for member and non-member commercial banks and employing a new methodology to construct network risk and balance-sheet risk measures for every commercial bank in the U.S. on the eve of the Depression. Doing so allows us to assess several previously unexplored issues concerning banking distress and the Great Depression: (1) we compute where bank-specific risk resided in the network prior to the Depression and measure how the pyramid-shape network contributed to ex ante systemic risk; (2) we show that a bank’s position in the network (as well as its solvency risk) mattered for its subsequent ability to survive the distress of the early 1930s; (3) we demonstrate that, in addition to its own balance sheet, banks faced risks from network neighbors during the banking distress of the early 1930s; and (4) we explore how feasible regulatory alternatives, such as widespread branch banking, appear to have had the potential to diffuse the concentration of risk inherent to the pyramid-shaped, network topology and thus potentially improve subsequent bank survivorship.

## 2 The Correspondent Banking Network

To construct network measures and to compute individual bank risk and systemwide risk, we build a data set with the linkages for all commercial banks and trusts in the United States in 1929, before the Great Depression begins. The precise links we observe are correspondent relationships, which are stated relationships between banks that are documented in bank directories. The most common and important reason a bank formed a correspondent relationship was to deposit funds into another bank, either for its own use or on behalf of customers. These interbank deposits were recorded on bank balance sheets as either “due to” or “due from,” representing the directional nature of the relationship. For example, when Bank A, “the respondent” (or bank that initiated the relationship), deposits funds into Bank B, its “correspondent” (the bank fulfilling that business need), Bank A records an asset on its balance sheet (“due from”) and bank B records a liability (“due to”).

---

<sup>3</sup>Because their selected sample is limited, [Calomiris et al. \(2019\)](#) focus only on the relationship between correspondent links and national bank distress in the 1930s, and are not able explore research questions that require the entire mapping of the correspondent network, such as how a bank’s position or its network neighbors’ position affected survival.

As described in the historical literature on banking, the correspondent network that existed in 1929, just prior to the Great Depression, had evolved over the nineteenth century – as the nation grew and population moved westward. Banks located in smaller towns and cities sought correspondent linkages with financial centers to carry out business on behalf of their customers as well as on their own account. It grew substantially in the 1860s, with the establishment of nationally-chartered banks and circulation of bank drafts as a national payments instrument ([James and Weiman, 2010](#); [Redenius, 2007](#); [James and Weiman, 2011](#)). Regulations contained in the national banking acts of the 1860s required “country banks” (those located in the hinterlands) with national bank charters to meet legal reserve requirements by keeping a portion of their reserves as cash in their vaults and the remainder (originally up to three-fifths) in correspondent banks in reserve or central reserve cities (larger cities dispersed throughout the country). State reserve requirements, which applied to banks that were chartered by state banking authorities, further reinforced the need for correspondent relationships. They required state-chartered banks to split their reserves between vault cash and interbank balances kept in banks located in large cities.

Given the incidence of banking panics in the second half of the 19<sup>th</sup> century ([Calomiris and Carlson, 2017](#); [Bordo and Wheelock, 2013](#); [Calomiris and Gorton, 1991](#); [Anderson et al., 2019](#)), drafters of the Federal Reserve Act of 1913 promoted the idea that the new central bank would reduce the likelihood of future panics by consolidating all interbank deposits in one of the system’s 12 regional banks (instead of having them scattered across hundreds of banks in numerous reserve cities), allowing a more efficient deployment of reserves when demand was high. However, by 1929, the first year we measure connections, only 10% of state-chartered commercial banks had joined the Federal Reserve System. Most small and medium-sized banks in the financial system opted not to join, and therefore satisfied their reserve requirements by keeping them in commercial banks located in larger cities. Even commercial banks that joined the Fed system deposited their excess reserves at correspondent banks in reserve and central reserve cities since commercial correspondent accounts paid a higher interest rate (typically 2%) than Federal Reserve Banks (0%). This was particularly true of reserve-city banks, which deposited their excess reserves in money-center banks in New York and Chicago. (See [Carlson and Wheelock \(2016\)](#), [Carlson and Wheelock \(2018\)](#),

Anderson et al. (2018), and Jaremski and Wheelock (2020) for additional discussion on how the founding of the Federal Reserve influenced the interbank market).

Despite the founding of the Fed, the reserve pyramid, which had its origins in 19<sup>th</sup>-century bank regulation, largely remained in tact in 1929, with the correspondent network’s structure supporting and maintaining it through the sizable interbank market (8% of total deposits). The failure of the Federal Reserve System to convince all state-chartered banks to switch their charters and join the system meant that, on the eve of the Depression, interbank deposits were concentrated in reserve and central reserve cities. As a result, and important for our network analysis, banks in these cities had the greatest number of respondents. That is, banks in New York and Chicago, the central reserve cities, were correspondents to the largest number of banks, often with respondents located in a large number of nearby and distant states. Banks in other large American cities, such as St. Louis, Philadelphia, and San Francisco, were correspondents to numerous banks, but usually established relationships that were more regional in nature, while banks in small towns and cities had the fewest relationships. In the next section of the paper, we discuss how this structure of interbank deposits fostered a pyramid-like network topology, as measured by correspondent relationships, and its implications for network stability.

### 3 Network Statistics Before the Banking Crisis

In order to measure systemic risk, we collected new data on correspondent relationships for the entire banking system in July 1929 (roughly 26,000 banks) – before the banking crises of the Great Depression began. Correspondent data, as well as information on each bank’s balance sheet and other characteristics (location, Federal Reserve membership, etc.), were hand-collected from the Rand McNally *Bankers Directory* (July 1929).<sup>4</sup> As noted in the introduction, an advantage of using network data on correspondent relationships is that they are stated, contractual relationships, rather than those inferred from data. As is the case with most empirical papers analyzing modern financial networks (e.g., Billio et al. (2012), Das et al. (2019)), we record whether a relationship between two banks existed –

---

<sup>4</sup>Appendix Section A.7 provides detailed information on our data sources and data-cleaning processes.

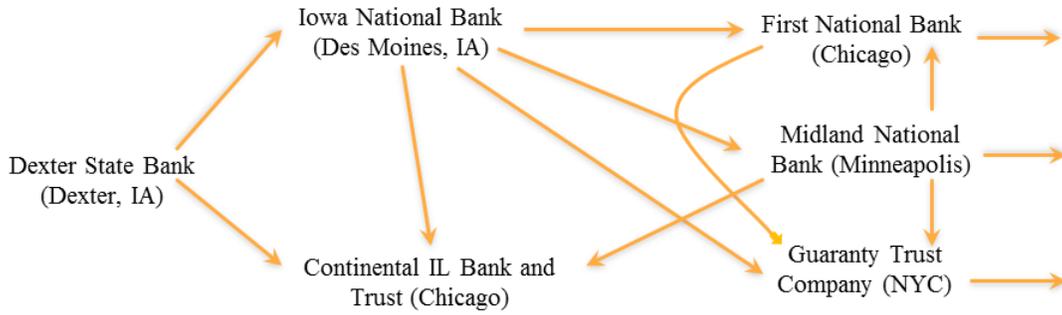


Figure 1: Illustrating correspondent relationships.

the extensive margin. Our historical sources do not provide systematic data on the intensive margin of a network relationship, such as the amount of money “due to” and “due from” for each respondent-correspondent relationship. This is not to deny that the intensive margin is also interesting. For example, in the case of our historical network, in addition to differences in flows of interbank deposits, the relative “importance” of correspondent relationships may have differed due to the quality or reliability of credit information or client services provided by particular correspondents, tempering how one interprets the measures of connectivity explored in this paper. Thus, like most studies exploring modern financial networks, a limitation of our data is that it forces us to treat the quality of connections as uniform.

Because Rand McNally precisely reports each correspondent relationship, we are able to code them directionally. We illustrate directionality in Figure 1. Dexter State Bank located in Dexter, Iowa, lists two banks as its correspondents in 1929: Iowa National Bank (Des Moines, IA) and Continental Illinois Bank and Trust (Chicago, IL). Iowa National Bank in turn lists First National Bank (Chicago, IL), Continental Illinois Bank and Trust, Midland National Bank (Minneapolis, MN), and Guaranty Trust Company (New York) as its correspondents. Figure 1 shows that even banks located in small towns connected directly to banks located in central reserve cities, thus concentrating network mass at the top of the pyramid.

We begin by describing the contours of the correspondent banking network before the banking panics of 1930-33 – a central feature of the Great Depression (Friedman and Schwartz, 1963; Wicker, 1996). We construct an adjacency matrix for the network for July 1929. The

matrix is defined as  $A(i, j)$ ,  $i, j = 1, 2, \dots, n$ , where  $n$  is the total number of banks or nodes in the system.  $A(i, j) = 1$  if respondent bank  $i$  has a relationship with correspondent bank  $j$ , else  $A(i, j) = 0$ . Since the correspondent-respondent relationship is directional,  $A$  is not symmetric,  $A(i, j) \neq A(j, i)$ . For the most part, the system in 1929 was one cluster, meaning you can get from one bank in the system to nearly any other bank through these interbank connections. We do not count self-connections, so we place zeros on the diagonal of the  $A$  matrix.

Table 1 describes the number of banks and connections in the network. In 1929, there were 25,684 banks (nodes) in the network and 70,679 interbank network connections. Figure 2 plots the locations of correspondent and respondent banks. If a bank is both a respondent and correspondent, it is plotted as a correspondent bank. The size of the dot displayed for correspondent banks is proportional to 1 plus the natural logarithm of the number of its connections. Country banks, the term used to describe banks located in small towns and cities, are close to their correspondents, which explains the widespread geographical coverage by correspondent banks located in larger cities.

Table 1: Number of banks and connections in the network.

Network	1929
Number of banks (nodes)	25,684
Correspondents	3,602
Network connections	70,679

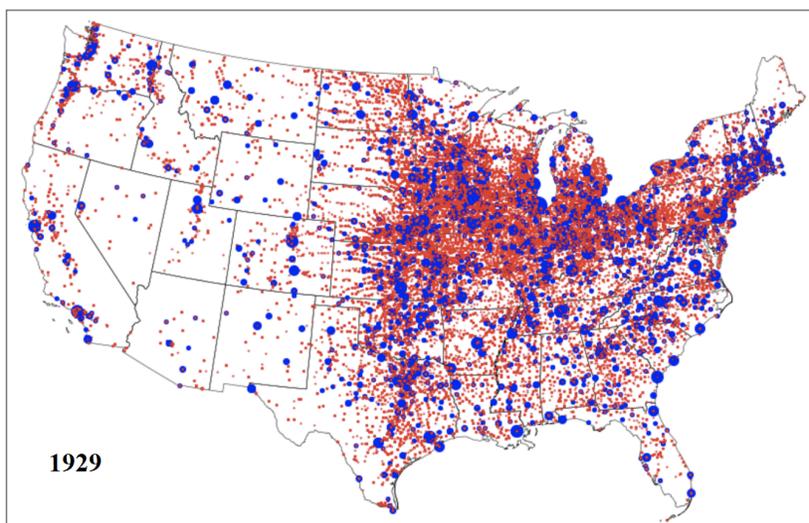


Figure 2: The location of all respondent banks (red) and correspondent banks (blue) sized by relative degree of each bank.

### 3.1 Degree Distribution

In examining the distribution of connections, most respondent banks have very few correspondents. On the other hand, correspondents had many respondents – sometimes in the thousands – a feature that has implications for the shape of the network. For 1929, the mean degree of nodes in the network is 5.5, with a standard deviation of 49.6.<sup>5</sup> The “degree distribution” of the nodes is shown in Figure 3. The upper left plot show nodes with degree less than 50 while the right plot show nodes with degree greater than 50. The distribution is extremely skewed: there are a large number of nodes with low degree and a few nodes with very high degree. This is characteristic of a scale-free network (Barabasi and Bonabeau, 2003).

We can also examine the degree distribution by plotting the (log) number of nodes against the (log) number of links (degree). The lower left panel of Figure 3 shows a log-log plot of the network in 1929, where extreme nodes with degree greater than 50 have been excluded. The quasi-linear and negative relationship of the log-log plot and the power law coefficient

---

<sup>5</sup>Note that the number of links per node is the number of links divided by the number of nodes, i.e.,  $70679/25684 = 2.75$ . Since each link connects two nodes, it accounts for two degrees. Therefore, the mean degree of nodes in the network is twice 2.75 or 5.5.

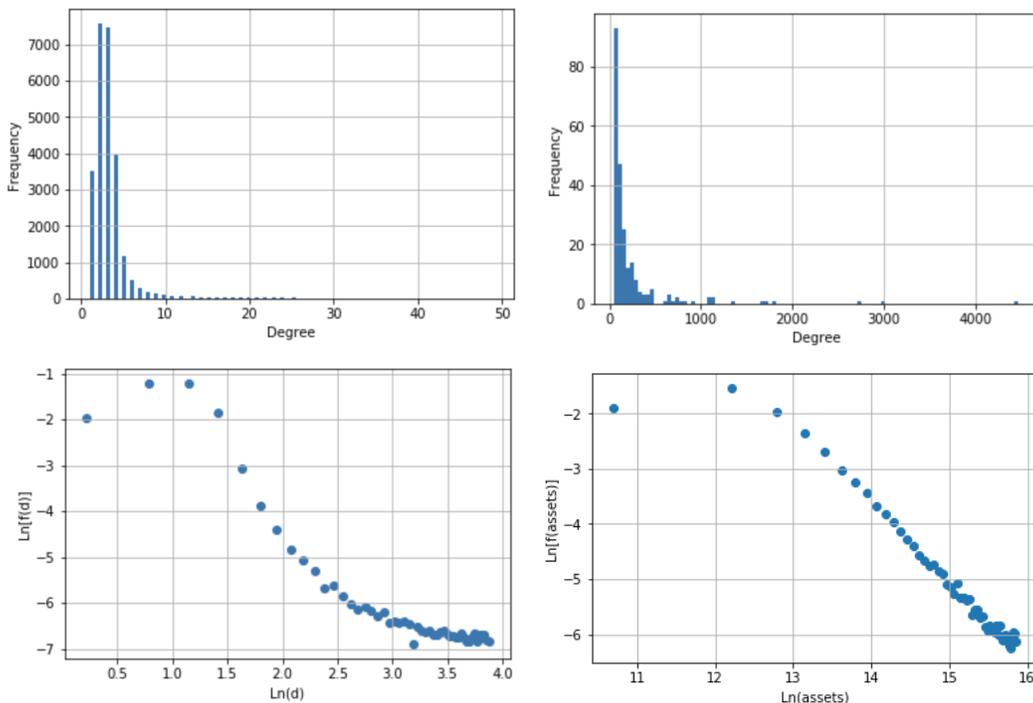


Figure 3: Top Panels: Degree distribution in 1929. The left panel plots nodes with degree less than 50 while the right panel plots nodes with degree greater than 50. Bottom Panels: Network Density and Assets. Log-log plot of the degree distribution in 1929 excluding extreme nodes with degree greater than 50. We plot log degree on the x-axis and log of the density function on the y-axis. For power-law densities, the function is  $f(d) = d^{-\alpha}$ . Taking logs we have  $\ln[f(d)] = -\alpha \ln(d)$ . The power law coefficient is  $\alpha = 1.65$ . The second plot shows the same for bank asset levels, with a power law coefficient of  $\alpha = 1.17$ .

( $\alpha = 1.65$ ) suggest bank nodes exhibit the usual shape, slope, and power law distribution characteristic of social networks (Barabasi, 2002; Barabasi and Bonabeau, 2003; Gabaix et al., 2003). Further, we compare the degree distribution to the asset size distribution of banks in 1929. The bottom right panel of Figure 3 shows a similar shape and slope as the degree distribution; however, the power law coefficient is lower ( $\alpha = 1.17$ ). To a large extent, the most well-connected banks were also the largest banks, fitting the modern policymaker’s definition of “systemically important.” However, bank size (assets) is not a proxy for interconnectedness (degree). We compared the distribution of the normalized log assets with that of normalized log degree using the Kolmogorov-Smirnov statistic, which shows a highly significant difference between the two distributions (p-value=0).

## 3.2 Centrality

Understanding which nodes have the greatest influence in the network, i.e., those that are the most critical, is of particular interest to policymakers who might want to identify “too big to fail” or systemically important financial institutions (SIFIs). To understand this feature of our network, we use two variants of the most common measure, centrality: degree and eigenvector centrality. We review these measures in depth in Appendix Section A.1.

Centrality can be measured by degree centrality or the number of connections a node has. However, degree centrality ignores the fact that a node with few connections may still have a huge influence if it is connected to a node with many connections. Eigenvector centrality is a more general formulation that takes this into account. This measure defines the centrality of a node as a function of the centrality of the nodes to which it is connected through the network adjacency matrix,  $A$  (Bonacich, 1987; Bonacich and Lloyd, 2001).

The banks with the highest eigenvector centrality in 1929 are shown in Table 2. In 1929, Continental Illinois Bank and Trust (Chicago) had the most linkages in the network at 4,474. About 17% of all banks connected to Continental Illinois in 1929. The most connected banks are located in New York City and Chicago, the central reserve cities. Banks in reserve cities like St. Louis, Philadelphia, and Cleveland are the next most connected, giving rise to the network’s pyramid topology. We review how these measures appear in different topological structures in the next section.

Some banks listed in Table 2 are notably the largest banks in the system. While there is alignment between the most connected and the biggest banks, the same cannot be said for the rest of the banks in the system, where connectivity and asset size differ substantially. On average, a bank’s percentile in the asset distribution is 24 percentage points different from its percentile in the degree distribution. This observed difference suggests that, across the entire distribution, size and connectivity are quite distinct.

Table 2: Top 10 banks by eigenvector centrality (normalized), along with their degree.

1929: Bank Name and Location	EigenCent	Degree
Continental Illinois Bank and Trust (Chicago, IL)	1.000	4,474
Chase National Bank (New York City, NY)	0.440	2,982
Central Hanover Bank and Trust (New York City, NY)	0.283	2,710
First National Bank of Chicago (Chicago, IL)	0.279	1,715
National City Bank (New York City, NY)	0.265	1,778
Guaranty Trust Company (New York City, NY)	0.226	1,652
First National Bank in St. Louis (St. Louis, MO)	0.181	1,113
Philadelphia National Bank (Philadelphia, PA)	0.181	1,127
Mercantile Commerce Bank and Trust (St. Louis, MO)	0.148	813
Union Trust Company (Cleveland, OH)	0.146	629

### 3.3 The Pyramid Structure of the Banking Network

Figure 4 displays a city-level network graph. It shows a very high density of connections in the central reserve cities of New York and Chicago as well as the medium-level of density in reserve cities (e.g., St. Louis, Kansas City, Philadelphia, Minneapolis, etc) – providing visual confirmation of the existence of the pyramid topology of the network in 1929.

To understand how topology and stability are related to each other, we conduct several experiments that compare our 1929 pyramid network (which is scale-free) to alternative scale-free and random graph topologies. Pyramid graphs are mostly tree-like even though they are not directed acyclic graphs (DAGs). Topologically, pyramid graphs are within the class of scale-free networks and are likely to have connections concentrated in a few nodes (Lasszlo-Barabasi and Albert, 1999; Barabasi and Bonabeau, 2003; Barabasi, 2002). In graph theory, the difference between these networks is determined by an examination of the degree distribution of the networks. Random networks are characterized by Gaussian-shaped degree distributions. Scale-free networks have degree distributions that follow a power law asymptotically and visually. The log-log plot (as in Figure 3) is linear with a negative slope equal to the power coefficient (see Lasszlo-Barabasi and Albert (1999)). Random networks are included in this study because they serve as a baseline control to which other networks can be compared; they are commonly included in studies of network topology. (See for example, Newman et al. (2002), Wang et al. (2014), Wijk et al. (2010), and Béliveau et al.

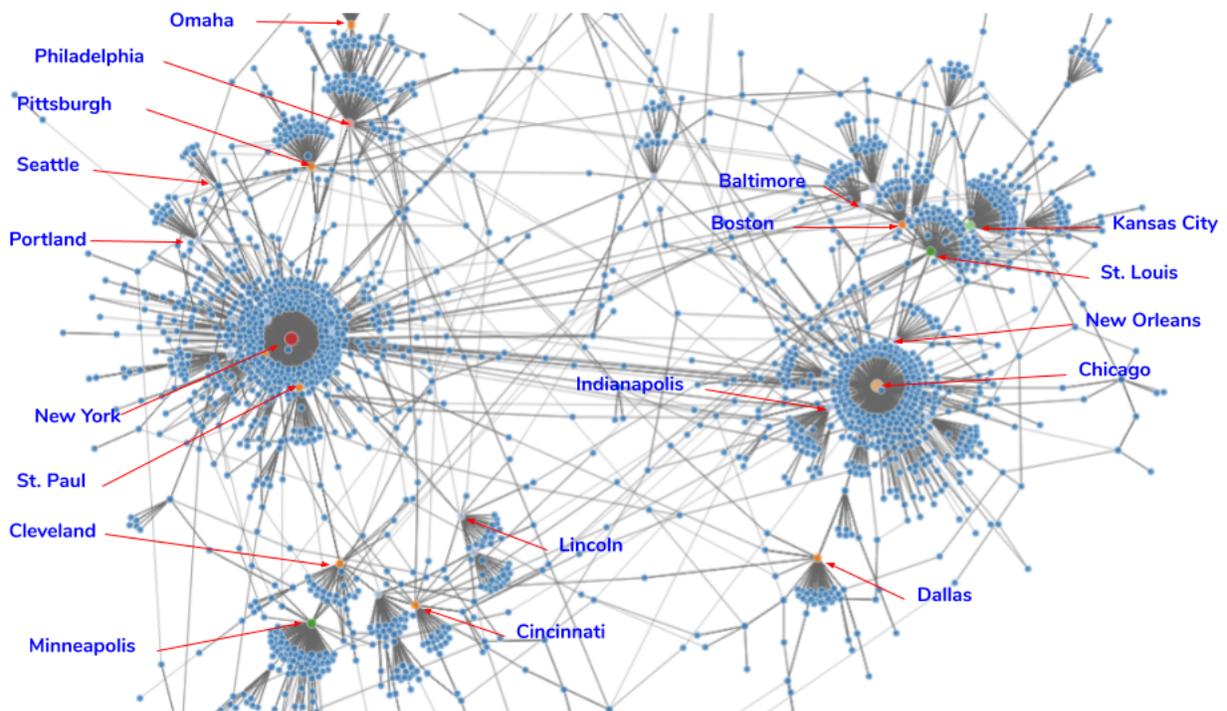


Figure 4: The network graph shown at the city-level. We only show nodes with at least 5 connections in order to make the graph easy on the eye. We have highlighted some of the major cities.

(2017).)

To understand whether the pre-existing topology of the correspondent network might have mattered for systemic stability, we use the `networkx` Python package to compare the diameter and fragility of the 1929 pyramid network to (i) a random graph network and to (ii) a simulated scale-free network, both with approximately the same number of nodes and mean degree. This ensures that the dimension of the networks is not a reason for differences in the comparison of the pyramid network of 1929 and the alternate networks, thus enabling us to isolate the network structure as the driver of differences in the simulations. For both (i) and (ii), we simulated 100 counterfactual networks, where the mean degree is 5.5 and the number of nodes is 25,684. We then average diameter and fragility results over these many realizations. These simulated networks also form the basis for further econometric analyses in Section 5, where we explore how these counterfactual networks impact the probability of bank survival.

“Diameter” is defined as the maximal shortest path between any pair of nodes in the cluster for a non-directional network. It is the max-min measure over all paths between all pairs of nodes. In essence, it measures how many banks it takes for a financial flow to spread from one edge of the network to the other, thus giving us some insight into how quickly counterparty “contagion” could occur. The bigger the diameter, the less likely it is that a local financial shock will become a global event in a banking system. We define the concept of “system fragility” as the mean squared degree sum (akin to the coefficient of variation, and if normalized, then it is the Herfindahl index) normalized by dividing by mean degree:  $F = E(d^2)/E(d)$ , where  $d$  is degree of each node. Fragility takes concentration more explicitly into consideration than diameter because a highly-concentrated network tends to have a greater risk of transmission. Hence, if a highly centralized node is compromised, the propensity for risk to spread to the other nodes is high.

Table 3 presents the results of the 1929 network and the counterfactual networks. The diameters are roughly comparable. In 1929, a bank in the system could get to any other bank in the system in 13 steps. Fragility, on the other hand, is dramatically higher in the pyramid-shaped network of 1929. We would expect a large difference between the pyramid and the

Table 3: The diameter and fragility of the networks.

Network	Diameter	Fragility	Eigenvector Centrality	
			total	maximum
1929 Pyramid	13	453	77	0.53
Random Graph	12	6.5	128	0.02
Scale-free (Barabasi type)	14	180	50	0.52

Eigenvector centrality is not normalized in this table to offer comparisons across the different network topologies.

random graph, making the scale-free network a more apt choice for comparison. However, even here, the fragility score for the observed 1929 network is much higher, suggesting that connections were even more concentrated than one would expect in a scale-free setting, implying that the 1929 system had a very high propensity for risk to spread. As discussed in Section 2, this shape evolved in response to 19<sup>th</sup> century regulation, which is why one can think of the existing network on the eve of the Depression, even if it was unintended consequence, as “fragile by design” (as coined by [Calomiris and Haber \(2014\)](#)).<sup>6</sup>

Table 3 also presents the total and maximum eigenvector centrality in each of the networks. We note that total centrality in a random graph network is substantially higher than for a scale-free and pyramid network, but the maximum value in the network is much smaller. In a random graph setting, an average node in the network is more central and important, whereas in the scale-free and pyramid settings, eigenvalue centrality is much more concentrated at the top nodes, leaving the average node with minimal centrality. Because eigenvector centrality captures important differences in network topology, we further explore it in econometric counterfactual exercises in Section 5.

Table 3 presents a macro perspective of overall network features to demonstrate how topology can matter for financial stability. In the pyramid structure, when just a few banks faced distress, the correspondent network could efficiently transfer funds from the center to places where banks were being run and satisfy the increase in demand for liquid funds.

---

<sup>6</sup>We note that the network that existed prior to the GFC was also scale-free. [Burdick et al. \(2011\)](#) document that the fragility of the network before the GFC was 172, which is similar to the value of our alternative scale-free network. However, these networks are difficult to compare given that the number of banks is much smaller in the GFC than in the Great Depression and the construction of networks is quite different.

Table 4: Summary statistics of our eigenvector centrality measure for subgroups of our sample.

Sample	Eigenvector Centrality		
	Mean	St. Dev.	Max
No Branching	0.006	0.011	1.000
Branching	0.005	0.008	0.440
Central Reserve City Banks	0.018	0.065	1.000
Reserve City Banks	0.012	0.017	0.181
Country Banks	0.005	0.005	0.052

A t-test confirms that the means for the different subgroups are statistically different from one another and that no branching, central reserve cities, and reserve cities are larger than their counterparts. “Country Banks” refer to those not located in reserve or central reserve cities.

However, the system could become overwhelmed if, during a panic, the demand for liquidity surged (Mitchener and Richardson, 2019) and positionally-important banks faced withdrawals from multiple respondents simultaneously. The pyramid topology was fragile as a result of policy and regulation. In addition to the regulation discussed in Section 2, involving the national banking era and founding of the Federal Reserve, laws prohibiting branch banking also may have played a role in reinforcing and buttressing the pyramid structure of the correspondent network. Interstate branch banking was forbidden in the United States, national banks were severely restricted from opening branches, and state-chartered banks were only allowed to operate branches if their state banking codes authorized them to do so. Many of the states with outright prohibitions on branch banking (23 altogether) were located in the Midwest and Plains states. As a result, small banks flourished in the U.S., with many effectively operating as local monopolies (White, 1983). Restrictions on branch banking meant that many small and medium-sized banks relied on correspondent banks in larger cities to provide services to customers and to meet state-mandated reserve requirements than would have otherwise been needed under a system of widespread branch banking.

Table 4 presents summary statistics of the eigenvector centrality variable for subgroups of our sample. The table demonstrates how regulatory environments are linked to the network’s topology. For example, the pyramid shape of the network was driven in part by state and federal regulations (Section 2), which permitted country banks to meet reserve requirements

by keeping them in banks located in larger cities. The higher values of eigenvector centrality for banks located in reserve and central-reserve cities noted in the table are a reflection of these regulations. Importantly, the statistics on branch banking offer a new observation: states that completely prohibited branching had higher average centrality scores than states that allowed some or full branching within the state. We return to this issue in Section 5, where we explore interactions between network topology and branch banking regulation using variation in branch-banking laws at the state level. Before we do so, we further develop a micro-perspective, creating bank-level measures that incorporate individual balance sheets, and further explore the extent to which network position and peer effects in the network matter for financial stability.

## 4 Quantifying Systemic Risk

### 4.1 Deriving Composite Risk Scores

In order to construct a measure of systemic “risk” rather than systemic “importance” (defined on only size and/or connectedness), it is necessary to quantify the credit quality of all commercial banks. Since credit ratings were non-existent for the tens of thousands of banks that were too small to be listed on the NYSE or regional exchanges, we employ financial ratios based on our balance sheet data available for all commercial banks in 1929. On the liabilities side, Rand McNally lists four items: paid-up capital, surplus and profits, deposits, and other liabilities. On the asset side of the ledger, Rand McNally also records four categories: loans and discounts, bonds and securities, miscellaneous assets, and cash and exchanges (due from banks).

We calculate a composite measure of credit risk ( $R$ ), computed as the product of inverse profitability ( $C$ ) and transformed leverage ( $L$ ). Inverse profitability is defined as:

$$C = a + \frac{1}{1 + BUF} \cdot b. \quad (1)$$

The key ratio here is  $BUF$ , which is a buffer stock of retained earnings. Specifically, this is surplus and profits divided by equity (the sum of surplus and profits and paid-up capital).  $BUF$  is close to a bank’s return on equity, but not identical to it, because Rand McNally

data does not report dividends paid out. However, economic historians have argued that  $BUF$  was a key ratio used by banks to expand operations or write off losses relative to the book value of net worth in the 1920s and 1930s (Carlson and Rose, 2015). One might claim that retained earnings and equity are the same, but in a system of book value accounting, retained earnings are a good proxy for profitability. Because  $BUF$  is bounded between 0 and 1, we set the scalars  $a = -8$  and  $b = 18$  so that  $C$  is then confined to a value between 1 and 10, and may be thought of as a mapping into a rating, with values inversely proportional to quality. The range of values is without loss of generality.

$C$  alone is an imperfect measure of credit risk. A low-equity bank could get a high  $BUF$  ratio, resulting in a lower value of  $C$  when the bank is actually quite risky because it operates at extreme leverage. To adjust for this, we multiply  $C$  by a transformed measure of leverage  $L$ :

$$L = \ln(1 + Assets/Equity). \tag{2}$$

$Assets$  is the sum of loans and discounts, miscellaneous assets, bonds and securities, and cash and exchanges (due from banks). In the above low-equity scenario, the low-equity bank would have high leverage, and thus would receive the appropriate higher score in our composite risk score (where an increasing score implies greater risk):

$$R = C \times L. \tag{3}$$

How do we interpret the units for  $R$ ? We note that  $R$  is analogous to credit risk in the Merton (1974) model, which is the inverse of distance-to-default (DTD), i.e.,  $R \propto \frac{1}{DTD} \approx \frac{\sigma Assets}{Assets-Debt} \approx \frac{\sigma Assets}{Equity}$ , where  $\sigma$  is volatility, i.e., business risk. In the Merton model, the reciprocal of DTD is a measure of credit risk: the higher this measure is, the more credit risk a firm has. Moody's KMV converts this measure into an expected default frequency (EDF), a variant of default probability, using an empirical mapping of DTD to defaults (see page 3, equation 6, and Section 3 of the Moody's white paper, Nazeran and Dwyer (2015)). It is important to note that the risk measure used by Moody's is simply: Business Risk  $\times$  Leverage =  $\sigma \times Assets/Equity$ , shown above. The risk measure we employ,  $R$ , is analogous, where business risk is proxied by an inverse measure of profitability,  $C$ , and leverage is proxied by  $L$ . Since  $R$  could also be used in the same way Moody's uses its measure (i.e., to

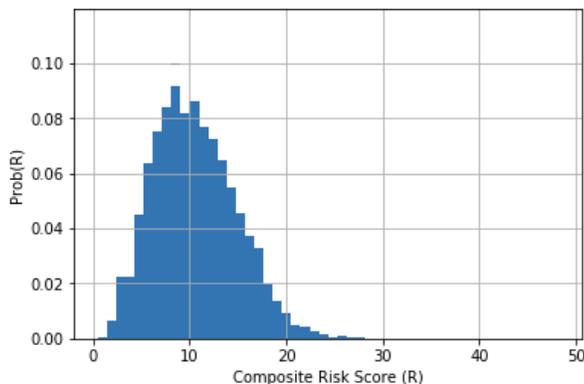


Figure 5: Histogram of  $R$  in 1929.

support an empirical mapping to expected default frequency), the units for  $R$  are the same as the reciprocal of DTD. That is,  $R$  may be thought of as a probability of default up to a scaling constant, or more generally, as a credit risk score.

We use  $R$  since there are no measures of asset volatility available for all banks in 1929. Further, because our sample includes all banks in the U.S., we are confined to working with the balance-sheet categories Rand McNally lists for assets and liabilities. In Appendix A.2, we compare  $R$  to several other ratios and measures of risk and show that  $R$  outperforms the other measures at predicting default. We also compute  $R$  for a sample of modern firms and demonstrate that  $R$  aligns with modern credit risk ratings. Thus, we use  $R$  as our main measure of internal credit risk. Figure 5 presents a histogram of  $R$  in 1929. Transformed leverage  $L$  has a mean value of 2.1, and most values are below 5. Given that the maximum value of  $C$  is 10, we expect a range for risk scores  $R$  that are well below 50. Indeed, most of our risk scores lie below 30. These ranges may be adjusted without loss of generality. The distribution is fairly normal, with a slight right skew. The transformations to  $L$  and the scalars employed in  $C$  have the net effect of de-emphasizing the outliers in the distribution.

Table 5 shows the average values of  $R$  for 1929. It is interesting to note, especially in comparison to the GFC, that the most connected banks have risk scores well below the mean. Prior to the banking crises of 1930-3, much of the balance-sheet risk within the system thus resided in medium- and smaller-sized banks, many of which were located in the Midwest and Plains states. Given that the top of the network appears stable on the eve

of the crisis, we next examine network risk by incorporating our network adjacency matrix with our composite risk measure.

Table 5: Composite credit risk averaged over degree subsamples in 1929.

Sample	$R$ in 1929
Full Sample	10.52
Banks with degree less than mean	10.56
Banks with degree greater than mean	9.97
Top 10 banks by degree	7.45

## 4.2 Systemic Risk Per Bank and Network-Neighbor Measures

To create a single measure of overall systemic risk, we combine the network adjacency matrix,  $A$ , with the composite risk score for each bank in the vector  $R$ , extending and modifying the metrics proposed in Das (2016) and Das et al. (2019). Systemic risk per bank,  $S$ , is defined as:

$$\begin{aligned}
 S &= \frac{1}{n} \cdot \sqrt{R^\top \cdot A \cdot R} \\
 &= \sqrt{\frac{R^\top}{n} \cdot A \cdot \frac{R}{n}} \\
 &= \sqrt{Q^\top \cdot A \cdot Q}
 \end{aligned} \tag{4}$$

where  $n$  is the number of banks in the system and superscript  $\top$  denotes the transpose of a vector or matrix. (Recall that  $R$  is an  $n$ -vector and  $A$  is a  $n \times n$  matrix. Thus,  $Q$  is an  $n$ -vector.) Division by  $n$  is a normalization used to measure systemic risk per financial institution. The systemic-risk measure,  $S$ , may be thought of as a network-weighted measure of composite credit risk in the banking system. Therefore, it has the same units as  $R$ . Appendix Section A.1 presents theory and examples of our systemic-risk measure.

$S$  is an appropriate measure for our historical setting because: (1) it captures both connectivity risk and credit risk; (2) it properly accounts for the directions of the correspondent-respondent relationships because  $A$  is not symmetric; (3) the functional form of  $S$  is linear and homogeneous which allows us to decompose the measure bank-by-bank and compute

individual bank contributions; and (4) it has attractive properties for prediction. Because  $S$  is constructed from raw data matrices, as opposed to being the result of regression output, it can be easily employed for prediction exercises without issues of estimation uncertainty.

To understand exactly where risk emanated from before the banking distress of the Great Depression, we can decompose our systemic-risk measure into the risk contribution of each bank. This decomposition of the scalar function  $S(R, A)$  is possible because the function is linear homogeneous in vector  $Q = [Q_1, Q_2, \dots, Q_n]^\top$ , the normalized value of  $R$ . We can apply Euler’s homogenous function theorem and obtain the *risk decomposition equation*:

$$S = \frac{\partial S}{\partial Q_1} \cdot Q_1 + \frac{\partial S}{\partial Q_2} \cdot Q_2 + \dots + \frac{\partial S}{\partial Q_n} \cdot Q_n. \tag{5}$$

Each partial derivative  $S_j = \frac{\partial S}{\partial Q_j}$  multiplied by  $Q_j$  is the risk contribution of bank  $j$ . We can calculate all derivatives  $S_j$  in closed form using the following vector derivative:

$$\frac{\partial S}{\partial Q} = \frac{1}{2S} [A \cdot Q + A^\top \cdot Q] \in \mathcal{R}^n \tag{6}$$

which gives an  $n$ -vector of derivatives  $S_j$ . Once we know the amount of risk that is contributed by each node, we can pinpoint the riskiest banks in the network in terms of their contribution to overall systemic risk.

With our fully identified adjacency matrix  $A$  and vector  $R$ , we can also explore “peer effects” stemming from neighbors in the network, and how a bank’s behavior may be related to the behavior of its network connections. In order to examine how “network neighbors” impact the risk of any node in the network, we create additional measures based on whether the adjacency matrix indicates an “in-link” relationship or an “out-link” relationship, where in-link refers to respondent connections and out-link refers to correspondent connections. For composite credit risk, the out-link relationship would be simply computed as  $A \cdot R \in \mathcal{R}^n$  (out-R). Correspondingly, the in-link relationship is given by the vector  $A^\top \cdot R \in \mathcal{R}^n$  (in-R). Summing these variables measures the total amount of balance-sheet risk emanating from a bank’s network neighbors, thus capturing risk coming from peers. These peer effects are a potentially important source of systemic risk that has yet to be explored by researchers.

Table 6 shows the percentage of systemic risk contributed by each of the top 20 risk-contributing banks as well as the total amount of network-neighbor risk for these banks (i.e.,

Table 6: Top 20 banks by percentage contribution to systemic risk in 1929

Bank Name and Location	SysRisk Percent (S%)	Total Network Neighbor R
Continental Illinois Bank and Trust (Chicago, IL)	3.23	50016
Chase National Bank (New York City NY)	1.56	29494
National Bank of the Republic of Chicago (Chicago, IL)	1.21	9390
First National Bank of Chicago (Chicago, IL)	1.19	18947
Commerce Trust Company (Kansas City, MO)	1.05	10401
National City Bank (New York City, NY)	1.04	17307
First National Bank (Minneapolis, MN)	0.97	14062
First National Bank in St. Louis (St. Louis, MO)	0.85	11760
Central Hanover Bank & Trust Company (New York City, NY)	0.76	27011
Guaranty Trust Company (New York City, NY)	0.74	15291
First National Bank of St. Paul (St. Paul, MN)	0.63	9382
First Wisconsin National Bank (Milwaukee, WI)	0.62	8552
National Stock Yards National Bank (National Stock Yards, IL)	0.61	4902
National Park Bank (New York City, NY)	0.60	13402
Mercantile Commerce Bank & Trust Company (St. Louis, MO)	0.59	8408
Fletcher American National Bank (Indianapolis, IN)	0.46	4409
Northwestern National Bank (Minneapolis, MN)	0.45	5162
Union Trust Company (Cleveland, OH)	0.45	6031
Fidelity National Bank and Trust Company (Kansas City, MO)	0.42	3793
Drovers National Bank (Chicago, IL)	0.41	5096

the sum of in-R and out-R). Comparing it to Table 2, we can see there is a considerable correspondence between the banks exhibiting the most systemic risk in 1929 and the banks showing the greatest centrality. However, an important additional feature of the 1929 network is that systemic risk is dispersed across the entire network. Taken together, the top 10 banks account for 12.6% of the total systemic risk and the top 20 banks account for 17.8%.

Further, Table 6 demonstrates the massive concentration of balance-sheet risk at top nodes in the network. On the eve of the Great Depression, Continental Illinois appeared healthy with a composite risk score of 9.16, well-below the average in the system. However, in 1933, the Reconstruction Finance Corporation (government-sponsored bank rescue program) became the largest stockholder in the bank (DBG, 1933). This outcome may be perplexing based on a simple examination of Continental’s balance sheet. However, once we map the entire network and compute network-neighbor risk, the reasoning becomes clear: 20% of the total balance-sheet risk in the system was linked to Continental through network neighbors, making the system ripe for network risk. Unlike the GFC, balance-sheet risk emanated

Table 7: Top 5 banks by percentage contribution to systemic risk under counterfactual topologies.

Banks	Random Network SysRisk Percent (S%)	Random Network Neighbor R	Scale-Free Network (Barabasi type) SysRisk Percent (S%)	Scale-Free Network (Barabasi type) Neighbor R
1	0.018	74	0.251	9994
2	0.018	87	0.205	6564
3	0.015	59	0.196	7568
4	0.014	74	0.155	3786
5	0.014	64	0.141	3506

from banks at the lower end of the size distribution, but made its way to the top through network connections. To see if balance-sheet risk is related across network neighbors, we calculated the correlation between a bank’s own balance-sheet risk and the average balance-sheet risk of its network neighbors. The correlation coefficient of 0.17 suggests a weak, positive relationship. Thus, network peers were not identical, unlike spatial peers which may have had correlated asset holdings.<sup>7</sup>

As a robustness check, we also recomputed the risk score in equation (3) using size weights based on bank assets. That is, we redefined  $R = C \times L \times w$ , where  $w$  is an asset-based weight, such that if the bank was in the top size decile by assets, then  $w = 3$ , and if it was between the 40<sup>th</sup> and 90<sup>th</sup> percentile, it was given a weight of  $w = 2$ . All banks below the 40<sup>th</sup> size percentile had weight  $w = 1$ . We recomputed the results in Table 6 and noticed no material changes. Therefore, the systemic-risk measure is robust to size-weighting of bank risk.

In Table 7, we examine whether systemic-risk and network-neighbor risk differ under the alternative topologies introduced in Section 3.3 and Table 3. The vector  $R$  remains the same in these calculations; only the adjacency matrix is changed to represent a different network of connections. Table 7 presents the top 5 banks by their systemic-risk contribution. Notably, the amount of risk residing at the top nodes in the network is substantially lower than in the observed 1929 pyramid network. Even relative to an alternative scale-free topology, network-neighbor and systemic risk are 5 and 13 times higher in the 1929 pyramid, respectively. In the next section, we explore the consequences of these differences in the topological concentration

<sup>7</sup>The correlation between a bank’s own balance-sheet risk and the average balance-sheet risk of banks in the same county is 0.63.

Table 8: Summary statistics of our network neighbor  $R$  measure for subgroups of our sample.

Sample	Network Neighbor $R$		
	Mean	St. Dev.	Max
No Branching	58.16	562.33	50016.09
Branching	52.45	493.12	29494.06
Central Reserve City Banks	740.38	3991.54	50016.9
Reserve City Banks	323.17	1013.93	14062.04
Country Banks	31.34	59.26	4902.54

A t-test confirms that the means for banks in central reserve and reserve cities are larger than the mean for country banks. The t-test fails to reject the null of a difference between the branching subgroups.

of risk for bank survival.

As shown in Table 4, the existing topology varied across certain regions of the United States, where eigenvector centrality was higher in areas that prohibited branching and in reserve and central reserve cities. In a similar vein, Table 8 shows the average of network neighbor  $R$  for subgroups of banks that correspond to different regulatory environments of the banking system. The table shows that network-neighbor, balance-sheet risk was concentrated in areas that prohibited branching and in reserve and central reserve cities. Again, this demonstrates a link between network characteristics and financial regulation that existed on the eve of the Depression, and suggests that the correspondent network may have had an oversized importance in some areas.

## 5 Network Features and Bank Survivorship

### 5.1 Probit Models and Covariate Construction

In this section, we match the bank-level measures of systemic risk and network position to micro-level information on commercial bank characteristics and bank survival in order to analyze several important and related questions. First, to what extent do the data support balance sheet, network position, and network neighbor measures in predicting whether banks survived the worst American banking crisis of the 20<sup>th</sup> century? Second, how large are the biases if network and network-neighbor measures are ignored? And third, what are the rel-

ative contributions of balance-sheet risk, network topology, and branch-banking regulations to a bank’s predicted probability of survival and what are the associated spatial patterns?

We use the sample of all commercial banks operating in 1929 to predict bank survivorship in 1934. The outcome variable,  $y_i$ , equals 1 if bank  $i$  survives the banking distress of the early 1930s and appears as a solvent financial institution in July 1934. Otherwise, it is set equal to 0. Banks are matched between 1929 and 1934 using their name, location, and routing number (i.e., a unique bank identification number). Data on bank status in 1934 are from Rand McNally (1934), from which we tabulate that 62% of commercial banks survived the distress of 1930-33 and the remainder of banks exited. Exits could be due to failure, voluntary liquidations, or mergers. (Our data on the full sample of commercial banks does not allow us to differentiate the reasons for bank exit, but in Appendix Section A.5, we consider specifications that take these differences into account using the sample of national banks.)

We consider several probit models, where in latent-variable form:

$$M_1 : y_i^* = x_{i1}^\top \beta_1 + \varepsilon_i \tag{7}$$

$$M_2 : y_i^* = x_{i1}^\top \beta_1 + x_{i2}^\top \beta_2 + \varepsilon_i \tag{8}$$

$$M_3 : y_i^* = x_{i1}^\top \beta_1 + x_{i2}^\top \beta_2 + x_{i3}^\top \beta_3 + \varepsilon_i. \tag{9}$$

The mapping from the latent  $y_i^*$  to the observed  $y_i$  is

$$y_i = 1\{y_i^* > 0\} \text{ for banks } i = 1, \dots, n, \tag{10}$$

and  $\varepsilon_i \sim N(0, 1)$ . The models differ by their covariates, where  $M_1$  includes standard balance sheet, location, and market information ( $x_{i1}$ ) that are common to bank-level models in the Great Depression literature (e.g., Calomiris and Mason (2003)).<sup>8</sup>  $M_2$  contains additional risk measures, including risk from network neighbors and spatial neighbors ( $x_{i2}$ ).  $M_3$  augments  $M_2$  with bank-specific network measures and interactions with regulation ( $x_{i3}$ ). We conduct

---

<sup>8</sup>More detailed balance sheet information is available for the subset of Federal Reserve member banks, however, there is a clear tradeoff when it comes to modeling systemic risk: 73% of the exits from the system were nonmembers. The results for the entire commercial banking system seem more relevant since our goal is to test whether systemic risk measures improve model fit rather than, for example, trying to present causal estimates of bank failures.

a Bayesian model comparison analysis to examine whether a financial system’s pre-existing network topology matters for bank survival (in our case the network that existed in July 1929) and to bring attention to the potential biases that may occur from the omission of various network effects.

Table 9 provides detailed descriptions of the covariates included in the three models, including measures capturing geographical differences in the banking industry and economic structure as well as bank regulation, much of which varied at the state level (Mitchener, 2005, 2007). For example,  $M_1$  includes measures for the size of the balance sheet, balance sheet ratios, the number of banks in the county, the share of deposits held at that bank as a ratio of total deposits in the county, Federal Reserve membership, indicators for whether the bank is located in a reserve city or central reserve city, county population, manufacturing establishments and acreage of cropland in the county, and Federal Reserve district indicators.  $M_1$  also includes a measure for the intensity of branch banking in the bank’s location (Carlson, 2004; Mitchener, 2005; Carlson and Mitchener, 2006). Following Carlson and Mitchener (2006), the branch banking intensity variable is defined as the share of banks operating branches in the state.<sup>9</sup>

$M_2$  includes the vector,  $x_{i2}$ , which adds covariates for a bank’s composite risk score ( $R$ ), the composite risk from spatial neighbors, and the composite risk from network neighbors, as discussed in Section 4.2.<sup>10</sup> Specifically, in-R ( $A^\top \cdot R$ ) and out-R ( $A \cdot R$ ) capture the total amount of composite credit risk held by a bank’s respondents and correspondents, respectively. Spatial risk is computed by summing the total amount of composite risk held by banks in the same county. Researchers have shown that local panics and bank runs featured prominently in the banking distress of the early 1930s (Wicker, 1980, 1996; Mitchener and Richardson, 2019, 2020), suggesting at least the possibility that some ex ante risk in 1929 might be spatial. One reason spatial risk may be important to control for is common asset-exposure, such as concentrated lending by banks to a particular industry or sector. For

---

<sup>9</sup>We define our branch banking measure at the state level since these regulations were implemented by state banking authorities and because existing research shows these regulations affected systemic stability at that level (Carlson and Mitchener, 2006, 2009; Mitchener, 2005).

<sup>10</sup>Our composite risk score encompasses leverage and retained earnings. We do not include these variables separately because Appendix Table A.2.1 shows that our measure dominates other commonly used measures. We do, however, include asset-side ratios in  $x_1$ .

example, in our sample period, agriculture had already been experiencing distress in the 1920s and could have been a factor in predisposing banks located in certain areas to failure (Temin, 1976; White, 1984; Calomiris and Mason, 2003). Correlated risk might arise due to non-correspondent connections among banks, such as business networks or bankers' social networks. Finally, institutional features of the banking system (the correspondent network) and branch-banking restrictions may have also predisposed networks to spatial risk, features we account for more explicitly as described below.

$M_3$  includes the vector,  $x_{i3}$ , namely measures for network position and systemic risk for bank  $i$  in 1929: eigenvector centrality and percent contribution to systemic risk. Additionally, we include an interaction term that allows us to investigate how regulations, which underlie branch banking systems, modify the correspondent network's impact on bank survival. We interact branch-banking intensity with eigenvector centrality. The tension that existed in the late 1920s between the correspondent network and branch banking is well-documented by contemporaries (Curtis, 1930; Nadler and Bogen, 1933). Many large correspondent banks took the lead in opposing branch banking as they were concerned that small banks would join a branch banking system instead of their correspondent business (Abrams and Settle, 1993). This interaction term allows us to examine how the presence of two different types of networks influenced survivability.<sup>11</sup>

---

<sup>11</sup>Further motivation for the interaction comes from Table 4, which shows that the means of eigenvector centrality in branching and non-branching areas are statistically different from one another.

Table 9: Definitions of the variables that enter specifications  $M_1$ ,  $M_2$ , and  $M_3$ .

Model	Variable	Definition
$M_1, M_2, M_3$	Loans/Assets	Ratio of loans and discounts to total assets
$M_1, M_2, M_3$	Bonds/Assets	Ratio of bonds and securities to total assets
$M_1, M_2, M_2$	LnAssets	Ln(Total Assets)
$M_1, M_2, M_3$	Fed Member	Indicator for Federal Reserve membership
$M_1, M_2, M_3$	Central Reserve City	Indicator for central reserve city location
$M_1, M_2, M_3$	Reserve City	Indicator for reserve city location
$M_1, M_2, M_3$	LnPopulation	Ln(County Population)
$M_1, M_2, M_3$	Manufact per capita	Manufacturing establishments in the county divided by county population <sup>†</sup>
$M_1, M_2, M_3$	Cropland per capita	Acreage of cropland in the county divided by county population
$M_1, M_2, M_3$	Branch Intensity	Ratio of number of banks operating branches in the state to total number of banks in the state
$M_1, M_2, M_3$	Num Banks in County	Number of banks operating in the county <sup>†</sup>
$M_1, M_2, M_3$	Deposits/County Deposits	Ratio of deposits held at the bank to the sum of all bank deposits in the county
$M_2, M_3$	Composite Risk	Composite Risk Score, $R$
$M_2, M_3$	Network Neighbor in-R	The sum of balance-sheet risk from network neighbors with inward links
$M_2, M_3$	Network Neighbor out-R	The sum of balance-sheet risk from network neighbors with outward links
$M_2, M_3$	County-R	The sum of balance-sheet risk from banks in the same county
$M_3$	EigenCent	Eigenvector Centrality, $C$
$M_3$	SysRisk Percent	Percent contribution to systemic risk ( $S$ )
$M_3$	Branch $\times$ EigenCent	Interaction between Branch Intensity and Eigenvector Centrality <sup>†</sup>

Note: County population, manufacturing establishments, and acres of cropland are from the 1930 *U.S. Censuses of Population, Manufacturing, and Agriculture*. <sup>†</sup>“Manufact per capita” is multiplied by 100 and “Branch  $\times$  EigenCent” is multiplied by 10 to rescale for numerical precision; “Num Banks in County” is divided by 10; “Network Neighbor in-R”, “Network Neighbor out-R”, and “County-R” are transformed by square-roots.

Peer effects have received considerable attention in applied microeconomics research. Notably, specifications  $M_2$  and  $M_3$  control for two types of peer effects – network and spatial – extending their use to the examination of financial crises. In the context of [Manski \(1993\)](#),  $M_2$  and  $M_3$  control for exogenous peer effects, i.e., if a bank’s spatial or network neighbors had high composite credit risk on the eve of the crisis, the bank’s own composite credit risk and probability of survival would be affected. [Manski \(1993\)](#) also discusses endogenous peer effects, meaning the propensity for an individual bank to survive varies within the survival of its peers. We find strong network and spatial results for these models as well; however,

given our primary objective is to consider which elements of the pre-existing network predict survival, we present these subsidiary results in Appendix A.4. Even so, because our analysis is based on the entire network adjacency matrix,  $A$ , and because we compute each bank’s balance sheet risk, our results are the first to explore how a bank’s network position and how network peer effects may have contributed to the Great Depression, and more broadly, whether such features play an important role in predicting outcomes of financial crises.

## 5.2 Estimation and Model Comparison

We estimate the models using both maximum likelihood and Bayesian Markov chain Monte Carlo (MCMC) methods. The estimation results are similar across both methods; however, the Bayesian results permit us to compute marginal likelihoods and posterior model probabilities (Greenberg, 2008). Because our main goal in this section is to compare models with and without measures for bank-specific systemic risk and peer effects, these features of the Bayesian approach allow us to understand how the data support these measures as predictors of bank survival and the potential biases that may stem from their omission.

Marginal likelihoods ( $\ln f(y|M_1)$  versus  $\ln f(y|M_2)$  versus  $\ln f(y|M_3)$ ) have several advantageous properties. They lead to finite sample model probabilities and provide a measure of *sequential out-of-sample predictive fit*, which makes better use of the data for model comparison. The latter of these is less well-known. Choudhary et al. (2017) show that for model  $M_l$ :

$$f(y|M_l) = \prod_{i=1}^n f(y_i | \{y_j\}_{j < i}, M_l) \tag{11}$$

$$= \prod_{i=1}^n \int f(y_i | \{y_j\}_{j < i}, \beta_l, M_l) \pi(\beta_l | \{y_j\}_{j < i}, M_l) d\beta_l. \tag{12}$$

Equation (11) represents the marginal likelihood as the product of  $n$  one-step-ahead sequential predictive densities, which follows from the law of total probability. Equation (12) then shows that these  $n$  one-step-ahead sequential predictive densities correspond to the cumulative out-of-sample prediction, where the fit of observation  $i$  is measured with respect to the posterior density,  $\pi(\beta|y)$ , based only on data up to the  $i$ th observation (not conditioning on anything after  $i$ ,  $\{y_j\}_{j \geq i}$ ). In-sample measures condition on the entire dataset, whereas

other out-of-sample measures typically require the researcher to use a subset of the data for estimation and the remainder for prediction. Thus, the results depend on which data were used for estimation. Marginal likelihoods, on the other hand, are invariant to rearranging the data.

For Bayesian estimation, we use the Accept-Reject Metropolis-Hastings (ARMH) algorithm to fit the model (Tierney, 1994; Chib and Jeliazkov, 2001). Following Chib and Jeliazkov (2001), we then use the building blocks of that algorithm to compute marginal likelihoods. The priors on  $\beta$  are centered at 0 with a standard deviation of 10, although the results are robust to various hyperparameters on the prior distributions as well as using a training sample. We conduct 11,000 MCMC iterations with a burn-in period of 1,000.

### 5.3 Results

Table 10 presents the parameter estimates and model comparison results for our three models. The data overwhelmingly support  $M_3$ : the specification has a posterior model probability of nearly 1 in comparison with  $M_1$  and  $M_2$ . Given that the models represent competing hypotheses about whether network and network-neighbor measures are important for modeling bank survival, we show the data are more likely to occur under  $M_3$ .<sup>12</sup>

---

<sup>12</sup>The odds in favor of network and network-neighbor measures over no network predictors is approximately  $3.6 \times 10^{139} : 1$ .

Table 10: Modeling Bank Survivorship.

	$M_1$	$M_2$	$M_3$
Intercept	-1.170 (0.178)	-0.200 (0.193)	-0.283 (0.199)
Loans/Assets	0.204 (0.080)	-0.038 (0.081)	-0.055 (0.082)
Bonds/Assets	1.282 (0.087)	1.022 (0.089)	0.991 (0.088)
LnAssets	0.183 (0.009)	0.156 (0.011)	0.161 (0.011)
Fed Member	0.027 (0.019)	0.038 (0.019)	0.042 (0.019)
Branch Intensity	0.713 (0.209)	0.831 (0.217)	1.049 (0.241)
Num Banks in County	-0.005 (0.004)	-0.019 (0.008)	-0.019 (0.008)
Deposits/County Deposits	0.352 (0.066)	0.455 (0.070)	0.458 (0.071)
Central Reserve City	-1.060 (0.102)	-1.008 (0.105)	-1.025 (0.109)
Reserve City	-0.357 (0.042)	-0.322 (0.043)	-0.307 (0.043)
LnPopulation	-0.097 (0.014)	-0.098 (0.016)	-0.094 (0.016)
Manufact per capita	0.470 (0.104)	0.520 (0.108)	0.538 (0.109)
Cropland per capita	-0.003 (0.001)	-0.003 (0.001)	-0.003 (0.001)
Composite Risk		-0.059 (0.002)	-0.059 (0.002)
Network Neighbor out-R		-0.012 (0.006)	-0.016 (0.007)
Network Neighbor in-R		0.000 (0.002)	-0.012 (0.003)
County-R		0.007 (0.005)	0.007 (0.005)
EigenCent			4.198 (1.787)
SysRisk Percent			2.537 (0.720)
Branch Intensity $\times$ EigenCent			-4.015 (2.230)
Fed. Dist. FE	Yes	Yes	Yes
$n$	24,761	24,761	24,761
Log-Marginal Likelihood	-15341.88	-15026.04	-15020.53
Numerical Standard Error	(0.008)	(0.008)	(0.019)
Posterior Model Probability	$2.75 \times 10^{-140}$	0.0041	0.9959

Posterior means and standard deviations (in parentheses) are presented for  $M_1$ ,  $M_2$ , and  $M_3$ . Maximum likelihood estimates and standard errors are presented in Appendix Table A.4.1. Appendix Section A.7 details our data cleaning procedures.

The parameter estimates on the variables common to multiple models are mostly stable across the specifications, but we note some interesting differences. In particular, in  $M_2$ , we find smaller posterior means (in absolute value) on Network Neighbor out-R and in-R. However, when we add controls for the centrality and systemic risk in  $M_3$ , the means increase dramatically (in absolute value), relative to the standard deviations. This finding reflects the fact that  $M_3$  disentangles positive effects of the correspondent network (e.g., profiting by lending respondents' interbank deposits to households and businesses and by

charging fees for services such as currency exchange, handling business transactions and trade documentation, and money transfers) from negative network effects (e.g., risk exposure to lots of respondents) by separately controlling for the total amount of balance-sheet risk stemming from network neighbors and high centrality. This result has an important implication, i.e., network-neighbor or peer effect measures suffer from biases if systemic-risk and centrality are not also taken into account. For example, network drawbacks increase if network neighbors are at a high risk of failure.

Focusing on results from  $M_3$ , the model best supported by the data, coefficients on bank-specific characteristics are consistent with findings in previous studies on financial crises and the Great Depression, more specifically. Table 10 shows that bank size (measured by LnAssets) and the ratio of Bonds/Assets are positively associated with bank survival. The number of banks in the county, a measure of local competition, has a negative impact on bank survival. The market share of the bank (measured by the share of county deposits held at that bank) has a positive impact, consistent with Calomiris and Mason (2003). Lower probabilities of survivorship are associated with the indicator variables for central reserve city and reserve city. These results align with Mitchener and Richardson (2019), who find that, during the banking panics of the early 1930s, banks in reserve cities and central reserve cities faced significant withdrawal pressure from banks located in the hinterland. On the other hand, being a Fed member improved a bank’s prospects of surviving the banking distress of the early 1930s, perhaps not surprising since nonmember banks constituted 73% of all exits. As shown in Table 10, the branch-banking-intensity variable is positive, which is consistent with Mitchener (2005, 2007), who finds that laws prohibiting branch banking raised suspension rates for banks. Interestingly, the interaction, Branch Intensity  $\times$  EigenCent, is negative, demonstrating that branch banking modifies the correspondent network’s impact on survivability. We explore this result further in the next section.

Both eigenvector centrality, measuring the importance of a bank’s position in the network, and a bank’s contribution to systemic risk are positively-signed. The latter result is likely picking up “too-big-to-fail” policy. Using hand-collected archival data from the National Archives (NARA RFC Card Index used in Vossmeier (2016)), we computed that

over 70% of the top systemic-risk contributors in Table 6 received large injections from the Reconstruction Finance Corporation’s (RFC) recapitalization program. The RFC was a government-sponsored bank rescue program established in 1932 that aided many banks. Jesse Jones, Chairman of the RFC, noted that the bank-rescue agency paid particular attention to large institutions in deciding which banks to grant assistance (Jones, 1951). The RFC application and examiner files do not comment on or report a bank’s correspondent network. Only in the case of large institutions, such as Continental Illinois, does Jesse Jones mention the bank was a “good correspondent” (Jones, 1951). It seems reasonable to conclude that this episode in financial history points to too-big-to-fail policies being applied to systemically risky banks.<sup>13</sup> Policies designed to rescue banks that were very interconnected in the correspondent network system are also noted in Gorton and Tallman (2018). Appendix Figure A.3.1 shows that most exits from the network were small to medium-sized banks, which is also noted in Wheelock (1995).

Table 10 also shows that composite risk score and network-neighbor’s composite risk score (both “in” and “out” measures) are negatively-signed, results that are quite intuitive. (The magnitude of these effects are explored in the next section.) The other balance sheet measures, Loans/Assets and Bonds/Assets, are formed from the asset side of the balance sheet, whereas composite risk is built from the liabilities side. When we include composite risk in the model, the Loans/Assets variable is no longer statistically different from zero. Interestingly, our spatial risk variable, County-R, also is not statistically different from zero. Taking our peer effect results together, we show that, in 1929, risk residing with network neighbors is a more revealing characteristic of per-crisis bank distress, relative to spatial neighbors. This is not to say that the spatial factors didn’t matter. As noted above, research on the Great Depression has established that spatial contagion occurred: runs were clustered in space and time, locally and regionally during panic periods. That is, beginning

---

<sup>13</sup>It is important to note that many banks receiving assistance from the RFC, outside of the top 20, also failed. Appendix Table A.4.2 shows that our network results are robust to specifications that include controls for RFC assistance. However, a full analysis of the RFC would require the researcher to disentangle the precise assistance that thousands of banks received from the RFC – from collateralized loans to preferred stock purchases – and confront sample-selection issues involved with applying for and receiving assistance. Such a study presents an interesting possibility for future research, but is clearly beyond the scope and aims of this paper.

with the first panic in the fall of 1930, spatial elements became quite important. However, in this paper, we do not focus on contagion or spatial dynamics during panic periods. Rather, our goal is to explore what pre-panic features of the banking system mattered for subsequent distress. Along these lines, our results point to risky network neighbors being better predictors of subsequent bank distress than, say, risky banks in the county.

Robustness checks, maximum likelihood estimates (with standard error adjustments), and information criteria for the probit specifications are provided in Appendix A.4. We additionally show in the appendix that the estimated coefficients are robust to including state fixed effects; however, the fixed effects lead to overfitting as the information criteria do not favor the specification. Therefore, the results presented in this main body are the strongest, and most supported by the data. Furthermore, Appendix A.5 presents an ordered specification in which we disentangle mergers and failures from the “exit” group. In the full sample, we coded exits as both mergers and failures, but these are clearly different outcomes. The data we have collected on state banks do not permit us to distinguish between different types of survival outcomes; however, for national banks, we are able to further distinguish between types of exits. The results for the ordered specification on the subsample of national banks align with the binary specification and show significant support in favor of a model that includes our bank-specific systemic-risk and network-neighbor measures to predict the probability of bank survival, merger, and failure.

### 5.3.1 Covariate Effects

We now consider an exercise that examines how balance sheet risk (both a bank’s own and its network neighbors), the network’s topology, and bank regulation (i.e., removing restrictions on state branch banking) contribute to a bank’s predicted probability of survival. We are not seeking causal results nor are we considering dynamic contagious effects. We seek to decompose a banks predicted probability of survival through a series of covariate effects to shed light on the relative contribution of each effect.

To do so, we define  $\delta$  as the expected difference in computed pointwise probabilities when  $x_i^\dagger$  (the original case) is changed to  $x_i^\ddagger$  (the altered case). Following [Jeliazkov and Vossmeier](#)

(2018), computation of these effects is done by marginalizing over the parameters with the posterior distribution and marginalizing over the sample with the empirical distribution of the covariates:

$$\delta = \int [\Pr(y_i = 1|x^\dagger, \beta) - \Pr(y_i = 1|x^\ddagger, \beta)] f(x) \pi(\beta|y) dx d\beta. \quad (13)$$

The ability to accommodate both estimation and sampling uncertainty here is an advantage of the Bayesian approach. In the text below, we present numerical results that are averaged over both the MCMC draws and units in the sample. Appendix Section A.6 presents histograms of the distributions of the average effects as a function of parameter uncertainty.

**Balance-Sheet Effects:** We consider covariate effects on individual bank balance-sheet risk and network-neighbor balance-sheet risk in order to understand how these variables (independently and jointly) contribute to the predicted probability of bank survival.<sup>14</sup>

For our first covariate effect, we analyze how a one-standard-deviation decrease in Composite Risk  $R$  (represented in  $x_i^\ddagger$ ) changes the probability of bank survival. We find that, holding all else fixed, decreasing individual bank composite risk by one standard deviation increases the predicted probability of bank survival by 8.35 percentage points (the probability of survival goes from 0.614 to 0.698). This equates to 2,065 more banks surviving, and suggests that much of the existing fragility in the system in 1929 was due to decisions of bank management.

For our second covariate effect, we examine how a one-standard-deviation decrease in network-neighbor risk changes the probability of bank survival. This change is represented in  $x_i^\ddagger$  in the variables Network Neighbor out-R and Network Neighbor in-R. However, if Network Neighbor in-R was 0 (not a correspondent), then we keep the value at 0. We find that, holding all else fixed, decreasing these variables by one standard deviation increases the probability of bank survival by 1.00 percentage points (the probability of survival goes from 0.614 to 0.624), equating to 248 more banks surviving. Thus, network-neighbors' balance-sheet risk is about 12% as important as a bank's own balance-sheet risk in predicting survival.

We also consider the joint effect of these two features of the 1929 network, represented

---

<sup>14</sup>Note that the posterior distribution of the spatial neighbor risk variable is centered at zero.

in  $x_i^\dagger$  by altering Composite Risk, R, Network Neighbor out-R, and Network Neighbor in-R. Holding all else fixed, we find that decreasing these variables by one standard deviation increases the probability of bank survival by 9.20 percentage points (the probability of survival goes from 0.614 to 0.706). Total bank exits between 1929 and 1934 was over 9,000; thus, this effect equates to 2,278 fewer exits (roughly equal to the sum of the two independent effects) – or approximately 24% of banks that exited – implying that balance sheet composite risk was a significant source of fragility in the 1929 network.

**Network Topology Effects:** We next focus on how network position and network topology contribute to the predicted probability of bank survival. Methodologically, this is a challenging question to address because these variables are linked to the network adjacency matrix,  $A$ , which forms the backbone of our eigenvector centrality, systemic risk, and in-network and out-network neighbor measures. Thus, we utilize the simulated counterfactual networks from Section 3.3 to explore various forms of  $A$  that have the same number of nodes and connections as the 1929 network. In the construction of the variables, the vector,  $R$ , remains the same. To account for a different topological structure, we change the adjacency matrix which, in turn, alters eigenvector centrality, systemic risk, and in-network and out-network neighbor measures. We average these variables over many simulations of  $A$ , so the results are not based on a particular realization. We also remove the effects stemming from reserve cities and central reserve cities because these are reflective of the pyramid topology. All of the other variables (balance sheet size, ratios, location, etc.) remain the same.

We consider two counterfactual scenarios: (i) how the probability of bank survivorship would change if the network had a random topology instead of a pyramid shape ( $x_{i, rn}^\dagger$ ) and (ii) how the probability of bank survivorship would change if the network had an alternative scale-free topology ( $x_{i, sf}^\dagger$ ). In the random network setting, we are essentially flattening the pyramid structure and giving each bank similar network importance, as demonstrated by the centrality statistics in Table 3. In the alternative scale-free example, the pyramid shape is not flattened and the distribution of centrality is roughly the same.

Holding all else fixed, we find that, if the correspondent network had a random topology, the probability of bank survival would be 4.50 percentage points higher in comparison to its

actual pyramid shape (i.e., the probability of survival goes from 0.614 to 0.659), amounting to an additional 1,114 banks surviving the Great Depression, or roughly 12% of those that exited the system. Had this alternative topological structure existed in 1929, it appears that the American banking system may have been less prone to crisis and more stable. For (ii), we find that, holding all else fixed, if the correspondent network had an alternative scale-free topology, the probability of bank survival is not statistically different from when the topology is pyramid shaped. This non-result follows from the fact that the pyramid structure of the correspondent network does not get flattened in this second counterfactual, i.e., the simulated network and the actual 1929 network are both scale-free.

**Regulation Effects:** The counterfactual network topologies reveal a great deal about the importance of the pyramid-shaped network in making the American banking system inherently fragile prior to the banking distress of the early 1930s. Of course, regulators did not have the authority to make it more stable ex ante by, say, re-engineering the entire banking system so it resembled a random-network topology. Instead, it might have been possible to change an existing regulation so as to dilute the pyramid-shape of the topology. Perhaps the most feasible change would have been to allow branch banking at the state level to flourish because, as Tables 4 and 8 show, areas with branching had lower levels of centrality and concentration of network neighbor risk. In our sample period, national prohibitions on branch banking existed; however, some states had enacted laws that permitted branch banking within a state. In several states, particularly California, branch banking had become widespread. Given our emphasis on network topology, we therefore explore how more widespread adoption of branch banking would have influenced the shape of the correspondent network and, in turn, the banking system’s stability on the eve of the Great Depression.

The legislative record and reports of state and national bank regulators from the 1920s document the tension between correspondent and branch-banking systems. Many large correspondent banks argued branch banking would “jeopardize their profitable correspondent business” (Abrams and Settle, 1993), believing that if branch banking were more widespread, small country banks would become part of the branching system, as opposed to seeking services from correspondents. They thus lobbied to prohibit branch banking. Curtis (1930)

(p. 179) states:

*The development of branch banking will also tend to disrupt the present relationship of city and country bank correspondents. Certain metropolitan banks will lose correspondents and business unless they, too, establish branches. But the relationship of parent and branch will be much more direct and intimate and dependable than that of correspondents and, therefore, the seasonal and panic demands for funds will be better cared for, and an economy will be realized because it will no longer be necessary to carry balances with correspondents.*

Because prohibitions on branch banking maintained the status quo, these regulations implicitly reinforced the correspondent network. For example, in a state like Illinois, which prohibited branch banking within its borders, Continental Illinois Bank and Trust had 1,357 correspondent relationships, just with banks based in the state (out of nearly 1800 banks in the state), let alone its 4,474 connections nationally. Further evidence comes from eigenvector centrality measures, where total centrality in Minnesota (no branch banking) far exceeded California (extensive statewide branch banking). We thus consider how the probability of bank survival would change if branch banking had been allowed to flourish within all states — a feasible regulatory counterfactual outcome given the contemporaneous debate as to whether the pyramid-shaped correspondent system had outlived its utility. In contrast to previous research which has focused on the benefits of branch banking for financial stability that arise from either portfolio diversification or industry consolidation (Carlson, 2004; Carlson and Mitchener, 2006), this exercise focuses on a different channel: how more widespread branch banking would have changed the topology of the correspondent network and how this might have, in turn, influenced systemic stability.

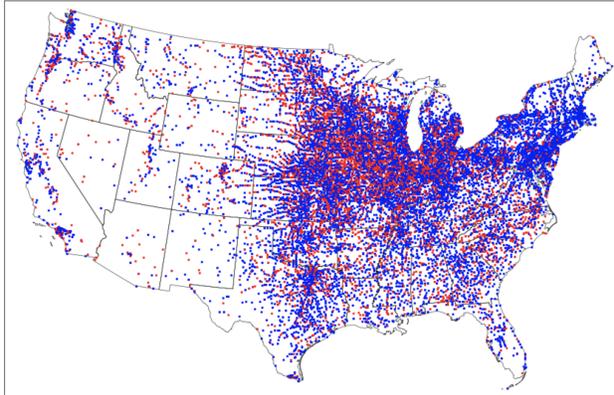
To model the alternative scenario of widespread branch banking, we set the branch-banking-intensity value for all regions of the U.S. to 40%, the highest state-level value in 1929. In  $x_{i,br}^\dagger$ , this change would be reflected in the branching intensity variable as well as the interaction with eigenvector centrality. Holding all else fixed, we find that, had branch banking been allowed to flourish at the state level, the probability of bank survival would have been 10.08 percentage points higher than what prevailed under restricted branching

(probability of survival goes from 0.614 to 0.715), equating to 2,496 more bank survivals or roughly 26% of those banks that closed. These competing network forces suggest that, had less weight been placed on the correspondent network before the crisis and more weight been placed on branch-banking systems, the fallout from the Great Depression may have been significantly less severe.

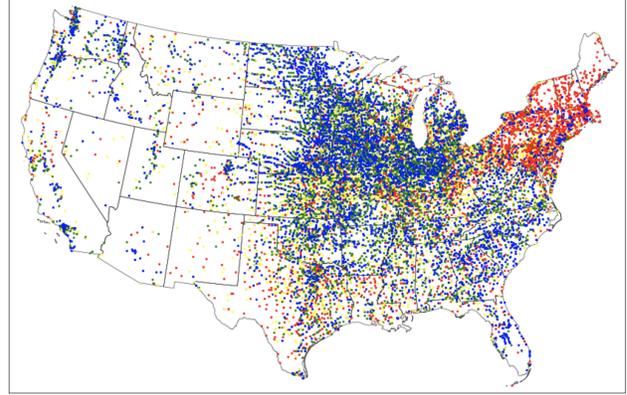
**Spatial Maps:** The covariate effects demonstrate the ways in which balance-sheet risk, network topology, and (related) regulations on branch banking contributed to a bank’s predicted probability of survival. But, as our discussion makes clear, the banking system was far from uniform in any of these respects (e.g., regulations varied at the state level and large cities were nodes of concentration in the correspondent network). Spatial variation in the existing features of the American banking system therefore may have had a differential impact on U.S. regions.

Figure 6 illustrates the variation in the intensity of different types of effects. Panel (a) presents the baseline, whether a bank existing in 1929 survived (blue) until July 1934. Panels (b), (c), and (d) reflect the outcomes when we decompose a bank’s predicted probability of survival into balance-sheet effects, network-topology effects, and branch-banking effects (as described above). The blue color represents locations that have the highest quartile positive change in the predicted probability of survival due to a particular effect, followed by green, yellow, and red, for the remaining three quartiles.

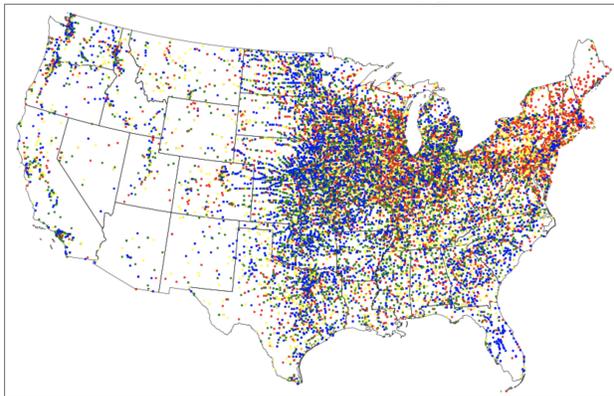
In panel (b), we see that improvements in balance-sheet risk (a one-standard-deviation decrease in R, in-R, and out-R) would have greatly benefited the Midwest. This finding is consistent with literature emphasizing that the waves of failures in midwestern cities, like Chicago and its suburbs (Calomiris and Mason, 1997; Esbitt, 1986), were related to portfolio problems. On the other hand, the Northeast would see a smaller change because survival probabilities were already high in that region (i.e., large concentration of blue in panel (a)), no doubt reflecting the stronger initial positions of bank portfolios in that region. In panel (c), we see that the benefits of moving to the counterfactual random topology would be widespread, touching each part of the country, but with the largest changes in big cities and in Michigan, Illinois, and Iowa. Lastly, in panel (d), we see that allowing branch banking



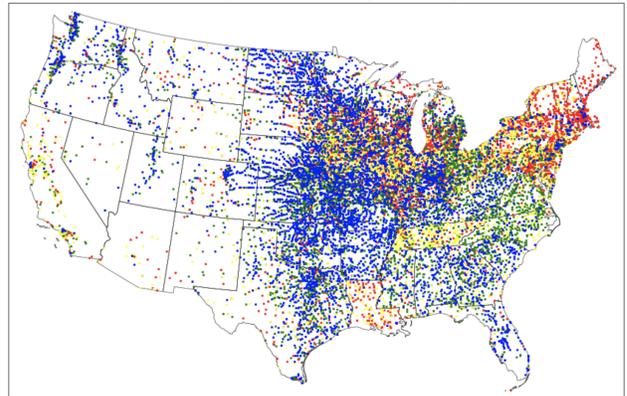
(a) 1934 bank survivals (blue) and closures (red).



(b) Balance-sheet shock.



(c) Topology shock.



(d) Branch-banking shock.

Figure 6: Map of the changes in the predicted probability of bank survival due to a shock. Highest to lowest quartiles for changes in survival probabilities are in blue, green, yellow, and red. The predicted probabilities for banks  $i = 1, \dots, n$  are averaged over the MCMC draws.

to flourish would have little effect in coastal locations like California or North Carolina, where branch banking was already strongly established, but would have greatly influenced survival probabilities in the heartland – where prohibitions on within-state branch banking were entrenched and where dilution of the correspondent network would deliver gains to systemic stability.

The American banking system of 1929 was fragmented in terms of regulatory, supervisory, and even monetary authority. As a result, if armed with the information presented in Figure 6, Fed policymakers might have tailored their responses to fit district needs (given the scope that each Federal Reserve District had in setting its own policies prior to the Banking Act of 1935 (Meltzer, 2003; Wheelock, 1991)). For example, a policymaker in the 7<sup>th</sup> Federal Reserve district (Chicago) might have pressured banks to move out of non-liquid real estate holdings or diversify away from lending to growers all producing a single crop, in order to reduce the concentration of balance-sheet risk in this region’s correspondents. On the other hand, policymakers in the 8<sup>th</sup> and 9<sup>th</sup> Federal Reserve districts (St. Louis and Minneapolis, respectively) might have pushed for liberalizing branch-banking laws. And for those policymakers that desired to reshape things on a national scale, they might have advocated more forcefully to end the correspondent system in 1913 – when the Federal Reserve System was created – and force all banks to join the Fed so as to reduce the pyramid-shape or scale-free topology of the network. Of course, optimal policy relies on many assumptions, and while our modeling by no means addresses all concerns, exploring these covariate effects by considering how they might have altered the decision making of contemporaries provides some additional perspective on the utility of our findings and methodology.

## 6 Conclusion

The Great Depression provides a laboratory for understanding how financial regulation can interact with network topology to predispose financial systems to crises. We combine new data on all correspondent network connections for commercial banks from July 1929 with a novel systemic-risk measurement approach that takes into account each bank’s network position, its risk, and its network-neighbors’ risk to shed light on how important these

features are in predicting subsequent banking distress. We show that all of these factors mattered significantly.

Our findings have important implications for policymakers and researchers interested in understanding risk in financial systems and modeling financial crises. For systemic-risk regulators, our results demonstrate that banking regulations can have unintended consequences for financial stability. In the late-19<sup>th</sup> century, the tiered structure of bank reserves was believed to be an effective way for banks to deploy funds when the seasonal demand for money was high, leading to a pyramid-shaped correspondent network. However, connections in the correspondent network became extremely concentrated over time. We show that the pyramid-shape of the interbank network concentrated risk in particular nodes and made the banking system in 1929 more fragile and prone to contagion risk in comparison to alternative network topologies. If the 1929 correspondent network instead had a random-network topology, *ex ante*, the American commercial bank system would have been more stable: nearly 1,100 more banks (roughly 5% of all banks or 12% of those that closed) would have survived the Great Depression. This is not to suggest that the pyramid-shaped correspondent network was inefficient for participating banks and their customers and failed to deliver benefits, but rather that the topology that proved useful in calm periods also exposed banks to greater risk in turbulent periods.

Banking regulations also buttressed the pyramid shape of the network. We show how feasible regulatory alternatives, such as widespread branch banking, appear to have had the potential to diffuse the concentration of risk inherent to the pyramid-shape and improve subsequent bank survivorship by 10 percentage points during the banking crisis of the early 1930s. However, the pyramid-network, a relic from the pre-Fed era, had become “locked in” through regulation. Its existence underscores the self-reinforcing nature of networks, showing that “history can matter” when there is the possibility of lock-in at a Pareto inferior equilibrium (Arthur, 1988, 1989; Farrell and Shapiro, 1989; Gallini and Karp, 1989; Klemperer, 1987a,b; Beggs and Klemperer, 1992).

We also show that some financial crises, like the Great Depression, are characterized by a preponderance of failures among small and medium-sized banks and that, before the crisis

of the 1930s started, default risk was highest for banks that were smaller and less connected. In such crises, upper-tail banks in the network can become exposed to risk emanating from lower parts of the distribution through network-neighbor linkages. Managing systemic risk thus requires accounting for the entire distribution of banks and how the shape of the network can transmit risks residing lower in the system to its upper reaches. Doing so brings focus to peer effects in financial networks. We demonstrate nontrivial biases in network-neighbor measures when controls for network position and systemic risk are omitted, and show that models that incorporate network position, systemic-risk, and network-neighbor variables are better supported by the data in the prediction of bank survival. Our research draws attention to the importance of incorporating peer effects into the study of financial networks and crises, and represents a first step in developing this promising line of research.

## References

- Abrams, B. and Settle, R. (1993), “Pressure-group influence and institutional change: Branch-banking legislation during the Great Depression,” *Public Choice*, 77, 687–705.
- Acemoglu, D., Ozdaglar, A., and Tahbaz-Salehi, A. (2015), “Systemic Risk and Stability in Financial Networks,” *American Economic Review*, 105, 564–608.
- Acharya, V. V., Pedersen, L. H., Philippon, T., and Richardson, M. (2017), “Measuring Systemic Risk,” *Review of Financial Studies*, 30, 2–47.
- Allen, F. and Gale, D. (2000), “Financial Contagion,” *Journal of Political Economy*, 108, 1–33.
- Allen, F., Babus, A., and Carletti, E. (2010), “Financial Connections and Systemic Risk,” Tech. Rep. w16177, National Bureau of Economic Research, Cambridge, MA.
- Anderson, H., Calomiris, C. W., Jaremski, M., and Richardson, G. (2018), “Liquidity Risk, Bank Networks, and the Value of Joining the Federal Reserve System,” *Journal of Money, Credit and Banking*, 50, 173–201.
- Anderson, H., Paddrik, M., and Wang, J. (2019), “Bank Networks and Systemic Risk: Evidence from the National Banking Acts,” *American Economic Review*, forthcoming.
- Anderson, H., Erol, S., and Ordonez, G. (2020), “Interbank Networks in the Shadows of the Federal Reserve Act,” *Working Paper*.
- Arthur, W. (1988), “Self-reinforcing Mechanisms in Economics,” in *The Economy as an Evolving Complex System*, Addison Wesley.
- Arthur, W. (1989), “Competing Technologies, Increasing Returns, and Lock-in by Historical Events,” *The Economic Journal*, 99, 116–131.
- Barabasi, A.-L. (2002), *Linked: The New Science of Networks*, Basic Books.
- Barabasi, A.-L. and Bonabeau, E. (2003), “Scale-Free Networks,” *Scientific American*, 288, 60–69.
- Barmouille, Y., Djebbari, H., and Fortin, B. (2009), “Identification of Peer Effects Through Social Networks,” *Journal of Econometrics*, 150, 41–55.
- Beggs, A. and Klemperer, P. (1992), “Multi-period Competition with Switching Costs,” *Econometrica*, 60, 651–666.
- Billio, M., Getmansky, M., Lo, A. W., and Pelizzon, L. (2012), “Econometric measures of connectedness and systemic risk in the finance and insurance sectors,” *Journal of Financial Economics*, 104, 535–559.
- Bisias, D., Flood, M., Lo, A. W., and Valavanis, S. (2012), “A Survey of Systemic Risk Analytics,” *Annual Review of Financial Economics*, 4, 255–296.

- Bonacich, P. (1987), “Power and Centrality: A Family of Measures,” *American Journal of Sociology*, 92, 1170–1182.
- Bonacich, P. and Lloyd, P. (2001), “Eigenvector-like measures of centrality for asymmetric relations,” *Social Networks*, 23, 191–201.
- Bordo, M. and Wheelock, D. (2013), “The Promise and Performance of the Federal Reserve as Lender of Last Resort 1914-1933,” in *A Return to Jekyll Island: The Origins, History, and Future of the Federal Reserve*, Cambridge University Press.
- Brock, W. A. and Durlauf, S. N. (2001), “Discrete Choice with Social Interactions,” *Review of Economics Studies*, 68, 235–260.
- Brownlees, C., Chabot, B., Ghysels, E., and Kurz, C. (2020), “Back to the Future: Back-testing Systemic Risk Measures during Historical Bank Runs and the Great Depression,” *Journal of Banking and Finance*, 113.
- Brunetti, C., Harris, J., Mankad, S., and Michailidis, G. (2019), “Interconnectedness in the Interbank Market,” *Journal of Financial Economics*, forthcoming.
- Burdick, D., Hernández, M. A., Ho, H., Koutrika, G., Krishnamurthy, R., Popa, L., Stanoi, I., Vaithyanathan, S., and Das, S. R. (2011), “Extracting, Linking and Integrating Data from Public Sources: A Financial Case Study,” *IEEE Data Eng. Bull.*, 34, 60–67.
- Béliveau, A., Goring, S., Platt, R. W., and Gustafson, P. (2017), “Network meta-analysis of disconnected networks: How dangerous are random baseline treatment effects?” *Research Synthesis Methods*, 8, 465–474.
- Caballero, R. and Simsek, A. (2013), “Fire Sales in a Model of Complexity,” *Journal of Finance*, 68, 2549–2587.
- Calomiris, C. and Gorton, G. (1991), “The Origins of Banking Panics: Models, Facts, and Bank Regulation,” *Financial Markets and Financial Crises*, pp. 109–174.
- Calomiris, C. and Haber, S. (2014), *Fragile by Design: The Political Origins of Banking Crises and Scare Credit*, Princeton University Press.
- Calomiris, C. and Mason, J. (1997), “Contagion and Bank Failures During the Great Depression: The June 1932 Chicago Banking Panic,” *American Economic Review*, 87, 863–883.
- Calomiris, C., Jaremski, M., and Wheelock, D. (2019), “Interbank Connections, Contagion, and Bank Distress in the Great Depression,” *SSRN Working Paper*.
- Calomiris, C. W. and Carlson, M. (2017), “Interbank networks in the National Banking Era: Their purpose and their role in the Panic of 1893,” *Journal of Financial Economics*, 125, 434–453.
- Calomiris, C. W. and Mason, J. R. (2003), “Fundamentals, Panics, and Bank Distress During the Depression,” *American Economic Review*, 93, 1615–1647.

- Carlson, M. (2004), “Are Branch Banks Better Survivors? Evidence from the Depression Era,” *Economic Inquiry*, 42, 111–126.
- Carlson, M. and Mitchener, K. (2006), “Branch Banking, Bank Competition, and Financial Stability,” *Journal of Money, Credit and Banking*, 38, 1293–1328.
- Carlson, M. and Mitchener, K. (2009), “Branch Banking as a Device for Discipline: Competition and Bank Survivorship during the Great Depression,” *Journal of Political Economy*, 117, 165–210.
- Carlson, M. and Rose, J. (2015), “Credit Availability and the Collapse of the Banking Section in the 1930s,” *Journal of Money, Credit and Banking*, 47, 1239–1271.
- Carlson, M. and Wheelock, D. (2016), “Interbank Markets and Banking Crises: New Evidence on the Establishment and Impact of the Federal Reserve,” *American Economic Review*, 106, 533–537.
- Carlson, M. and Wheelock, D. (2018), “Did the Founding of the Federal Reserve Affect the Vulnerability of the Interbank System to Contagion Risk,” *Journal of Money, Credit and Banking*, 50, 1711–1750.
- Chib, S. (1995), “Marginal Likelihood from the Gibbs Output,” *Journal of the American Statistical Association*, 90, 1313–1321.
- Chib, S. and Jeliazkov, I. (2001), “Marginal Likelihood from the Metropolis-Hastings Output,” *Journal of the American Statistical Association*, 96, 270–281.
- Choudhary, V., Currim, I., Dewan, S., Jeliazkov, I., Mintz, O., and Turner, J. (2017), “Evaluation Set Size and Purchase: Evidence from a Product Search Engine,” *Journal of Interactive Marketing*, 37, 16–31.
- Curtis, J. (1930), “The Turning Financial Worm,” in *Chain, Group and Branch Banking*, H.W. Wilson Co.
- Das, S. R. (2016), “Matrix Metrics: *Network-Based Systemic Risk Scoring*,” *The Journal of Alternative Investments*, 18, 33–51.
- Das, S. R., Kim, S., and Ostrov, D. (2019), “Dynamic Systemic Risk Networks,” *Journal of Financial Data Science*, 1, 141–158.
- Dasgupta, A. (2004), “Financial Contagion through Capital Connections: A Model of the Origin and Spread of Bank Panics,” *Journal of the European Economic Association*, 2, 1049–1084.
- DBG (1933), “RFC largest Stockholder in Chicago’s Biggest Bank,” *Daily Boston Globe*, December 20, 1933.
- Dupont, B. (2017), “Bank Networks and Suspensions in the 1893 Panic: Evidence from the State Banks and their Correspondents in Kansas,” *Financial History Review*, 24, 265–282.

- Elliott, M., Golub, B., and Jackson, M. O. (2014), “Financial Networks and Contagion,” *American Economic Review*, 104, 3115–3153.
- Esbitt, M. (1986), “Bank Portfolios and Bank Failures During the Great Depression: Chicago,” *Journal of Economic History*, 46, 455–462.
- Farrell, J. and Shapiro, C. (1989), “Optimal Contracts with Lock-in,” *American Economic Review*, 79, 51–68.
- Freixas, X., Parigi, B. M., and Rochet, J.-C. (2000), “Systemic Risk, Interbank Relations, and Liquidity Provision by the Central Bank,” *Journal of Money, Credit and Banking*, 32, 611–638.
- Friedman, M. and Schwartz, A. J. (1963), *A Monetary History of the United States, 1867-1960*, Princeton University Press.
- Gabaix, X., Gopikrishnan, P., Plerou, V., and Stanley, H. E. (2003), “A theory of power-law distributions in financial market fluctuations,” *Nature*, 423, 267–270.
- Gallini, N. and Karp, L. (1989), “Sales and Consumer Lock-in,” *Economica*, 56, 279–294.
- Gorton, G. and Tallman, E. (2018), *Fighting Financial Crises: Learning from the Past*, University of Chicago Press.
- Greenberg, E. (2008), *Bayesian Econometrics*, Cambridge University Press.
- Griliches, Z. (1961), “A Note on Serial Correlation Bias in Estimates of Distributed Lags,” *Econometrica*, 29.
- Griliches, Z. (1967), “Distributed Lags: A Survey,” *Econometrica*, 35.
- Heitfield, E., Richardson, G., and Wang, S. (2017), “Contagion During the Initial Banking Panic of the Great Depression,” *NBER Working Paper 23629*.
- Iyer, R. and Puri, M. (2012), “Understanding Bank Runs: The Importance of Depositor-Bank Relationships and Networks,” *American Economic Review*, 102, 1414 – 1445.
- James, J. A. and Weiman, D. F. (2010), “From Drafts to Checks: The Evolution of Correspondent Banking Networks and the Formation of the Modern U.S. Payments System, 1850-1914,” *Journal of Money, Credit and Banking*, 42, 237–265.
- James, J. A. and Weiman, D. F. (2011), “The National Banking Acts and the Transformation of New York City Banking During the Civil War Era,” *The Journal of Economic History*, 71, 338–362.
- Jaremski, M. and Wheelock, D. (2020), “The Founding of the Federal Reserve, the Great Depression and the Evolution of the U.S. Interbank Network,” *Journal of Economic History*, 80, 69–99.

- Jeliazkov, I. and Vossmeier, A. (2018), “The Impact of Estimation Uncertainty on Covariate Effects in Nonlinear Models,” *Statistical Papers*, 59, 1031–1042.
- Jeliazkov, I., Graves, J., and Kutzbach, M. (2008), “Fitting and comparison of models for multivariate ordinal outcomes,” in *Bayesian Econometrics*, vol. 23 of *Advances in Econometrics*, pp. 115–156, Emerald Group Publishing Limited.
- Jones, J. H. (1951), *Fifty billion dollars: my thirteen years with the RFC, 1932-1945*, Macmillan Co., New York, OCLC: 233209.
- Keele, L. and Kelly, N. (2006), “Dynamic Models for Dynamic Theories: The Ins and Ours of Lagged Dependent Variables,” *Political Analysis*, 14, 186–205.
- Klemperer, P. (1987a), “The Competitiveness of Markets with Consumer Switching Costs,” *Rand Journal of Economics*, 18, 138–150.
- Klemperer, P. (1987b), “Markets with Consumer Switching Costs,” *Quarterly Journal of Economics*, 102, 375–394.
- Lasszlo-Barabasi, A. and Albert, R. (1999), “Emergence of Scaling in Random Networks,” *Science*, 286, 509–512.
- Manski, C. F. (1993), “Identification of Endogenous Social Effects: The Reflection Problem,” *Review of Economics Studies*, 60, 531–542.
- Markose, S., Giansante, S., and Shaghghi, A. (2012), “‘Too Interconnected to Fail’ Financial Network of US CDS Market: Topological Fragility and Systemic Risk,” *Journal of Economic Behavior and Organization*, 83, 627–646.
- Meltzer, A. (2003), *A History of the Federal Reserve, Volume 1: 1913-1951*, University of Chicago Press.
- Merton, R. C. (1974), “On the Pricing of Corporate Debt: The Risk Structure of Interest Rates,” *The Journal of Finance*, 29, 449–470.
- Mitchener, K. (2005), “Bank Supervision, Regulation, and Instability During the Great Depression,” *Journal of Economic History*, 65, 152–185.
- Mitchener, K. and Richardson, G. (2020), “Contagion of Fear,” *NBER Working Paper 26859*.
- Mitchener, K. J. (2007), “Are Prudential Supervision and Regulation Pillars or Financial Stability? Evidence from the Great Depression,” *Journal of Law and Economics*, 50, 273–302.
- Mitchener, K. J. and Richardson, G. (2019), “Network Contagion and Interbank Amplification during the Great Depression,” *Journal of Political Economy*, April.
- Mizrucki, M. and Neuman, E. (2008), “The Effect of Density on the Level of Bias in the Network Autocorrelation Model,” *Social Networks*, 30, 190–200.

- Mizrucki, M. and Neuman, E. (2010), “Structure and Bias in the Network Autocorrelation Model,” *Social Networks*, 32, 290–300.
- Nadler, M. and Bogen, J. (1933), *The Banking Crisis: The End of an Epoch*, New York: Dodd, Mead and Co.
- Nazeran, P. and Dwyer, D. (June 2015), “Credit Risk Modeling of Public Firms: EDF9,” *Moody’s Analytics, Modeling Methodology*.
- Newman, M. E. J., Watts, D. J., and Strogatz, S. H. (2002), “Random graph models of social networks,” *Proceedings of the National Academy of Sciences*, 99, 2566–2572.
- Redenius, S. A. (2007), “Designing a national currency: antebellum payment networks and the structure of the national banking system,” *Financial History Review*, 14.
- Richardson, G. (2007), “The Check is in the Mail; Correspondent Clearing and the Collapse of the Banking System, 1930 to 1933,” *The Journal of Economic History*, 67, 643–671.
- Temin, P. (1976), *Did Monetary Forces Cause the Great Depression?*, New York: W.W. Norton.
- Tierney, L. (1994), “Markov Chains for Exploring Posterior Distributions,” *The Annals of Statistics*, 22, 1701–1728.
- Vossmeier, A. (2016), “Sample Selection and Treatment Effect Estimation of Lender of Last Resort Policies,” *Journal of Business and Economic Statistics*, 34, 197–212.
- Wang, W., Neuman, E., and Newman, D. (2014), “Statistical Power of the Social Network Autocorrelation Model,” *Social Networks*, 38, 88–99.
- Wheelock, D. (1991), *The Strategy and Consistency of Federal Reserve Monetary Policy, 1924-1933*, Cambridge University Press.
- Wheelock, D. (1995), “Regulation, Market Structure, and the Bank Failures of the Great Depression,” *Federal Reserve Bank of St. Louis Review*, 77, 27–38.
- White, E. (1984), “A Reinterpretation of the Banking Crisis of 1930,” *Journal of Economic History*, 44, 119–183.
- White, N., E. (1983), *The Regulation and Reform of the American Banking System, 1900-1929*, Princeton University Press, Princeton, N.J.
- Wicker, E. (1980), “A Reconsideration of the Causes of the Banking Panic of 1930,” *Journal of Economic History*, 40, 571–583.
- Wicker, E. (1996), *The Banking Panics of the Great Depression*, Cambridge University Press, Cambridge.
- Wijk, B. C. M. v., Stam, C. J., and Daffertshofer, A. (2010), “Comparing Brain Networks of Different Size and Connectivity Density Using Graph Theory,” *PLOS ONE*, 5, e13701, Publisher: Public Library of Science.

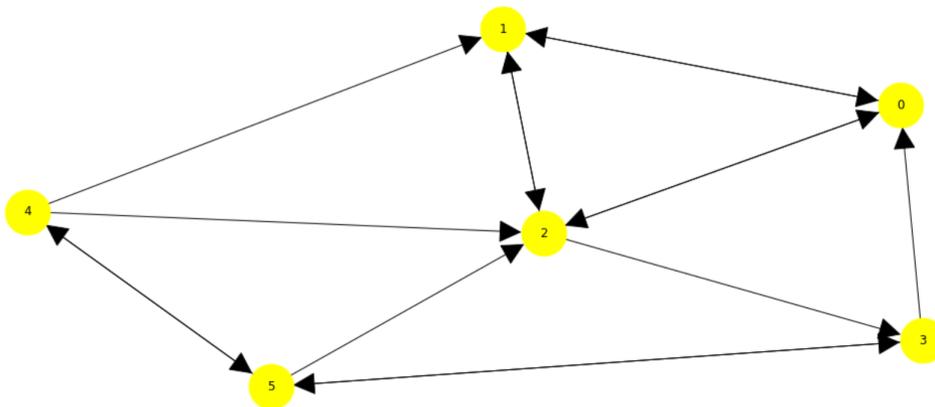


Figure A.1.1: A network of 6 nodes and 15 links.

## A Appendix

### A.1 Theory and Examples

In this section, we introduce network and systemic-risk measures that are employed to understand the 1929 network. We begin with a simple example network in order to obtain a better understanding of network topology and its relation to our systemic-risk metric. Figure A.1.1 shows a network of just 6 banks (nodes) and 15 links. It is a directed network, so that a relationship can run both ways, and the network is unweighted to correspond to the dataset in the paper, i.e., all links have unit weight and point from the bank placing money (the respondent) to the one receiving it (the correspondent). Some links are uni-directed and have one arrowhead and others are bi-directed and have two arrowheads. Nodes are numbered from 0 through 5.

The network may be represented by an “adjacency matrix” shown in equation (14) displaying the connections between banks. (Banks are numbered 0 through 5.)

$$A = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix} \quad (14)$$

Understanding which nodes have the greatest influence in the network, i.e., those that are the most critical, is of particular interest to policymakers who might want to identify “too big to fail” or systemically important financial institutions (SIFIs). To understand this feature of our network, we use two variants of the most common measure, centrality: degree and eigenvector centrality.

*Degree:* Centrality can be measured by the number of connections a node has, i.e., its

“degree.” Of course, degree centrality ignores the fact that a node with few connections may still have a huge influence if it is connected to a node with many connections. Degree centrality may be varied by computing degree to a chosen depth, say three levels deep, a parameter chosen at the discretion of the modeler.

For the six nodes in our example network, the degrees are  $\{5, 5, 7, 4, 4, 5\}$  for a total of 30 (twice the number of links, 15, because each link adds a degree to two nodes). We see that node #2 has the greatest degree. Each node’s degree can be apportioned into in-degree (links coming in) versus out-degree (links going out). In the context of the paper, we may think of nodes with a greater proportion of in-degree as correspondent banks and nodes with greater out-degree proportion as respondent banks.

*Eigenvector centrality* is a more general formulation (Bonacich, 1987; Bonacich and Lloyd, 2001), where centrality ( $c_i$ ) of a node is defined as a function of the centrality of the nodes to which it is connected, through the network adjacency matrix,  $A$ . This leads to a circular system of  $n$  simultaneous equations:

$$c_i = \sum_{j=1}^n A_{ij}c_j, \quad \forall i = 1, 2, \dots, n. \quad (15)$$

This system of equations may be written in matrix form such that:

$$\lambda c = A \cdot c \quad (16)$$

where  $\lambda$  is a scalar quantity,  $c$  is a vector of size  $n$ , and as before,  $A \in \mathcal{R}^{n \times n}$ . Equation (16) is an eigensystem, and one solution to this system of equations is the principal eigenvector in an eigenvalue decomposition of adjacency matrix,  $A$ . This is known as the centrality vector, which contains  $n$  components,  $c_i, i = 1, 2, \dots, n$ . Eigenvalue centrality is equivalent to degree centrality computed with infinite degree depth.

A computation of eigenvector centrality for the example network gives the following result for the six nodes:  $\{0.5727, 0.5034, 0.5493, 0.2982, 0.0653, 0.1541\}$ . Notice that node #0 has the highest eigenvector centrality even though node #2 has the highest degree. The first three nodes derive their influence from their position in the network relative to the other nodes. They have relatively more in-degree. Since the banking network in the Great Depression was known to be pyramidal, nodes with a higher extent of in-degree will be higher up in the pyramid.

*Systemic Risk Score:* We combine the network adjacency matrix,  $A$ , with a composite risk score for each bank in a vector denoted  $R$ , to create a single measure of overall systemic risk, extending and modifying the metrics proposed in Das (2016) and Das et al. (2019).  $R$  is a vector of credit quality score for each bank, where a higher score means poorer quality. Our approach allows one to empirically estimate system-wide “exposure” despite not knowing everything we might want about the financial network. (For example, data on balance-sheet linkages between financial institutions is often opaque or incomplete, both historically and today.) Composite systemic risk per bank,  $S$ , is thus defined as:

$$S = \frac{1}{n} \cdot \sqrt{R^\top \cdot A \cdot R}$$

$$\begin{aligned}
&= \sqrt{\frac{R^\top}{n} \cdot A \cdot \frac{R}{n}} \\
&= \sqrt{Q^\top \cdot A \cdot Q}
\end{aligned} \tag{17}$$

where  $n$ , as noted before, is the number of banks in the system, and superscript  $\top$  denotes the transpose of a vector or matrix. (Recall that  $R$  is an  $n$ -vector and  $A$  is a  $n \times n$  matrix. Thus,  $Q$  is an  $n$ -vector.) Division by  $n$  is a normalization used to measure systemic risk per financial institution. Since the elements of  $R$  and  $A$  are all non-negative,  $S \geq 0$ .

Equation (17) implies that systemic risk, as denoted by scalar quantity,  $S$ , increases if the elements of  $R$  (individual bank risk) increase, holding  $n$  and  $A$  constant. Likewise, *ceteris paribus*, if the elements of  $A$  (interconnectedness of banks) increase, systemic risk per bank also increases. The systemic-risk measure,  $S$ , may be thought of as a network-weighted measure of composite credit risk in the banking system. If there are no network linkages, then  $A$  is a zero matrix and  $S = 0$ . We note that this measure can be decomposed into two components: (1) the amount emanating from individual bank risk and (2) that from network effects. If we replace the network matrix,  $A$ , with the identity matrix and recompute  $S$ , i.e.,  $S' = \sqrt{\sum_{i=1}^n R_i^2}$  then this measures composite emanating risk from individual bank risk only. The ratio  $S/S' \geq 1$  is a measure of systemic risk arising from network effects.

We examine how this systemic risk score changes with risk levels,  $R$ . Assume that bank  $i$  can either have low risk ( $R_i = 1$ ) or high risk ( $R_i = 2$ ). If all banks were low risk, the vector of credit quality is  $R = [1, 1, 1, 1, 1, 1]$ . Using this vector, we compute a lower bound for systemic risk, i.e.,

$$S = \frac{1}{n} \sqrt{R^\top \cdot A \cdot R} = 0.6455$$

But if all banks were high risk, i.e.,  $R = [2, 2, 2, 2, 2, 2]$ , then the systemic risk per bank would be 1.2910 (an upper bound, assuming no change in network structure). This is exactly double because the function  $S(R, A)$  is linear homogenous in  $R$ . These values are the lower and upper bounds for  $S$ . We can examine the effect of increasing the risk of banks with low eigenvector centrality, i.e., letting  $R = [1, 1, 1, 2, 2, 2]$ , in which case  $S = 0.9428$ . But, if we raise the risk of high centrality banks, such that  $R = [2, 2, 2, 1, 1, 1]$ , then  $S = 1.0274$ , and is higher as expected. Both these values lie between the upper and lower bounds computed earlier.

## A.2 R as a Measure of Risk

Credit ratings are typically employed as the measure for internal balance-sheet risk. However, this is not available for our sample, leading us to make use of the eight asset and liability categories we have from each bank’s balance sheet. As a result, we develop the measure  $R$  which is the product of transformed leverage and inverse profitability risk score. To understand how well  $R$  captures risk, we compare  $R$  to several standard balance sheet measures that are often employed in banking studies. We mimic the balance sheet measures used in Calomiris and Mason (2003). However, since we are studying all banks, instead of limiting the sample to Federal Reserve members, we are unable to replicate all of the measures used in that study as our balance sheet categories are not as granular. For comparison, we look at how well each measure fits a probit model bank survival. The outcome variable for the probit model equals 1 if the bank appears in 1934 and 0 otherwise.

Table A.2.1 presents the estimate of each risk measure, and the log-marginal likelihood estimate and BIC of a model with only the tested risk measure and an intercept. The composite risk score  $R$  has the highest marginal likelihood and lowest BIC, thus providing evidence that the data support this as a measure of risk, relative to the 6 other measures tested.

Table A.2.1: The log-marginal likelihood estimate and BIC associated with each model where only an intercept and the risk measure are used as covariates in a probit model of bank survival.

Model and Risk Measure	Estimate	MargLik	BIC
<b>1. Composite Risk, R</b>	<b>-0.073 (0.002)</b>	<b>-15807.93</b>	<b>31588</b>
2. Loans/Deposits Ratio	-0.610 (0.029)	-16360.28	32715
3. Non-Cash Assets/Total Assets Ratio	-1.341 (0.043)	-16144.65	32268
4. Loans/Total Assets Ratio	-0.927 (0.044)	-16350.35	32696
5. Bonds/Total Assets Ratio	1.442 (0.044)	-16091.55	32162
6. ROA, Surplus & Profits /Total Assets	5.801 (0.163)	-16155.61	32268
7. Total Assets/Equity	0.009 (0.002)	-16520.83	33030

To further establish  $R$  as a valid measure of credit risk, we employed data on 1,304 credit ratings issued from major rating agencies (e.g., Standard and Poors) on U.S. firms traded on the NYSE or Nasdaq.<sup>15</sup> For this sample, we construct our composite credit risk measure  $R$  and compare it to the issued credit rating. Table A.2.2 shows the average value of  $R$  over the rating subgroups. We can see that  $R$  grows as the credit quality worsens, which demonstrates that our credit rating proxy aligns with actual credit ratings. We also see improved performance in using  $R$ , relative to  $C$  or  $L$ .

<sup>15</sup>Data sourced from <https://www.kaggle.com/agewerc/corporate-credit-rating>.

Table A.2.2: Credit ratings and  $R$  displayed for a sample of modern U.S. firms.

Rating	$n$	$C$	$L$	$R$
AAA, AA	83	7.382	2.773	19.015
A	337	9.162	3.080	27.049
BBB	492	11.028	3.185	34.379
BB	266	12.082	3.181	39.347
B	113	12.362	3.664	45.479
CCC, C	13	12.228	3.846	45.826

### A.3 Bank Exits

Between 1929 and 1934, over 9,000 banks exited the network, which is largely attributable to failures and mergers. Figure A.3.1 presents a histogram of all bank exits as a function of (log) total assets in 1929. For comparison purposes, the 25<sup>th</sup>, 50<sup>th</sup>, and 75<sup>th</sup> percentiles of the asset distribution for all banks (not just exits) are displayed by the red dashed lines. Clearly, most of the banks that exited the network were smaller than the median-sized bank in the system.

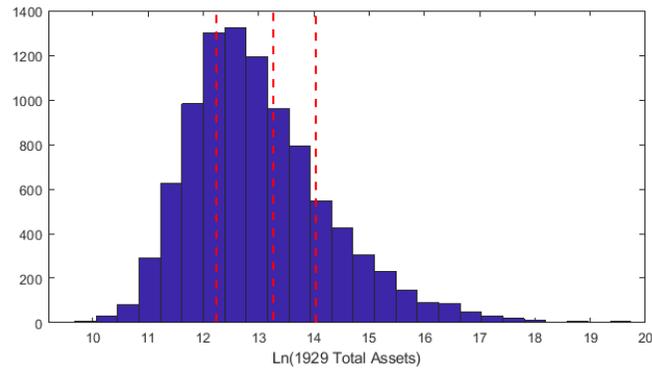


Figure A.3.1: Histogram of bank exits from the network by 1934 as a function of (log) total assets in 1929. The red dashed lines represent the 25<sup>th</sup>, 50<sup>th</sup>, and 75<sup>th</sup> percentiles of the asset distribution of the entire banking population (not just exits) in 1929.

## A.4 Robustness

Table A.4.1 shows the robustness of our results to additional specifications and estimation methods. We present the maximum likelihood results for  $M_1$ ,  $M_2$ , and  $M_3$ , which align closely with the posterior means from the MCMC estimation. Additionally, we report the information criteria associated with each model. The model rankings from the information criteria are almost the same as the marginal likelihood rankings; however, models  $M_2$  and  $M_3$  are very close based on the *BIC*, whereas the marginal-likelihood strongly favors  $M_3$ . Given that the marginal likelihood can be interpreted as a measure of *sequential out-of-sample predictive fit*, unlike the *BIC*, we focus on the  $M_3$  results. Column  $M_3$ -S.E. reports the MLE results with robust standard errors clustered at the county level, which do not change the statistical significance of our parameter estimates.

The column  $M_4$  reports the results for a model that includes state fixed effects. The main results are preserved, but the information criteria suggest overfitting.<sup>16</sup> Therefore, we focus our analysis on  $M_3$ . While omitted variables are always a concern in large-scale analyses, the overfitting that stems from the models with state fixed effects suggests that our balance sheet, correspondent network, county, peer, and Federal Reserve measures are comprehensive at explaining bank survival, outside of state geography.

Lastly,  $M_5$  reports the results for a model that includes endogenous peer effects. In his discussion of the “reflection problem,” Manski (1993) introduces exogenous and endogenous peer effects. In our context, the endogenous effect may arise because the propensity for an individual bank to survive varies within the survival of its network neighbors. Typically, the endogenous peer effects present significant challenges for modeling and estimation. In the standard linear-in-means model, identification problems arise from the perfect collinearity between the expected mean outcome of the group and its mean characteristics (Manski, 1993; Barmouille et al., 2009). However, in a binary choice model, like those in equations (7)–(9), this identification problem does not exist because of the nonlinear link function between group means and group behaviors, equation (10) (Brock and Durlauf, 2001).

We define the average survival of network neighbors as:

$$\text{Network Neighbor Average } y_i = \frac{1}{p_i}(A_{*,i}^\top \cdot y) + \frac{1}{q_i}(A_{i,*} \cdot y), \quad (18)$$

where  $p_i$  and  $q_i$  represent the number of respondents and correspondents for bank  $i$ , respectively. By including this variable, one may be concerned with serial correlation bias. The time series literature notes that this bias is quite small for low levels of autocorrelation (Griliches, 1961, 1967; Keele and Kelly, 2006). The network literature expands on this work by demonstrating that the bias is nearly 0 for low density, sparse networks (Mizrucki and Neuman, 2008, 2010). With our number of connections equalling 70,696 and number of banks equalling 25,684, our network is extremely sparse with density near 0, implying that serial correlation bias would be very minimal in our setting and supporting the consistency of our results. We include the average survival of banks in the same county to capture spatial peers. The results for  $M_5$  confirm that both spatial peer effects and network peer effects

---

<sup>16</sup>In the MCMC estimation of this model, the marginal likelihood is -15075.04, giving the model with state fixed effects a posterior model probability of 0 and agreeing with the *BIC*.

influence bank survival. These variables are not included in the main specifications of the paper because they are not based on 1929 values and instead incorporate forward-looking information, which obfuscates the predictive purpose of our work.

**Reconstruction Finance Corporation:** As mentioned in Section 5.3, a government-sponsored rescue program, the Reconstruction Finance Corporation (RFC), was established in 1932 with the objective of assisting weak banks. To ensure our network results are robust to a specification that controls for RFC assistance, we input a variable for the amount of RFC assistance into  $M_3$ . Because the RFC data need to be hand-collected from records in the National Archives, we restrict our sample to five states – Alabama, Arkansas, Michigan, Mississippi, and Tennessee – the sample employed in [Vossmeyer \(2016\)](#).

Table A.4.2 presents the results. Many of our findings are robust to a specification that controls for RFC assistance – negative effects from network neighbor risk, positive effects from network position, and positive branching effects but a negative interaction with eigenvector centrality. We note some differences in statistical significance, but that is likely due to the 5 state sample, where none of these states allow full branching and none of the banks have a central network position in a central reserve city.

These results should be taken with caution. Because the RFC didn't exist until 1932 (after the first banking panics) and because the selection process to receive RFC assistance was nonrandom (banks had to apply for assistance and governing authorities had to approve assistance), the results in this specification suffer from sample selection issues. [Vossmeyer \(2016\)](#) notes large biases in a study of RFC assistance that does not properly control for the multiple selection mechanisms and treatment-response outcomes. A full analysis of the RFC and the correspondent network would require the researcher to disentangle dynamics in the timing of the assistance and confront sample-selection issues involved with applying for and receiving assistance. Such a study presents an interesting possibility for future research.

Table A.4.1: Maximum likelihood estimates and standard errors are presented for  $M_1$ ,  $M_2$ , and  $M_3$ . Column  $M_3$ -S.E reports the MLE results with robust standard errors clustered at the county level. Column  $M_4$  reports the results for a model with state fixed effects. Column  $M_5$  reports the results for a model with endogenous peer effects. \*\*\* implies significance at the 1% level, \*\* implies significance at the 5% level, \* implies significance at the 10% level.

	$M_1$ -MLE	$M_2$ -MLE	$M_3$ -MLE	$M_3$ -S.E.	$M_4$	$M_5$
Intercept	-1.169 (0.188)***	-0.205 (0.209)	-0.289 (0.214)	(0.295)	-0.568 (0.404)	-2.844 (0.234)***
Loans/Assets	0.205 (0.085)**	-0.037 (0.086)	-0.053 (0.087)	(0.095)	-0.139 (0.090)	-0.202 (0.093)**
Bonds/Assets	1.282 (0.092)***	1.023 (0.095)***	0.993 (0.095)***	(0.104)***	0.899 (0.099)***	0.569 (0.101)***
LnAssets	0.183 (0.010)***	0.156 (0.011)***	0.160 (0.012)***	(0.013)***	0.163 (0.012)***	0.148 (0.013)***
Fed Member	0.027 (0.020)	0.038 (0.020)*	0.042 (0.021)**	(0.024)*	0.043 (0.021)**	0.028 (0.022)
Branch Intensity	0.712 (0.224)**	0.831 (0.228)**	1.067 (0.261)**	(0.314)**		0.165 (0.280)***
Num Banks in County	-0.005 (0.004)	-0.018 (0.008)**	-0.019 (0.008)**	(0.004)**	-0.014 (0.008)	-0.070 (0.008)***
Deposits/County Deposits	0.352 (0.070)***	0.453 (0.075)***	0.457 (0.075)***	(0.079)***	0.484 (0.077)***	1.088 (0.095)***
Central Reserve City	-1.060 (0.109)***	-1.008 (0.114)***	-1.021 (0.114)**	(0.163)***	-0.908 (0.118)**	-0.217 (0.116)*
Reserve City	-0.356 (0.045)***	-0.322 (0.046)***	-0.307 (0.047)***	(0.072)***	-0.317 (0.048)***	-0.327 (0.048)***
LnPopulation	-0.097 (0.014)***	-0.098 (0.017)***	-0.094 (0.017)***	(0.024)***	-0.095 (0.019)***	-0.096 (0.018)***
Manufact per capita	0.469 (0.111)***	0.521 (0.114)***	0.541 (0.114)***	(0.168)***	0.136 (0.126)	0.038 (0.121)
Cropland per capita	-0.003 (0.001)**	-0.003 (0.001)**	-0.003 (0.001)**	(0.001)**	-0.002 (0.001)	-0.003 (0.001)**
Composite Risk		-0.059 (0.002)***	-0.059 (0.002)***	(0.002)***	-0.059 (0.002)***	-0.056 (0.002)***
Network Neighbor out-R		-0.012 (0.008)	-0.016 (0.008)**	(0.0008)**	-0.011 (0.008)	-0.004 (0.009)
Network Neighbor in-R		0.000 (0.002)	-0.012 (0.004)**	(0.005)***	-0.013 (0.004)***	-0.012 (0.004)***
County-R		0.007 (0.005)	0.007 (0.005)	(0.006)	0.007 (0.006)	0.053 (0.005)
EigenCent			4.393 (1.952)**	(1.978)**	5.890 (2.020)***	2.538 (2.068)
SysRisk Percent			2.497 (0.807)***	(0.996)**	2.412 (0.816)***	2.287 (0.833)***
Branch Intensity $\times$ EigenCent			-4.370 (2.424)*	(2.513)*	-3.877 (2.600)	-1.736 (2.498)
Network Neighbor average y						0.063 (0.028)**
County average y						3.151 (0.056)***
Fed. Dist. FE	Yes	Yes	Yes		Yes	Yes
State FE	No	No	No		Yes	No
$n$	24,761	24,761	24,761		24,761	24,761
$BIC$	30646.76	29992.74	30000.46		30138.56	26146.35
LnLikelihood	-15201.97	-14854.73	-14843.42		-14679.84	-12906.25

Table A.4.2: RFC Specification.

	$M_3$ with RFC Amount
Intercept	-1.727 (0.794)**
Loans/Assets	-0.617 (0.259)**
Bonds/Assets	0.293 (0.305)
LnAssets	0.192 (0.045)***
Fed Member	-0.184 (0.074)**
Branch Intensity	3.168 (0.654)***
Num Banks in County	-0.132 (0.070)*
Deposits/County Deposits	0.261 (0.243)
Reserve City	-0.224 (0.195)
LnPopulation	0.036 (0.073)
Manufact per capita	0.267 (0.530)
Cropland per capita	0.032 (0.020)
Composite Risk	-0.066 (0.008)***
Network Neighbor out-R	-0.041 (0.028)
Network Neighbor in-R	-0.032 (0.015)**
County-R	-0.016 (0.024)
EigenCent	15.629 (10.065)
SysRisk Percent	7.158 (5.972)
Branch Intensity $\times$ EigenCent	-18.365 (11.258)*
ln(RFC Amount)	0.070 (0.005)***
Fed. Dist. FE	Yes
$n$	2,366

Probit model that includes a variable for RFC Assistance. Maximum likelihood estimates and standard errors (in parentheses) are presented.

## A.5 Ordered Specification

To obtain results for the population of commercial banks, our outcome variable combined failures and mergers, but these outcomes clearly differ. We lack detailed information on survival outcomes of state banks; however, for national banks, we are able to further distinguish between these types of exits using data from the *Annual Report* of the Comptroller of the Currency (1929-1934). All national banks are, by definition, members of the Federal Reserve System, and based on the full-sample results above, we know there are differences among Fed and non-Fed members; hence, the smaller sample is not representative and does not achieve our goal of understanding the entire mapping of the interbank system. The sample of national banks is nevertheless useful for testing the robustness of our model comparison results.

To do so, we again use three models with covariates that are the same as  $M_1$ ,  $M_2$ , and  $M_3$ . However, the outcome variable is now ordered and defined as:

$$y_i = \begin{cases} 3 & \text{Bank Survives} & \gamma_3 > y_i^* \leq \gamma_2 \\ 2 & \text{Bank Merges} & \gamma_2 > y_i^* \leq \gamma_1 \\ 1 & \text{Bank Fails} & \gamma_1 > y_i^* \leq \gamma_0, \end{cases} \quad (19)$$

and  $\varepsilon \sim N(0, 1)$ , i.e., an ordered probit model. 18% of the sample fail, 14% merge, and 68% survive. The models are estimated by Gibbs sampling methods (Algorithm 2 in Jeliazkov et al. (2008)), and the MLE parameter estimates align. The priors on  $\beta$  are centered at 0 with a standard deviation of 10 and the results are based on 5,000 MCMC draws with a burn in of 1,000. Marginal likelihood calculations follow from Chib (1995) and Jeliazkov et al. (2008).

Table A.5.1: Model comparison results for the three ordered models.

	$M_1$	$M_2$	$M_3$
$n$	7,372	7,372	7,372
Log-Marginal Likelihood	-7504.02	-7296.38	-7294.22
Posterior Model Probability	$7.67 \times 10^{-92}$	0.115	0.885

Table A.5.1 shows the model comparison results for the ordered probit models. As was true for the full population of commercial banks, a comparison of marginal likelihood reveals strong support for  $M_3$ , with a posterior model probability of 88.5%. These results align with the main findings in the paper and point to an additional important implication: a specification including our new measures not only improves the prediction of bank failures, but also those for bank mergers, an important observed outcome in most financial crises.

## A.6 Distribution of Shocks

In Section 5.3.1, we present shocks to the banking system, including balance-sheet shocks, network shocks, and regulation shocks. The numerical results presented in that section are averaged over both the MCMC draws and the sample to accommodate estimation and sampling uncertainty. Figure A.6.1 presents histograms of the distributions of the average effects as a function of parameter uncertainty. The balance sheet (R and in-out measures), random network, and branch banking shocks are statistically different from 0, whereas the scale-free network shock is not.

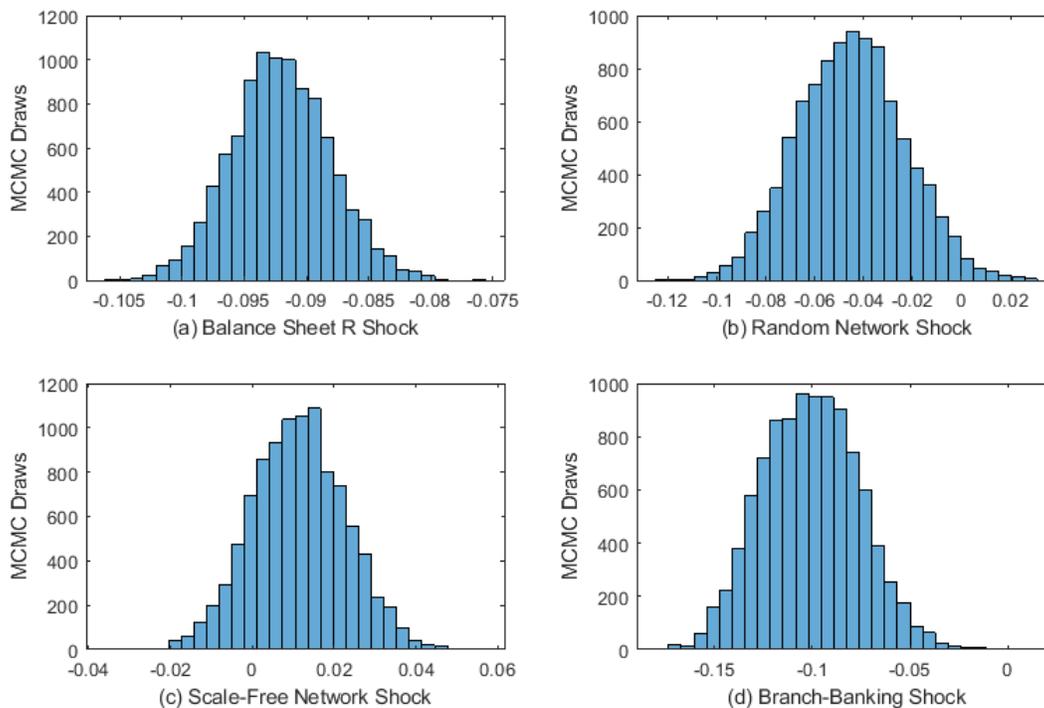


Figure A.6.1: Distributions of the average effect from a shock as a function of parameter uncertainty: (a) Balance Sheet, R, Shock, (b) Random Network Shock, (c) Scale-Free Network Shock, and (d) Branch-Banking Shock.

## A.7 Data Sources and Data Cleaning

Our primary data source is the *Rand McNally Bankers Directory*. We digitized the July 1929 issue, leading us to our sample of financial institutions. The directory reports the balance sheets of all bank and trust companies, including private, state, and national banks. It does not include information on building and loan associations. In constructing the sample, we omit any bank that was operating on a restricted basis, in receivership or liquidation, or going through a consolidation or reorganization. In the larger cities, *Rand McNally* often reports information on non-depository institutions (investment and securities companies) and foreign banks. We omit these observations.

In constructing the network data, we remove rows and columns of the adjacency matrix where banks did not have relationships with another bank. This drops 5% of the banks in the sample. However, of these, only 181 banks had financial data, so we drop less than 1% of the usable data. (Financial data was often missing for the reasons listed in the first paragraph.) Of the 181 observations without network links but with financial data, most are smaller, private banks or non-standard banks (e.g., Morris Plan Companies, mutual savings, title companies, etc). The total number of unique banks in the network is 25,684.<sup>17</sup>

While network metrics are computed for all 25,684 nodes in the network, balance-sheet data were less populated and some nodes had incomplete or inaccurate data. We omit banks that had potential errors reported on their balance sheet (e.g., cases where total assets and total liabilities were dramatically different). Using the total nodes data, the following balance sheet characteristics or ratios attributes are more limited than total nodes (with the number of available observations reported in parentheses): assets (25,067), equity (25,302), and assets and equity (25,061). Lastly, in the econometric analyses, we merge in county data and other bank characteristics. The county data are from the 1930 U.S. Censuses of Population, Manufacturing, and Agriculture. The other bank characteristics (Fed membership, loans, and bonds) are obtained from *Rand McNally*. The branching intensity variable is from [Carlson and Mitchener \(2006\)](#). The sample that includes the full set of network, balance sheet, county, and other information is 24,761. We computed network statistics where we set  $R$  to the mean for banks with missing or inaccurate balance sheet information. Doing so does not materially change our results.

---

<sup>17</sup>The FDIC reports that in 1929 the total number of charter-holding banks in 1929 was 25,733. The median value of total assets for networked banks is 501,000. For the 181 non-networked banks, it is 346,000.