Daniel Fricke, Tarik Roukny

Generalists and specialists in the credit market

Article (Accepted version) (Refereed)

Original citation:
Fricke, Daniel and Roukny, Tarik (2018) *Generalists and specialists in the credit market.* Journal of Banking and Finance. ISSN 0378-4266

DOI: 10.1016/j.jbankfin.2018.04.014

Generalists and Specialists in the Credit Market

Daniel Fricke, Tarik Roukny

PII: S0378-4266(18)30087-6
DOI: 10.1016/j.jbankfin.2018.04.014
Reference: JBF 5335

To appear in: *Journal of Banking and Finance*

Received date: 30 August 2016
Revised date: 25 October 2017
Accepted date: 16 April 2018

Generalists and Specialists in the Credit Market

Daniel Fricke* **Tarik Roukny†**

Abstract

In this paper, we propose a method to analyze the structure of the credit market. Using historical data from Japan, we explore banks' lending patterns to the real economy. We find that generalist banks (with diversified lending) and specialist banks (with focused lending) coexist, and tend to stick to their strategies over time. Similarly, we also document the coexistence of generalist and specialist industries (based on their borrowing patterns). The observed interaction patterns in the credit market indicate a strong overlap in banks' loan portfolios, mainly due to specialist banks focusing their investments on the very same generalist industries. A stylized model matches these patterns and allows us to identify economically meaningful sets of generalist banks/industries. Lastly, we find that generalist banks are not necessarily less vulnerable to shocks compared to specialists. In fact, high leverage levels can undo the benefits of diversification.

Keywords: bank lending, portfolio theory, fire sales, diversification, systemic risk.

JEL Classification: G11, G21, G28, G32

*Corresponding author. University College London; London School of Economics, Systemic Risk Centre; and University of Oxford, Saïd Business School. Email: danielfrickeoxford@gmail.com

†Massachusetts Institute of Technology; Université Libre de Bruxelles; Fonds National de la Recherche Scientifique. Email: roukny@mit.edu

1 Introduction

Understanding the structure of the credit market is paramount to ensure the role of the banking system as an efficient liquidity and credit allocation mechanism. Failure to manage and regulate the banking system can in fact have disastrous externalities, as exemplified by experiences of financial crises (Hoggarth et al. (2002); Dell'Ariccia et al. (2008)). The shape of credit interactions between banks and the real economy is, therefore, a key element in the analysis, but little is known about how this credit network looks like. Do most banks hold diversified loan portfolios and therefore provide liquidity to firms from all industries of the economy? Are there specialists that focus their portfolio on a small number of industries? If so, do different specialists focus on different industries? Are there significant differences in the risk profiles of these institutions? We tackle these questions in this paper.

Whether banks should diversify their loan portfolios or focus on a small number of industries is an important yet open research question. For what follows, we define *generalist banks* as those banks that diversify their loan portfolios across many different industries, thereby interacting with a very heterogeneous set of firms. We also define *specialist banks* as those banks that hold more concentrated portfolios and only interact with firms from a relatively small subset of industries.[1] Like other investors, banks face a trade-off in choosing their diversification levels. On the one hand, generalist banks should be, through the benefits of diversification, less vulnerable to firm- or industry-specific shocks. On the other hand, gaining industry-specific expertise, e.g., via more efficient screening and monitoring of particular types of firms, is valuable to banks (Stomper (2006)). By focusing on relatively

[1] Both the theoretical and the empirical literature offer mixed recommendations on whether banks should be generalists or specialists. On the theoretical side, Diamond (1984) finds that the benefits from delegated monitoring are maximized when banks are completely diversified, whereas Stomper (2006) shows that generalists and specialists coexist in equilibrium. Winton (1999) assumes that the gains from diversification depend on the riskiness of the bank, finding that medium-risk banks should diversify, while low- and high-risk banks should specialize. On the empirical side, Acharya et al. (2006) find that the predictions of Winton (1999) appear to hold for Italian banks, while Hayden et al. (2007) find the exact opposite relationship for German banks.

few types of businesses, specialist banks might therefore be able to improve their performance at the cost of becoming more vulnerable to industry-specific shocks. In addition, banks face time-varying external constraints (e.g., regulatory) that may further encourage either diversification or specialization.[2] From this perspective, it is not surprising that banks' diversification levels are found to be rather heterogeneous empirically (e.g., Acharya et al. (2006); Hayden et al. (2007); Tabak et al. (2011)). Nevertheless, little is known about the typical pattern of interactions between banks and the real economy, and the prevalence of generalists and specialists in these systems.

A related question is whether specialist banks are indeed *special*, i.e., do specialist banks tend to occupy niches? In a world where specialist banks possess comparative advantages in dealing with firms from certain industries (e.g., via gaining superior information about their counterparties), they should focus their activities on specialist industries where few other banks are present. In this regard, recent research highlights the role of overlapping portfolios as a potential source of systemic risk (Wagner (2011); Caccioli et al. (2014); Greenwood et al. (2015); Glasserman and Young (2015)). The idea is that, by holding common assets, banks are not only prone to the same direct shocks, but also to systemic asset liquidations of other banks. Little is known, however, about how overlapping bank portfolios are in the real world.

This paper fills these gaps by proposing a method to identify and analyze the coexistence of generalists and specialists in detail. We apply our method to a dataset containing Japanese banks' industrial loan portfolios over the period 1980–2013. Given the strong wave of institutional changes that have taken place over this sample period, the Japanese banking system is a particularly interesting case study.

As an illustration for our main finding, a network representation of the credit interactions between banks and the real economy is shown in Figure 1 for one particular year. In the Figure, banks (industries) are shown as nodes on the left (right) and a connection is drawn

[2]For example, many regulatory frameworks impose upper limits on a bank's exposure to individual borrowers (e.g., BIS (2014)).

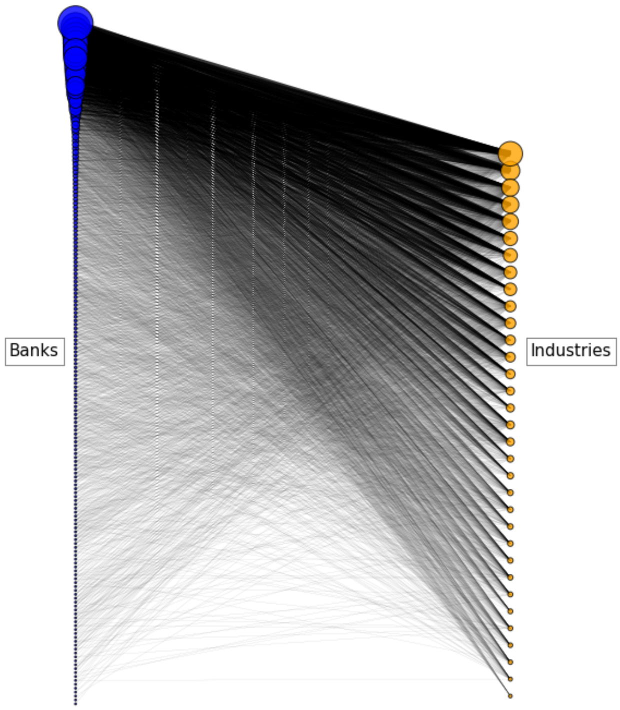

Figure 1: The Japanese credit network in 1980 (binary version). Banks are depicted as blue nodes on the left and industries as yellow nodes on the right. Banks and industries are connected through borrowing and lending activities on the credit market - these links are shown in black. Nodes are arranged according to their number of connections.

between two nodes if a bank provided loans to firms from that particular industry. First, it is clear that a similar dichotomy of generalists and specialists can be applied to both the credit supply side (banks, nodes on the left), and the credit demand side (industries, nodes on the right): in Figure 1 nodes are sorted according to their number of connections, such that generalists (specialists) are closer to the top (bottom) of the Figure. The coexistence of generalists and specialists is hardly surprising, but their interaction patterns are: Figure 1 shows that specialist banks and specialist industries rarely interact with each other (nodes closer to the bottom of Figure 1 are almost not connected to one another). Hence, specialist banks mainly interact with firms from generalist industries, and similarly specialist industries mainly interact with generalist banks. This implies a significant overlap in banks' loan portfolios: most specialist banks focus their lending activities on the very same parts of the real economy. These findings suggest that there is nothing special about specialist banks.

We introduce a stylized model of generalists and specialists that matches the patterns shown in Figure 1. The two main model assumptions are that generalists (specialists) should have as many (few) connections as possible. We relate the data to the idealized network structure that would prevail if these assumptions were true, allowing us to classify generalists and specialists over time. The model fit is significant and generally superior relative to various network randomizations. Interestingly, the fit slowly deteriorates over time as a result of a progressive shrinking of the size of the Japanese banking system. However, this trend is remarkably weak when taking into account the dramatic institutional changes that took place over the sample period (culminating in the so-called 'Big Bang'). Our findings remain qualitatively similar when using a continuous version of our generalist-specialist model, where each node has its own affinity to be a generalist.

The architecture of the credit network is very persistent over time, largely due to the fact that nodes' strategy profiles (i.e., being a generalist or a specialist) are very stable. In this regard, we also explore features that distinguish generalists from specialists using additional node-specific information. For banks, we find that bank type and size are the most important determinants. For industries, we find that, while size is also of major predictive power, additional factors, such as geographical constraints, play an important role. Lastly, we calculate banks' vulnerability to liquidation shocks (à la Greenwood et al. (2015)) and analyze how a bank's riskiness depends on its network position. We find that generalist banks tend to be significantly less vulnerable, at least when they are not highly leveraged. Hence, banks' position in the credit network can be informative about their riskiness from a systemic perspective.

Our paper makes several contributions. First, we introduce different models of generalists and specialists that allow us to describe the interactions between banks and the real economy in a simple way. In fact, this paper is the first to analyze the peculiar interaction patterns in the credit market. Our methodology classifies nodes into generalists and specialists solely based on the observed credit network. In line with recent findings for interbank networks (Craig and von Peter (2014)), generalist banks and industries form the core of the system and their activity accounts for a large part of the monetary flows in the economy. Our approach helps policymakers and regulators understand and monitor the evolution of the structure of the banking system, particularly so in the presence of changes in institutional features.

Second, the fact that specialists tend to concentrate their activities on generalists opposes the idea of specialists' expert knowledge in isolated niches. Indeed, the relatively small number of specialist-specialist interactions indicates a strong overlap in banks' loan portfolios, since specialists invest where generalists invest as well. While a significant portfolio overlap alone does not rule out the possibility that specialist banks do possess superior information, we find no evidence that specialist banks outperform generalists.

Third, the observed interaction pattern between banks and the real economy carries important implications for the literature on the determinants of the number of credit relationships per firm (Guiso and Minetti (2010)). The way banks specialize into certain industries makes firms' industry affiliations an important determinant for the number of bank relationships (Ongena and Yu (2016)).

Lastly, as a first step towards understanding network formation, we relate nodes' network position and their individual characteristics. In line with the expectation that different bank/industry types are likely to build different patterns of links, we find that a small set of node-specific variables reliably predicts whether a node will be a generalist.

The remainder of the paper is structured as follows: section 2 provides the necessary networks background for the generalist-specialist model which is introduced in section 3. A continuous version of the model is developed in Section 4. Section 5 contains a brief overview of the Japanese commercial banking system. In section 6, we present the empirical results on the generalist-specialist model, explore the features of generalists, and relate network position with banks' performances and systemic vulnerabilities, respectively. Section 7 summarizes the main results and concludes. The Appendix contains extensive additional information, most importantly robustness checks on the application of the generalist-specialist model.

2 The Credit Network

The credit network consists of two distinct sets of nodes: the first set contains a total number of n^B nodes (banks), and the second set n^I nodes (industries). A connection (link) exists between a bank and an industry when there is a credit relationship between two nodes.[3] Technically, the network is bipartite (also called two-mode in the social networks literature, see Jackson (2008)), since links can only arise between banks and industries.

To be precise, we represent banks' industrial loan portfolios as a weighted incidence matrix of dimension $(n^B \times n^I)$, which we will denote as $\mathbf{W}$. An element $w_{i,j}$ of this matrix represents the total value of credit extended by bank i to firms from industry j at a given point in time (time-indices dropped for convenience). The value of $w_{i,j}$ can thus be seen as a measure of link intensity. From the weighted matrix $\mathbf{W}$ we obtain the binary incidence or adjacency matrix that will be of major interest in everything that follows. We denote this matrix as $\mathbf{A}$, where $a_{i,j} = 1$ if $w_{i,j} > 0$, and $a_{i,j} = 0$ otherwise. In other words, this matrix informs on the link existence between a bank and an industry. We will refer to matrix $\mathbf{A}$ as the *credit network*. Note that each node needs have at least one link (minimum degree of one) in order to be considered as part of the network at any point in time. Technically, $(\sum_j a_{i,j}) > 0 \ \ \forall i$, and $(\sum_i a_{i,j}) > 0 \ \ \forall j$. The total number of links in the credit network is denoted as m:

$$m = \sum_{i=1}^{n^B} \sum_{j=1}^{n^I} a_{i,j}.$$

In our dataset we observe yearly snapshots of the credit network (matrix $\mathbf{A}$). Our focus

[3]Note that the credit network is aggregated in the sense that banks interact with firms, which themselves are affiliated with an economic industry.

on the bank-industry network, rather than on more micro-level bank-firm interactions, is justified by the fact that banks, like other investors, are likely to seek diversification benefits by investing in different industries. Similar to Ibragimov et al. (2011), we think of the n^I different industries as risk classes. In addition, there is an inherent asymmetry in the size of counterparties on the two sides: banks' business model involves dealing with a relatively large number of firms, whereas firms are well-known to interact with few banks at any given point in time.[4] Therefore, the total number of banks that provide loans to a whole industry is an indicator of how diversified the funding of a given industry is. Lastly, exploring the effects of industry-specific shocks, rather than firm-specific shocks, is also in line with recent work on input-output networks (e.g., Acemoglu et al. (2012)).

3 The Generalist-Specialist Model

In this section we propose a stylized model of bank-industry credit interactions. The model is an extension of the core-periphery model of Borgatti and Everett (2000) - which was successfully applied to unipartite interbank networks (e.g., Craig and von Peter (2014)) - for the case of bipartite networks.

In the model, nodes (banks and industries, respectively) can be either of two types: generalists or specialists. In line with the idea that *generalists* are highly diversified, we require that they have as many links as possible. Hence, generalists interact with all nodes from the other set; for example, a generalist bank extends loans to firms from all kinds of industries and a generalist industry borrows from all kinds of banks.[5] On the other hand, *specialists* should have as few links as possible and will only interact with a smaller subset of

[4]For example, Detragiache et al. (2000) report that the median number of bank relations for small businesses in the US and Italy is 2 and 5, respectively, and the share of firms with only one bank relationship is 44.5% and 11%, respectively.

[5]Note that this definition does not require that individual firms from a given generalist industry should be generalists as well, i.e., borrow from a large number of banks. In fact, we do not know a priori whether a given firm is part of a generalist or specialist industry given the industry-level definition of generalists and specialists on the borrowing side.

nodes. As we will show below, the stylized credit network that arises under these assumptions shares many of the features shown in Figure 1.

An additional advantage of the model is that it classifies generalists and specialists, without having to specify an arbitrary cutoff value in terms of what makes a generalist.[6] In section 4, we develop a more flexible version of the generalist-specialist model, where each node has its own affinity to be a generalist.

3.1 Setup

We seek to decompose the two sets of nodes (n^B banks and n^I industries) into subsets of generalists and specialists, respectively. Economically, we make the following assumptions about generalists and specialists:

Assumption 1 *Generalist banks (industries) should interact with as many industries (banks) as possible.*

Assumption 2 *Specialist banks (industries) should interact with as few industries (banks) as possible.*

Assumption 1 states that generalists are maximally connected, such that

$$\sum_j a_{i,j} = n_g^I + n_s^I = n^I \quad \text{if} \quad i \in B_g$$

$$\sum_i a_{i,j} = n_g^B + n_s^B = n^B \quad \text{if} \quad j \in I_g,$$

where B_g and I_g are the sets of generalist banks and industries of size n_g^B and n_g^I, respectively. Given that each node must be connected to at least one other node, the second assumption

[6]For example, one might define generalists as the x most highly connected (or diversified) nodes, but it is unclear what a good value for x would be. Our approach is model-/data-driven and does not rely on such an ad-hoc cutoff value.

implies that

$$\sum_j a_{i,j} = n_g^I \quad \text{if} \quad i \in B_s$$

$$\sum_i a_{i,j} = n_g^B \quad \text{if} \quad j \in I_s,$$

where B_s and I_s are the sets of specialist banks and specialist industries respectively. Based on these results, after sorting generalists and specialists accordingly, the following idealized pattern matrix ($\mathbf{A}^*$) for a 'pure' generalist-specialist segmentation arises:

$$\mathbf{A}^* = \begin{pmatrix} \mathbf{GG} & \mathbf{GS} \\ \mathbf{SG} & \mathbf{SS} \end{pmatrix} = \begin{pmatrix} \mathbf{1} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{pmatrix}, \tag{1}$$

where $\mathbf{1}$ and $\mathbf{0}$ denote submatrices of ones and zeros, and G stands for generalists and S for specialists (where we drop the B and I superscripts for convenience), respectively. For example, the **GG**-block (of dimension $n_g^B \times n_g^I$) contains the subset of highly interconnected generalist banks and generalist industries. Under our assumptions, all blocks except for the **SS**-block should be maximally connected since generalists will be connected to all generalists and specialists from the other set. On the other hand, the **SS**-block (of dimension $n_s^B \times n_s^I$) contains the two sets of specialists which should be minimally connected (i.e., contain as few links as possible). Figure 2 shows an illustration of a small credit network according to Eq. (1). Note that generalists are connected with all nodes from the other side, while specialists only interact with the generalists of the other side. Thus, specialists from the two different sets of nodes are unconnected and in this sense, the model should be able to match the empirical patterns in Figure 1.

3.2 Optimization

In the following, we use the discrete generalist-specialist framework to classify banks/industries as generalists and specialists, respectively. This classification can be summarized by two 'gen-

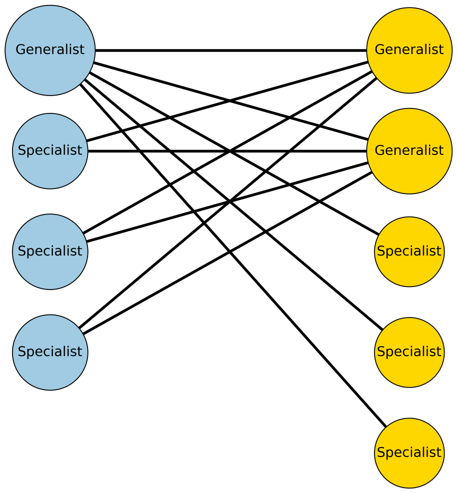

Figure 2: Illustration of the network structure implied by the generalist-specialist model in Eq. (1). Generalist nodes from each side interact with all nodes from the other side, while specialists only interact with generalists.

eralist level' vectors, γ^B and γ^I (of length n^B and n^I, respectively). For a generalist bank, $\gamma_i^B = 1$, and 0 otherwise. Similarly, $\gamma_j^I = 1$ for a generalist industry, and 0 otherwise.

We aim at finding the optimal generalist-level vectors, i.e., partitions of generalists and specialists for which the observed network is as close as possible to the idealized pattern matrix in Eq. (1). In line with previous work on unipartite networks (Craig and von Peter (2014)), we measure the 'fit' of the corresponding generalist-specialist structure as the total number of inconsistencies between the observed network and the idealized pattern matrix $\mathbf{A}^*$ of the same dimension. Residuals are obtained by counting the errors in each of the four blocks of Eq. (1) and aggregating over the blocks. For the general version of the generalist-specialist model, the aggregate errors of the individual blocks can be written as

$$\mathbf{E}(\gamma^B, \gamma^I) = \begin{pmatrix} E_{GG} & E_{GS} \\ E_{SG} & E_{SS} \end{pmatrix}, \tag{2}$$

where γ^B and γ^I are the generalist-level vectors, as defined above. According to the idealized pattern matrix above, we require all but the specialist-specialist block to be maximally connected. Therefore any missing link in those blocks is counted as an error, whereas the SS block should be minimally connected and we count any existing link in this block as an error.

The total error score (e) simply aggregates the errors across all blocks, normalized by the total number of links in the network, m, to facilitate comparison over time. Formally this can be written as

$$e(\gamma^B, \gamma^S) = \frac{E_{GG} + E_{GS} + E_{SG} + E_{SS}}{m} \tag{3}$$

with $e(\cdot)$ being a function of the generalist level vectors since every possible partition is associated with a particular value of e. In the absence of a generalist-specialist structure, i.e., in a system with only specialists, the normalized error score would take a value of one (see Appendix D for details), such that a 'significant' generalist-specialist structure should display much lower error scores. We minimize the error score using a plain vanilla genetic algorithm with the typical elements (reproduction, crossover, mutation, and election).[7]

4 The Continuous Generalist-Specialist Model

Our baseline model, namely the partition-based approach of the discrete generalist-specialist model presented above, might appear somewhat restrictive. A reasonable alternative is to consider a continuous model in which each node has its own affinity to be a generalist. This extension also allows for weighted interaction matrices (i.e., the weighted credit network $\mathbf{W}$, rather than $\mathbf{A}$). In the following, we briefly develop a continuous model which is based on the singular value decomposition (SVD), where the generalist-levels for banks and industries then correspond to the leading left- and right-eigenvectors, respectively.

[7]We experimented with different parametrizations and found the results to be very stable, assuring us that we indeed manage to find the global minimum in a few seconds. More details on the algorithm are available upon request from the authors.

4.1 Setup

We aim at obtaining two generalist vectors, γ^B and γ^I, where $1 \geq \gamma_i^B \geq 0 \; \forall i \in \{1, \cdots, n^B\}$ and $1 \geq \gamma_j^I \geq 0 \; \forall j \in \{1, \cdots, n^S\}$, with idealized pattern matrix

$$\mathbf{W}^* = \gamma^B . \times \gamma^I, \tag{4}$$

where . denotes element-wise multiplication. This matrix should approximate the observed data matrix as closely as possible. Note that the continuous model is less restrictive compared to the discrete model presented above, as it explicitly allows for specialist-specialist interactions. These should, however, be weaker than connections in the other blocks.

4.2 Estimation

We rely on a SVD to estimate the continuous generalist-level vectors. Let $\mathbf{W}$ be a real $(n^B \times n^I)$ matrix with $n^B \geq n^I$ (of rank n^I). Matrix $\mathbf{W}$ can be decomposed as

$$\mathbf{W} = \mathbf{\Gamma^B \Sigma \Gamma^I}, \tag{5}$$

where $(\mathbf{\Gamma^B})^\mathbf{T}\mathbf{\Gamma^B} = (\mathbf{\Gamma^I})^\mathbf{T}\mathbf{\Gamma^I} = \mathbf{\Gamma^I}(\mathbf{\Gamma^I})^\mathbf{T}$, $()^T$ denoting the transpose, and $\mathbf{\Sigma} = diag(\sigma_1, \sigma_2, \cdots, \sigma_n^I)$, where $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n^I \geq 0$. Matrix $\mathbf{\Gamma^B}$ consists of n^B orthonormalized eigenvectors associated with the n^B largest eigenvalues of $\mathbf{W}(\mathbf{W})^\mathbf{T}$, and the matrix $\mathbf{\Gamma^I}$ consists of the orthonormalized eigenvectors of $(\mathbf{W})^\mathbf{T}\mathbf{W}$. The singular values are the diagonal elements of $\mathbf{\Sigma}$, i.e., the non-negative square roots of the eigenvalues of $(\mathbf{W})^\mathbf{T}\mathbf{W}$.

Similar to the approach of Boyd et al. (2010), we can approximate the observed network using only the first k singular values. Here we set $k = 1$ and define the generalist-level of banks and industries as the corresponding eigenvectors of σ_1, such that

$$\mathbf{W}^* = \gamma^B \sigma_1 \gamma^I, \tag{6}$$

where $\gamma^B = \mathbf{\Gamma_1^B}$ and $\gamma^I = \mathbf{\Gamma_1^I}$. We evaluate the fit of the model using a R^2 measure

$$R^2 = 1 - \frac{SS(\mathbf{W}^* - \mathbf{W})}{SS(\mathbf{W} - \langle \mathbf{W} \rangle)}, \tag{7}$$

where $SS(\cdot)$ denotes the sum of all squared elements of the argument and $\langle \cdot \rangle$ denotes the average across all observations. Hence, maximizing R^2 is equivalent to minimizing $SS(\mathbf{W}^* - \mathbf{W})$, which is exactly the what the SVD does.[8]

5 Institutional Background

Before turning to the empirical results, we provide a brief overview of the institutional features of the Japanese banking system and its historical evolution.[9] As Uchida and Udell (2010) point out, there are several reasons why the Japanese banking system deserves to be studied in depth.

1. Japan is the world's third largest economy in terms of GDP and the banking system is an essential part of this economy.

2. Similar to several other important developed economies, such as Germany, Japan has historically been a banking-oriented financial system.

3. Like other countries, the Japanese banking system has been in a major transition since the bursting of the asset price bubble in the early 1990s. In fact, Hoshi and Kashyap (2010) highlight the analogies between the 1990s Japanese banking crisis and the 2008 US financial crisis. While certain features of Japan's financial system are certainly unique (such as the 'main bank' system), there are general lessons to be learned from its analysis.

[8]Note that the R^2 trivially equals 1 if we used all singular values in the approximation.

[9]A detailed description of the Japanese financial system and its history can be found in Itō (1992) and Hoshi and Kashyap (2004).

4. We crucially rely on reliable mirco-level data, and focusing on Japan is useful, since bank-firm interactions are recorded in commercial databases. Accessing data for other countries is much more cumbersome, as these are collected by supervisory institutions (e.g., Germany) or not available at all (e.g., US).

In the following, we provide some background on several of the above points.

5.1 The Japanese Financial System

Historically, the Japanese banking system has been segmented, mainly due to legally mandated specializations for different types of banks, and it still retains some of its original features. Japanese banks are segmented into different bank types, most importantly city banks, regional banks, tier-2 regional banks, long-term credit banks, and trust banks. These are the bank types we focus on in this paper.[10]

City banks have nationwide branches and provide wholesale lending to large corporate customers, accept individual deposits, and offer consumer loans. These banks dominate most segments of the domestic market, and are active internationally. Increased competition between banks, however, led city banks to also interact with small- and medium-sized firms (Shin and Kolari (2004)).

Regional banks (or tier-1 regional banks) are much smaller in scale than city banks and tend to have a regional focus. They primarily service small, regional firms, but also individuals. Most of their lending is directed to small- and medium-sized firms. Similar to city banks, regional banks are allowed to have nationwide branches, but the total number and location of these branches has to be approved by the supervisory authority.

[10] Of lesser importance are foreign banks, Shinkin banks, and credit cooperatives. Foreign banks' market share in the Japanese corporate lending market is traditionally very low. Shinkin banks are cooperative institutions, which conduct their banking businesses within their respective local area. Due to their mutual form, Shinkin banks provide services to their members, which are normally small- or medium-sized enterprises, and individuals. Credit Cooperatives conduct all their activities within their respective prefecture. Also note that both Shinkin banks and Credit Cooperatives are not considered commercial banks (BIS (2001)) and operate under a different judicial framework (Uchida and Udell (2010)). As such, we do not consider them in the rest of our analysis.

Tier-2 regional banks were initially established as mutual (Sogo) banks, but were transformed into regional banks under the 1992 Banking Act. These banks are smaller in scale than tier-1 regional banks, and their activities are normally confined to the prefecture in which their respective head offices are located.

Long-term credit banks supply, as their name suggests, long-term private credit. The key feature that distinguishes this bank type from city and regional banks is the long-term nature of their assets and liabilities. With the collapse of the Long-term Credit Bank of Japan in the early 2000s, this bank type went out of existence.

Finally, *trust banks* offer both financing and asset management services. They receive and manage funds on behalf of their depositors, where the investments are typically longer-term.

With increased deregulation in the 1980s, different bank types started competing with each other. Furthermore, the bursting of the asset price bubble in the early 1990s and its long-lasting impact on banks' balance sheets led to a restructuring of the entire banking system. Consolidation and numerous bank failures ultimately concluded the Japanese 'Big Bang' in the early 2000's, and nowadays the five remaining city banks (so-called Mega Banking Groups) dominate large parts of the market. Also, geographical segmentation is still likely to play a role, in particular for relationship loans where physical proximity is a major determinant for active interactions. For example, using data on a large Belgian bank, Degryse and Ongena (2005) find that the median distance between the bank and its borrowers is 1.40 miles.

5.2 The 'Main Bank' System

Banking relationships are far more important in Japan compared to many other countries. Japanese firms strongly rely on bank debt, although market-based financing has become more important since the deregulation period in the 1980s. The relationships between banks and firms, however, are much deeper than in many other countries. Firms typically have a main bank, which is not only the biggest lender, but often also holds equity shares in the firm and may also have representatives on the firms' corporate board.[11] Hence, relationships are generally very long-term oriented; for example, Uchida et al. (2008) report an average duration of Japanese bank-firm credit relationships of 30 years. The main bank is particularly important during times of distress, when it can require changes in the firm's management and its board of directors.[12]

5.3 Comparing the Japanese Banking System with the US and Germany

Let us provide a brief comparison between the Japanese commercial banking system and two other major banking systems: Germany and the United States. This comparison is useful for taking a dynamic perspective on the generalist-specialist model in the sense that the Japanese commercial banking system (like many other banking systems) underwent substantial changes over the last 30 years. Table 1 reports the main features for each of the three examples.

The German banking system is well-known for its *three-pillar approach.* This setup explicitly encourages the coexistence of generalists (mainly *private banks*) and specialists (cooperatives or *Volksbanken*, and public savings banks or *Sparkassen*), with the latter facing geographical restrictions regarding their activities (Hüfner (2010); Goddard et al. (2010)).

[11]Traditionally, different groups of banks and firms used to be part of the same *keiretsu* group. The *one-set-policy* of keiretsus led to their presence in every industry with a limited amount of firms in each industry in order to avoid intra-group competition (see Gerlach (1992) for an extensive overview). These groups constitute a particular feature of the Japanese system, but their importance has been repeatedly put to question (e.g., Miwa and Ramseyer (2002)). In our framework, we should expect keiretsu banks to act as generalists given that they are closely connected with firms from all kinds of industries within their group. Nevertheless, given the relatively small number of keiretsu banks (between 6 and 9 over our sample period, see Gerlach (1992)) this feature is likely to have only limited explanatory power in the overall structure of the banking system, in particular when it comes to the structure of specialist banks' loan portfolios.

[12]The literature has uncovered several dark sides of these close relationships. For example, firms may have trouble finding alternative funding sources, when its main bank is in distress, and the main bank could use its inside information to extract excessive rents from the firm. Moreover, Japanese banks misallocated credit by 'evergreening' loans to the weakest firms (Peek and Rosengren (2005); Caballero et al. (2008)).

	Japan	US	Germany
Number of banks (1985) (2006)	 140 117	 14,407 7,279	 4,739 2,050
Banking Assets/GDP (2000)	 1.31	 0.75	 1.19
Financing source	Banking	Market	Banking
Special features	1. Main bank - Keiretsu 2. Historically fragmented 3. Deregulation (90s)	1. Historically state-limited 2. Deregulation (80s)	1. Public credit services Volksbanken & Sparkassen 2. European integration

Table 1: Comparison between the Japanese, the US, and the German banking system. Sources: *Oxford Handbook on Banking, European Central Bank, Deutsche Bundesbank, FDIC and IMF International Financial Statistics (obtained from Demirgüç-Kunt and Maksimovic (2002)).*

This rigid hierarchical structure was implemented in the 19th century and is still in place today. Similar to Japan, banks in Germany are allowed to own shares of firms and take part in firms' decision processes (Frankel et al. (1991)). The German banking system is Europe's most widely populated system: the number of active banks was around 2,050 in 2006.

In contrast, the US banking system has been marked by a set of strict rules from the early 20th century which apply to all commercial banks (Jayaratne and Morgan (2000); Morgan et al. (2004); Acharya et al. (2011)). In essence, banks were limited in their geographical scope (i.e., branching limits and inter-state interdiction) and activity (i.e., separation between commercial and investment business).[13] This particular setting gave rise to a large number of small banks, which is vastly dominated by community banks.[14] The 1980s and 90s were marked by important deregulation efforts which ultimately removed limitations on the branching capacity and banking activity.[15] As a result, large waves of mergers and acquisitions allowed large multi-state (even multi-national) universal banks to emerge, while the number of active banks shrunk by half (Calomiris (2006)). The US also differs from

[13]The McFadden Act of 1927 explicitly prohibited interstate branching of commercial banks. The Glass-Steagall Act of 1933, among other things, prohibited investment banking activities of commercial banks.

[14]DeYoung (2010) shows that up until 1980s the total number of banks was relatively constant with values around 14,000 banks, 95% of which were community banks with less than $1 Bill. of assets.

[15]The Riegle-Neal Act of 1994 partially repealed the McFadden Act while the Gramm-Leach-Bliley Act of 1999 repealed the Glass-Steagall Act.

Japan and Germany in that it has a market-oriented financial system, such that traditional bank-firm credit interactions are generally of lesser importance.

Overall, this historical and international perspective shows that different regulatory environments can indeed have an impact on the structure of credit markets in the sense that they might encourage banks to be generalists or specialists to a certain extent. As such, the German system pro-actively encourages the existence of specialist banks (i.e., the Sparkassen) with limited business activity while the branching and activity deregulations in the US allowed for the emergence of large, generalists banks. Interestingly, the Japanese case appears to lie somewhere between the German and the US systems due to both its historically segmented structure, coupled with more recent waves of deregulation.

6 Empirical Analysis

As explained above, we are interested in the industrial loan portfolios of banks, i.e., we aggregate banks' loan exposures to the level of economic industries. While the persistence at the micro-level (bank-firm) suggests that there should also be high persistence at the aggregated level (bank-industry), we mainly focus on the cross-sectional distribution of generalists and specialists in the credit network. Our sample starts in 1980, corresponding to a highly segmented banking system, and ends in 2013, thus covering both the deregulation period and various financial crises. Hence, the time dimension is also of interest given the historical evolution of the Japanese banking system.

6.1 Data

Our analysis crucially relies on detailed data on bank-firm loan interactions. Like many other studies (e.g., Caballero et al. (2008); Ono et al. (2014); Peek and Rosengren (2005); Shin and Kolari (2004)), we use data from the Nikkei NEEDS database in the following.[16]

[16]More details can be found online: `https://www.nikkeieu.com/needs/needs_data.html`.

The database provides extensive accounting and loan information for all listed companies in Japan. Thus, our dataset exhibitis a sample bias in the sense that we only observe the borrowing patterns of listed companies; in order to addressed this limitation, we performed extensive robustness checks in this regard, which are discussed below and in Appendix E. Since 1996, the sample also includes firms traded in the JASDAQ (OTC market). Most importantly, the *Corporate Borrowings from Financial Institutions* data contain information on firms' outstanding loan volumes from each lender at the end of the firm's fiscal year. The data are based on survey data compiled by Nikkei Media Marketing, Inc. and are classified as short-term (up to 1 year) and long-term borrowing (more than 1 year). We use the sum of short- and long-term borrowing in everything that follows. The sample period covers the years 1980–2013. Most firms' fiscal years end in March, and for the sake of simplicity we refer to calendar years in what follows.

We complement the loan data with additional characteristics of the banks and firms from the *Corporate Financial Information* and the *Corporate Attribute* parts of the NEEDS database (most importantly balance sheet characteristics and industry affiliations). Unfortunately, the database contains only the most recent industry affiliation for each firm, such that these affiliations are likely to be most accurate for the most recent sample period. Nevertheless, we still include all years in our sample since the observed structures are generally very persistent. Our final sample includes 179 banks, 4,502 firms, and 34 industries.[17]

6.2 Summary Statistics

Table 2 and Figure 3 provide some summary statistics for our dataset. The total loan volume is on the order of 5 trillion Japanese Yen, which on average corresponds to roughly 12% of nominal GDP. The left panel of Figure 3 shows that the number of active banks in the

[17]See Appendix A for the complete list of industries and Appendix E for additional analyses based on different levels of industry granularity.

Summary statistics					
General	Mean	Median	Min.	Max	Std.
Total loan volume*	5.07	4.42	3.23	7.54	1.55
Number of banks	129.4	135	104	143	12.6
Number of firms	2,066.2	2,263.5	1,386	2,774	543.0
Number of industries	32.65	32.00	31	34	0.92
Firms per industry	63.21	35	5	463	71.93
Bank-firm network					
Links	18,699	18,206	12,308	26,033	3,748.3
Bank-degree	144.5	49.00	5	2,116	25.76
HHI-bank	0.105	0.065	0.003	0.877	0.103
Firm-degree	9.05	7	1	104	8.04
HHI-firm	0.297	0.243	0.017	1	0.205
Bank-industry network					
Links	2,283.8	2.389	1,752	2,698	290.7
Bank-degree	17.65	18	1	34	9.12
HHI-bank	0.168	0.120	0.032	1	0.132
Industry-degree	69.95	68	11	140	29.44
HHI-industry	0.081	0.064	0.031	0.517	0.050

Table 2: Summary statistics for the yearly bank-firm and bank-industry networks. 'Links' denotes the number of connections in the corresponding network. 'Degree' denotes the number of links per node. 'HHI' denotes the normalized Hirschman-Herfindahl Index of lending and borrowing concentration, for banks and firms/industries, respectively. This is defined as the squared sum of normalized portfolio weights in the corresponding weighted credit network (matrix $\mathbf{W}$ for bank-industry connections). Note: $* =$ in trillion Yen.

sample tends to decrease over time, with an average value of 129.[18] Moreover, the number of active banks steadily declines after the second half of the 1990s, a period that corresponds to a change of policy from the Japanese government to allow financial institutions to fail given that it could no longer find any 'white knight' institution strong enough to acquire those in serious distress (Woo and Kanaya (2000)). The number of firms is much larger, with an average value of 2,066. As the right panel of Figure 3 illustrates, the number of active firms jumps from 1,734 in 1995 to 2,523 in 1996. This structural break is solely due to the fact that the Nikkei NEEDS database covers JASDAQ companies from 1996 onwards. Therefore, since the jump in the active number of firms could have an effect on some of the results, we always check the robustness of our findings for different subsamples. The left panel of Figure 3 shows that the number of active industries is very stable over time (average value of 32). Given the long sample period under study, Appendix C takes a closer look at the dynamics of bank-firm interactions, finding that, despite an increased usage of stock-market financing, the typical micro-level interactions are very stable.

Table 2 also shows summary statistics for the basic bank-firm networks and the bank-industry credit networks which are the main focus of this paper. The average bank interacts with 145 firms from 18 different industries (bank-degree). Similarly, the average firm borrows from 9 banks (firm-degree) and the average industry interacts with 70 banks (industry-degree). These averages, however, hide significant heterogeneity in nodes' lending/borrowing patterns. In fact, Table 2 shows that some banks focus their investments on only one industry, while others interact with firms from all 34 industries. The same is true for industries' borrowing patterns: some industries borrow from as few as 11 banks, while others receive funding from basically the entire banking system. Overall, these results indicate that a crucial ingredient of the generalist-specialist model - the existence of heterogeneous borrowing and

[18]We define active banks as those banks with at least 5 firm loan relationships (minimum degree of 5) in a given year and active industries need to consist of at least 5 borrowing firms in a given year. We experimented with different cutoffs and find that the qualitative results remain unaffected.

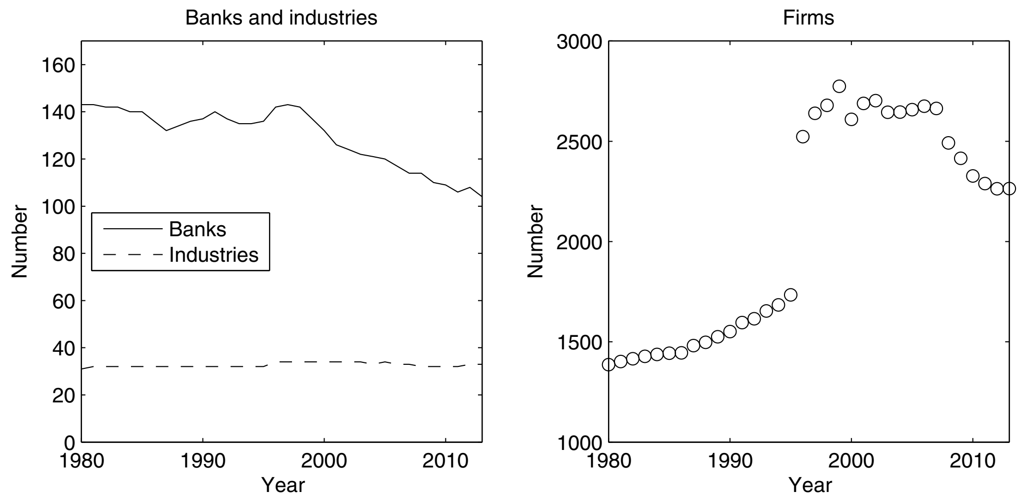

Figure 3: Number of active banks and industries (left panel) and number of active firms (right panel). Note: the structural break in the number of firms from 1995 to 1996 is solely due to inclusion of more firms in the Nikkei database from that time onwards.

lending patterns - is confirmed by our dataset.

6.3 Generalist-Specialist Model

In this section, we present the results from fitting the discrete generalist-specialist model to the Japanese data separately for each year. Before turning to the time dimension of the results, we start out by focusing on one particular snapshot of the credit network in order to illustrate the spirit of our approach and the 'typical' finding. Figure 4 shows the both the weighted credit network and the binary adjacency matrix for the year 1990 (left and center panel), and the corresponding idealized pattern matrix as defined in Eq. (1) (right panel). In the Figure, rows correspond to banks, columns correspond to industries, and a black dot indicates a link between a given pair of nodes.

As we will show in more detail below, the fit of the model is significant, but not perfect. In particular, we do observe a number of specialist-specialist interactions (in the bottom right part of the actual network) that are absent in the stylized generalist-specialist model.

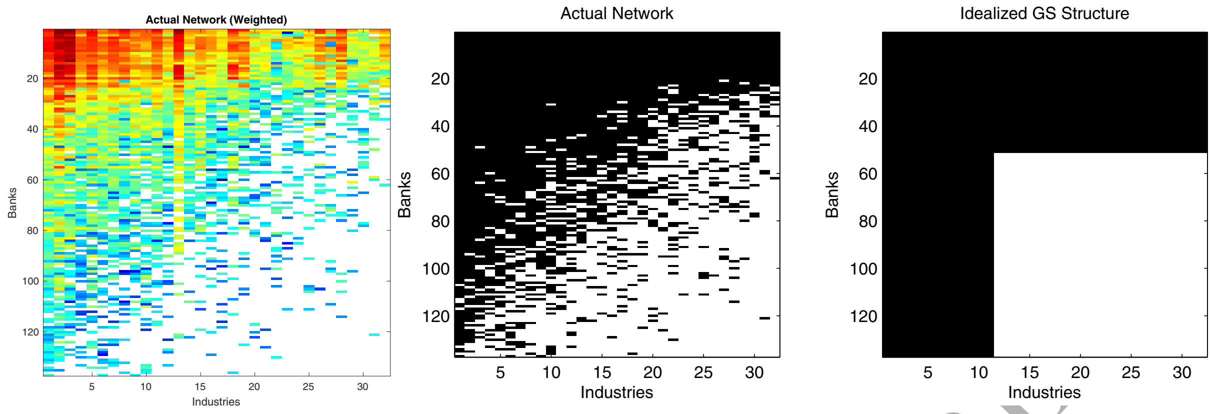

Figure 4: Illustration of the actual credit network matrix (left: weighted, center: binary) and the idealized generalist-specialist structure (right) in 1990. Generalist nodes are sorted first according to the optimal partition vectors. Rows correspond to banks, columns correspond to industries, and a black dot indicates a link between two nodes. Note: the left panel shows log-transformed credit volumes for better visibility.

This result illustrates a limitation of the binary generalist-specialist model, namely that it has difficulties matching asymmetric structures with very broadly distributed diversification strategies.

Temporal Evolution. Figure 5 shows the absolute number of generalists and specialists over time. The left panel shows that the number of generalist banks was relatively stable for the first half of the sample period, and has been decreasing afterwards: there were 62 generalist banks in 1980 and roughly 37 in the post-2000 period. In order to check whether this decline is mainly driven by the negative trend in the total number of active banks, the solid line shows the fraction of generalist banks over time (defined as the number of generalist banks relative to the total number of active banks). Clearly, this was much more stable: before 2000 roughly 40% of the banks were generalists, while the values are closer to 35% afterwards. Hence, the number of generalist banks appears to decline roughly proportionate to the size of the system except for some volatility around the global financial crisis of 2008.

The right panel of Figure 5 show the same decomposition for the industries. The absolute

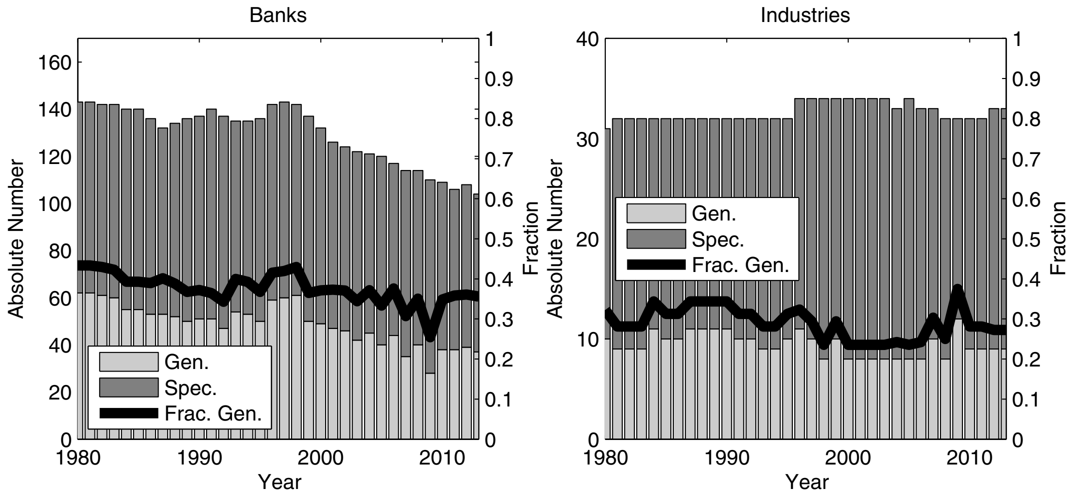

Figure 5: Number of generalist/specialist banks (left) and industries (right). The solid lines show the relative fraction of generalists over time (right y-axis).

number of generalists was very stable over the sample period, with a typical value of 10 generalist industries. Similarly, the fraction of generalist industries was very stable as well (average value close to 30%). In summary, these results indicate that the relative abundance of generalists and specialists is quite stable over the sample period.

Economic importance. How important are generalists and specialists economically? Table 3 illustrates the importance of interactions between generalists and specialists. The left column shows the density (number of existing links relative to the maximum number of possible links) for each block. We see that the generalist-generalist block is almost fully connected since the density is close to 100%. Similarly, the off-diagonal blocks are also well-connected with densities around 70%, while the specialist-specialist block has an average density of 21%. The right column shows that, on average, the interactions between generalist banks and generalist industries amount to 62% of the entire loan volumes (see also the left panel of Figure 4). Interestingly, while we do observe quite a few links in the SS-block, these links are of minor economic importance from a system-wide perspective, as they account for only

1% of the entire loan volumes. Hence, while specialist-specialist interactions are not as rare as imposed by our model, the underlying loan volumes are indeed tiny relative to all other blocks.

		Industries			
		Gen	Spec	Gen	Spec
Banks	Gen	99.2%	77.6%	62.0%	34.1%
	Spec	71.9%	21.3%	2.4%	1%
		Density		**Volume**	

Table 3: Economic importance of generalists and specialists. Left: density in each of the different blocks (number of existing links relative to maximum possible number of links). Right: actual loan volumes in each block relative to the total volumes. Values are calculated separately for each year (1980 until 2013) and then averaged over time.

Persistence in strategies. As argued in previous sections, we should expect substantial persistence in nodes' strategy profiles. Table 4 shows that this is indeed the case. More precisely, the Table shows the transition probabilities between the different strategies (generalist, specialist, and inactive) for both banks and industries, respectively. The left panel shows the results for a transition period of 1 year, and the right panel for a period of 5 years. Overall, we see a lot of persistence in nodes' strategy profiles, since the diagonal elements are generally very large. For example, there is a 94% (95%) chance for are generalist bank (industry) to remain generalist in the following year. The values are slightly lower when we look at a transition period of 5 years (84% and 91% for banks and industries, respectively) but the general picture remains unaffected. Hence, banks' strategies are sticky, highlighting the fact that the coexistence of both generalists and specialists can be interpreted as an equilibrium phenomenon - we do not find evidence that the typical specialist bank tends to transition towards becoming a generalist over time (Wagner (2010)).

Model fit. How good is the fit of the generalist-specialist model? Figure 6 shows the total error score and the contribution of the different blocks to this total. We find that the error

Banks	Generalist$_{t+1}$	Specialist$_{t+1}$	Inactive$_{t+1}$	Generalist$_{t+5}$	Specialist$_{t+5}$	Inactive$_{t+5}$
Generalist$_t$	93.7%	5.6%	0.7%	84.2%	11.5%	4.3%
Specialist$_t$	2.9%	94.3%	2.7%	4.8%	85.5%	9.6%
Inactive$_t$	0%	2.8%	97.2%	0%	9.0%	91.0%
Industries	Generalist$_{t+1}$	Specialist$_{t+1}$	Inactive$_{t+1}$	Generalist$_{t+5}$	Specialist$_{t+5}$	Inactive$_{t+5}$
Generalist$_t$	95.2%	4.8%	0%	90.8%	9.2%	0%
Specialist$_t$	1.7%	97.8%	0.5%	3.4%	95.1%	1.5%
Inactive$_t$	0.9%	4.5%	94.6%	2.1%	11.5%	86.5%

Table 4: transition probabilities for banks (top) and industries (bottom) between the three possible configurations, namely generalist, specialist, or inactive. Values are shown for transition periods of 1 year (left panel) and 5 years (right panel).

score is always well below 1, the value that would arise in the absence of a generalist-specialist structure. Note that the error score steadily increases throughout the sample period, though at a slow rate (average increase of 0.3% per year). Hence, the fit of the generalist-specialist model deteriorates over time. Given that our sample period covers major events that should have had an impact on the structure of the system, it is quite remarkable that none of those events appear to be associated with a structural break in the error score. For example, the increase of the total error score from 2008 to 2009 is 0.031, while the standard deviation of the full time series is 0.032 - in other words, not an extreme event by any means.

In order to explore this finding in more depth, Figure 6 also shows that most of the errors indeed come from specialist-specialist interactions, as one would expect from the comparison between the actual network and the idealized generalist-specialist structure in Figure 4. This suggests that specialist-specialist interactions have become significantly more likely over the sample period. The relative increase of specialist banks' interactions with specialist industries appears to be in line with the idea that deregulation tends to lead to less hierarchical (or less segmented) banking systems. Nevertheless, the structure of the credit network appears to change very slowly, likely due to the long-term nature of bank-firm interactions in Japan.

Significance. As a next step, we want to test whether the observed generalist-specialist structure of the credit network is significant. In order to do this, we need to define different

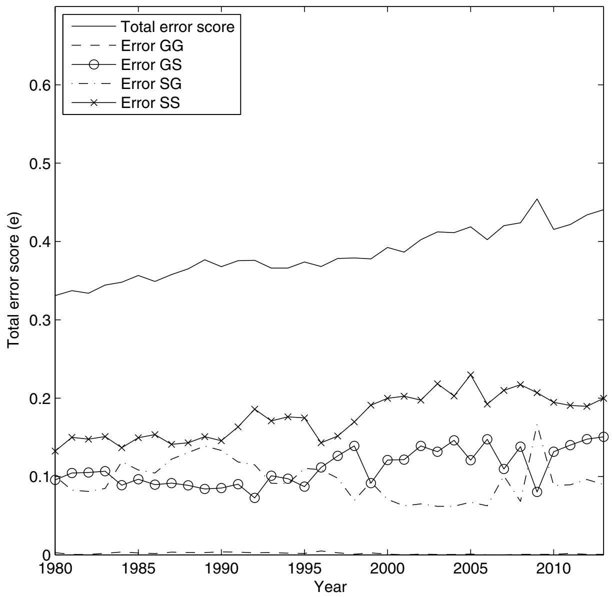

Figure 6: Error score of the generalist-specialist model for the observed networks, defined as $(E_{GG} + E_{GS} + E_{SG} + E_{SS})/m$. For all but the SS-block, the absence of a link is counted as an error, while in the SS-block the presence of a link is counted as an error. The plot also shows each block's contribution to this total. For example, the contribution of the specialist-specialist block is calculated as E_{SS}/m.

randomization procedures (or null models), which tell us how the credit network looks like if connections were formed at random. Comparing the results for the actual network with the synthetic networks then gives us an indication of whether reasonably simple probabilistic models are able to replicate the observed interaction patterns and, more importantly for our purposes, whether the generalist-specialist model tends to yield similar results.

Null models are randomized versions of the actual network, where certain characteristics are kept fixed to make them comparable. The first null model is a completely random network:

1. Erdös-Renyi (ER) random network: the probability of a link between any bank-industry pair is equal to ρ, where ρ is the observed density of the network. Since all interactions have the same probability, we expect a poor fit of the generalist-specialist

model (see Appendix D for details on the expected error score in ER random networks).

The ER random network is the simplest possible null model, since it ignores any preferences in the borrowing/lending decisions. We also explored a multitude of more elaborate null models, but here we restrict ourselves to those inspired by the Balls-and-Bins model of Armenter and Koren (2014).[19] The basic idea in all of these models is as follows: each bank interacts with a certain number of firms. This is the number of balls per bank. For each bank, we throw these balls into bins, which are the industries. The probability of a ball ending up in a given bin (i.e., of drawing a connection) depends on the relative bin sizes, and there are different ways to define these. We experimented with the following three cases:

2. Homogeneous bin size ('Balls+Bins: Homogeneous'): in this case, each ball has the same probability of ending up in a given bin. Note that this null model is related to the ER random network above, with the main difference being that it allows for heterogeneity across banks in terms of the number of connections.

3. Bin size proportional to the number of firms per industry ('Balls+Bins: Nfirms'): in this case, larger industries are more likely to attract more links.

4. Bin size proportional to the total loan volumes per industry ('Balls+Bins: Volume'): in this case, industries with larger loan volumes will attract more links.

Note that these null models fix each banks' number of interactions in the bank-firm network, rather than the total number of links in the bank-industry credit network. Hence, the randomized networks based on the balls-and-bins models generally display different densities compared to the actual networks, while the ER random networks have the exact same density as the observed ones.[20]

[19]Results for other null models generally show that the observed error score is always well below that of the randomized networks. The only exception is the so-called configuration model, which predicts significantly fewer specialist-specialist interactions compared to what we observe in the data. More details are available

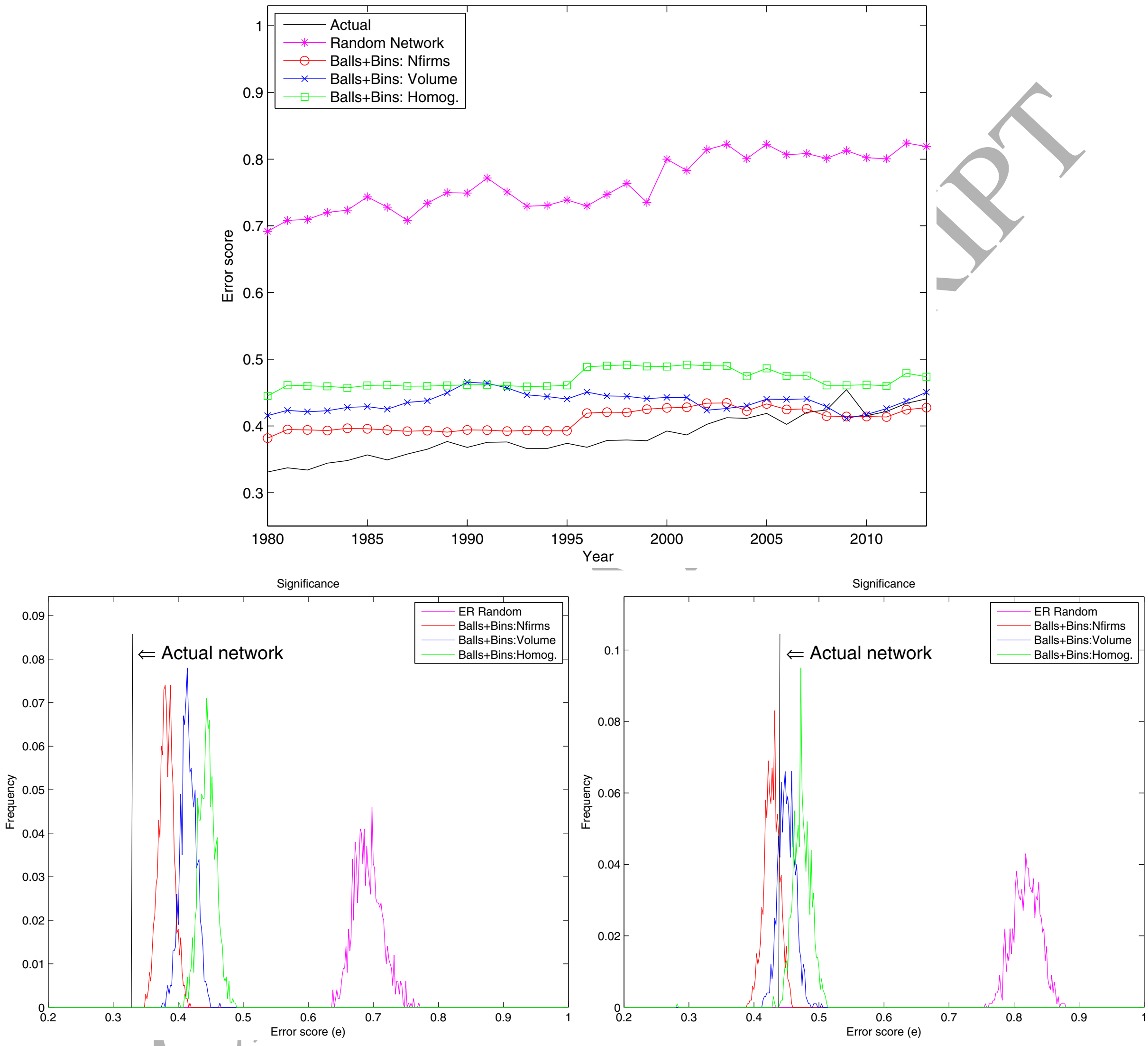

Figure 7: Error score and null models. Top: average error score for the different null models (averaged over 1,000 realizations for each null model). Bottom: histograms of null model error scores for 1980 (left) and 2013 (right) - based on the same 1,000 realizations used in the top panel.

The top panel of Figure 7 shows the error scores for the actual network and the average error scores for 1,000 synthetic networks of the different null models (we focus on the cross-sectional distribution of these in more detail below). We see that, until around 2006 the actual networks generally yield lower error scores compared to all null models considered here. Not surprisingly, the ER random networks produce by far the largest error scores. Among the balls-and-bins models, the homogeneous model performs worst, as it generally predicts a much larger error score compared to the actual network. Lastly, the heterogeneous balls-and-bins models perform much better at the end of the sample period, in the sense that they produce error scores close to the actual network.

The bottom panels of Figure 7 show the cross-sectional distributions of the error scores from the different null models for two different years (1980 and 2013, respectively). Clearly, in both cases the ER random network is a very poor description of the actual network. On the one hand, we see that in 1980 (left panel) the balls-and-bins models generate very similar error score distributions, but with significantly higher means compared to the observed error score. On the other hand, the results for 2013 (right panel) show that the error score from the actual network is not statistically different from the three balls-and-bins models.

In summary, these results show that the observed generalist-specialist architecture cannot be fully captured by simple generative models for most of the sample period. Nevertheless, the balls-and-bins models appear to match certain characteristics of the observed interaction patterns. This is consistent with Armenter and Koren (2014), who reproduced patterns of product-/firm-level trade flows across export destinations based on similar models.

Continuous Model. In section 4 we also introduced a continuous version of the generalist-specialist model. Here we briefly present some results from this alternative specification,

---

upon request from the authors.

[20]Up until 1996, the balls-and-bins models generally yield networks with lower densities compared to the observed ones. Afterwards, the densities became much more similar, suggesting that the jump in the number of firms from 1995 to 1996 affects the performance of the balls-and-bins models.

which suggest that the limitations of the discrete model are modest. Note that we apply the model both to the binary interaction matrix $\mathbf{A}$ and the weighted credit network $\mathbf{W}$; in the latter case, we log-transform the data as $\tilde{\mathbf{W}} = \log(1 + \mathbf{W})$ in order to mitigate skewness.

The left panel of Figure 8 shows the fit of the continuous model both for the binary and the weighted network matrix. The dynamics of the fit are very similar in both cases, but using the weighted network matrix consistently leads to a better fit. This is due to the fact that the weighted network contains more information on the heterogeneity of link weights. For the binary network, the R^2 is around 0.47 pre-1996 and closer to 0.40 afterwards, while the values for the weighted networks are 0.53 and 0.45, respectively. As before, there appears to be a structural break in the data in the late 1990s. Overall, however, the continuous version of the generalist-specialist model performs reasonably well, since it generally explains more than 40% of the observed variance.

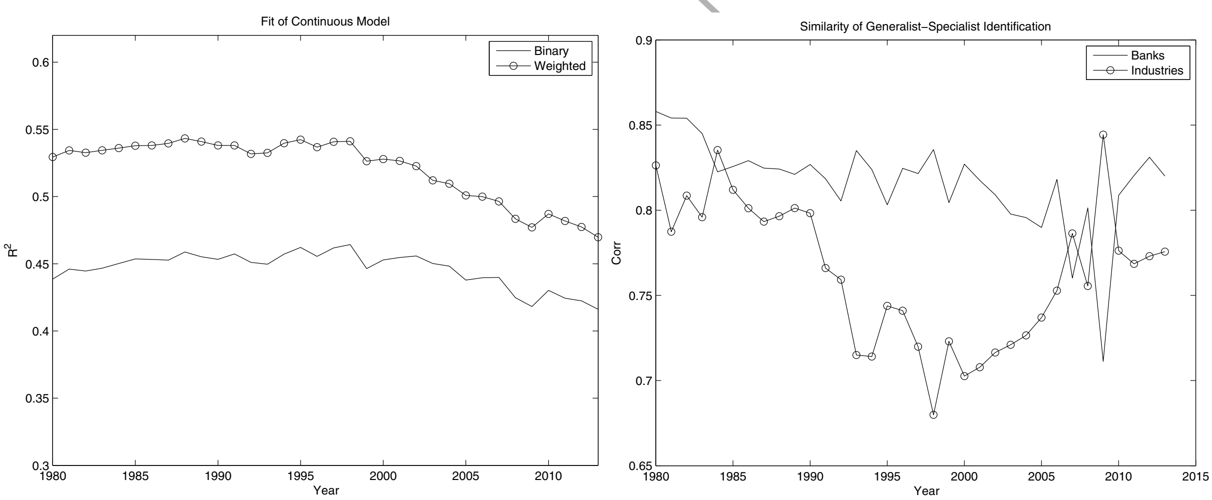

Figure 8: Results for the continuous generalist-specialist model. Left: fit for binary and weighted network (R^2). Right: similarity of discrete and continuous generalist-specialist partition vectors, measured as the correlation between the generalist-level vectors in any year. (Note: in the right panel we use the binary adjacency matrix; we find very similar results when using the weighted network matrix.)

The right panel of Figure 8 shows the similarity between the generalist-level vectors derived from the discrete and the continuous model, using the binary network. (We obtain very similar results for the weighted network.) For each year, we plot the correlation between the corresponding vectors for the banks and industries, respectively. The correlation is around 0.8 for both banks and industries, suggesting that the results are comparable to those from the discrete model.

Robustness Checks. We have performed a large number of robustness checks on the generalist-specialist model. For the sake of brevity, we leave details to Appendix E and only summarize the most important findings.

- **Listed versus unlisted companies.** An important issue is that our dataset does not capture the entire universe of Japanese businesses, but only the subset of listed companies. In principle, this sample bias could have an effect on the observed architecture of the credit network: it might be that specialist banks (those concentrating on a small number of industries) will not be classified as specialists anymore when including very small, unlisted companies to the sample. This concern is valid if two conditions are satisfied: (1) these small firms are affiliated with specialist industries; and (2) they mainly borrow from small, specialist banks.

  Unfortunately, given the nature of our dataset, we cannot directly test these two conditions. However, we implemented various in-sample tests that make use of the fact that our dataset covers a significant number of very small companies (often satisfying standard criteria of being small-to-medium enterprises, or SMEs). In particular, in the light of the two conditions above, we document in Appendix E.1 that these small firms tend to (1) tend to be affiliated with generalist industries; and (2) mainly borrow from generalist banks. Interestingly, the latter finding is by no means peculiar about the Japanese financial system, but broadly in line with the contemporaneous empirical literature on SME financing around the world.[21]

[21] The convention in much of the classic banking literature, namely that small banks should be better

The above two pieces of evidence, thus, suggest that adding small firms to the analysis is unlikely to dramatically affect the results of the generalist-specialist model; we further support this assertion by re-applying the model when excluding the smallest firms from the analysis. In this case, we find that the fit of the model *deteriorates* compared to the baseline analysis. Overall, these results indicate that the credit network is likely to display a generalist-specialist structure even when using more exhaustive datasets.

- **Intensive versus extensive margin.** Our main analysis focused more on the extensive margin (i.e, whether a link is present or not) than on the intensive margin (i.e., the intensity of links). As such, we find it quite remarkable that the simple, binary version of the generalist-specialist model provides such a good fit. In order to take a closer look at the extensive margin, we also presented some results from the continuous version of the generalist-specialist model, which were generally consistent with those reported for the baseline model. In addition, Appendix E.2 contains some robustness checks that disregard relatively small credit interactions, finding that the model fit is not strongly affected compared to the baseline scenario.

- **Aggregation and industry granularity.** The aggregation level of the credit network, in terms of the granularity of the industry classification, affects the fit of the generalist-specialist model. In Appendix E.3 we show that there is a negative relationship: a more (less) granular industry classification makes it less (more) likely for banks to be connected with firms from that a given industry. In this sense, with a more coarse-grained industry classification it is easier to detect generalists, and thus achieve a better fit.

suited to lend to small, opaque businesses (e.g.,e Berger et al. (2005)), has been questioned by a number of recent empirical studies (e.g., Berger et al. (2007); De la Torre et al. (2010); Berger and Black (2011); Bonfim and Dai (2017); Berger et al. (2014)). Broadly speaking, these studies show no dominance of small banks over large banks in the borrowing activity of small firms. Explanations include technological progress (Berger and Udell (2006); Frame et al. (2001); Frame and Woosley (2004)), deregulation (DeYoung et al. (2011)), and changes in the organizational structure of large banks (Canales and Nanda (2012)).

- **Moving time windows.** In the baseline analysis, we draw a connection based on whether a given bank provides loans to firms from a given industry in a given year. In Appendix E.4, we relaxed this definition using longer time windows (between 2 and 10 years), finding that the fit of the model slightly improves with longer time windows.

Discussion. The results presented so far suggest that the idealized generalist-specialist structure in Eq. (1), despite being based on strong assumptions, allows us to understand the structure and the evolution of a credit market.

It should be clear that the relative abundance of generalist and specialist banks, and the stickiness of their strategy profiles, depend on both the regulatory framework and the historical background of a given banking system. In section 5 we compared the Japanese banking system with those of the US and Germany, respectively, and we discussed how different regulatory approaches can have an effect on the relative abundance of generalist and specialist banks. As described in Section 5 as well, the Japanese banking system was traditionally characterized by a hierarchical structure consisting of different bank types with generalist and specialist purposes. Indeed, our framework confirms this segmentation. Interestingly, the major events that took place in Japan during our sample period (e.g., large sequence of deregulation and several financial crises) did not translate themselves into structural breaks in the fit of our model. It thus appears that the evolution of the Japanese banking system was much more progressive over time.

Note also that the idealized generalist-specialist structure in Eq. (1) predicts that the set of specialist banks would be connected to the very same industries, such that the overlap in banks' portfolios would be maximal. Clearly, the observed networks display certain deviations from this prediction, but the good fit of the generalist-specialist model suggests that portfolio overlap is indeed very strong.

This raises the question of why banks hold such similar loan portfolios. The theoretical literature suggests several possibilities, which may not be mutually exclusive. First, bank owners may invest in correlated assets because, due to limited liability, they do not internalize the costs of a joint failure (Acharya and Yorulmazer (2008)). Hence, banks want to increase the likelihood of failing simultaneously in order to induce a regulator to bail them out. Second, by aiming at holding more diversified portfolios, banks may become more similar as an unintended side effect (Wagner (2010)). Third, herding by investors will tend to result in institutions taking on similar exposures. Such herding may arise for psychological reasons, reputational concerns, but may also be rooted in performance evaluation as managers will not be fired if they under-perform jointly with their peers (Rajan (2005)). It is very difficult to directly test the empirical relevance of each of these theories and we therefore leave this exercise as an interesting avenue for future research.

6.4 Predicting Nodes' Network Position

Our notion of generalist-specialist interactions appears to capture a structural feature of the credit network. In a similar fashion as for core-periphery interbank networks (Craig and von Peter (2014)), this partition is derived from the pattern of credit interactions only. In the following, we show that node-specific features help predict whether a given bank or industry is a generalist or specialist, respectively. As pointed out by Craig and von Peter (2014), this is important because it allows to predict the network position of banks and industries using data that should be readily available (e.g., balance sheet information), and it shows that certain features systematically relate to being a generalist.

More precisely, we use a probit framework where the binary dependent variables are the generalist-levels (i.e., $\gamma^B_{i,t}$ and $\gamma^I_{j,t}$), for banks and industries, respectively. We estimate

$$\text{Prob}(\gamma^B_{i,t} = 1 | X_{i,t}) = \Phi(X^T_{i,t}\beta^B), \tag{8}$$

$$\text{Prob}(\gamma^I_{j,t} = 1 | X_{j,t}) = \Phi(X^T_{j,t}\beta^I), \tag{9}$$

separately for banks and industries; Φ represents the cumulative normal distribution, X denotes the set of control variables (always including a constant), and β^B/β^I the corresponding parameter vectors. In the following, we always include data for the full sample period and show marginal effects rather than the parameter estimates. A marginal effect of a given independent variable is the corresponding partial derivative of the prediction function, which we evaluate at the mean of the variable. In other words, the reported coefficients tell us the effect of a 1-unit change of a given variable, relative to its mean, on the probability of being a generalist. Due to the jump in the number of firms from 1995 to 1996, we also run the regressions using only data from 1996 onwards as a robustness check (see Appendix G).[22] Lastly, Appendix H shows that results from standard OLS regressions for the generalist-levels from the continuous version of the model remain qualitatively similar.

6.4.1 Industries

Control Variables. We begin with the industries. In this case, we use the following control variables:

- TotalAssets: (natural logarithm of) the sum of total assets of all active firms in a given industry.
- TotalLoans: (natural logarithm of) the sum of the total loan volume of each industry as reported in the loan data.
- IntrinsicSize: defined as log(TotalAssets - TotalLoans) and measures the size of a industry, excluding the borrowing activities reported in the data.
- NumberFirms: (natural logarithm of) the number of active firms in a given industry.
- Employees: (natural logarithm of) the number of employees in a given industry.

[22]We also ran the logit model as an alternative specification. Finally we included various time-, bank-, and industry-specific fixed effects, none of which alter our main findings. Results on these robustness checks are available upon request from the authors.

- Leverage: this is defined as log(TotalLiabilities/Equity) and is based on book values.
- Current asset ratio: this is defined as (CurrentAssets/TotalAssets), with CurrentAssets being all assets that can be converted into cash within a year.
- Hirschmann-Herfindahl Index, HHI: (natural logarithm) of firms' squared relative borrowing volumes within each industry.[23]
- IndustryGeography: relative number of banks with headquarters located in the same geographical area as the firms' headquarters from any given industry. Measures how easily banks can be reached by different industries. There are 47 geographical areas based on the 'JIS Codes of Administrative Divisions of Japan' (see Appendix B for a complete list).
- Interest spread: defined as the difference between the uncollateralized overnight call rate (interbank market) and the Bank of Japan's official discount rate, measured in percent. This variable controls for general funding conditions.
- GDP growth: growth of gross domestic input, measured in percent. This variable controls for macroeconomic conditions.

Regression results. Table 5 shows the results.[24] Clearly, we cannot simultaneously include all of the above-mentioned control variables due to severe collinearity issues. Therefore, we proceed in steps. First, larger industries (as measured by either total assets, total loan volume, intrinsic size, number of firms, or number of employees) are more likely to be generalists. Note, however, while all size measures are highly significant, the overall fit of these specifications is far from perfect (R^2 between 0.25 and 0.45), suggesting that size is important, but not the only feature that plays a role.

[23]Defining the HHI based on firms' total assets yields qualitatively similar results.

[24]Appendix F shows a list of industries that are identified as generalists at least once over the sample period.

Gen. Industries	(1)	(2)	(3)	(4)	(5)	(6)
log(TotalAssets$_{t-1}$)	0.2648***	–	–	–	–	0.1573***
	(0.0148)					(0.0264)
log(TotalLoan$_{t-1}$)	–	0.2770***	–	–	–	–
		(0.0170)				
log(IntrinsicSize$_{t-1}$)	–	–	0.2465***	–	–	–
			(0.0145)			
log(NumberFirms$_{t-1}$)	–	–	–	0.3031***	–	–
				(0.0153)		
log(Employees$_{t-1}$)	–	–	–	–	0.2283***	–
					(0.0130)	
log(Leverage$_{t-1}$)	–	–	–	–	–	0.1774***
						(0.0343)
CurrentAssetRatio$_{t-1}$	–	–	–	–	–	1.2350***
						(0.1400)
log(HHI$_{t-1}$)	–	–	–	–	–	-0.0610***
						(0.0194)
IndustryGeography	–	–	–	–	–	0.0055***
						(0.0009)
Interest Spread$_{t-1}$	–	–	–	–	–	-0.0121
						(0.0158)
GDP$^{\text{nom.}}$ growth	–	–	–	–	–	-0.0024
						(0.0051)
Pseudo-R^2	0.362	0.450	0.326	0.420	0.247	0.649
Obs.	1,056	1,056	1,056	1,056	1,056	897

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 5: probit model for generalist industries. Data from 1980 - 2013. The Table reports the marginal effects for the control variables evaluated at the means (robust standard errors in parentheses).

For the other characteristics, it turns out that generalist industries are more leveraged. This is not surprising since their borrowing activity mechanically increases their leverage ratio. Interestingly, generalist industries also hold more liquid assets as shown by the strong positive effect of the current asset ratio. Note that the current asset ratio can also be seen as a measure of asset non-specificity (see Strömberg (2000)), where lower values correspond to more specific assets. Hence, these results are consistent with the existing literature on asset specificity, which suggests that firms from industries with less specific assets (higher current asset ratio) should find it easier to obtain funding from the banking system.[25] Not

[25]The idea is that a firm with very specific assets is likely to experience lower liquidation values in case of default (see for example, Williamson (1988), and Shleifer and Vishny (1992)), such that its assets cannot be easily redeployed for other purposes. If banks take this into account in their funding decision,

surprisingly, more concentrated industries (those with larger HHI) are significantly less likely to be generalists. Also we find that generalist industries tend to be geographically closer to a larger number of banks. Regarding the macroeconomic indicators, both interest spreads and GDP growth turn out to be insignificant. Hence, it seems that the state of the economy as a whole does not affect industries' borrowing patterns, in good times as in bad times.

6.4.2 Banks

Control Variables. For the banks we use the following control variables:

- TotalAssets: (natural logarithm of) banks' balance sheet size.
- TotalLoans: (natural logarithm of) banks' loan volumes.
- IntrinsicSize: this is defined as log(TotalAssets - TotalLoans).
- Bank type: we include dummies for different bank types, namely for city banks, and tier-2 regional banks. Note that (Tier-1) regional banks are the most common bank type. We expect the bank type to be very important in the regressions, since the Japanese banking system used to be strongly segmented and some bank types were required to be generalists, while others were set up to be much more specialized.
- Cash ratio: fraction of cash holdings relative to total assets (measured in %).
- Net interbank position: interbank assets minus interbank liabilities relative to total assets (measured in %).
- Hirschmann-Herfindahl Index, HHI: (natural logarithm of) sum of squared weights of a bank's industrial loan portfolio.
- Leverage: (natural logarithm of) banks' leverage ratios.

asset specificity should be negatively related to the availability of bank lending such that we would expect asset-specific industries to be less likely to be generalists.

- Bank geography: number of industries located in the same geographical area as a given bank. Measures how easily banks can reach firms from different industries.

- Interest spread and GDP growth are as defined above.

Gen. Banks	(1)	(2)	(3)	(4)	(5)
log(TotalAssets$_{t-1}$)	0.5409***	–	–	–	0.7303***
	(0.0164)				(0.0394)
log(TotalLoans$_{t-1}$)	–	0.4355***	–	–	–
		(0.0131)			
log(IntrinsicSize$_{t-1}$)	–	–	0.5210***	–	–
			(0.0157)		
D(City bank)	–	–	–	0.7367***	–
				(0.0647)	
D(Tier-2 regional bank)	–	–	–	-0.6158***	–
				(0.0197)	
CashRatio$_{t-1}$	–	–	–	–	-0.0951***
					(0.0190)
log(HHI$_{t-1}$)	–	–	–	–	-0.3647***
					(0.0259)
log(Leverage$_{t-1}$)	–	–	–	–	-0.0042
					(0.0524)
NetInterbank$_{t-1}$	–	–	–	–	0.0150***
					(0.0056)
BankGeography$_{t-1}$	–	–	–	–	-0.0122***
					(0.0014)
Interest spread$_{t-1}$	–	–	–	–	0.1433***
					(0.0236)
GDP$^{nom.}$ growth	–	–	–	–	0.0351***
					(0.0063)
Pseudo-R^2	0.502	0.657	0.478	0.293	0.681
Obs.	4,094	4,094	4,094	4,286	2,979

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 6: probit model for generalist banks. Data from 1980 - 2013. The Table reports the marginal effects for the control variables evaluated at the means (robust standard errors in parentheses).

Regression results. Table 6 reports the results.[26] Similar to the regressions for the industries, we cannot include all control variables simultaneously, and we again proceed in steps. First, we find that each of the size proxies (total assets, total liabilities, and intrinsic

[26] Appendix F shows a list of banks that are identified as generalists at least once over the sample period.

size) are strongly positively significant. Hence, larger banks are more likely to be generalists. Note that size again does not explain everything; just as for the industry regressions, the fit of the model is far from perfect (R^2 between 0.48 and 0.66), suggesting that other factors are likely to play a role as well.

Second, we find that the bank type indicators are important control variables. City banks, the largest banks in the Japanese banking system, are significantly more likely to be generalists. In contrast, the smaller (tier-2) regional banks are much more likely to be specialists.[27] This result confirms that the institutional framework does affect whether some banks are generalists or specialists.

Finally, the last column includes additional bank-specific characteristics and also controls for macroeconomic conditions. We find that generalist banks are significantly more diversified (lower HHI), tend to hold less cash, but do not seem to use significantly more leverage over the whole sample period. In Appendix G we show the results for pre- and post-1996, respectively, finding that prior to 1996, generalist banks were significantly more leveraged, but the reverse is true after 1996. This is likely due to the severe banking crisis which made it necessary for larger banks to reduce their leverage (or made them more likely to default and therefore become inactive). In the next section, this feature will play an important role in the analysis of banks' vulnerabilities to systemic asset liquidations.

Somewhat surprisingly, the number of industries present in the same geographical area as the banks (BankGeography) turns out to be negatively significant. Hence, banks appear to be *less* likely to be generalists if there are many industries close to a bank's headquarters. This result may be driven by the fact that we only observe the geographical location of the banks' headquarters, but have no information on their different branches or local offices, which could serve as more accurate indicators for a bank's geographical spread.[28] Regarding the

[27] In Appendix G, we report the same Table for two different sample periods (pre- and post-1996) for which we find similar results as in Table 6. Hence, bank types and bank strategies (i.e., generalist or specialist) remain strongly related throughout our 34-years sample period.

[28] The result is robust to using more granular geographical indicators, e.g., post codes.

macroeconomic factors, a larger interest spread increases the probability of being a generalist for all banks. Similarly, macroeconomic growth also has a positive effect. Essentially all of these results are robust to using data only for the regional banks (the most abundant bank type in Japan), see Appendix G.

6.4.3 Are Specialist Banks More Profitable?

The good fit of the generalist-specialist model suggests a strong overlap in Japanese banks loan portfolios. This overlap implies that banks' performances should be strongly correlated, since most banks tend to interact with firms from the same industries. Clearly, different banks will not necessarily interact with the exact same firms - for example, a specialist bank, by exclusively focusing its lending activity on a particular type of firms, might be able to identify higher-quality firms relative to a generalist bank (Stomper (2006)). If this was the case, this should translate itself into banks' performances and generalist banks should be significantly less profitable compared to specialist banks. We test this hypothesis in the following.[29]

We use two performance measures, namely return on assets (ROA, defined as Net Income/Total Assets), and return on equity (ROE, defined as Net Income/Total Equity).[30] We run the following regressions:

$$\pi^B_{i,t} = \beta \mathbf{X}_{i,t} + \epsilon_{i,t}, \tag{10}$$

where $\pi^B_{i,t}$ is a measure of bank profitability (ROA and ROE, respectively), and $\mathbf{X}_{i,t}$ includes the set of control variables. The most important control variable is γ^B, i.e., banks' gener-

[29]Note that the existing empirical literature on the relationship between performance and diversification is anything but settled (see Tabak et al. (2011) for an overview), with some papers finding that specialist banks indeed outperform generalists (e.g, Acharya et al. (2006); Tabak et al. (2011)) and others finding the opposite result (e.g., Hayden et al. (2007)).

[30]Both ROA and ROE are winsorized at the top and bottom 1% to avoid the impact of extreme outliers. This does not affect the qualitative results.

	ROA			ROE		
	All	1980-1995	1996-2013	All	1980-1995	1996-2013
γ_t^B	-0.0005	-0.0090	-0.0524	0.6798	-0.0984	-0.6835
	(0.0272)	(0.0209)	(0.0580)	(0.9586)	(0.4966)	(2.1019)
log(Leverage$_{t-1}$)	-0.1539***	0.1944***	-0.1821***	-8.0093***	11.4000***	-14.8117***
	(0.0281)	(0.0288)	(0.0507)	(0.9903)	(0.6846)	(1.8345)
Bank FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
adj. R^2	0.271	0.353	0.262	0.222	0.510	0.189
Obs.	4,094	2,127	1,967	4,094	2,127	1,967

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 7: Are generalist banks less profitable? Results from OLS regressions of bank profitability (ROA and ROE, respectively) against generalist levels, γ^B, for different subperiods (robust standard errors in parentheses).

alist levels. If specialist banks were indeed more profitable relative to generalist banks, the parameter on γ^B should be significantly negative. We include bank's lagged leverage ratios in the regressions in order to control for banks' riskiness. In order to control for unobserved heterogeneity and possible time trends, we always include bank- and year-fixed effects. We run the regressions separately for different sample periods.

Results. The results can be found in Table 7. We see that the parameter on γ^B is not significantly different from zero, irrespective of the performance measure and the time period under study.[31] Therefore, specialist banks do not appear to outperform generalist banks, suggesting that they are unlikely to possess informational advantages or at least they do not appear to benefit from them.

Of course, bank profitability is driven by many other factors than just the diversification level of banks' loan portfolios. Given that we include bank FEs in the regressions, however, we implicitly control for all these other factors. Hence, diversification levels do not appear to be informative about Japanese banks' performances.

[31] The relationship with γ^B remains insignificant when dropping leverage as a control variable. Alternatively, when dropping the bank FEs we find that the relationship becomes negatively significant for both performance measures only for the pre-1996 period (at 1% level for ROA, and 10% level for ROA, unreported result).

6.4.4 Are Generalist Banks Less Vulnerable to Systemic Liquidations?

The probit regressions in Table 6 were quite informative in terms of predicting which banks are likely to be generalists. However, we did not learn very much regarding whether generalist banks are more or less risky compared to specialist banks, in particular from a systemic perspective. The last and final step is therefore to test whether generalist banks are more or less vulnerable to systemic asset liquidations.

Here we calculate the vulnerability indicators of Greenwood et al. (2015), who present a simple model of bank deleveraging. Their main model assumptions are:

1. banks target their leverage,

2. banks hold their investment portfolio weights constant,

3. asset sales/purchases generate price impact.

The basic idea is that, in response to a negative return on their asset portfolios, banks will have to sell assets in order reach their target leverage. These asset liquidations will occur proportional to the actual portfolio weights and will, due to less than perfectly liquid asset markets, generate an additional second-round impact on prices.[32] These price impact will then affect other banks who are exposed to the liquidated assets themselves, thereby leading to additional asset liquidations. Within this simple framework, Greenwood et al. (2015) propose several bank-specific and aggregate vulnerability indicators. In the following, we will focus on one specific indicator, namely *indirect vulnerability*. Bank i's indirect vulnerability with respect to a given shock is defined as the percentage of its equity that is wiped out by other banks' deleveraging. In our application, we parametrize the system as follows:

- portfolio size is set to the observed value of each bank's loan portfolio.

[32]The model is only concerned with direct and indirect effects of selling pressure on other market participants and ignores broader macroeconomic effects. Ali Anari and Mason (2005) show that asset liquidations of failed banks can have a sizable impact on output in the short- to medium-term.

- Leverage, as defined above, is set to the observed value for each bank.[33]
- All assets have the same price impact parameter of $10^{(-12)}$, and we set all cross-price impacts to zero.[34]

In the following, we explore systematic shocks in which we impose an initial shock of -1% on all outstanding loan amounts. For each year, we apply the model of Greenwood et al. (2015) separately and calculate banks' indirect vulnerabilities. Given the benefits of diversification, we expect generalist banks to be less vulnerable to these systemic asset liquidations.

Note that both Greenwood et al. (2015) and Duarte and Eisenbach (2014) apply the same model to various asset classes of banks' asset portfolios (including corporate loans), while we focus exclusively on corporate loans here. Clearly, these types of securities are less liquid compared to other instruments, but are, as pointed out by Drucker and Puri (2009), in fact traded on reasonably liquid secondary markets.

Results. The results are shown in Table 8. We regress banks' indirect vulnerabilities on their contemporaneous generalist levels (results are almost identical when using lagged values), lagged leverage ratios, and bank- and time-fixed effects. In order to additionally acknowledge a potential change in the sign of the relationships, we run separate regressions for the different subperiods.

Looking at the whole sample period, we find that generalist banks tend to be less vulnerable compared to specialists.[35] This is to be expected, given that their higher diversification

[33] In their sample of large EU-banks, Greenwood et al. (2015) reduce the impact of extremely high leverage values by using a cutoff value for leverage of 30, which affects roughly 20% of their sample banks. We experimented with various (data-driven) cutoffs, where we winsorized the top $x\%$ of the leverage values. This does not affect any of the qualitative results reported in the main text.

[34] In this exercise, it only matters that price impact is positive; it can be seen as a scaling parameter for the vulnerability indicator, but it does not affect the sign of the relationship with banks' generalist levels.

[35] We also experimented with idiosyncratic shocks, where we shock one industry (rather than all industries) at a time and then averaged the vulnerability estimates across the different shock scenarios. Given the linearity of the model, it turns out that the vulnerabilities are almost identical for systematic and idiosyncratic shocks, such that all of the results are robust to this alternative shock scenario.

	Indirect Vulnerability		
	All	1980-1995	1996-2013
γ_t^B	-0.0509***	-0.0251	-0.1202***
	(0.0195)	(0.0234)	(0.0269)
log(Leverage$_{t-1}$)	0.7110***	0.9371***	0.4326***
	(0.0741)	(0.0374)	(0.0866)
Bank FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
adj. R^2	0.798	0.802	0.845
Obs.	4,094	1,995	2,099

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 8: Are generalist banks less vulnerable to systemic liquidations? Results from OLS regressions of bank's indirect vulnerabilities on their generalist levels, γ^B (robust standard errors in parentheses). Bank i's indirect vulnerability is defined as the percentage of its equity that is wiped out by other banks' deleveraging, after a negative shock of -1% on all asset returns (Greenwood et al. (2015)).

levels should make them less prone to shocks. However, these results are driven by the post-1996 period. In fact, during the first half of the sample period there is no significant difference between generalists and specialists in terms of their vulnerabilities. This result is mainly driven by the fact that generalist banks tended to be highly leveraged in the pre-1996 period, while the opposite was true in the post-1996 period (see Appendix H). In fact, Table 8 shows that higher leverage ratios are associated with higher IVs, but the strength of the relationship has decreased over time. Hence, it appears that high leverage can undo some benefits of diversification, at least in the model of Greenwood et al. (2015).

7 Conclusions

In this paper, we propose a method to analyze the structure of the credit market and apply it to the case of Japanese banks' loan portfolios. We find that a stylized model of generalists and specialists captures the main features of the credit network. This indicates that the interactions between banks and the real economy can be described by a reasonably simple structure, where generalist banks interact with firms from all kinds of industries and specialist banks tend to focus their lending on generalist industries. Quite remarkably, despite the fact that the Japanese banking system underwent substantial changes in its institutional features over the sample period, there is no obvious structural break in the fit of the model. Rather, we observe a very slow deterioration of the fit.

Our findings suggest several interesting avenues for future research. First and foremost, it is important to ask whether other banking systems display a similar generalist-specialist architecture. In line with the empirical finding of a general core-periphery structure of interbank networks, we expect that the generalist-specialist architecture is not unique to the Japanese banking system. Given that our dataset exhibits a sample bias (i.e., we only observe borrowing of listed companies), we look forward to applications of the model to more complete datasets. Another important question is how the structure of the credit network affects systemic risk as a whole. Is the observed generalist-specialist architecture more or less vulnerable from a systemic perspective? Finally, the characteristics of generalist banks match those of money-center banks in interbank networks. It is of utmost importance to explore interactions between these networks, both theoretically and empirically, and incorporate them in broader macroeconomic models.

References

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

Acharya, V. V., Hasan, I., and Saunders, A. (2006). Should banks be diversified? Evidence from individual bank loan portfolios. *Journal of Business*, 79(3):1355–1412.

Acharya, V. V., Imbs, J., and Sturgess, J. (2011). Finance and efficiency: Do bank branching regulations matter? *Review of Finance*, 15:135–172.

Acharya, V. V. and Yorulmazer, T. (2008). Information contagion and bank herding. *Journal of Money, Credit and Banking*, 40(1):215–231.

Ali Anari, J. K. and Mason, J. (2005). Bank asset liquidation and the propagation of the U.S. great depression. *Journal of Money, Credit and Banking*, 37(4):753–773.

Armenter, R. and Koren, M. (2014). A balls-and-bins model of trade. *American Economic Review*, 104(7):2127–51.

Berger, A. N. and Black, L. K. (2011). Bank size, lending technologies, and small business finance. *Journal of Banking & Finance*, 35(3):724–735.

Berger, A. N., Goulding, W., and Rice, T. (2014). Do small businesses still prefer community banks? *Journal of Banking & Finance*, 44:264–278.

Berger, A. N., Miller, N. H., Petersen, M. A., Rajan, R. G., and Stein, J. C. (2005). Does function follow organizational form? evidence from the lending practices of large and small banks. *Journal of Financial economics*, 76(2):237–269.

Berger, A. N., Rosen, R. J., and Udell, G. F. (2007). Does market size structure affect competition? the case of small business lending. *Journal of Banking & Finance*, 31(1):11–33.

Berger, A. N. and Udell, G. F. (2006). A more complete conceptual framework for sme finance. *Journal of Banking & Finance*, 30(11):2945–2966.

BIS (2001). Group of ten report - consolidation in the financial sector. Standard documents, Bank for International Settlements (BIS).

BIS (2014). Supervisory framework for measuring and controlling large exposures. Standard documents, Bank for International Settlements (BIS).

Bonfim, D. and Dai, Q. (2017). Bank size and lending specialization. *Economic Notes*, 46(2):329–379.

Borgatti, S. P. and Everett, M. G. (2000). Models of core/periphery structures. *Social Networks*, 21(4):375–395.

Boyd, J. P., Fitzgerald, W. J., Mahutga, M. C., and Smith, D. A. (2010). Computing continuous core/periphery structures for social relations data with MINRES/SVD. *Social Networks*, 32(2):125–137.

Caballero, R. J., Hoshi, T., and Kashyap, A. K. (2008). Zombie lending and depressed restructuring in Japan. *American Economic Review*, 98:1943–1977.

Caccioli, F., Shrestha, M., Moore, C., and Farmer, J. D. (2014). Stability analysis of financial contagion due to overlapping portfolios. *Journal of Banking & Finance*, 46:233–245.

Calomiris, C. W. (2006). *U.S. bank deregulation in historical perspective*. Cambridge University Press.

Canales, R. and Nanda, R. (2012). A darker side to decentralized banks: Market power and credit rationing in sme lending. *Journal of Financial Economics*, 105(2):353–366.

Craig, B. and von Peter, G. (2014). Interbank tiering and money center banks. *Journal of Financial Intermediation*, 23(3):322–347.

De la Torre, A., Pería, M. S. M., and Schmukler, S. L. (2010). Bank involvement with smes: Beyond relationship lending. *Journal of Banking & Finance*, 34(9):2280–2293.

Degryse, H. and Ongena, S. (2005). Distance, lending relationships, and competition. *Journal of Finance*, 60(1):231–266.

Dell'Ariccia, G., Detragiache, E., and Rajan, R. (2008). The real effect of banking crises. *Journal of Financial Intermediation*, 17:89–112.

Demirgüç-Kunt, A. and Maksimovic, V. (2002). Funding growth in bank-based and market-based financial systems: evidence from firm-level data. *Journal of Financial Economics*, 65(3):337–363.

Detragiache, E., Garella, P., and Guiso, L. (2000). Multiple versus single banking relationships: Theory and evidence. *Journal of Finance*, 55(3):1133–1161.

DeYoung, R. (2010). Banking in the United States. In *The Oxford Handbook of Banking*. Oxford University Press, UK.

DeYoung, R., Frame, W. S., Glennon, D., and Nigro, P. (2011). The information revolution and small business lending: The missing evidence. *Journal of Financial Services Research*, 39(1-2):19–33.

Diamond, D. W. (1984). Financial intermediation and delegated monitoring. *Review of Economic Studies*, 51(3):393–414.

Drucker, S. and Puri, M. (2009). On loan sales, loan contracting, and lending relationships. *Review of Financial Studies*, 22(7):2835–2872.

Duarte, F. and Eisenbach, T. M. (2014). Fire-sale spillovers and systemic risk. Staff report, Federal Reserve Bank of New York.

Frame, W. S., Srinivasan, A., and Woosley, L. (2001). The effect of credit scoring on small-business lending. *Journal of money, credit and banking*, pages 813–825.

Frame, W. S. and Woosley, L. (2004). Credit scoring and the availability of small business credit in low-and moderate-income areas. *Financial Review*, 39(1):35–54.

Frankel, A. B., Montgomery, J. D., Friedman, B. M., and Gertler, M. (1991). Financial structure: an international perspective. *Brookings Papers on Economic Activity*, 1991(1):257–310.

Gerlach, M. L. (1992). *Alliance capitalism: The social organization of Japanese business*. Univ of California Press.

Glasserman, P. and Young, H. P. (2015). How likely is contagion in financial networks? *Journal of Banking & Finance*, 50:383–399.

Goddard, J., Molyneux, P., and Wilson, J. O. (2010). Banking in the European Union. In *The Oxford Handbook of Banking*. Oxford University Press, UK.

Greenwood, R., Landier, A., and Thesmar, D. (2015). Vulnerable banks. *Journal of Financial Economics*, 115(3):471–485.

Guiso, L. and Minetti, R. (2010). The structure of multiple credit relationships: Evidence from U.S. firms. *Journal of Money, Credit and Banking*, 42(6):1037–1071.

Hayden, E., Porath, D., and von Westernhagen, N. (2007). Does diversification improve the performance of German banks? Evidence from individual bank loan portfolios. *Journal of Financial Services Research*, 32(3):123–140.

Hoggarth, G., Reis, R., and Saporta, V. (2002). Costs of banking system instability: some empirical evidence. *Journal of Banking & Finance*, 26(5):825–855.

Hoshi, T. and Kashyap, A. (2004). *Corporate financing and governance in Japan: the road to the future*. MIT Press.

Hoshi, T. and Kashyap, A. K. (2010). Will the US bank recapitalization succeed? eight lessons from japan. *Journal of Financial Economics*, 97(3):398–417.

Hüfner, F. (2010). The German banking system: lessons from the financial crisis. OECD working paper.

Ibragimov, R., Jaffee, D., and Walden, J. (2011). Diversification disasters. *Journal of Financial Economics*, 99(2):333–348.

Itō, T. (1992). *The Japanese Economy*. MIT Press.

Jackson, M. O. (2008). *Social and economic networks*. Princeton University Press.

Jayaratne, J. and Morgan, D. P. (2000). Capital market frictions and deposit constraints at banks. *Journal of Money, Credit and Banking*, 32(1):74–92.

Miwa, Y. and Ramseyer, J. M. (2002). The fable of the keiretsu. *Journal of Economics & Management Strategy*, 11(2):169–224.

Morgan, D. P., Rime, B., and Strahan, P. E. (2004). Bank integration and state business cycles. *Quarterly Journal of Economics*, 119(4):1555–1584.

Ongena, S. and Yu, Y. (2016). Firm industry affiliation and multiple bank relationships. *Journal of Financial Services Research*, forthcoming.

Ono, A., Hasumi, R., and Hirata, H. (2014). Differentiated use of small business credit scoring by relationship lenders and transactional lenders: evidence from firm-bank matched data in japan. *Journal of Banking & Finance*, 42:371–380.

Peek, J. and Rosengren, E. S. (2005). Unnatural selection: perverse incentives and the misallocation of credit in Japan. *American Economic Review*, 95(4):1144–1166.

Rajan, R. G. (2005). Has financial development made the world riskier? NBER working paper.

Shin, G. H. and Kolari, J. W. (2004). Do some lenders have information advantages? Evidence from Japanese credit market data. *Journal of Banking & Finance*, 28(10):2331–2351.

Shleifer, A. and Vishny, R. W. (1992). Liquidation values and debt capacity: A market equilibrium approach. *The Journal of Finance*, 47(4):1343–1366.

Stomper, A. (2006). A theory of banks' industry expertise, market power, and credit risk. *Management Science*, 52(10):1618–1633.

Strömberg, P. (2000). Conflicts of interest and market illiquidity in bankruptcy auctions: Theory and tests. *The Journal of Finance*, 55(6):2641–2692.

Tabak, B. M., Fazio, D. M., and Cajueiro, D. O. (2011). The effects of loan portfolio concentration on Brazilian banks' return and risk. *Journal of Banking & Finance*, 35:3065–3076.

Uchida, H. and Udell, G. (2010). Banking in Japan. *Oxford Handbook of Banking*, pages 903–940.

Uchida, H., Udell, G. F., and Watanabe, W. (2008). Bank size and lending relationships in Japan. *Journal of the Japanese and International Economies*, 22(2):242–267.

Wagner, W. (2010). Diversification at financial institutions and systemic crises. *Journal of Financial Intermediation*, 19(3):373–386.

Wagner, W. (2011). Systemic liquidation risk and the diversity–diversification trade-off. *Journal of Finance*, 66(4):1141–1175.

Williamson, O. E. (1988). Corporate finance and corporate governance. *The Journal of Finance*, 43(3):567–591.

Winton, A. (1999). Dont put all your eggs in one basket? Diversification and specialization in lending. Working paper.

Woo, D. and Kanaya, A. (2000). The Japanese banking crisis of the 1990's: Sources and lessons. IMF working paper.