

Saïd Business School  
Research Papers



---

March 2016

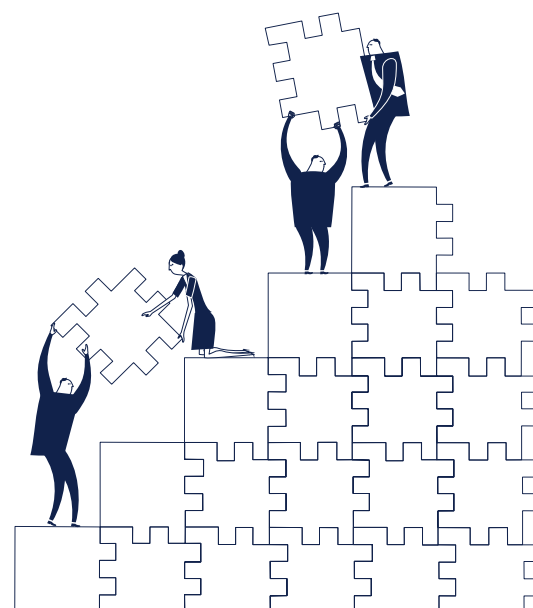
# Generalists and Specialists in the Credit Market

Daniel Fricke

*Saïd Business School, University of Oxford; Institute for New Economic Thinking at the Oxford Martin School; Kiel Institute for the World Economy*

Tarik Roukny

*Université Libre de Bruxelles (ULB)*



---

Saïd Business School RP 2016-11

The Saïd Business School's working paper series aims to provide early access to high-quality and rigorous academic research. Oxford Saïd's working papers reflect a commitment to excellence, and an interdisciplinary scope that is appropriate to a business school embedded in one of the world's major research universities..

This paper is authorised or co-authored by Oxford Saïd faculty. It is circulated for comment and discussion only. Contents should be considered preliminary, and are not to be quoted or reproduced without the author's permission.

# Generalists and Specialists in the Credit Market

Daniel Fricke\*      Tarik Roukny†

## Abstract

In this paper, we explore the cross-section of Japanese banks' industrial loan portfolios over the last 34 years. We show that banks with diversified lending (generalists) and banks with focused lending (specialists) coexist. Similarly, industries with diversified borrowing (generalists) and industries with focused borrowing (specialists) also coexist. Interestingly, specialist banks and specialist industries rarely interact, suggesting significant overlap in banks' loan portfolios. We introduce a model to describe these interaction patterns and identify a persistent and economically meaningful set of generalist banks/industries. We find that size is an important determinant for being a generalist. Finally, we show that generalist banks tend to be less vulnerable compared to specialist banks.

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

**JEL Classification:** G11, G20, G21, G28, G32

---

\*Corresponding author. Saïd Business School and Institute for New Economic Thinking, University of Oxford; email: daniel.fricke@sbs.ox.ac.uk

†Université Libre de Bruxelles; Fonds National de la Recherche Scientifique; email: troukny@ulb.ac.be

# I Introduction

Whether banks should diversify their loan portfolios or focus on a small number of industries where they have special expertise remains an open research question. In fact, banks often face conflicting incentives which may encourage either diversification or specialisation.<sup>1</sup> On the one hand, banks that extend loans to firms from many different economic industries should be, through the benefits of diversification, less affected by firm- or industry-specific shocks. On the other hand, there is no doubt that gaining industry-specific expertise, e.g., via the screening and monitoring of a particular type of firm, is valuable but costly to banks (e.g., Stomper (2006)). By focusing on relatively few types of businesses, banks might therefore be able to improve their performance, but will be more vulnerable to industry-specific shocks.

In 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 interact only with firms from a relatively small subset of industries.

Neither the theoretical nor the empirical literature offer a unanimous recommendation on whether it is optimal for banks to be generalist or specialist.<sup>2</sup> Not surprisingly, existing empirical evidence indicates that banks' levels of diversification can be quite heterogeneous (e.g., Acharya et al. (2006) for Italy, and Hayden et al. (2007) for Germany), but little is known about the prevalence of generalists and specialists in these systems. We seek to fill this gap.

In this paper, we explore the characteristics of generalist and specialist banks in

---

<sup>1</sup>For example, many regulatory frameworks have imposed upper limits on a bank's exposure with regards to individual borrowers (e.g., BIS (2014)).

<sup>2</sup>On the theoretical side, Diamond (1984) finds that the benefits from delegated monitoring are maximized when banks are completely diversified, whereas Stomper (2006) uses an equilibrium model to show that generalists and specialists coexist. Winton (1999) uses a model where the gains from diversification depend on the riskiness of the bank, finding that medium-risk banks should diversify as much as possible, while low- and high-risk banks should be specialized. On the empirical side, the results are mixed with regard to whether diversification is beneficial for individual banks. For example, 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.

more depth using detailed data on Japanese banks' industrial loan portfolios over the period 1980 - 2013. By the same token, we are also interested in analysing the interaction patterns for the borrowing side of the credit network. For this purpose, we introduce a *generalist-specialist model* that allows us to identify the different types of banks and industries based on the interactions between these actors in the credit market.

A major advantage of our generalist-specialist model is that we can explicitly study the heterogeneity of these groups' lending and borrowing portfolios respectively. One interesting question in this regard is whether specialists tend to occupy niches, e.g., to what extent focused banks tend to invest in certain industries where few other banks are present. In other words: are specialist banks indeed *special*? We would expect this to be the case, since otherwise there is little room for gaining superior information relative to all competitors with the very same strategy. In fact, competition should lower the expected profits from gathering this particular information. Our major finding is that specialist banks are not special at all, since they tend to interact with the very same generalist industries.

Figure 1 provides a graphical network representation of the interactions between Japanese banks and industries from our data in the year 1980, and illustrates the main results of this paper. First, we find that a similar dichotomy of generalists and specialists can be applied to both the credit supply side, i.e., banks (nodes on the left), and the credit demand side, i.e., industries (nodes on the right). In Figure 1 nodes are sorted according to their number of connections, such that generalists are sorted above the specialists. It is apparent even to the naked eye that highly diversified banks coexist alongside much more specialised ones and a similar result holds for the industries' borrowing partners. Somewhat surprisingly, however, when we explore the connection patterns in the credit network in Figure 1, it also becomes clear that specialist banks tend to interact with generalist industries, and similarly that specialist industries tend to interact with generalist banks. Hence, the set of specialist banks tends to concentrate their loan portfolios on the very same economic

industries, which implies a significant overlap in their loan portfolios. The theoretical literature suggests different reasons for such a strong portfolio overlap (see, e.g., Rajan (2005); Acharya and Yorulmazer (2008); Wagner (2010)), but it is not clear which of these theories are most relevant in the case of banks' loan portfolios. Again, a similar remark can be made about the borrowing patterns of specialist industries: they tend to borrow from the same set of generalist banks, suggesting that these industries might be rather vulnerable to the funding behavior of a small set of banks. Overall, there are very few interactions between specialist banks and specialist industries - nodes closer to the bottom of Figure 1 are almost unconnected.

In order to better understand the patterns present in Figure 1, we introduce a stylized model of generalists and specialists that captures the observed interaction patterns in the credit network. The basic model assumptions are that generalists (specialists) will have as many (few) links as possible. We relate the data to the idealized network structure, which allows us to classify generalists and specialists for any given interaction matrix. The fit of the model is significant: the number of inconsistencies between the model and the observed (binary) credit data ranges between 30-45%, which is significantly low compared to various randomization methods. Interestingly, the fit of the model deteriorates over time, probably due to the fact that the Japanese banking system underwent various crisis periods and significant adjustments in the institutional features of the financial system (culminating in the so-called 'Big Bang'). Nevertheless, both the number of generalist banks and generalist industries as identified by our model, appear to be relatively stable over time. Moreover, given the high persistence of bank-firm interactions in Japan, it is not surprising that individual economic actors consistently remain either generalists or specialists over time. In this regard, we show that the strategy profiles (i.e., being a generalist or a specialist) can be predicted using additional information on the banks and industries, respectively. 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, are important as well. Finally,

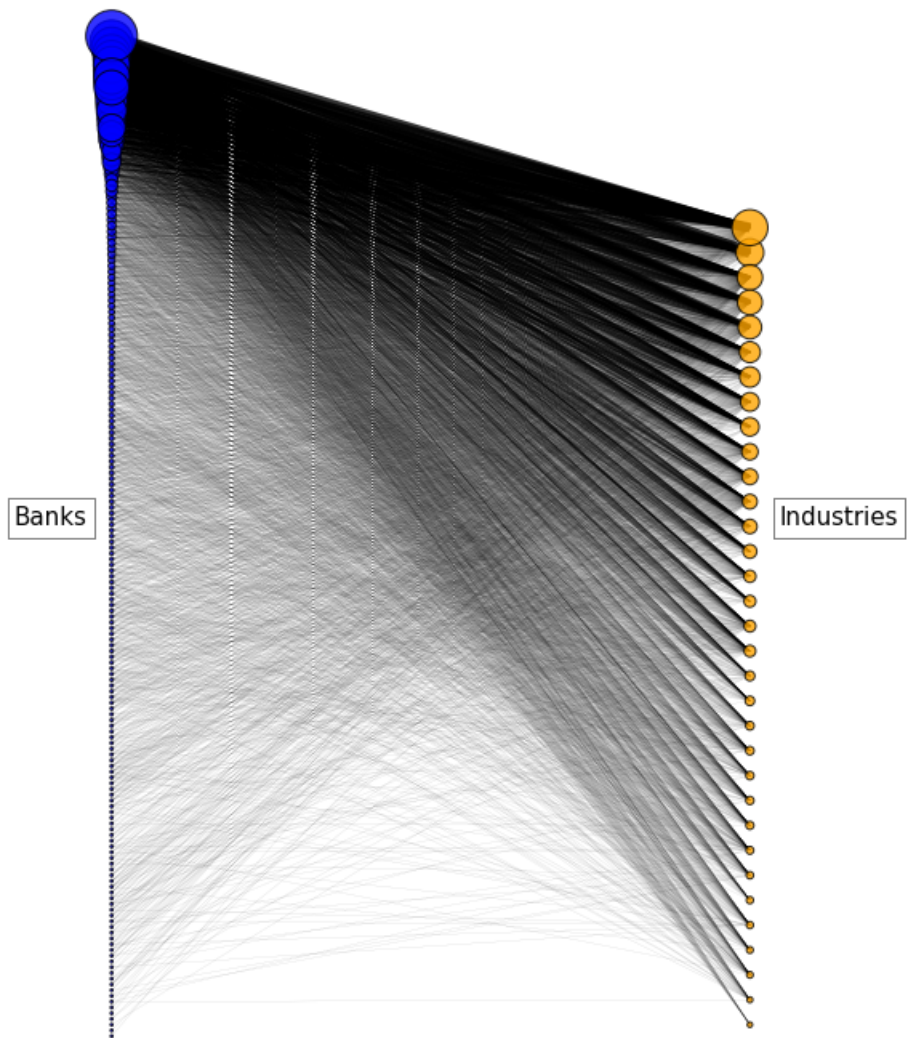


Figure 1: The Japanese credit network in 1980 (binary version). Banks are depicted as blue nodes on the left and industries as yellow nodes on the right. Banks and industries are connected through borrowing and lending activities on the credit market - these links are shown in black. Nodes are arranged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

we run a vulnerability analysis on the banks and analyse how their position in the credit network affects their riskiness. Using the model of Greenwood et al. (2015), we show that, compared with specialist banks, generalist banks are systematically less vulnerable to industry-specific shocks, at least for the second half of the sample period. This suggests that diversification indeed makes generalist banks less prone to such idiosyncratic shocks. Interestingly, it turns out that generalist banks are also less vulnerable than specialist banks with regards to common shocks that affect all industries.

Our work makes several contributions. First, our model of generalists and specialists allows us to describe the interactions between banks and the real economy in a very simple way. Our methodology classifies nodes strategies (i.e., generalist or specialist) based on the observed interaction matrix for both the liquidity supply and demand sides in a structural way. In line with recent findings for interbank networks (see Craig and Von Peter (2014)), generalist banks and generalist industries are central to the structure of the system and their activity accounts for a large part of the monetary flows in the economy.

Second, the fact that specialists tend to concentrate their activity on generalists opposes the idea of expert knowledge acquired in isolated niches, i.e., *specialists are not that special*. Moreover, the relatively small number of specialist-specialist interactions highlights a strong overlap in banks' loan portfolios, which is important from a systemic perspective since specialist banks will be prone to the very same shocks (see Caccioli et al. (2014); Cai et al. (2014); Greenwood et al. (2015)).<sup>3</sup> As such, our work also contributes to the literature on the homogenization of the banking system and the associated risks (see Wagner (2010); Wagner (2011); Fricke (forthcoming)). Interestingly, we do not observe that all banks tend to become generalists over time - rather, even at the end of our sample period in 2013, we still can distinguish between generalists and specialists in our data.

---

<sup>3</sup>We aim at exploring the relationship between portfolio overlap and systemic risk in future research. The basic idea is that the observed investment structure and the subsequent high level of portfolio overlap renders the system more vulnerable under distressed conditions.

Third, the observed interaction pattern between banks and economic industries carries important implications for the literature on the determinants of the number of credit relationships per firm.<sup>4</sup> Our findings suggest that the way banks specialize into certain industries makes firm-industry affiliations an important determinant for the number of bank relationships (see Ongena and Yu (2013)). Hence, firms from generalist industries are likely to have significantly more bank relationships compared to firms from specialist industries; in simple terms, they have more options to find banks that are willing to lend to them.

Lastly, the econometric part of this paper provides a link between nodes' network position and their individual characteristics. In line with the view that different types of banks are likely to build different patterns of links, we find that a small set of bank-specific variables reliably predict whether banks tend to diversify or focus. Ever since the seminal contributions of both Allen and Gale (2000) and Freixas and Rochet (2000), financial contagion in interbank networks has been studied in depth. However, our findings suggest that we need to look at broader financial networks, which also take banks' lending to the real economy into account.

The remainder of this paper is structured as follows: section II provides the necessary background for the generalist-specialist model in section III. In section IV we provide a brief overview of the Japanese banking system. In section V, we present the empirical results on the generalist-specialist model, explore the determinants of being a generalist, and also show that generalist banks tend to be less vulnerable based on the vulnerability indicators of Greenwood et al. (2015). Section VI summarizes the main results and concludes.

---

<sup>4</sup>Guiso and Minetti (2010) provide an overview of the various explanations for why firms should optimally borrow from more than one bank. For example, in Detragiache et al. (2000), a firm borrows from multiple banks in order to insure itself against negative liquidity shocks affecting its main bank.



## II The Credit Network

The credit network consists of two distinct sets of nodes, where the first set contains a total number of  $n^B$  nodes (banks), and the second set a total of  $n^I$  nodes (industries). A link exists between a bank and an industry when there is a credit relationship between two nodes.<sup>5</sup> 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.

We represent the credit network as a rectangular matrix of size  $(n^B \times n^I)$ , which we will denote by  $\mathbf{W}$ . An element  $w_{i,j}$  of this matrix represents the total value of credit extended by bank  $i$  to industry  $j$  within a given period.<sup>6</sup> 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 adjacency matrix that will be of interest in the following. We denote this matrix as  $\mathbf{A}$ , where  $a_{i,j} = 1$  if  $w_{i,j} > 0$  and 0 otherwise. In other words, this matrix informs on the link existence between a bank and an industry.<sup>7</sup> The total number of links in the market is denoted as  $m$ :

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

In the following, we will focus on the yearly, binary adjacency matrices,  $\mathbf{A}$ . Hence, for each calendar year we observe a snapshot of the outstanding loan volumes of banks with respect to the different industries. Our focus 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. If a bank only interacts with firms from a given industry, the total number of counterparties of a bank can be a poor measure of diversification.

---

<sup>5</sup>Note that the credit network is aggregated in the sense that banks interact with firms, which themselves are affiliated with some economic industry.

<sup>6</sup>We drop time subscripts in the following.

<sup>7</sup>In international trade,  $\mathbf{A}$  corresponds to the extensive margin, while the non-zero elements of  $\mathbf{W}$  would correspond to the intensive margin, see Armenter and Koren (2014).

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.<sup>8</sup> 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. Exploring the effects of industry-specific shocks, rather than firm-specific shocks, is also in line with recent work on input-output networks, see Acemoglu et al. (2012).

### III The Generalist-Specialist Model

How should one classify generalists and specialists? In this section we propose a stylized model of bank credit interactions that is governed by two sets of actors for both banks and industries, respectively. *Generalists* interact with all actors from the other set of nodes; for example, a generalist bank extends credit to all the different industries and a generalist industry receives funding from all banks. *Specialists* only interact with a small subset of actors from the other set of nodes. In particular, they only interact with generalists from the other side. Hence, a specialist bank will be active in the generalist industries and a specialist industry will only receive funding from the generalist banks. Figure 2 illustrates our model with a small two-mode network made of only generalist and specialist agents. Note that generalists are connected with all nodes from the other side while the specialists only interact with the generalists of the other side. Thus, specialists from the two different sets of nodes are not connected with each other.

Our model is an adjusted version of the core-periphery model of Borgatti and Everett (2000) that was successfully applied to unipartite banking networks (e.g., Craig and Von Peter (2014)). Here we use a similar approach for the bipartite credit

---

<sup>8</sup>For example, Detragiache et al. (2000) report that the median number of bank relations for small businesses in the U.S. and Italy is 2 and 5, respectively, and the share of firms with only one bank relationship is 44.5% and 11%.

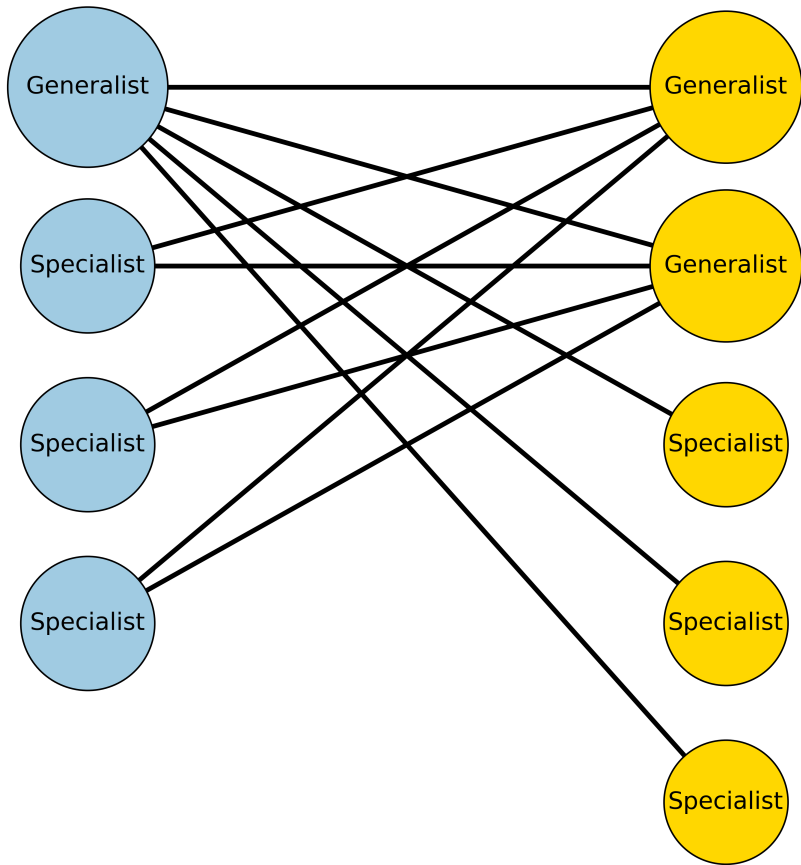


Figure 2: Illustration of the network structure implied by the generalist-specialist model. Generalist nodes from each side interact with all nodes from the other side, while specialists only interact with generalists.

networks of banks and industries. The discrete model partitions banks and industries such that generalist (specialist) nodes are maximally (minimally) connected to each other.<sup>9</sup> In line with the idea that generalists are highly diversified, we require that they have as many links as possible. On the other hand, specialists should have as few links as possible.

We should stress that one major advantage of our approach is that we do not have to specify an arbitrary cutoff value for defining generalists and specialists. For example, one might say that generalists are the  $x\%$  of nodes with the highest degree, but it is unclear what a good value for  $x$  would be. Instead, our approach for the classification of generalists and specialists is data-driven and does not rely on such ad-hoc cutoff values.

## A Setup

We have two sets of nodes,  $n^B$  banks and  $n^I$  industries, and we seek to decompose these into generalists and specialists, respectively. In order to identify the  $n_g^B$  generalist banks and  $n_g^I$  generalist industries, we aim at sorting the binary adjacency matrix,  $\mathbf{A}$ , such that we have the generalist-generalist region as a 1-block in the upper left part (of dimension  $n_g^B \times n_g^I$ ) and the specialist-specialist region as a 0-block in the lower right part (of dimension  $(n^B - n_g^B) \times (n^I - n_g^I)$ ). From an economic perspective, we make the following assumptions to identify 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.*

---

<sup>9</sup>The concept of discrete group membership can be extended by considering generalists and specialists as opposite ends of a continuum. In Appendix D, we present a continuous version of the model, where banks and industries can have heterogeneous ‘generalist levels’, respectively. In summary, we find that the discrete and the continuous model tend to give very similar results, even when we use the weighted credit networks,  $\mathbf{W}$ , rather than the binary adjacency matrices,  $\mathbf{A}$ .

Assumption 1 implies that generalists are connected with everyone, such that

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

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

where  $B^g$  and  $I^g$  are the sets of generalist banks and generalist industries respectively. Given that all nodes must have at least one link, i.e.,  $\sum_i a_{i,j} > 0 \quad \forall i$  and  $\sum_j a_{i,j} > 0 \quad \forall j$ , the second assumption therefore implies that

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

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

where  $B^s$  and  $I^s$  are the sets of specialist banks and specialist industries respectively. These results show that the following idealized pattern matrix ( $\mathbf{A}^*$ ) for a ‘pure’ generalist-specialist segmentation is in line with the above assumptions:

$$\mathbf{A}^* = \begin{pmatrix} \mathbf{GG} & \mathbf{GS} \\ \mathbf{GS} & \mathbf{SS} \end{pmatrix} = \begin{pmatrix} \mathbf{1} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{pmatrix}, \quad (1)$$

where  $\mathbf{1}$  and  $\mathbf{0}$  denote submatrices of ones and zeros,  $\mathbf{GG}$  stands for generalist-generalist interactions,  $\mathbf{SS}$  for specialist-specialist interactions, and so on.

The  $\mathbf{GG}$ -block contains the subset of highly interconnected banks and industries, while the  $\mathbf{SS}$ -block contains the specialists which should have as few links as possible. Our above assumptions determine that the off-diagonal blocks should be 1-blocks as well (each generalist being connected to all specialist nodes).

## B Optimization

In the following, we use the discrete generalist-specialist framework and adapt it to classify each bank and industry as a generalist or specialist, respectively. This classification can be summarized by two ‘generalist level’ vectors  $\gamma^B$  and  $\gamma^I$  of length  $n^B$  and  $n^I$ , respectively. For a generalist bank,  $\gamma_i^B = 1$ , and zero otherwise. Similarly,  $\gamma_j^I = 1$  for generalist industries, and zero 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 (see 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 simply 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 with arbitrary off-diagonal blocks, the aggregate errors of the individual blocks can be written as

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

where  $\gamma^B$  and  $\gamma^S$  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. On the other hand, the specialist-specialist block should be minimally connected, such that we count any existing link in this block as an error.

The total error score ( $e$ ) then simply aggregates the errors across the relevant blocks, normalized by the total number of links in the network. Formally this can be written as

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

with  $e(\cdot)$  being a function of the generalist level vectors since every possible partition is associated with a particular value of  $e$  and  $m$  being the total number of links in the credit network of banks and industries. Note that, by setting all nodes as specialists, the error score takes a value of one (see Appendix C for details). Hence, a ‘significant’ generalist-specialist structure should display an error score significantly smaller than one, while a perfect generalist-specialist structure would yield a value of zero. Below we will compare the observed error score with various null models in order to assess the significance in more depth.

In the following, we find the optimal generalist-level vectors using a genetic algorithm. Convergence is rapidly obtained and results are robust to changes in the GA parameters.<sup>10</sup> We also experimented with alternative optimization methods and all of the results shown below are not sensitive in this regard.

## IV Institutional Background

Before presenting empirical results, we provide a brief overview of the institutional features of the Japanese banking system. Clearly, a detailed description of the Japanese financial system and its history is beyond the scope of this paper. We refer the interested reader to Itō (1992); Hoshi and Kashyap (2004); Uchida and Udell (2010).

As Uchida and Udell (2010) point out, there are several reasons why the Japanese banking system is particularly interesting and 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.

---

<sup>10</sup>Details on the genetic algorithm and its implementation are available upon request from the authors.

3. The Japanese banking industry has several interesting idiosyncratic features related to the Japanese corporate environment such as the ‘main banking’ system.
4. 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 U.S. financial crisis. While certain features of Japan’s financial system are certainly unique, there are general lessons to be learned from its analysis.
5. We crucially rely on reliable micro-level data, and focusing on Japan is useful, since bank-firm interactions are recorded in public databases (which can be accessed for a subscription fee). 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., U.S.).

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

## *A The Japanese Financial System*

Historically, the Japanese banking system has been segmented, mainly due to legally mandated specialisations for different types of banks, and it still retains some of its original characteristics. 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.<sup>11</sup>

City banks have nationwide branches and provide wholesale lending to large corporate customers, accept individual deposits, and offer consumer loans. These banks

---

<sup>11</sup>Of lesser importance are Shinkin banks, credit cooperatives, and foreign banks. 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. Foreign banks’ market share in the Japanese corporate lending market is traditionally very low.



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. More than half of their lending is typically directed to small- and medium-sized firms. Similar to city banks, the regional banks are allowed to have nationwide branches, but the total number of branches and their location has to be approved by the supervisory authority.

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 the tier-1 regional banks, and 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 both sides of their balance sheet. With the collapse of the Long-term Credit Bank of Japan in the early 2000s, this bank type does not exist anymore as of today.

Finally, trust banks offer both financing and asset management services. They receive and manage funds on behalf of the money's owners, where the investments are typically longer-term.

With increased deregulation in the 1980s, different bank types started to compete with each other. The bursting of the asset price bubble in the early 1990s and its long-lasting impact on the 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.<sup>12</sup>

## *B The ‘Main Bank’ System*

Banking relationships are far more important in Japan compared to many other countries. Japanese firms rely much more strongly 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 sitting on the firms’ corporate board.<sup>13</sup> Hence, relationships are generally very long-term oriented; for example, Uchida et al. (2008) report an average duration of bank-firm relationships in Japan 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. However, the literature has also uncovered several dark sides of these close relationships. First, firms may have trouble finding alternative funding sources, when its main bank is in distress. Second, the main bank has a lot of inside information and may extract excessive rents from the firm. Finally, as documented by Peek and Rosengren (2005) and Caballero et al. (2008), Japanese banks misallocated credit by ‘evergreening’ loans to the weakest firms.

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 are more interested in the cross-sectional distribution of generalists and specialists in the credit network. Clearly, we expect larger banks to be more diversified, while smaller regional banks should have more concentrated portfolios.

---

<sup>12</sup>For example, using data on a large Belgian bank, Degryse and Ongena (2005) find that the median physical distance between the bank and its borrowers is 1.40 miles.

<sup>13</sup>Traditionally, groups of banks and firms used to be part of the same *keiretsu* group, but the importance of these groups has been repeatedly put to question, see for example Miwa and Ramseyer (2002).

## V Empirical Analysis

### A Data

Our main data source is the Nikkei NEEDS database, which provides extensive accounting and loan information for listed companies in Japan.<sup>14</sup> 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 includes firms listed in Japanese stock markets, but since 1996 also includes firms that are traded in the OTC market (or JASDAQ as it called nowadays). Most firms’ fiscal years end in March, and for the sake of simplicity we will refer to calendar years in everything that follows. We complement the data with additional characteristics of the banks and firms obtained from the ‘Corporate Financial Information’ and the ‘Corporate Attribute’ parts of the NEEDS database and our final sample covers all years between 1980 and 2013. Unfortunately, the dataset contains only the most recent industry codes for each of the firms. Therefore, industry affiliations are likely to be most accurate for the most recent years in the sample.<sup>15</sup> Nevertheless, we still include all years in our sample since the observed structures are generally very persistent.

### B Summary Statistics

Table 1 provides a few summary statistics for the data. We define active banks as those banks with at least 5 loan relationships (minimum degree of 5) in a given year and active industries need to consist of at least 5 firms in a given year.<sup>16</sup> The number

---

<sup>14</sup>More details can be found here [https://www.nikkeieu.com/needs/needs\\_data.html](https://www.nikkeieu.com/needs/needs_data.html).

<sup>15</sup>A complete list of all industries in the dataset can be found in Appendix A.

<sup>16</sup>We experimented with different cutoffs and find that the qualitative results remain unaffected.

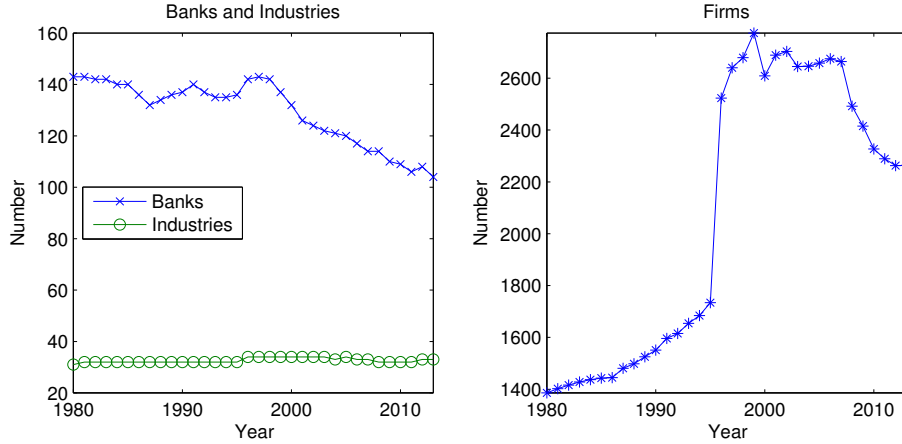


Figure 3: Number of active banks and industries (left panel) and number of active firms (right panel).

of active banks in the sample varies over time, with an average value of 129. In fact, as in many other banking systems, the number of active banks has been steadily declining from the mid-1990s onwards, see the left panel of Figure 3. The number of active firms is, not surprisingly, 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-listed companies from 1996 onwards. Therefore, since the jump in the active number of firms could have an effect on some of the results, we check the robustness of our findings for the post-1995 period. The number of active industries is very stable over time, with an average value of 32. We also see some heterogeneity in the number of firms per industry. While the average value is around 63, we observe values as low as 5 (our lower cutoff value) and as large as 463. The total loan volume is on the order of 5 trillion Japanese Yen.

Table 1 also shows summary statistics for the basic bank-firm networks and the aggregated bank-industry credit networks which are the main focus of this paper. Banks interact with 144 firms on average (bank-degree), but there is a lot of heterogeneity in these values and we will illustrate the distribution of bank-degrees in more detail below. Interestingly, firms borrow from 9 banks on average (firm-degree),

<b>Summary statistics</b>					
<i>General</i>	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
<i>Bank-firm network</i>					
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
Links	9.05	7	1	104	8.04
Firm-degree	0.297	0.243	0.017	1	0.205
<i>Bank-industry network</i>					
Links	2,283.8	2,389.0	1,752.0	2,698.0	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 1: 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 simply 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.

which is a quite a large value compared to the values reported by Detragiache et al. (2000) for the U.S. and Italy, respectively. It may well be that, by focusing on relatively large, listed firms, there is an upward bias in the number of borrowing partners for the firms. Finally, for the bank-industry credit network, we also observe a lot of heterogeneity in the bank- and industry-degrees, respectively. On average, banks interact with firms from 17 different industries, while industries receive funding from 70 banks.

### **B.1 Evidence of Coexisting Strategies**

We take a closer look at the distributional aspects of bank-industry interactions, mainly in order to highlight the coexistence of different diversification strategies. Figure 4 shows the distribution of the number of industries that banks connect to (bank-degrees) for the entire sample period. In order to facilitate the comparison of values from different time periods, we normalize the bank-degrees by their maximum possible value for each period (the number of active industries in a given year). The top left panel shows the results for the entire sample period, while the top right and bottom panel show the results when using only data for the periods 1980 - 1995 and 1996-2013, respectively. For all periods, the histograms show a continuous spectrum of interaction patterns, with the coexistence of high, medium and low numbers of interactions.

Figure 5 shows the same distributional plots for the industries. The results deviate from the bank case, since most industries receive funding from an intermediate number of banks. Nevertheless, we still observe that some industries receive funding from very few banks and others receive funding from many banks.

We see that the distribution of degrees is very broad for both banks and industries. Hence, it is neither the case that all banks are very diversified, nor that they are all very focused. Rather, it turns out that various behavioral strategies can be observed, suggesting that generalists and specialists indeed tend to coexist. In what follows we show that, despite the existence of intermediate strategies, our simple and intuitive

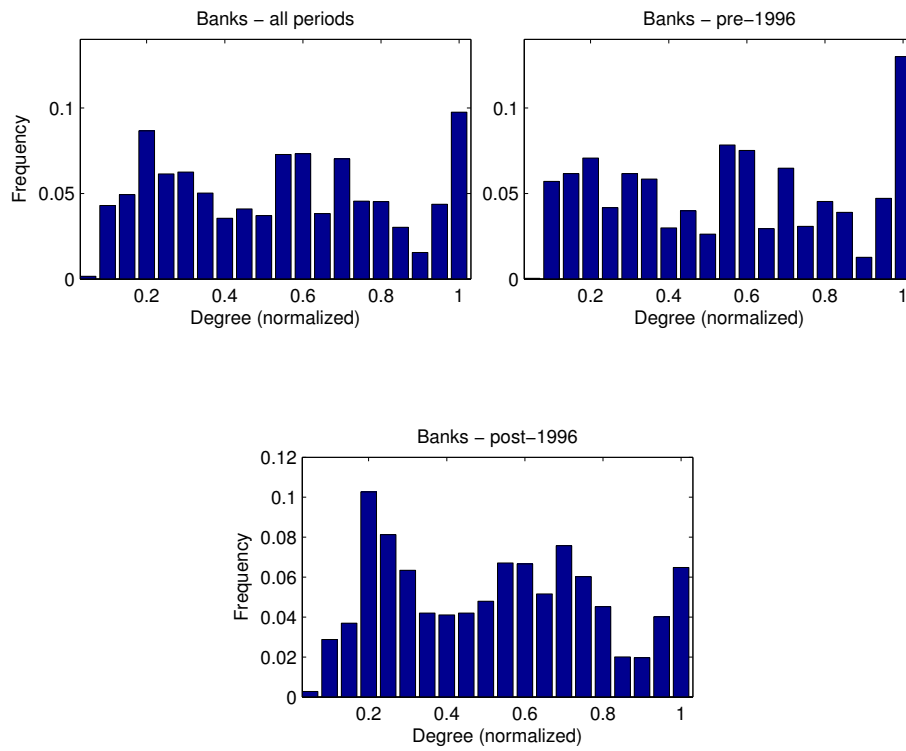


Figure 4: Histogram of normalized bank-degrees (number of connections). Top left: all observations. Top right: data before 1996. Bottom: data after 1996.

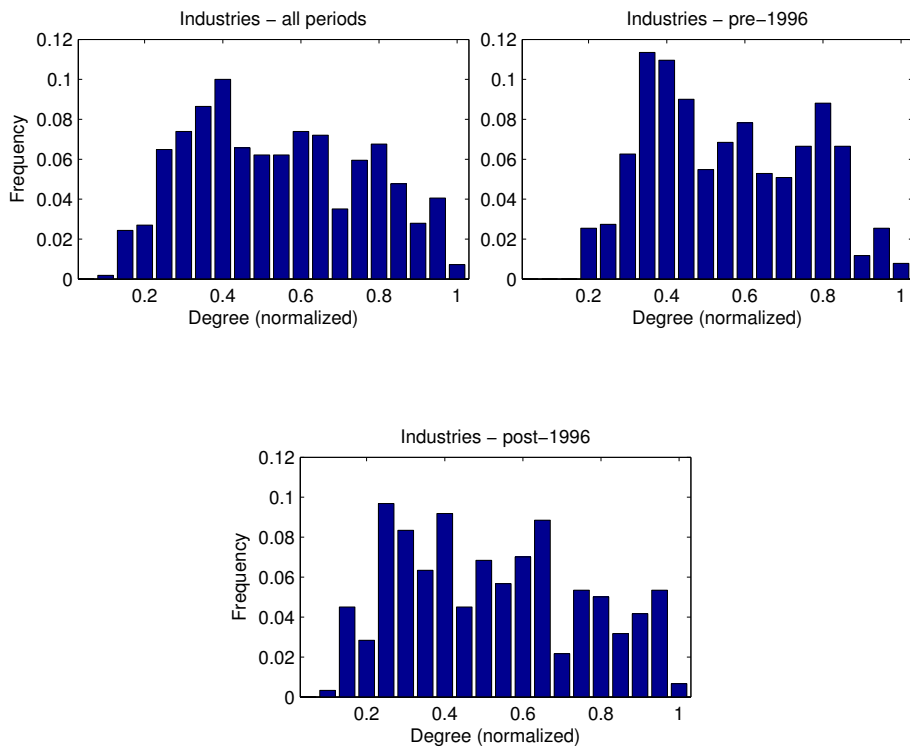


Figure 5: Histogram of normalized industry-degrees (number of connections). Top left: all observations. Top right: data before 1996. Bottom: data after 1996.



generalist-specialist dichotomy does a good job in capturing significant features of the interaction patterns of the credit market.

## B.2 Evidence of Rare Specialist-Specialist Interactions

The next step is to get an understanding of the interaction patterns between banks and industries. For this purpose, Figure 6 shows the (normalized) *average nearest neighbor degrees* (ANND) for banks and industries, respectively, in 1990.<sup>17</sup> For each bank, ANND measures the average degree (number of banks) of the industries that this bank interacts with. In the left panel, we plot the ANND for each bank against its normalized degree, i.e., relative to the maximum possible value. The right panel shows the same statistics for the industries. In both cases there is a clear negative relationship between the degree and the ANND.<sup>18</sup> This indicates that low degree banks (specialists) tend to interact with higher degree industries (generalists), and high degree banks (generalists) tend to interact with lower degree industries (specialists). Similarly, low degree industries (specialists) tend to interact with higher degree banks (generalists), while high degree industries (generalists) tend to interact with lower degree banks (specialists). Overall, these results indicate that interactions between specialist banks and specialist industries are indeed rare, since otherwise there would not be a negative relationship. Therefore, one of the main implications of the generalist-specialist model, namely a sparse specialist-specialist block, seems to be supported by this preliminary data analysis.

## C Generalist-Specialist Model

**Main Results.** Here we present the main results from fitting the generalist-specialist model for each year between 1980 and 2013. Figure 7 shows the number of generalist banks and industries over time, respectively, where the left panel shows the absolute number of generalist banks and industries, and the right panel shows the relative

---

<sup>17</sup>The results are very stable over time and there is no particular reason why we show the values for this year.

<sup>18</sup>In networks jargon, the interaction pattern is disassortative (see Newman (2010)).

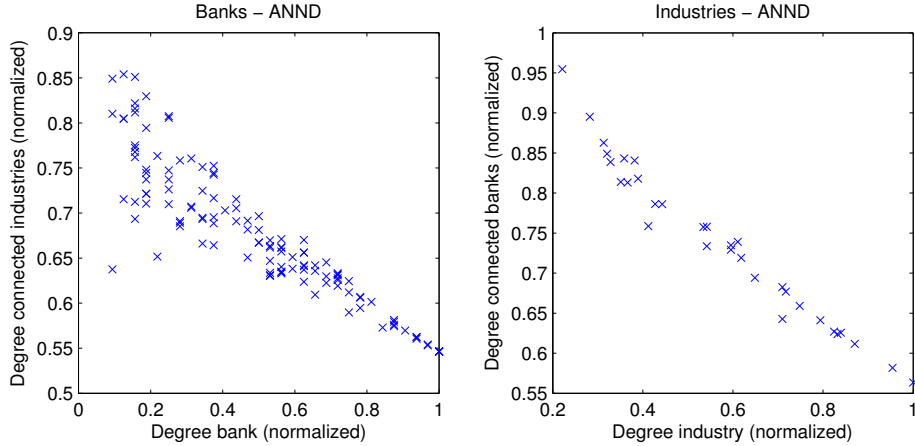


Figure 6: Average nearest neighbor degree (ANND) in 1990. For each bank, ANND measures the average degree of the industries it is connected to. It is defined similarly for the industries. We normalize both the degrees and ANND by the maximum admitted value.

number (normalized by the number of active nodes in each set, respectively). We see that the number of generalist banks was relatively stable for the first half of the sample period, and has been decreasing after that. While there were around 62 generalist banks at the beginning of the sample period, the values are closer to 37 for the post-2000 period. Interestingly, the relative number of generalists was generally much less affected: before 2000 roughly 40% of the banks were generalists, while the values are closer to 35% afterwards. For the industries, there were around 10 generalists over most of the sample period. Interestingly, the relative size of the set was close to 33% before the mid-1990s, and slightly decreased afterwards to around 25%.

Table 2 highlights the economic importance of generalists and specialists. First, we see that the density in the generalist-generalist block (i.e., the number of actual links, relative to the number of possible links) 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 only 20%. Second, we see that interactions between generalist banks and generalist industries contribute to only 20% of all links, but amount to more than 60% of the entire loan volume. Interestingly,

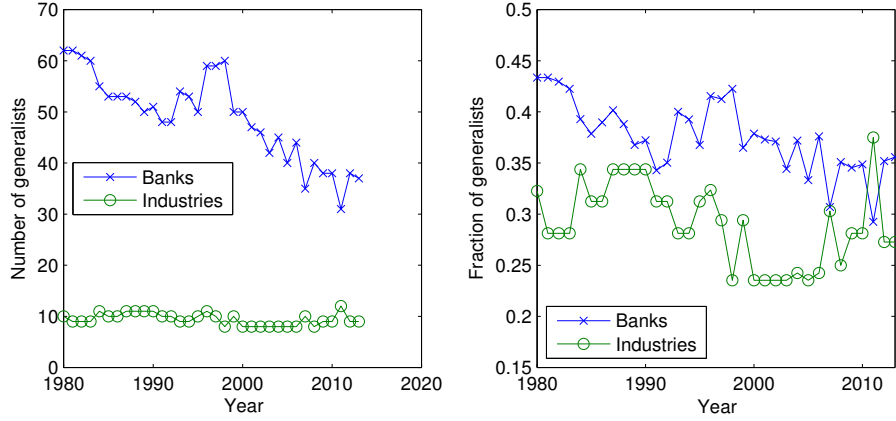


Figure 7: Number of generalist banks and industries, absolute (left) and relative (right).

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 amount to only 1% of the entire loan volume.

		Density		Links		Volume	
(Banks)	(Industries)	Generalist	Specialist	Generalist	Specialist	Generalist	Specialist
	Generalist	Specialist	99.2%	77.6%	20.2%	39.7%	62.0%
		71.9%	21.3%	23.0%	17.1%	2.4%	1%

Table 2: Economic importance of generalists and specialists. Left: density in each of the different block (number of existing links relative to maximum possible number of links). Center: actual number of links in each block relative to the total number of links. Right: actual loan volumes in each block relative to the total volumes.

Figure 8 shows the total error score (left) and the contribution of the different blocks to this total (right). First, the error score is always well below 1, i.e., the value we would obtain without any generalists. The error score follows a steady but slow increase throughout the sample period, such that the fit of the model slightly deteriorates over time. From the right panel we see that most of the errors come from existing links between specialists (SS-block) and that a large part of the increased error score in the mid-1990s was due to the SS- and GS-blocks.

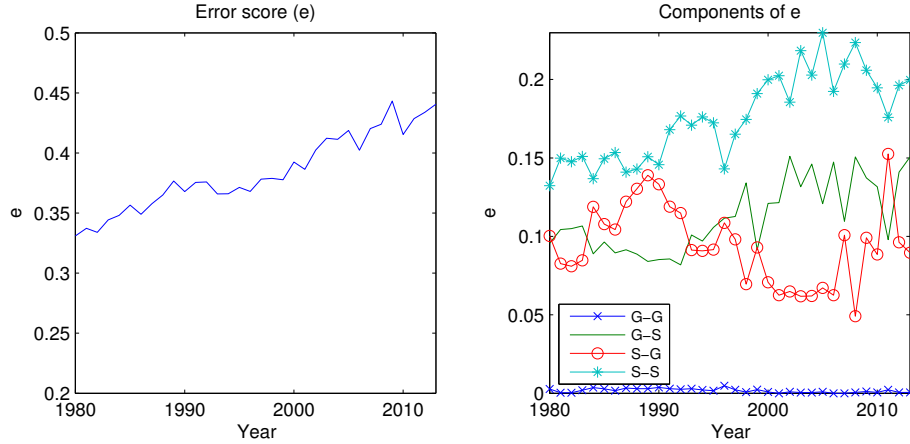


Figure 8: Error score of the generalist-specialist model for the observed networks. Left: total error score. Right: contribution of the different blocks.

Figure 9 illustrates the adjacency matrix of the actual credit network for the year 1990 (left panel), and the corresponding idealized pattern matrix as defined in Eq. (1) (right panel). Clearly, in line with the above results, we observe quite a few links in the SS-block. This illustrates the main limitation of the binary generalist-specialist approach, namely that it has difficulties ‘matching’ asymmetric structures and a very broad distribution of diversification levels as shown above. Nevertheless, we also explore a continuous version of the generalist-specialist model (see Appendix D), which explicitly allows for specialist-specialist interactions. In this case, we find results that are largely comparable to those from the discrete model presented here, suggesting that the limitations of the discrete model are not severe.

Table 3 reports the transition probabilities between the different states (generalist, specialist, and inactive) from one year to the next. In general, there is a lot of persistence in generalist/specialist identifications, since the diagonal elements are large. For example, we see that there is a 94% (96%) chance for a generalist bank (industry) to remain generalist in the following year.

**Significance.** As a next step, we aim at exploring whether the observed network structure is significant. In order to do this, we need to define null models, which

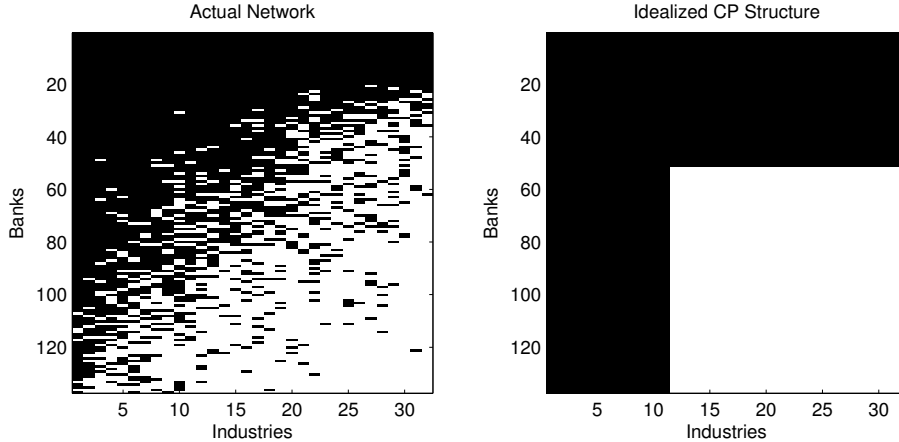


Figure 9: Illustration of the network matrix and the idealized generalist-specialist structure in 1990. Generalist nodes are sorted first according to the optimal partition vectors. Black dots indicate links.

	<b>Banks</b>			<b>Industries</b>		
	Generalist <sub>t</sub>	Specialist <sub>t</sub>	Inactive <sub>t</sub>	Generalist <sub>t</sub>	Specialist <sub>t</sub>	Inactive <sub>t</sub>
Generalist <sub>t-1</sub>	94.3%	5.0%	0.7%	96.7%	3.3%	0%
Specialist <sub>t-1</sub>	2.6%	94.6%	2.8%	1.0%	98.5%	0.5%
Inactive <sub>t-1</sub>	0%	2.8%	97.2%	0.9%	4.5%	94.6%

Table 3: State transition probabilities for banks and industries between the three possible configurations (generalist, specialist, or inactive).

tell us how the credit network looks like when nodes form connections 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 we keep certain characteristics fixed in order to facilitate the comparison between the actual and the synthetic networks. The first null model is a completely random network:

1. Erdős-Renyi (ER) random network: the probability of a link between any bank-industry pair is equal to  $p$ , where  $p$  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 the Appendix C for details on the expected error score in ER random networks).

The ER random network is the most simplistic null model one can think of, but clearly very unrealistic from an economic perspective, since all nodes connect with the very same probability. Therefore, we have 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).<sup>19</sup> 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. We aim at throwing the balls into the bins, which are the industries. The probability of a ball ending up in a given bin depends on the size of the bin, and there are different ways to define these bin sizes. 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.
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 same density as the observed ones.<sup>20</sup>

---

<sup>19</sup>Results for other null models generally show that the observed error score is always well below that of the randomized networks. More details are available upon request from the authors.

<sup>20</sup>Up until 1996, the balls-and-bins models generally yield networks with lower densities compared to the observed ones, such that they predict fewer links between banks and industries. Afterwards, however, the densities become more similar. This suggests that the jump in the number of firms from 1995 to 1996 appears to have some effect on the performance of the balls-and-bins models.

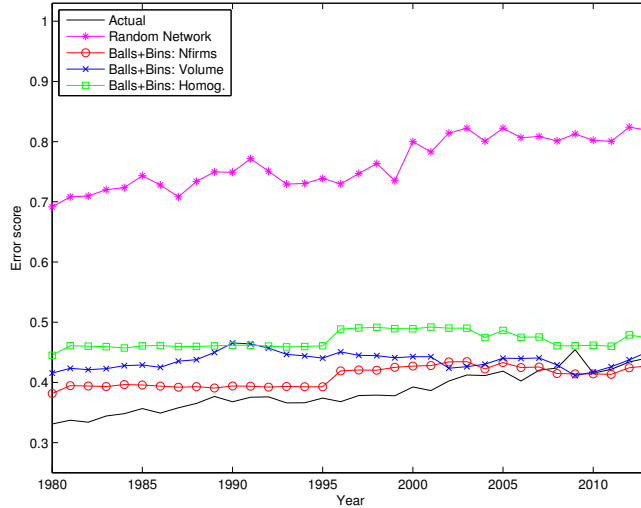


Figure 10: Error score and null models. Here we show the average values over 1,000 network realizations of the corresponding null model.

Figure 10 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 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 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. Interestingly, the heterogeneous balls-and-bins models perform much better at the end of the sample period, in the sense that they produce error scores closer to the actual network.

Finally, Figure 11 shows the distributions of the error scores from the different null models for the year 1980 and compares these values with the actual error score. Clearly, the ER random network is a very poor description of the actual network. On the other hand, we see that the balls-and-bins models generate very similar error score distributions, but with significantly higher means compared to the observed error score. Figure 12 shows the results from the same exercise for the year 2013. In this case, the error score from the actual network is not statistically different from the three balls-and-bins models. This suggests that, despite their simplicity, the balls-

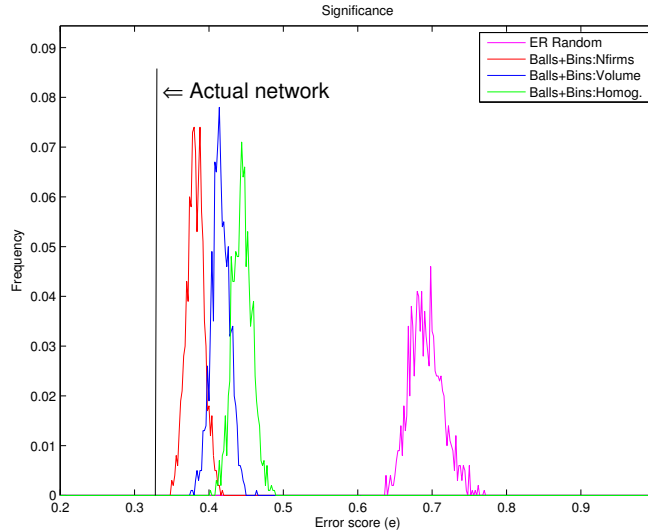


Figure 11: Error score and null models for 1980 - histograms for 1,000 network realizations of the corresponding null model.

and-bins models appear to capture certain characteristics of the observed network structure. These results are in line with Armenter and Koren (2014), who found that the balls-and-bins model does a good job in reproducing the pattern of product- and firm-level trade flows across export destinations.

**Discussion.** The results presented so far imply that there is a very pronounced overlap in the loan portfolios of Japanese banks. Note 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, our results showed that the observed network displays certain deviations from this prediction, but the good fit of the generalist-specialist model suggests that portfolio overlap is indeed very strong.

Why do banks hold so 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, see Acharya and Yorulmazer (2008). Hence, banks want to increase the likelihood of failing simultaneously in order to induce a regulator to bail



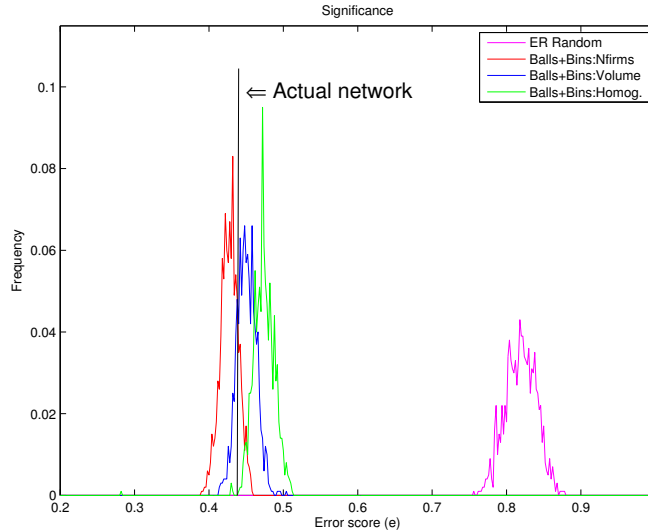


Figure 12: Error score and null models for 2013 - histograms for 1,000 network realizations of the corresponding null model.

them out. In addition to size constraints, there is also a limit to this overlap since competition effects decrease lending margins in some industries/regions. Second, by aiming at holding more diversified portfolios, banks may become more similar as an unintended side effect, see Wagner (2010). Third, herding by managers (or the traders employed by the institutions) 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, see Rajan (2005). Empirically, it is of course difficult to directly test the relevance of each of these theories. We believe that this is an interesting avenue for future research, but an exploration of this question is beyond the scope of this paper.

#### *D Identifying Features of Generalists*

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, see 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 will be a generalist or a specialist. As pointed out by Craig and Von Peter (2014), this is important because (1) it allows to predict the network position of banks and industries using data that should be readily available (e.g., balance sheet information), and (2) it shows that certain features systematically relate to being a generalist. Our dataset contains information on several characteristics of individual nodes, and we use some of these variables to predict whether a node ends up as a generalist.

To be precise, we use a Probit framework where the binary dependent variables are  $\gamma_{i,t}^B$  and  $\gamma_{j,t}^I$  (i.e., generalist levels), for banks and industries, respectively. We estimate

$$Prob(\gamma_{i,t}^B = 1|X_{i,t}) = \Phi(X_{i,t}^T\beta^B), \quad (4)$$

and

$$Prob(\gamma_{j,t}^I = 1|X_{j,t}) = \Phi(X_{j,t}^T\beta^I), \quad (5)$$

separately for banks and industries;  $\Phi$  represents the cumulative normal distribution,  $X$  denotes the set of control variables (always including a constant), and  $\beta^B$ ,  $\beta^I$  the corresponding parameter vectors. In the following, we always show marginal effects (evaluated at the means of the variables) rather than the parameter estimates. In the basic specification, we include data for the full sample period; 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 F). We also used the Logit model as an alternative to the Probit and found very similar results. Finally we analysed various time-, bank-, and industry-specific fixed effects, none of which alter our main findings.<sup>21</sup>

---

<sup>21</sup>Results on these robustness checks are available upon request from the authors.

## D.1 Industries

**Control Variables.** We start out with presenting the results for 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  $\log(\text{TotalAssets} - \text{TotalLoans})$  and measures the size of a industry, excluding the borrowing activities reported in the data.
- **FirmsPerIndustry:** (natural logarithm of) the number of active firms in a given industry.
- **Employees:** (natural logarithm of) the number of employees in a given industry.
- **Leverage:** this is defined as  $\log(\text{TotalLiabilities}/\text{Equity})$ , based on book values.
- **Current asset ratio:** this is defined as  $(\text{CurrentAssets}/\text{TotalAssets})$ , with CurrentAssets being all assets that can be converted into cash within a year. As such, this variable can be seen as an (inverse) measure for the riskiness of the firms from any industry.
- **IndustryGeography:** 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. We define geographical areas based on 47 distinct '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. This variable controls for macroeconomic conditions.

<b>Generalist Industries</b>	(1)	(2)	(3)	(4)	(5)	(6)
log(TotalAssets <sub>t-1</sub> )	0.2624*** (0.0147)	-	-	-	-	-
log(TotalLoan <sub>t-1</sub> )	-	0.2770*** (0.0170)	-	-	-	-
log(IntrinsicSize <sub>t-1</sub> )	-	-	0.2465*** (0.0145)	-	-	-
log(FirmsPerIndustry <sub>t-1</sub> )	-	-	-	0.3031*** (0.0153)	-	-
log(Employees <sub>t-1</sub> )	-	-	-	-	0.2283*** (0.0130)	-
log(Leverage <sub>t-1</sub> )	-	-	-	-	-	0.2914*** (0.0321)
CurrentAssetRatio	-	-	-	-	-	1.3449*** (0.1204)
IndustryGeography	-	-	-	-	-	0.0146*** (0.0014)
Interest Spread <sub>t-1</sub> (in %)	-	-	-	-	-	-0.0460** (0.0198)
GDP <sup>nom.</sup> growth (in %)	-	-	-	-	-	-0.0123** (0.0056)
Pseudo- $R^2$	0.362	0.450	0.326	0.420	0.247	0.539
Obs.	1,056	1,056	1,056	1,056	1,056	897

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

Table 4: 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).

- GDP growth: growth of gross domestic input, measured in percent. This variable controls for macroeconomic conditions.

**Regression results.** Table 4 reports the main regression results. 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. 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. Combining the two results on leverage and current

assets indicates that it is not clear whether generalist industries are more or less risky overall. We also find that generalist industries tend to be geographically close to a large number of banks. Regarding the macroeconomic indicators, both interest spreads and GDP growth turn out to be negatively significant.

For the sake of completeness, in Appendix E we report an overview of the industries that are identified as generalists at least once over the sample period. Looking at the labels of the different industries, we can get a feeling of where some of the heterogeneity in the number of firms per industry comes from.

## D.2 Banks

**Control Variables.** We now turn to the results for the banks. In this case, 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  $\log(\text{TotalAssets} - \text{TotalLoans})$ .
- 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.
- Systemicness: based on Greenwood, Landier, and Thesmar (2015). The systemicness of bank  $i$  measures its contribution to aggregate vulnerability, which itself is defined as the percentage of aggregate bank equity that would be wiped out by bank deleveraging, given some shock to asset returns. See the next section for more details.
- Cash ratio: fraction of cash holdings relative to total assets (measured in %).

<b>Generalist Banks</b>	(1)	(2)	(3)	(4)	(5)	(6)
log(TotalAssets <sub>t-1</sub> )	0.5431*** (0.0164)	–	–	–	–	–
log(TotalLoans <sub>t-1</sub> )	–	0.4400*** (0.0134)	–	–	–	–
log(IntrinsicSize <sub>t-1</sub> )	–	–	0.5230*** (0.0157)	–	–	–
D(City bank)	–	–	–	0.7759*** (0.0674)	–	–
D(Tier-2 regional bank)	–	–	–	-0.5281*** (0.0189)	–	–
log(Systemicness <sub>t-1</sub> )	–	–	–	–	0.4158*** (0.0141)	–
CashRatio <sub>t-1</sub> (in %)	–	–	–	–	–	-0.1270*** (0.0134)
log(Leverage <sub>t-1</sub> )	–	–	–	–	–	-0.1982*** (0.0358)
NetInterbank <sub>t-1</sub>	–	–	–	–	–	0.0466*** (0.0029)
BankGeography <sub>t-1</sub>	–	–	–	–	–	-0.0091*** (0.0009)
Interest spread <sub>t-1</sub>	–	–	–	–	–	0.0663*** (0.0141)
GDP <sup>nom.</sup> growth (in %)	–	–	–	–	–	0.0097** (0.0039)
Pseudo- $R^2$	0.507	0.662	0.483	0.294	0.592	0.127
Obs.	4,094	4,094	4,094	4,286	4,094	2,979

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

Table 5: 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).

- Net interbank position: interbank assets minus interbank liabilities relative to total assets (measured in %).
- Leverage: (natural logarithm of) banks' leverage ratios.
- 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.

**Regression results.** Table 5 reports the results from our probit regressions for the banks. 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 size) are strongly positively significant. Hence, larger banks are more likely to be generalists. Second, we find that the bank type indicators are important control variables. City banks, the largest banks in the Japanese banking system, turn out to be much more likely to end up as generalists, whereas the smaller (tier-2) regional banks are significantly less likely to be generalists. Third, we find that higher systemicness (in the sense of Greenwood et al. (2015)), is a strong predictor for generalists. Finally, we use additional bank-specific characteristics and also control for macroeconomic conditions in the very last column. We find that generalist banks appear to hold less cash and seem to use less leverage. Hence, in terms of riskiness these results are not straightforward. 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.<sup>22</sup> Regarding the macroeconomic factors, a larger interest spread increases the probability to be in the a generalist for all banks. Similarly, macroeconomic growth also has a positive effect. Most of these results are very stable when using data only for the regional banks (tier-1 and tier-2), see Appendix. Similar to the case of industries, we report a list of generalist banks in the Appendix, along with their probability to be generalists and their bank type.

### D.3 Are Generalist Banks More Vulnerable?

The regression results from the previous section were quite informative in terms of predicting which banks (and industries) are generalists. However, we did not learn

---

<sup>22</sup>The result is robust to using alternative, more granular, geographical indicators.

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 use banks' generalist levels as an explanatory variable for their vulnerability. In order to do this, we calculate the vulnerability indicators of Greenwood et al. (2015), who present a simple model of bank deleveraging. The main 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 to target their leverage. These asset liquidations will occur proportional to the actual portfolio weights and will, due to less than perfectly liquid secondary markets, generate an additional negative effect on prices.<sup>23</sup> These price changes will then affect other banks holding some of the liquidated assets themselves, potentially leading them to liquidate assets as well. Note that both Greenwood et al. (2015) and Duarte and Eisenbach (2014) apply the 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, as pointed out by Drucker and Puri (2009), corporate loans are in fact traded on reasonably liquid secondary markets.

Within this simple framework, Greenwood et al. (2015) propose several bank-specific and aggregate vulnerability indicators. *Aggregate vulnerability* is defined as the percentage of aggregate bank equity that would be wiped out by bank deleveraging, given some shock to asset returns. This indicator omits the direct impact of the shock on net worth, emphasizing only the spillovers across banks. The *systemicness* of bank  $i$  is simply defined as a bank's contribution to aggregate vulnerability. The

---

<sup>23</sup>The model of Greenwood et al. (2015) is only concerned with direct and indirect effects of selling pressure on other market participants and ignores broader macroeconomic effects. As shown by Anari et al. (2005), however, the asset liquidations of failed banks can have a sizable impact on output in the short- to medium-term.



*indirect vulnerability* of bank  $i$  with respect to a given shock is the impact of this shock on its equity through the deleveraging of other banks. Finally, the *direct vulnerability* of bank  $i$  measures the direct impact of these shocks on its portfolio. Here we are interested the bank-specific indicators, namely direct and indirect vulnerability, and systemicness.

We parametrize the system as follows:

- portfolio size is set to the observed value of each bank’s loan portfolio,
- leverage, as defined above, is set to the observed value for each bank,<sup>24</sup>
- all assets have the same price impact of  $10^{-12}$ , and we set all cross-price impacts to zero.<sup>25</sup>

In the following, we explore the following two scenarios:

- (a) **Idiosyncratic shocks.** We shock one asset (i.e., industry) at a time by imposing a return of -10% on the outstanding loan amounts, while setting all other returns to zero. In this case, the final vulnerability estimates are the averages across the different shock scenarios.
- (b) **Systematic shocks.** We shock all assets by imposing a return of -1% on all outstanding loan amounts.

For each year, we apply the model of Greenwood et al. (2015) separately for the two scenarios and record banks’ direct and indirect vulnerabilities.

---

<sup>24</sup>In their sample of large EU-banks, Greenwood et al. (2015) observe extremely high leverage values. In order to reduce the impact of these outliers, they use 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 (with  $x$  equal to 5% or 1% for example). This does not affect any of the qualitative results reported in the main text.

<sup>25</sup>In this exercise, it only matters that price impact is positive. The actual parameter value can be seen as a scaling parameter, but we found that it does not affect the sign of the relationship between whether a bank is a generalist and its level of vulnerability.

**Results.** Table 6 reports the regression results for the two scenarios where we predict banks' vulnerability measures using the lagged generalist levels (we obtain very similar results when using contemporaneous observations) based on a simple linear OLS regression with year fixed effects. In order to acknowledge a potential change in the sign of the relationships, we run three separate regressions for each scenario, namely (1) using all observations, (2) using only data before 1996, and (3) using data only after 1996. It turns out that the results are practically identical for the two different shock scenarios: for the indirect vulnerability, we see that being a generalist is strongly negatively significant when using all observations. However, these results are driven by the post-1996 period, since the relationship is in fact positively significant in the pre-1996 period. For direct vulnerability, we see comparable results similar results both in terms of the signs and significance of the parameters, apart from the pre-1996 period where the relationship is insignificant.

These results suggest that, compared with specialists which cluster in the same industries, generalist banks tend to be less vulnerable both with respect to industry-specific and common shocks, at least for the latter half of the sample period. To some extent, these results are driven by the fact that generalist banks tend to be less leveraged (as shown in the regressions in Table 5) and that their higher levels of diversification indeed make them less prone to the shock scenarios we consider here. Overall, these findings indicate that generalists appear to enjoy certain diversification benefits and how banks position themselves in the credit network can be informative about their riskiness.

## VI Conclusions

In this paper, we explored the cross-sectional structure of Japanese banks' industrial loan portfolios during the years 1980 - 2013. Our main results are as follows: generalist banks (diversified lending) and specialist banks (focused lending) coexist. Similarly, generalist industries (diversified borrowing) and specialist industries (fo-

**Scenario (a): Industry-Specific Shocks**

	Indirect Vulnerability			Direct Vulnerability		
	All	1980-1995	1996-2013	All	1980-1995	1996-2013
Generalist <sub>t-1</sub>	-0.0575*** (0.0145)	0.1128*** (0.0189)	-0.2256*** (0.0211)	-0.0714*** (0.0087)	-0.0148 (0.0106)	-0.1272*** (0.0135)
Constant	9.2102*** (0.0089)	9.4101*** (0.0119)	9.0174*** (0.0128)	-2.4500*** (0.0054)	-2.2797*** (0.0067)	-2.6128*** (0.0082)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
adj. $R^2$	0.555	0.278	0.586	0.419	0.249	0.169
Obs.	4,094	1,995	2,099	4,094	1,995	2,099

**Scenario (b): Common Shocks**

	Indirect Vulnerability			Direct Vulnerability		
	All	1980-1995	1996-2013	All	1980-1995	1996-2013
Generalist <sub>t-1</sub>	-0.0575*** (0.0145)	0.1128*** (0.0189)	-0.2256*** (0.0211)	-0.0714*** (0.0087)	-0.0148 (0.0106)	-0.1272*** (0.0135)
Constant	9.2170*** (0.0089)	9.4169*** (0.0119)	9.0243*** (0.0128)	-1.2661*** (0.0054)	-1.1166*** (0.0067)	-1.4091*** (0.0082)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
adj. $R^2$	0.555	0.278	0.586	0.390	0.249	0.186
Obs.	4,094	1,995	2,099	4,094	1,995	2,099

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

Table 6: Are generalist banks more vulnerable? Results from OLS regressions, where we regress the different vulnerability measures against the lagged generalist levels (including time FEs).

cused borrowing) also coexist. Most surprisingly, it turns out that most specialist banks tend to concentrate their lending activities on the very same generalist industries, while specialist industries receive funding almost exclusively from generalist banks. Using a simple generalist-specialist model, we can reasonably well describe the observed asymmetric structure of the credit network over the entire sample period. While the model shows a deteriorating fit over time, the fit is significantly better compared to alternative benchmark models. Most importantly, the model allows us to identify a highly persistent, and economically meaningful set of generalist banks and industries in a structural way. These findings imply that the interactions between banks and the real economy can be described by a reasonably simple structure. We believe that incorporating this structure of interactions in broader macroeconomic models is of utmost importance to disentangle the feedback effects between the banking system and the real economy.

While we are somewhat agnostic with regards to the micro-mechanisms leading to the observed network architecture, the coexistence of generalist and specialist banks is in line with existing theoretical models (see Stomper (2006)) and our econometric analysis is an important step towards understanding certain key characteristics of generalists and specialists. We find that size is an important factor for being a generalist, for both banks and industries, but we also highlight additional relevant factors. For example, due to the segmentation of the Japanese banking system, bank type (e.g., city banks) is a strong predictor. For industries, we find that both leverage and geographical proximity to banks are very important factors. Finally, regarding the stability implications of generalist banks, we compute the vulnerability indicators of Greenwood et al. (2015), and find that generalist banks tend to be less vulnerable compared to specialists for the second half of the sample period. Hence, how banks position themselves in the credit network has important implications for their riskiness.

Our findings suggest several interesting avenues for future research. First and foremost, we believe that our findings generalize to many other banking systems and

we would like to see the empirical evidence for other countries. Second, an immediate question is how important the structure of banks' investments are for systemic risk as a whole? A huge literature on fire sales shows that strong portfolio overlap might reduce the probability of contagion spirals, but when contagion occurs, large parts of the banking system go extinct at the same time. The observed generalist-specialist structure might therefore be rather vulnerable to large industry-specific shocks. Third, a micro-founded theoretical model that accounts for both banks' lending decisions and firms' or industries' borrowing decisions would be of utmost importance in order to explore different regulatory measures that might help mitigating systemic risk. Lastly, the determinants for generalist banks appear to be very similar to those for the interbank network (see Craig and Von Peter (2014)). This suggests that money center banks play a key role in different types of financial networks, and it is of utmost importance to test this hypothesis empirically and explore the implications of this finding for systemic stability.

## 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. and Yorulmazer, T. (2008). Information contagion and bank herding. *Journal of Money, Credit and Banking*, 40(1):215–231.
- Allen, F. and Gale, D. (2000). Financial contagion. *Journal of Political Economy*, 108(1):1–33.
- Anari, A., Kolari, J., 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.
- BIS (2014). Supervisory framework for measuring and controlling large exposures. Standard documents, Bank of International Settlements (BIS).

- Borgatti, S. P. and Everett, M. G. (2000). Models of core/periphery structures. *Social Networks*, 21(4):375–395.
- 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.
- Cai, J., Saunders, A., and Steffen, S. (2014). Syndication, interconnectedness, and systemic risk. *Working paper*, SSRN.
- Craig, B. and Von Peter, G. (2014). Interbank tiering and money center banks. *Journal of Financial Intermediation*, 23(3):322–347.
- Degryse, H. and Ongena, S. (2005). Distance, lending relationships, and competition. *Journal of Finance*, 60(1):231–266.
- Detragiache, E., Garella, P., and Guiso, L. (2000). Multiple versus single banking relationships: Theory and evidence. *Journal of Finance*, 55(3):1133–1161.
- 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.
- Freixas, Xavier, B. P. and Rochet, J.-C. (2000). Systemic risk, interbank relations and liquidity provision by the central bank. *Journal of Money, Credit and Banking*, 32(3):611–638.
- Fricke, D. (2016). Has the banking system become more homogeneous? evidence from banks’ loan portfolios. *Economics Letters*.
- 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 US 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.

- 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.
- 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.
- Miwa, Y. and Ramseyer, J. M. (2002). The fable of the keiretsu. *Journal of Economics & Management Strategy*, 11(2):169–224.
- Ongena, S. and Yu, Y. (2013). Firm industry affiliation and multiple bank relationships. *Working paper*, SSRN.
- 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* - No. 11728.
- 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.
- Stomper, A. (2006). A theory of banks’ industry expertise, market power, and credit risk. *Management Science*, 52(10):1618–1633.
- 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.
- Winton, A. (1999). Don’t put all your eggs in one basket? Diversification and specialization in lending. *Working paper*.