Specialization in Banking *

Kristian Blickle[†] Cecilia Parlatore[‡] Anthony Saunders[§]

July 2024

Abstract

Using supervisory data on the loan portfolios of large US banks, we document that these banks specialize by concentrating their lending disproportionately in a few industries. This specialization is consistent with banks having industry-specific knowledge, reflected in reduced risk of loan defaults, lower aggregate charge-offs, and higher propensity to lend to opaque firms in the preferred industry. Banks attract high-quality borrowers by offering generous loan terms in their specialized industry, especially to borrowers with alternative options. Banks focus on their preferred industry in times of instability and relatively lower tier 1 capital as well as after sudden surges in deposits.

*We would like to thank the editor Antoinette Schoar, the associate editor, and two anonymous referees for their guidance. We would further like to thank Tobias Berg, Allen Berger, Darrell Duffie, Quirin Fleckenstein, Linda Goldberg, Amanda Heitz, Sebastian Hillenbrand, Victoria Ivashina, Elena Loutskina, Stephan Luck, Daniel Paravisini, Matt Plosser, Philipp Schnabl, John Sedunov, Johannes Stroebel, Dominik Supera, Jonathan Wallen, and Toni Whited as well as seminar/conference participants at the Federal Reserve Bank of Philadelphia, the Federal Reserve Bank of New York, the Federal Reserve Bank of San Francisco, New York University Stern School of Business, the IMF, the American Bankers Association (ABA), Temple University, Villanova, the University of Sydney, FINEST, MoFiR Workshop on Banking, the FDIC, Third Conference on the Interconnectedness of Financial Systems, Financial Intermediation Research Conference (FIRS), the Society for Economic Dynamics (SED), and the Western Finance Association (WFA) for helpful comments and suggestions. The opinions expressed in this paper are those of the authors and do not necessarily represent those of the Federal Reserve Bank of New York or the Federal Reserve System. Any errors are our own.

[†]Federal Reserve Bank of New York

[‡]NYU Stern School of Business, CEPR, and NBER

[§]NYU Stern School of Business

Introduction

Banks are traditionally tasked with selecting high-quality borrowers and monitoring their adherence to loan covenants. However, borrower selection and loan monitoring require the costly acquisition of information. Economies of scale and deeper experience may be built up through the specialization of lending to certain industries. After all, repeated interactions with individual borrowers have been shown to improve a bank's knowledge of a borrower.[1] Similarly, recurrent lending in a specific industry can enable banks to evaluate the business models or collateral of borrowers in that industry better. In this paper, we investigate loan portfolio specialization across industries among large U.S. banks. We further analyze how this specialization is associated with reduced information asymmetries between lender and borrower in the lender's industry of specialization, resulting in better loan performance.

Specialization can be a nebulous concept and we explore its implications using a variety of definitions. Most simply, however, specialization can be thought of as the degree to which a bank is "over-invested" in an industry relative to a "diversified" lending portfolio, which would be based solely on the relative size (i.e. the total borrowing) of each industry. Accordingly, specialized banks direct a greater proportion of their C&I lending to a particular industry – at the expense of other industries – than would be expected under pure diversification. Using this and other measures, our paper presents several novel – and surprising – facts based on such specialization, categorized broadly into five groups.

First, we document the existence and persistence of specialization. We use detailed supervisory data sets that cover the loan portfolios of large US banks to show that even these entities concentrate disproportionately on a few industries. Figure 1 shows the average over-investment (in percentage points) of the 40 stress-tested banks in the US – i.e. the nation's largest banks. Over-investment (which we refer to as "excess" specialization below) is computed as the difference between the share of a bank's portfolio in an industry and

[1] For a discussion of relationship lending see: Bernanke (1983), James (1987), Petersen and Rajan (1995), Berger and Udell (1995), or Degryse and Ongena (2005).

the investment that would be expected under full diversification. We separately look at a bank's favorite, second favorite, and all other industries. As the figure shows, the average bank directs 9% more of its total C&I lending to its single "favorite" industry (measured here at the two-digit NAICS level) than would be expected if it were diversified. Notably, banks only have one or two preferred industries, which remain stable over time. We show that this specialization is a conscious bank choice – reaffirmed each quarter through new loan origination – and not the accidental by-product of select large loans with long maturity.

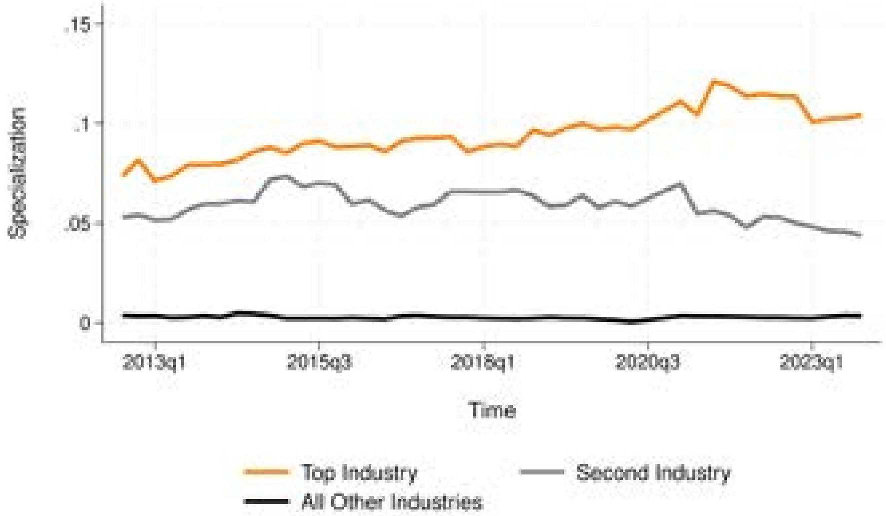
Second, we find that this type of specialization is consistent with informational advantages that facilitate better ex-ante screening and ex-post monitoring of loans. Specifically, loans made by banks specialized in a borrower's industry are less likely to become non-accruing or be downgraded before maturity/renegotiation. Using both bank internal risk ratings and supervisory data on the monitoring activity of lenders, we show that this performance is likely driven by both superior screening and superior monitoring. Further evidence that specialization helps reduce information asymmetries between a lender and a borrower comes from specialized lenders being more likely to lend to smaller and unknown borrowers.[2]

Third, we show that banks attract and keep good borrowers in an industry in which they are specialized by offering competitive terms. We find that specialization positively correlates with larger loans, made at lower rates, and with a longer maturity. This difference is most pronounced for unsecured loans and holds especially for firms with access to alternative sources of funding, suggesting that specialization by large banks may be partly a necessary answer to increased loan competition.

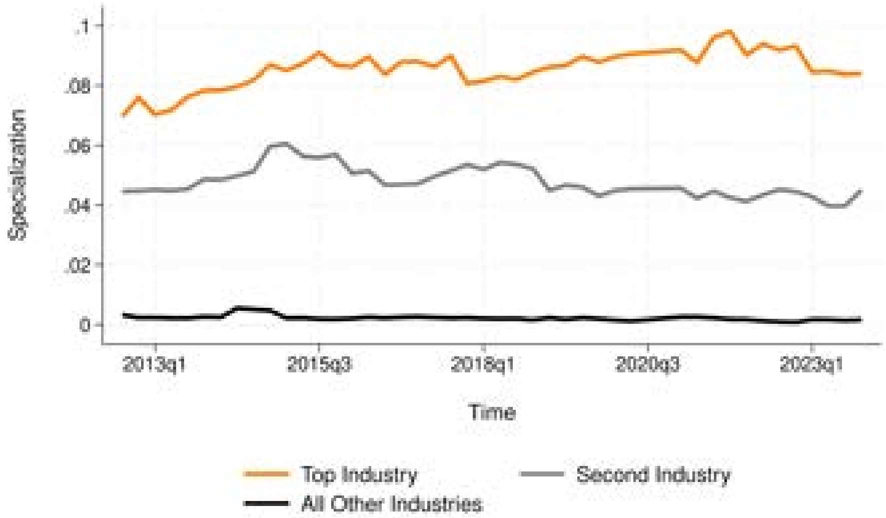
Fourth, we show that banks weigh the benefits of specializing with its costs and risks and cannot specialize indefinitely. On the one hand, the loans available to be held by banks (from a regulatory standpoint) are finite. On the other hand, concentration may increase bank risk in the event of a sector-specific downturn. During our primary sample period, which covers

[2]Opaque smaller borrowers are most likely to be credit constrained due to informational asymmetries (Beck et al. (2006a) or Gertler and Gilchrist (1994)), usually requiring them to pledge expensive collateral to circumvent these constraints (Barro (1976), Hart and Moore (1994), Stiglitz and Weiss (1981)). We show that specialized lenders may reduce borrower opacity through information gained about an industry.

Figure 1: ***Excess Specialization***



(a) Excess Specialization



(b) Excess Specialization (Volume Weighted)

Note: This figure shows the degree to which banks in our data are "over-invested" in their "favorite", second favorite, and all other industries. A perfectly diversified bank is one that is invested in accordance with industry size, with larger industries (relative to all C&I lending) receiving proportionately more loans. We use excess specialization, which is defined as $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}} - \frac{LoanAmount_{s,t}}{\sum_s LoanAmount_{s,t}}$; for term loans in our sample. A favorite industry is one in which a bank is the most "over invested". Panel (a) depicts unweighted calculations of average specialization while panel (b) weights our specialization measures according to committed loan amounts before averaging.

2011q3-2023q3, we find that more specialized banks (i.e. banks with a greater focus on their top industry) earn slightly lower returns. However, these banks also charge off fewer loans in aggregate, even during the COVID-19 pandemic.[3] This corroborates earlier evidence by Beck et al. (2022) who, using stock data, find that individual and systemic bank risk is lowered by specialization. Furthermore, we find that banks appear to concentrate on their industry of specialization in times of lower Tier1 capital or when the share of risk-weighted assets is high, implying a trade-off at the bank level between rent-seeking and stability.

Finally, by looking at periods with a sudden and quasi-unsolicited increase in deposits – such as occurred during the COVID-19 outbreak – we find that banks do not reshuffle their loan portfolios proportionally: they are less likely to decrease lending to their industry of specialization when funds are tight and more likely to increase lending to it after unplanned fund inflows. Moreover, using COMPUSTAT data, we show that this preference has real effects on firm growth, with borrowers of specialized lenders benefiting. This finding highlights that banks are not fungible, given their specialization, and deposit reshuffling within the system can have real, firm-level consequences.

Our results are crucial to understanding the effect of a reallocation of deposits among lenders on the distribution of credit across firms and industries, especially given the increased fluctuations in deposits within and into the banking sector (see Appendix Figure A.11). For instance, there was an aggregate shift in deposits from smaller banks and money market funds (MMF) to large banks, which started just before COVID-19 and accelerated during the crisis, while large banks began losing deposits again in 2022, something that accelerated in the weeks following the initial Silicon Valley Bank (SVB) turmoil.

The analyses in this paper are primarily based on the FR Y-14 Q-H archive, which tracks all C&I loans over 1 million USD for all stress-tested US banks. Our data is the closest thing to a credit registry for the United States and encompasses around 75% of corporate lending. Unfortunately, we observe only originated loans and not loan applications. As such, our regressions reflect ex-post equilibrium outcomes. We are careful about interpreting our

[3] We do not have data on a sector-specific crisis. Such an event where loans in a specific industry must be written off wholesale, might prove especially destabilizing to a specialized lender. This risk remains one of the omitted concerns in our study.

results as causal. Nevertheless, we can account for a host of loan and bank characteristics in all regressions to ensure that the patterns we identify are not the result of omitted variables. We account, for example, for the degree to which a bank has captured an industry. If a bank captures a majority lender stake in a sector in an attempt to extract monopoly rents, we may accidentally measure specialization where none holds.[4] While specialization is correlated with industry capture, our results on specialization hold despite of – and not because of – a bank's role in an industry. We also account for the individual relationships that may exist between a borrower and a bank, which we may otherwise mistake for industry-level specialization. Finally, we saturate our specifications with a variety of industry-time, bank-time, or even firm-time fixed effects.[5]

Where possible, we address further concerns about potential loan-level differences. For instance, we account for loan purpose, loan type, and even bank-internal risk ratings. These risk ratings are standardized across all banks in our sample as part of the stress testing process and therefore highly comparable across loans and banks. Even in the face of detailed controls, endogeneity concerns about borrower-bank selection remain. We address these openly throughout the paper and perform a variety of robustness analyses to highlight the stability of our results. Unfortunately, we are unable to observe bond underwriting activity in our data. It is possible that bond underwriting activities also display similar specialization dynamics to loan origination, which would allow banks to gain additional insights into industries.[6] This remains an omitted variable concern in our case.

We use the Syndicated National Credit registry (SNC) to supplement our Y14 data in two ways. Although it is less detailed than the Y14 data and covers fewer smaller loans,

[4]Monopoly on information can translate into market power even in a competitive banking industry (see: Broecker (1990), Sharpe (1990) and Rajan (1992), Riordan (1993)). Specialization can also be a byproduct of banks' search for monopolistic rents. Which is the case is important for regulation purposes. We, explore both dimensions.

[5]This is somewhat akin to a Khwaja and Mian (2008)-style control, ultimately holding constant any firm-time effects. Only very large firms borrow from more than one lender in a given period. Hence, we do not take this as our primary specification.

[6]Neuhann and Saidi (2018) show that universal banks, with their securities underwriting activities, were able to reduce information asymmetries regarding their borrowers and deepen financial contracting opportunities.

the SNC data covers many small lenders. This allows us to analyze the differences between lenders concentrated into a single industry out of a lack of options, and truly specialized lenders that reap the benefit of repeated interactions. Secondly, the SNC covers a longer period. This allows us to assuage concerns that our results are the product of the recent low-rate period and also hold during a general recession.

Our work contributes to several important strands of the banking literature. First, we offer new and important evidence in the longstanding debate between specialization and diversification. Theoretically, diversification can be associated with a reduction in the exposure to local and idiosyncratic shocks (Boyd and Prescott (1986), Diamond (1984)). Acharya and Yorulmazer (2007) argue that banks may wish to hold correlated investments to minimize the direct effects of negative information shocks on borrowing costs. Empirically, Tabak et al. (2011) document better bank performance and lower risk in more diversified banks in Brazil. Westernhagen et al. (2004) argue that bank concentration can be a risk at the aggregate level. However, diversification can also increase the correlation across banks' portfolios, increasing the risks of contagion and the probability of systemic crises (Allen and Gale (2000),Haldane and May (2011), Yellen (2013), Goldstein et al. (2020)). Acharya et al. (2006) find that bank diversification is not associated with superior returns or safer portfolios. Greenwood et al. (2015) argue that fire sales can propagate through a financial system if lenders hold common portfolios. Instead, concentration can be associated with lower risk (consider: Beck et al. (2022), Beck et al. (2007) or Beck et al. (2006a)). While the size of their balance sheets may incentivize banks to acquire significant amounts of information amounts of knowledge about an industry through specialization (see (Marquez 2002)), the natural incentives to specialize and capture an industry must be weighed against possible risks. We offer evidence that shows that in stable times, banks with specialized portfolios may see fewer charge-offs in their industry of specialization. However, we also show that banks earn lower returns, trading off stability for profitability.

There are many dimensions in which a bank can specialize. For instance, early work by

Carey et al. (1998) suggests lenders specialize by borrower type and contract terms, specialization in small bank business models is discussed in Blickle (2020), and specialization in collateral is discussed in Gopal (2019). Giometti and Pietrosanti (2022) show that specialized lenders offer less restrictive loan contracts. More closely related to our work, Paravisini et al. (2020) develop an approach to identify bank specialization in lending and show that Peruvian banks specialize across export markets. This specialization has real economic effects on their borrowers. We show that banks are not only currency-specialized. Instead, we provide novel evidence to show that even large US banks specialize by concentrating on certain industries. Moreover, detailed supervisory data allows us to both make a compelling link between this specialization and bank information acquisition as well as show the trade-off faced by a bank when specializing.

Our paper also contributes to the wider literature on banking and bank business models in general. Broadly speaking, there are two main theories of banking (see Battacharya and Thakor (1993) for a survey of the theories of financial intermediation). On the one hand, banks provide liquidity and maturity transformation to their depositors by issuing demandable deposits and investing in longer-term loans. Under this view, risk-averse banks will choose to diversify their loan portfolio to maximize risk-sharing among their depositors and minimize the risk of bank runs (see Diamond and Dybvig (1983) and Allen and Gale (1998)). On the other hand, banks may have an informational advantage in lending, by screening and monitoring loans (Diamond (1984)). Loutskina and Strahan (2011) show that concentrated lenders focus on information-intensive loans while Frattaroli and Herpfer (2022) show that borrowers from the same bank – especially opaque borrowers – are more likely to enter a business partnership as banks facilitate information transfer. In this context, increasing returns to information acquisition push banks towards holding specialized loan portfolios. Understanding how specialization impacts bank riskiness is therefore key to understanding the interplay between risk-related capital regulation and loan portfolio specialization.

The remainder of this paper is organized as follows. Section 1 discusses theoretical motivations *for* and presents the primary hypotheses tested *in* this paper. Section 2 describes the data used. Section 3 showcases specialization in the banks of our sample and offers summary statistics. In Section 4 we discuss our Methodology. In Section 5 we explore the performance of loans made by specialized banks. Section 6 describes loan terms offered to borrowers. Finally, Section 7 discusses aggregate results at the industry and bank levels and Section 8 presents results for several extensions. Section 9 concludes.

1 Hypotheses Development

Before diving into our empirical analysis, we briefly discuss the theoretical links between banks' information sets and loan portfolios. This section's goal is to lay out simple hypotheses by leveraging the existing literature.

Many complementary theories seek to justify why banks exist (see Battacharya and Thakor (1993) and Gorton and Winton (2003) for a review of the early literature on the role of banks). On the one hand, banks allow agents to smooth their consumption by providing insurance against idiosyncratic consumption shocks (Diamond and Dybvig (1983) and Allen and Gale (1998)). On the other hand, banks allow investors to reduce their exposures to risk either by monitoring (Diamond (1984)) or by screening (Leland and Pyle (1977), Campbell and Kracaw (1980), Boyd and Prescott (1986), Blickle et al. (2022)) the loans that they originate more efficiently than individual depositors.

Monitoring and screening require banks to invest in information technologies that feature increasing returns to scale. To see this, one can think of a bank's information acquisition and loan issuance decisions as a portfolio choice with endogenous learning as in vanNieuwerburgh and Veldkamp (2010). In this context, a bank learns only about risks it chooses to hold in its portfolio.[7] More concretely, it is not optimal for a bank to learn about the mining industry if it expects to make only information technology loans. Put differently, the more a bank expects to lend in an industry, the higher the benefit from learning about it. Conversely,

[7]This logic is also present in Parlatore and Philippon (forthcoming) in the context of stress test design.

the more a bank knows about an industry, the higher the incentives to lend in it. These increasing returns to scale imply that when information acquisition is costly or there are information processing constraints, a bank will specialize its learning and its loan portfolio will be tilted toward industries about which it knows more.

In the cross-section, if all banks were equally informed, one should expect them to hold, on average, the same portfolios. They would simply hold the industry-size weighted loan portfolio, where the share of lending to each industry is simply a function of the size of that industry. Asymmetries in the quality of information among banks, however, lead to heterogeneous portfolios across banks. More precisely, one would expect banks with banks with informational advantages in an industry to lend relatively more in it than those less informed about it, (see Broecker (1990), He et al. (2023), Blickle et al. (2024)). These observations lead us to our first hypothesis:

Hypothesis 1: Banks with informational advantages will be under-diversified (i.e. not hold the industry portfolio) and specialize their loan portfolios.

Informational advantages also grant banks an edge in the competition for good loans to relatively safer borrowers.[8] The more a bank knows about a particular industry, the better the ex-ante screening and ex-post monitoring of loans in that industry in its portfolio. Hence, one would expect banks with informational advantages to have better loan performance in the industries in which they specialize.

Information is not the only driver of banks' under-diversifed portfolios. If banks are constrained by entry costs or balance sheet limitations, then specialization would not be an active bank choice. In particular, we would not observe the valuable repeat interactions with borrowers in a particular industry. Rather, specialization in this context would be a mechanical outcome arising from the inability of the bank to diversify more. If this were the case, there should be no systematic difference in the performance of loans made by specialized and non-specialized banks. Our second hypothesis focuses on these systematic

[8]See Blickle et al. (2024), Goldstein et al. (2024), He et al. (2023) and He et al. (2024) for formal models of competition among differentially informed banks.

differences in expected performance.

Hypothesis 2: Bank specialization driven by informational advantages should be correlated with better loan performance in an industry in which a bank specializes. Constraint-driven concentration is different from specialization – particularly in smaller entities – and should not be associated with better loan performance.

A question that arises in the context of loan selection under specialization is why specialized banks simply do not make loans that are rated as "high risk" by the market – and that pay a commensurate rate – but that the specialist can see are actually of a higher "true" quality. Ultimately, on a sliding scale of quality, one might expect specialized banks to experience the same default rates as non-specialized banks but earn higher returns as they "correctly" pick and monitor lucrative but somewhat riskier borrowers. The simple and intuitive answer is risk weights. Risky loans – or loans to borrowers that the market and bank supervisors view as risky – will carry high-risk weights. A single specialized bank will be unable to convince regulators or rating agencies and would face the penalty of having to hold significantly higher capital against "risky" loans. Ultimately, our assumptions on loan performance are identified within the type of loans that are typically held by banks. This limit to the loans banks can hold leads to our third hypothesis:

Hypothesis 3: If banks face risk constraints on loans, bank specialization driven by informational advantages is associated with more stable loan portfolios (i.e. fewer defaults) in the industry in which a bank specializes.

Specialized banks having lower aggregate risk is generally corroborated by recent studies. Beck et al. (2006b) show that systemic crises can be less pronounced in concentrated banking systems, and Beck et al. (2022) show that individual bank risk decreases with specialization. Similarly, Wagner (2010) and Greenwood et al. (2015) argue that commonality in bank portfolios, such as could arise if every bank held the same diversified portfolio, make systemic crises more likely.[9]

[9]Similarly, as part of a broader model, Allen et al. (2012) show that banks with a clustered asset structure are likely to default together.

Increasing returns in information acquisition and the resulting loan performance theoretically push banks towards ever-increasing specialization in their lending portfolios. However, beyond the above-discussed limits to the type of loans banks can hold, there are costs related to specialization that will limit the extent to which banks over-invest in an industry. The most direct cost is literal. Specialized banks must attract desirable borrowers with attractive loan terms. Otherwise, the neutral borrower would be indifferent in their lender choice. This implies the specialized bank may grant lower interest rates and larger or longer maturity loans than a less informed competitor. Saidi and Streitz (2020) show that concentrated lenders offer lower interest rate loans to their borrowers while Ross (2010) also shows that the dominant bank in an industry offers favorably low rates.[10] These generous terms reduce profitability. Hence, there may be a trade-off between bank profit and bank stability from fewer aggregate defaults in ordinary times. Instances in which a bank suffers from lower-than-usual capital may trigger a drive for specialization. This discussion brings us to our final hypothesis.

Hypothesis 4: In ordinary times a bank may gravitate towards additional specialization, especially when stability is valuable to said bank, at the cost of profitability.

Finally, it is worth noting that banks may need to concern themselves with the concentration of risk that results from a concentrated portfolio; an under-diversified specialized bank may be over-exposed to industry-specific shocks. This can lead to excessive bank losses even after small industry-specific negative shocks, and even triggering wider financial crises. We do not observe such an event in our data between 1995 and 2023, however.

2 Data

Our primary data comes from the FR Y-14 Q,maintained by the Federal Reserve and used in support of the stress testing of major financial institutions. The data includes various details

[10] Work by Asker and Ljungqvist (2010) suggests that competitive terms may be in response to the borrower fear that competitive information about them is leaked to its rivals by a concentrated bank that lends to many firms in a sector.

on every bank that has ever been subject to stress tests. We specifically use the sub-database H.1, which contains detailed quarterly information on the C&I loans of reporting banks. Reporting institutions must file all loans they hold with a total balance-sheet commitment of more than 1 million USD in a given quarter. In the sample period between 2012:Q2 and 2023:Q2, we observe 40 banks that report several million loan observations. We keep observations for which we observe the amount, maturity, and interest rate. We naturally remove observations with interest rates or maturities that are likely the result of coding errors (i.e. negative or beyond reasonable coding ranges). In our cleaned sample, we thus focus on about 297,000 term loans, with a total of over 3,000,000 loan-quarter observations.

Unlike other commercially available databases, which cover a subset of the market or specialize in syndicated lending, our data contains highly detailed information on over 75% of *all* C&I lending in the United States (by USD volume) during the sample period.[11] Moreover, it includes both syndicated and non-syndicated loans. This, in particular, allows us to look at the differential impact of specialization on larger loans, which may include multiple syndicate members, compared with smaller loans, which are issued and held by a single bank.

Banks report a large set of characteristics for each loan that areuseful for our analysis. Loan characteristics include the type of loan (credit line vs. term loan), total committed amount, total drawn amount, interest rate, whether a loan is collateralized or unsecured, loan maturity, a loan's risk rating as rated by the bank and reported to Federal Reserve examiners, as well as whether a loan has become non-performing or is in arrears. Besides loan characteristics, the data contains additional information on borrower characteristics. These include borrower name, location, and most importantly, industry. We use a borrower's 2- and 4-digit NAICS industry classifications to define specialization, which we discuss in detail below.

The average loan size in our sample is 4.5 million, which is right-skewed toward a few very large loans. Our data is reported in thousands of USD and logged. As can be seen in

[11]See Blickle et al. (2020) for a discussion of the shortcomings of Dealscan data.

Table 1: ***Summary Statistics of Key Variables***

	N	Mean	SD	Top-Industry	Other-Industry	Diff
Log Amount	296,951	8.42	1.16	8.50	8.43	0.07***
Interest Rate	296,951	3.80	1.70	3.61	3.83	-0.21***
Maturity Remaining	296,951	24.51	21.05	19.84	18.40	1.4***
Unsecured	296,951	0.18	0.38	0.13	0.19	-0.06***
Loan Becomes Non-Performing	296,951	0.04	0.19	0.03	0.04	-0.01***
Industry Capture	296,951	0.09	0.098	0.086	0.073	0.013***
Bank-Firm Interactions	296,951	9.32	37.2	6.93	9.69	-2.7***

Note: This table shows summary statistics for loans in our sample. We count each bank-loan combination only once, on the date when it is first observed in our data (this may be a different date from the loan's first origination date for some early loans). Log size is based on the natural logarithm of the committed exposure, scaled by 1000 USD. The interest rate is the un-adjusted cost of the loan, measured in percent. Maturity is measured in quarters remaining. Unsecured is a dummy that takes the value of 1 if the loan is not secured with any type of collateral. "Non-performing" is also a dummy that takes the value of 1 if the loan falls in arrears, has negative maturity, or is otherwise in default. Industry capture measures the degree to which a given bank accounts for all lending to the two-digit industry of the borrower – it is measured as a share of all lending to that industry in the Y14 data. Finally, bank-firm interactions are a count of the number of times a given borrower and lender ever interacted in the Y14 data. The mean values of each variable data are split by whether a loan is made in a lender's favorite (i.e. "top") industry or not.

Table 1 the average size of our logged loans is 8.4. The average interest rate for loans in our sample period is 3.8%. Around 18% of loans use no form of collateral. Non-performing loan is a dummy that takes the value of 1 if a loan is either flagged as non-accruing or in default by the reporting bank or if the loan is more than 89 days in payment arrears or if the loan has any negative maturity because some amount of the principal or interest has not been re-payed at loan end. Only just over 4% of term loans ever become non-performing during their time in the sample.

The degree to which an industry has been captured by a single bank is around 9% in our sample – meaning the average bank accounts for 9% of C&I lending to a certain two-digit industry, as defined within the Y14 sample. The average number of interactions between a bank and a borrower is 9. This figure naturally rises towards the end of our sample and is driven by larger institutions. The regressions below use both past bank-borrower interactions (where the average is 4.2) and total interactions to capture a "relationship" between the bank and borrower.

3 Measuring and Documenting Specialization

In this section, we define our measure and document bank specialization. We start by defining our measures of specialization. Then, we take our measures to the data and look at summary statistics and specialization patterns across and within banks over time.

3.1 Measuring Specialization

We measure bank specialization in an industry in several ways. Our first measure of specialization is also our simplest. We use the share of a bank's C&I portfolio invested in a single industry or sector. We define industry according to 2-digit NAICS codes and sectors according to 4-digit NAICS codes.[12] However, because some industries are naturally larger than others banks may invest larger shares of their C&I portfolios in larger industries. For example, more funds are lent to manufacturing businesses than to warehousing businesses in any given year. Hence, a perfectly diversified bank would lend according to the total amount an industry/sector borrows in a given period. As such, it may mechanically lend much more to one industry than another.

Our second measure of specialization, which we refer to as "excess" specialization, takes this into account. Specifically, we calculate the difference between the share of bank b's C&I portfolio invested in a single industry/sector s and the 'diversified share' that a perfectly diversified bank would invest in that sector. This diversified share is based on the total lending directed towards the industry/sector s in a given period, relative to all C&I lending, across all sectors in the period – see Equation 1. For example, a bank that directs 10% of its portfolio in the warehousing industry in 2012q3 when this industry accounts for 5% of total C&I lending in 2012q3 would be over-invested in that industry by 5% and have an excess specialization of 5%. We use excess specialization as our baseline measure of specialization

[12]There are 24 industries based on 2-digit NAICS codes and 301 sectors based on 4-digit ones.

in most of our analyses.

$$Excess\ Specialization \equiv \frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}} - \frac{LoanAmount_{s,t}}{\sum_s LoanAmount_{s,t}} \tag{1}$$

Alternatively, we can calculate a "relative" specialization measure, similar to the one used in Paravisini et al. (2020) (see the expression in (2)). In this case, our fictive bank with 10% of its portfolio invested in the warehousing industry in 2012q3 and excess specialization of 5% is over-invested by a factor of 2. A potential shortcoming of this measure is that it can introduce large right tails. After all, a bank that invests 10% of its portfolio in a small industry that accounts for only 1% of all C&I lending would be over-invested by a factor of 10. Extreme tails are especially common when using 4-digit industries at a quarterly frequency. The question is whether the extreme right tail provides meaningful information.

$$Relative\ Specialization \equiv \frac{\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}}}{\frac{LoanAmount_{s,t}}{\sum_s LoanAmount_{s,t}}} \tag{2}$$

A fourth alternative way to measure specialization is to take the opposite extreme and collapse the entire tail of our data. We can define a bank as being specialized in an industry if said industry is its "favorite" – i.e. if it has over-invested most heavily in that industry. As such, we have a binary definition of specialization that abstracts from the degree to which a bank is over-invested. We show results for each of these measures in the paper below.

3.2 Documenting Specialization

We begin our analysis by documenting that our sample of large stress-tested banks specialize. Figure 1 (discussed briefly above) shows the average bank's excess specialization in its "top" 2-digit industry (i.e. the industry in which it is most specialized), its second most preferred industry, and all other industries. We remove lending to financial services from our data. If a bank does not lend in a particular industry, that industry is not part of that bank's portfolio, which explains why the average excess share in all other industries is non-negative.

Figure 1 shows that the share of the average bank's portfolio in its most favored industry is substantially larger than would be expected from a pure diversification standpoint. The mean bank invests an average of 9% more of its portfolio in its most favored industry (over our sample period) than a fully diversified bank. Moreover, banks have – on average – one or two preferred industries in which they are over-invested to a significant degree. In all other industries, they are either not invested or invested following diversification expectations.

Since we use several definitions of specialization throughout our analyses, it behooves us to test how our various measures correlate. In Figure A.1, which is part of an extensive discussion on specialization in Section A in our Online Appendix, we show that there exists a high degree of correlation between excess and relative specialization, between excess specialization and the share of a bank's portfolio in an industry, as well as between specialization calculated using held loans and specialization calculated using new loans originated in that period. This implies that, while there may be differences in degrees, the industries in which we consider a bank to have specialized are not dependent on the definition used. Also, the correlation between specialization calculated using held loans and new loans is important, as it shows that a bank reaffirms its specialization through the issuance of new loans on a continued basis (more on this below).

Table 2 shows summary statistics for excess and relative specialization and the banks' portfolio shares. We use 2-digit NAICS industries and 4-digit NAICS sectors.We split our data into a bank's most favored industry and all other industries. This split reflects a binary definition of specialization we also use later in the paper. We can see that a bank is over-invested in its favored industry by 9% (as seen graphically above). This corresponds to a relative over-investment factor of 3.5, implying it is investing over three times more in its favorite industry than expected under full diversification.

Conversely, the average bank is not over-invested in all other industries, suggesting a significant disparity in preferences among banks. At the 4-digit sector level, this disparity is

Table 2: ***Summary Statistics of Specialization***

	Specialization Type	Top Industry Mean	SD	25-pct	75-pct	All Other Industries Mean	SD	25-pct	75-pct
Two Digit	Relative Specialization	3.56	1.30	2.37	5.12	1.26	0.71	0.80	1.56
	Excess Specialization	0.09	0.05	0.07	0.14	0.01	0.03	-0.01	0.02
	Portfolio Share	0.18	0.14	0.11	0.20	0.05	0.04	0.03	0.07
Four Digit	Relative Specialization	5.92	1.89	6.71	6.96	1.97	1.71	0.79	4.59
	Excess Specialization	0.04	0.01	0.03	0.05	0.01	0.02	-0.00	0.01
	Portfolio Share	0.07	0.05	0.06	0.07	0.03	0.08	0.00	0.02

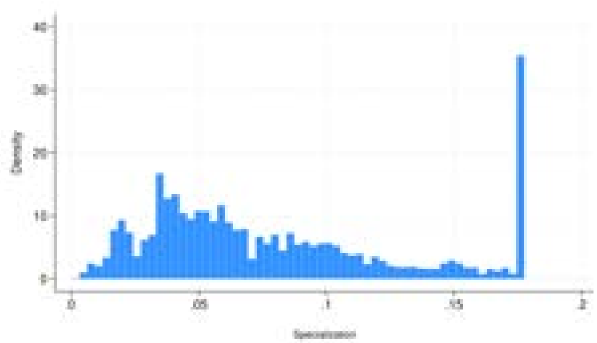
Note: This table shows summary statistics for various specialization measures at the 2-digit industry and 4-digit sector level. Specialization is defined as the degree to which a bank is over-invested in an industry, relative to a perfectly diversified portfolio. A diversified portfolio is one based solely on the size of an industry relative to all C&I lending. We show relative (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}} / \frac{LoanAmount_{s,t}}{\sum_s LoanAmount_{s,t}}$), excess (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}} - \frac{LoanAmount_{s,t}}{\sum_s LoanAmount_{s,t}}$) specialization and portfolio share (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}}$). We split data by whether an industry/sector is a bank's most preferred "top" industry (as measured by the respective specialization measure) or not.

even more pronounced. The relative over-investment at the 4-digit level is 4% (reflecting an over-investment factor of 5.9) for a bank's most preferred sector.

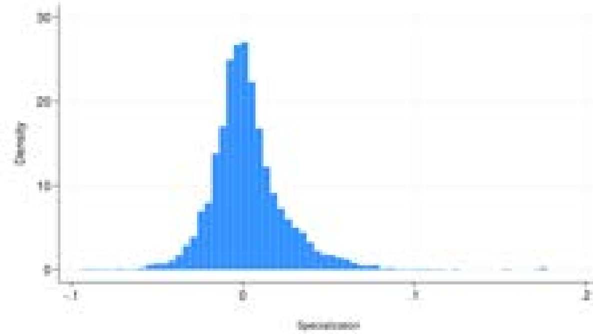
In Figure 2 we show the distribution of bank specialization at the two-digit level. In Panel (a) we show a histogram of banks' excess specialization in their top industries – i.e. the industry in which a bank has the highest degree of excess specialization. In Panel (b) we show a histogram of bank excess specialization in all other industries. As can be seen in Panel (a), there is significant heterogeneity in the degree to which a bank specializes in its favorite industry. Some banks are extremely specialized while others have very little excess specialization. We explore some of this variation when analyzing the determinants of specialization below. On the other hand, the distribution of specialization in non-favored industries is somewhat right-tailed but generally normal. This distribution is closer to one that we would expect if lending occurred without regard for an industry preference.

Across banks, Panel (c) in Figure 2 shows that specialization is negatively associated with bank size across banks. While large banks are less over-invested than smaller ones, they still have a preferred industry. Moreover, for the very large banks in our sample, even a small degree of excess specialization indicates a large dollar amount invested in a single

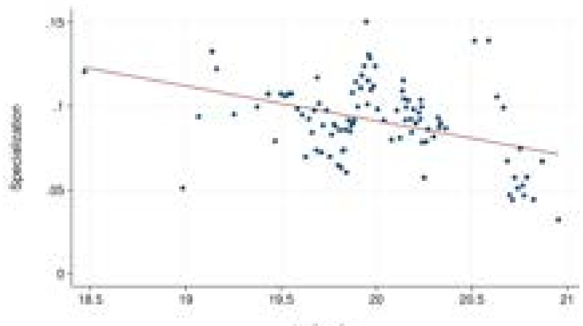
Figure 2: ***Variance of ExcessSpecialization and Bank Size***



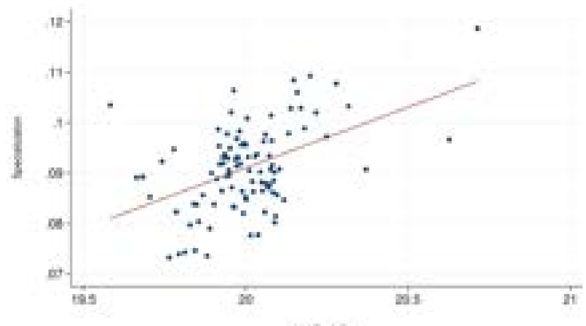
(a) Favorite Industry.



(b) All Other Industries.



(c) Specialization vs Bank Size



(d) Specialization vs Bank Size – within Bank

Note: This figure plots the distribution of excess specialization for the banks in our data in panels (a) and (b). Excess specialization is measured as the degree to which a bank is over-invested in an industry. A perfectly diversified bank is one that is invested by industry size, so that excess is $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}} - \frac{LoanAmount_{s,t}}{\sum_s LoanAmount_{s,t}}$. We split our sample into a bank's favored industry (panel (a)) and all other industries (panel (b)). Panels (c) and (d) relate specialization to bank size using bin scatters, where each bin represents at least 5 observations. Panel (c) accounts for bank category type and panel (d) accounts for bank fixed effects.

industry. Within bank, Panel (d) in the same figure shows a positive correlation between specialization and size. This suggests that bank growth is associated with a "doubling down" in its preferred industry.

We explore the relationship between specialization and other bank characteristics extensively in Section A in the Appendix. A few points are worth highlighting here. First, as Figure A.2 shows the most pronounced correlation is between the share of a bank's portfolio invested in C&I lending and a bank's specialization in its favorite industry. Higher shares of a bank's portfolio in C&I relate to higher degrees of specialization, a relationship lacking for any other bank asset category. As such, lending to corporations is associated with some degree of industry specialization.

Second, in terms of stability, we find that past specialization is highly predictive of future specialization by a given bank in an industry. In table A.1, we show that specialization in the previous quarter predicts current specialization almost perfectly – even when using newly originated loans. This holds for a bank's favorite industry and all other industries to which it lends. While not all banks retain the same preferred industry during their time in our sample, even the ones with lower consistency in their preferred industry tend to cycle between 2-3 preferred industries. This evidence suggests that specialization is not the result of some (randomly assigned) large loans maturing.[13]

Finally, we present summary statistics that relate loan characteristics to bank specialization. Table 1, shows that specialization is associated with differences in loan terms and loan performance. Loans in favored industries are on average larger, have lower rates and a longer maturity, and – perhaps most importantly – are less likely to become non-performing.

This is corroborated in Figure 3. Panel (a) shows the average difference in loan size

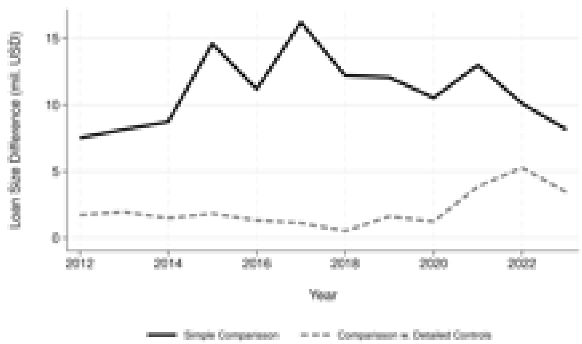
[13] A way to assuage concerns that bank specialization is the result of random lending would be to look at how specialization would develop if banks lent more randomly. We do this in Appendix A. We allocate loans in our data to banks randomly, based on bank and loan size. We then re-calculate our measure of specialization. In Figure A.4 we can see that the replication of Figure 1 under these assumptions looks very different to our data. There is little stability in a bank's favorite industry over time. Instead, a random allocation of loans to banks leads to very low levels of specialization, with high variability across time. This is strong evidence in favor of the fact that banks are choosing to invest in industries, rather than lending randomly.

granted by a bank in its most specialized/favorite industry and all others. We see that a bank's preferred industry receives larger loans. This holds – though to a lesser extent – if we account for loan characteristics such as its purpose, its rating, and the industry in question. We see similar patterns when looking at loan rates (in Panel (b)) as well as the share of loans in any period that becomes non-performing (Panel (c)).[14] If we account for loan controls, the difference in performance grows, with loans by specialized banks becoming even safer. Finally, we can see that the relationship between non-performance and specialization is fairly linear in Panel (d). The more specialized the bank, the less likely a loan is to be non-performing.

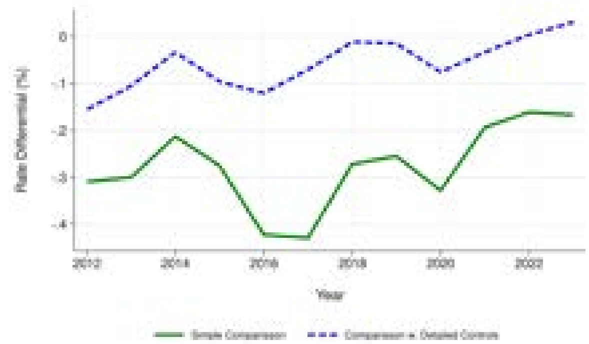
We test whether bank specialization in its favorite industries relates to balance sheet characteristics. In Figure A.3 and Table A.3 in the Appendix, we see that specialization relates positively to banks having a greater share of risky assets and negatively to banks having greater Tier 1 capital. This implies that banks are actively seeking the stability and familiarity of specialization in risky times and branching out into new industries in times of greater stability.

[14]Non-performance is defined as a borrower missing payments, a loan having negative maturity as the outstanding amount has not been repaid, or slipping into outright default.

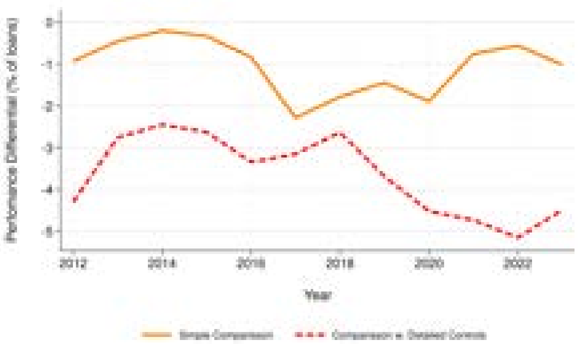
Figure 3: *Excess Specialization, Loan Terms, and Loan Performance*



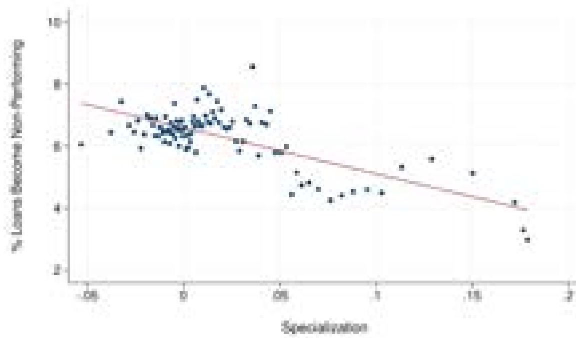
(a) Loan Size – Specialized vs Non-Specialized



(b) Interest Rates – Specialized vs Non-Specialized



(c) Loan Performance – Specialized vs Non-Specialized



(d) Loan Performance vs. Specialization

Note: This figure shows the difference between the size (panel (a)), the rates charged (panel (b)), and the likelihood of becoming non-performing (panel (c)) for loans granted by banks to their favorite industry vs. all other industries. Solid lines are raw differences and dashed lines account for loan rating, purpose, and industry*time. A bank's favorite industry is the one in which it is most over-invested (i.e. where $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}} - \frac{LoanAmount_{s,t}}{\sum_s LoanAmount_{s,t}}$ is largest for term loans in our sample). Panel (d) relates the propensity of loans to become non-performing to bank specialization in the borrower's industry.

4 Methodology

In this section, we discuss the methodology employed in the empirical exercises in this paper.

4.1 Specialization and Loan Performance

A natural question is how specialization impacts loan performance and, ultimately, the performance of specialized banks. If specialization is associated with a greater ability to ex-ante select or ex-post monitor high-quality loans, then one would expect the performance of loans by specialized banks to outperform their non-specialized counterparts. To assess this in our data, we regress loan performance on the industry specialization of the bank that granted it. Our baseline specifications take the following form:

$$\begin{aligned} NonPerformance_{l,i,b,s,T} = \beta_0 + \beta_l Specialization_{b,s,t} + \beta_2 \mathbf{X}_{l,b} + \beta_3 Relationship_{i,b} + \\ \beta_4 ShareZip_{i,z,t} + \xi_{b,t} + \sigma_{s,t} + \phi_{loanriskrating} + \omega_{loanpurpose} + \epsilon_{l,i,b,s,t} \end{aligned} \quad (3)$$

This specification relates the performance of loan l to firm i that is operating in industry/sector s and located in zip code z, granted by bank b at any time over its maturity T to the specialization of bank b in industry/sector s at t, the period when the loan is originated.

Non-performance occurs if a loan ever becomes 90 days past due, has any amount outstanding past the maturity date, or is otherwise tagged as non-accruing. The dependent variable is forward-looking, taking the value of 1 if the loan becomes non-performing at any point in the future. Hence, we relate a bank's specialization at loan origination to future loan performance. Our variable of interest is the degree to which a bank is specialized, i.e. over-invested, in an industry at a given point in time. Our premise is that the greater the specialization, the more knowledge gained and the greater the future loan performance. We use excess specialization, discussed above, at the 2-digit industry level as a baseline. We corroborate our results using relative specialization, the share of a bank's C&I portfolio in a single industry, or a binary variable for a bank's favorite industry at the 2- and 4-digit levels.

We include several key controls in all our specifications. These include loan terms – such as size, rate, maturity, and whether the loan is collateralized with one of 6 types of collateral (Real estate, Marketable securities, Accounts receivable, Fixed assets, Blanket lien, Other, or whether it is unsecured). We also include relationship variables that capture the number of times the bank and firm have interacted in our data. Previous relationships may, by themselves, build knowledge about a borrower's quality that is unrelated to industry specialization. We wish to disentangle the direct effect of information a bank has on a single borrower from the experience a bank may have in an entire industry. We acknowledge that our data begins after many banks have established relationships with firms. As such, we also include future bank-firm interactions as a rough proxy for the overall propensity of the bank to lend to the firm in question. We further include loan type, and loan purpose fixed effects.

Beyond identifying any informational advantage a bank may have in the industry in which it specializes, we are interested in whether any performance differential is the result of a bank's ability to monitor a loan post-origination or whether it is the result of a superior ability to pick loans ex ante. To help disentangle these two possibilities, we use a bank's internal ratings as well as the loan's interest rate at origination. In specifications that account for initial ratings/the rate paid, superior subsequent loan performance could arguably be the result of bank monitoring.[15] After all, the loan's ex-ante quality at origination should be reflected in a combination of interest rate and internal rating. The bank's rating is not known to the market but communicated to examiners in the context of stress testing. Thus, it reflects verifiable information that may be secret but has to be objective. It cannot be a reflection of the bank's industry knowledge if this cannot be communicated to stress testing officials. Hence, the rating cannot reflect the bank's expectations around its monitoring.

Finally, we use bank-time and industry-time fixed effects in all regressions. Bank-time fixed effects allow us to show differences in loan performance observed within a bank, by com-

[15] Heitz et al. (2023) find compelling evidence of banks actively monitoring loans in the construction sector.

paring loans made to industries in which the bank is differently specialized. With industry-time fixed effects, we can show differences in loan performance observed within an industry, by comparing loans made by banks that are differently specialized. In an extension, we use firm-time fixed effects; we compare loans made to the same firm by two differently specialized banks. However, since the sample of firms that borrow from multiple banks in a period is small and non-random, this is seen as indicative evidence only.

Accounting for Industry Capture and Geography Banks with a high degree of specialization have a significant amount of capital invested in a single industry and may capture a significant share of said industry. Many have documented an increase in bank concentration in the loan market over the past years (consider: Fernholz and Koch (2016) or Laeven et al. (2016)). Therefore, the specialization observed could be unrelated to a bank's knowledge about an industry and driven by a bank's rent-seeking behavior. Through decreased competition and higher market power, banks that capture an industry may be able to extract high rents from captive companies. Moreover, capture may exacerbate asymmetric information problems and prevent the entry of competitive new entrants (see Cetorelli and Strahan (2006) and Bikker and Haaf (2002)). To test our hypothesis and gauge the effects of knowledge-driven specialization, divorced from any rent-seeking behavior by banks, we control for a bank's industry capture. We define capture as the share of the Herfindahl-Hirschman index (HHI) of industry s that is accounted for by bank b at time t. HHI is a relatively common measure of competition. As such, the degree to which a single bank affects the competitiveness of an industry is a good measure of the degree to which it has captured that industry. An industry with only one bank will be perfectly captured by that bank. Similar to our measure of specialization discussed above, the measure is continuous and bounded between 0 and 1, making it easy to interpret. There naturally exists a high degree of correlation between industry capture and specialization. However, given that these industries are defined at the two-digit NAICS code, they are extremely large. As a consequence, even large banks can specialize without necessarily capturing an industry.

Lastly, bank specialization in certain industries may have grown from regional concentration. After all, certain industries might have originally clustered in certain areas. To account for this, we also define a "regional share" variable that accounts for the share of a bank's lending in a geography. We use either zip z or state g as a geographic indicator. We determine the location of investment with the headquarters of the borrower so that this measure can be seen as a rough approximation. Nevertheless, it would capture regional hubs – such as automotive in Detroit – that might have grown alongside banks initially operating in that area.

4.2 Specialization and Loan Characteristics

We next seek to understand how specialized banks may attract borrowers they deem valuable. To do so, we relate a bank's specialization in an industry to the loan characteristics it offers in the industry in question. Our baseline regressions take the same form as the regressions discussed above in Equation 3.

The primary loan characteristics used as left-hand-side (LHS) variables are log loan amount, the interest rate, loan maturity, and whether a loan is secured or unsecured.[16] We readily acknowledge that all the loan characteristics are simultaneously determined. Therefore, we use those loan characteristics that are not employed as the dependent variables as additional independent regressors. That is to say, in regressions focusing on the correlation between loan size and specialization, for example, we control for the loan's interest rate, its time to maturity, and whether it is secured. We observe only equilibrium outcomes and, as such, we can measure only correlations as opposed to causal relationships.

Ultimately, we cannot observe an individual firm's loan demand. To the extent that it is either time-invariant or driven by aggregate trends, fluctuations in a firm's loan demand

[16] Bander and Lewis (1986) document that higher interest rates lead to more aggressive product market strategies. Cetorelli and Gambera (2001) point out that bank concentration and product market competition could be correlated.

would be captured by firm fixed effects in the specifications that employ them. However, while adding firm fixed effects accounts for some component of a firm's loan demand, it also implies that our regression coefficients are identified among firms obtaining multiple loans from different banks in the same narrow period. It is worth noting that this specific subset of firms may not be fully representative. Therefore, we are careful when interpreting our results.

4.3 Specialization and Bank Performance

If banks offer beneficial loan terms to attract good borrowers, as suggested by our hypothesis, this should manifest as lower aggregate profitability. We test this by collapsing our data to the bank-time level to identify the average specialization of a given bank. We then combine this data with information from Y9-C data, which is maintained by the Federal Reserve System and tracks key balance sheet and income data of individual bank holding companies. This combined data allows us to relate aggregate bank performance to specialization.

If bank specialization is driven by informational advantages, we would expect banks with a high degree of specialization to have a safer more stable loan portfolio overall. Similarly, if banks are able to choose and manage the loans to their preferred industry better, we would expect loans by these banks to "buck" negative industry trends, and loans by specialized banks should perform better than loans by non-specialized banks, even in a downturn.

To test whether specialized banks buck negative industry trends, we relate a loan's performance to the aggregate performance of all loans in its industry. We then interact the degree of the lender's specialization with a measure of aggregate industry performance to determine whether specialization can assuage aggregate issues.

4.4 Extension I: Specialization and SME Lending

It has been well documented that large banks are somewhat less willing to lend to SMEs (see Berger et al. (1998), Strahan and Weston (1998), Peek and Rosengren (1998), Berger et al. (2005), or Berger and Udell (2008)). This may in part be a consequence of the fixed cost of each loan contract, which becomes unattractive in the case of large banks that lend to large borrowers. It may also be related, however, to the opacity of small firms and the physical distance between loan officers and SME borrowers that prevents the buildup of soft information. Strahan (2017) specifically points out that lager banks –like those in our sample – prefer lending based on "hard" information.

If specialization is driven by the bank's incentives to acquire industry-specific knowledge, we would expect specialized banks to be more willing to lend to small firms. If a bank has an advantage in assessing firms in an industry, it should be able to assess small firms in that industry better than competitors and ultimately be more willing to engage in lending to these opaque borrowers. We test this proposition with a regression that relates the propensity of a bank to lend to small firms to its specialization in the respective industry. The regression takes the same form as those above, except that we replace our outcome of interest with a dummy for whether the loan in question is a small loan (less than 3 mil. USD) or a loan to a small firm (less than 25 mil. USD in assets).

4.5 Extension II: Deposit Growth, Specialization, and Real Effects

In an additional extension, we use the sudden change in aggregate deposits held at large banks during the COVID-19 outbreak. As Figure A.11 shows, deposits at large banks rose suddenly in 2020. We exploit this rise in deposits, as it was exogenous to the actions of any individual bank at that time. First, we test whether banks are more likely to lend new – and unsolicited – deposits to their industry of specialization. We collapse our data to the bank-industry-time level and merge in FR2900 data on weekly bank deposit changes. We relate deposit shocks to the degree to which a bank is specialized in each industry.

Next, we relate the bank deposit shock to real effects by combining our bank data with firm-level information from COMPUSTAT. We limit our sample to those firms that have borrowed from a Y14 lender within a five-year period (2018 to 2022) at least once. We relate loan growth to firms and the subsequent growth of these firms to the expansion of bank deposits (see Appendix I for more details).

5 Specialization and Loan Performance

5.1 Baseline Results

Table 3 shows that a loan is less likely to become non-performing – over the time it is observed in our sample – if made by a bank more specialized in the borrower's industry. Looking at Column (1), we can see that a loan in a bank's most preferred industry is around 1.3% less likely to be non-performing than a loan in any other industry.[17] Given the average non-performance rate of loans in our sample of 4%, this effect is both statistically and economically meaningful. If loan ratings are standardized on a scale of 1 to 10, then we would observe similar performance differences between loans in the best rating group (i.e. group 1) and loans 6 rating buckets lower (i.e. rating group 7).

If we include bank*time and industry*time fixed effects, the coefficient drops slightly, remaining economically and statistically meaningful. If we further account for the interest rate paid by the borrower as well as the loan's size as indicators of observable loan riskiness (Column (3)), we see that the effect is again only marginally diminished. As such, the specialized bank outperforms that market's expectation of a loan (i.e. the interest rate charged for a given size).

In Column (4) we include the bank's ratings for the loan at the time the loan is originated.

[17]This is computed by taking the delta of mean specialization in a preferred industry (0.09) and mean specialization in all other industries (0.01) multiplied by the coefficient of 1.6.

Table 3: ***Specialization and Loan Performance***

	(1)	(2)	(3)	(4)	(5)
	Loan ever becomes non-performing				
Excess Specialization	-0.156***	-0.139***	-0.121***	-0.091***	-0.098***
	[0.016]	[0.014]	[0.014]	[0.013]	[0.013]
Interest rate			0.014***	0.007***	0.007***
			[0.000]	[0.000]	[0.000]
Log loan amount			0.000	0.001***	0.001**
			[0.000]	[0.000]	[0.000]
Share of Portfolio in ZIP					-0.013*
					[0.007]
Number of past interactions (relationship)					-0.000
					[0.000]
Future Interactions (relationship)					-0.001***
					[0.000]
General Fixed Effects	Purpose, Time				
Specific Fixed Effects	Bank	Industry*Time, Bank*Time			
Loan Rating Fixed Effects	No	No	No	Yes	Yes
Collateral Fixed Effects	No	No	No	No	Yes
Mean of dependent variable	0.04	0.04	0.04	0.04	0.04
R^2	0.011	0.037	0.046	0.14	0.14
N	298,043	298,043	298,043	296,951	296,951

Note: This table shows the coefficients of interest for equation:

$$NonPerformance_{l,i,b,s,T} = \beta_0 + \beta_l Specialization_{b,s,t} + \beta_2 \mathbf{X}_{l,b} + \beta_3 Relationship_{i,b} + \beta_4 ShareZip_{i,z,t} \xi_{b,t} + \sigma_{s,t} + \phi_{loanriskrating} + \omega_{loanpurpose} + \epsilon_{l,i,b,s,t}$$

It regresses whether loan l to firm i in quarter t operating in sector/industry s and located in zip code z, which is made by bank b ever becomes non-performing in future periods on bank b's specialization in industry s. "Non-performing" is a dummy that takes the value of 1 if the loan ever falls in arrears, has negative maturity, or is otherwise in default after it is first observed in our data. Specialization is defined as the degree to which a bank is over-invested in an industry, relative to a perfectly diversified portfolio. A diversified portfolio is one based solely on the size of an industry relative to all C&I lending. We use "excess" (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}} - \frac{LoanAmount_{s,t}}{\sum_s LoanAmount_{s,t}}$)) specialization at the two-digit industry level. Columns include additional controls and fixed effects as specified. Column (1) includes only time-invariant bank fixed effects as well as purpose and quarter*year. Columns (2)-(5) additionally include industry*time and bank*time fixed effects. We account for the interest rate paid in (3)-(5), the loan's rating at first observation (4)-(5), and the size of the loan (4)-(5). We include dummies for different collateral types pledged in Column (5), including real estate, marketable securities, accounts receivable, fixed assets, blanket lien, and "all other". Our data is focused on the first observation of a loan and contains only term loans. Standard errors are clustered at the bank-industry-year level while *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

The coefficient is 25% smaller under this specification than the coefficient in Column (3). To some extent then, specialized banks pick firms they consider to be safe.[18] In Appendix B, we show the relationship between various borrower characteristics and specialization. We can see that more specialized lenders are more likely to lend to safer borrowers in A.5.[19] However, this relationship is somewhat noisy. As such, even when we include initial ratings in our regressions, our effect of interest remains economically and statistically meaningful. Ratings in our setting, as discussed above, are risk groups for the purpose of examinations and stress testing. They reflect observable information and will likely not include all of a bank's specialized knowledge about the borrower. They are further unlikely to include a bank's private assessment of its monitoring ability in its favorite industry.

In Column (5), we include additional controls. First, we include fixed effects for the various types of collateral that borrowers may pledge. These include fixed assets, business lien, accounts receivable, real estate, "other", cash flow, and whether a loan is "unsecured". We additionally include the share of the lender's C&I portfolio invested in the borrower's zip code. This captures regional concentration that may relate to specialization. Certain industries have a greater propensity to cluster in certain regions, partly due to historical reasons. We want to avoid accidentally entangling our measure of industry specialization with a bank's ability to discern good borrowers in certain regions – perhaps because of their physical presence in these areas. Finally, we include a measure of the number of times a bank and a borrower have interacted or will interact. Combined, this is a measure of relationship lending as we want to disentangle a bank's expertise with a single borrower from its expertise in an industry. Overall, our coefficient of interest remains unchanged when compared with Column (3). Industry specialization by banks is a key determinant of

[18]Of course, an alternative explanation is that banks do not report fully accurate risk ratings for loans in specialized industries, as discussed in Behn et al. (2022) If they overestimate risks because these are based not on soft information but on hard metrics, we cannot disentangle monitoring from loan selection.

[19]Specialization is associated with lending to smaller firms, despite the aggregate trend of large banks becoming increasingly less likely to lend to smaller borrowers. Section B discusses this in detail. Large specialized banks are more likely to make loans to smaller firms or make smaller loans, in almost all periods of our data. This is additional evidence that specialization can overcome opacity.

future loan performance.

An important question that arises in the context of the above discussion is the degree to which different forms of specialization – such as regional specialization, borrower relationships, etc. – individually play a role in determining loan performance. In Table 4 we look at each form of specialization independently, before “horseracing” the various measures in the same regression. We additionally include the “share” of a bank’s portfolio in a single industry as a measure of specialization, since the larger a bank’s participation in a given industry, the higher the incentives the bank has to acquire knowledge in it.

In Column (1) of Table 4 we see the effect of our baseline specialization measure without including any other form of specialization – such as regional concentration or relationship lending. The effect is very similar to the effect measured above in Table 3. In the face of risk controls that include loan rating and the interest rate paid, the addition of other measures of specialization has a marginal effect on our coefficient of interest (Column (5)). In Column (2) we use the share of a bank’s portfolio invested in an industry as our measure of specialization. We find that the effect of industry concentration on whether loans become non-performing is similar to the effect of specialization; a one standard deviation increase in concentration would reduce the chance of loan non-performance by 70 basis points. This measure is highly correlated with our baseline measure of specialization, especially when including industry-time fixed effects. Given that we winsorize the former but not the latter, there are mechanical differences that arise.

We next analyze the degree to which regional specialization – i.e. the share of a bank’s C&I portfolio in a single zip code – affects loan performance. The measure is calculated using the borrower’s head quarters (HQ) location. Nevertheless, we see that the relationship between regional concentration and eventual loan issues is negative and significant. This is unaffected by whether we use specialization at the state level (see Appendix Table A 17). Finally, we use relationship lending – i.e. the count of bank-firm interactions – as our measure of interest. As in all the cases above, the relationship is negative and significant.

Table 4: *Alternative Specialization Measures and Loan Performance*

	(1)	(2)	(3)	(4)	(5)	(6)
	Loan ever becomes non-performing					
Excess Specialization	-0.097***				-0.096***	-0.007***
	[0.013]				[0.026]	[0.002]
Share of Portfolio in industry		-0.089***			-0.020	-0.002
		[0.012]			[0.017]	[0.004]
Share of Portfolio in ZIP			-0.015**		-0.012	-0.001
			[0.007]		[0.007]	[0.00]
Borrower Relationship				-0.001***	-0.001***	-0.003***
				[0.000]	[0.000]	[0.001]
Fixed Effects	Bank*Time, Loan Purpose, Loan Rating					
Industry Time FE	Yes	Yes	Yes	Yes	No	No
Controls	Loan Rate, Size, Maturity, Bank Industry Capture, Collateral					
Standardized Coefficients	No	No	No	No	No	Yes
Mean of dependent variable	0.05	0.04	0.04	0.04	0.04	0.04
R^2	0.16	0.16	0.16	0.16	0.17	0.17
N	296,951	296,951	296,951	296,951	296,951	296,951

Note: This table shows the coefficients of interest for equation:

$$NonPerformance_{l,i,b,s,T} = \beta_0 + \beta_l Specialization_{b,s,t} + \beta_2 \mathbf{X}_{l,b} + \\ \xi_{b,t} + \sigma_{s,t} + \phi_{loanriskrating} + \omega_{loanpurpose} + \epsilon_{l,i,b,s,t}$$

It regresses whether loan l to firm i in quarter t operating in sector/industry s and located in zip code z, which is made by bank b ever becomes non-performing in future periods on bank b's specialization in industry s. "Non-performing" is a dummy that takes the value of 1 if the loan ever falls in arrears, has negative maturity, or is otherwise in default after it is first observed in our data. Specialization is defined in several different ways. First, we use the degree to which a bank is over-invested in an industry, relative to a perfectly diversified portfolio. A diversified portfolio is one based solely on the size of an industry relative to all C&I lending. We use excess specialization (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}} - \frac{LoanAmount_{s,t}}{\sum_s LoanAmount_{s,t}}$)) at the two-digit industry level. Second, we use the share of a bank's portfolio invested in a single industry (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}}$). Third, we use the share of a bank's portfolio invested in a single zip code (i.e. $\frac{LoanAmount_{b,z,t}}{\sum_s LoanAmount_{b,z,t}}$). Finally, we use the relationship between a bank and a borrower, measured as the number of interactions the two have over our entire dataset. Column (6) makes use of standardized coefficients. All columns contain bank*time, collateral type (including real estate, marketable securities, accounts receivable, fixed assets, blanket lien, and "all other"), and loan purpose fixed effects. We additionally account for the interest rate paid, the loan's rating at first observation, and the size of the loan. Columns (1) -(4) include industry*time fixed effects. Our data is focused on the first observation of a loan and contains only term loans. Standard errors are clustered at the bank-industry-year level while *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

We horse-race all measures against one another in Column (5) and interpret the regression coefficients as standardized in Column (6). To interpret the coefficients on both excess specialization and portfolio concentration, we exclude the industry*time fixed effects in these columns. A one standard deviation increase in specialization would lead to a nearly 1% decrease in the chance of a loan becoming non-performing even in this saturated specification. The horse race reveals that our measure of specialization is not picking up another type of concentration. Moreover, while portfolio concentration and specialization are highly correlated, leading to issues of multicollinearity when estimating regressions including them both, each are valuable measures of specialization in their own right.

In Table 5 we show only the coefficient of interest for 21 variations of the main regression, focusing primarily on alternative definitions of specialization. Each cell represents an individual regression, for which we are displaying only the coefficient on our specialization measure. First, we use relative as well as excess specialization (see Equations 1 and 2) at the 2-digit industry and 4-digit sector levels. We next use the share of a bank's C&I portfolio in an industry – also at the 2- and 4-digit NAICS level – as well as a dummy for a bank's favorite 2-digit industry. A favorite industry is one in which a bank is most over-invested, as judged by the degree of excess specialization. We include purpose, bank*year-quarter, industry*year-quarter, and collateral type fixed effects in all columns. Columns (2) and (3) include interest paid by the borrower and Column (3) includes loan ratings at first observation. The first row shows the same coefficients already discussed above. All other rows show the alternative definitions of specialization.

As can be seen, our key result is independent of the definition of specialization that we employ. Going from any industry to a lender's most preferred industry – in relative 2-digit specialization rather than excess specialization – is associated with a 1.2% reduction in the likelihood that a loan becomes non-performing. This is highly similar to the 1.3% discussed above. The effects are statistically and economically significant at the four-digit level as well. However, the impact of relative specialization is somewhat smaller at this more granular level. This is because a lender is much more likely to have "multiple favorite sectors" that are related to the same 2-digit industry at the 4-digit-NAICS level. This raises the average specialization that we measure for sectors that are not a bank's favorite. Nevertheless, the effect of moving from a bank's favorite four-digit sector to any other four-digit sector to which the bank lends is associated with a 40 basis points drop in the likelihood of a loan becoming non-performing. Finally, our binary measure of a bank's favorite industry corroborates our analyses thus far.

It is worth noting that portfolio share measures are not winsorized whereas excess specialization measures are (at the 98 and 2 percent levels). This implies slightly different distributional tails, which explains some of the deviation between the two measures even if we include industry*time fixed effects. However, these results allow us to say that our findings are not driven by the sample construction.

5.2 *Specialized vs. Constrained Lending*

To highlight the connection between results observed and information acquisition resulting from bank specialization, we perform a series of additional tests. We primarily wish to show that banks are not specialized because they are lending-constrained. If a bank has sufficient capital to make only one loan to a borrower, the bank would be fully specialized in the borrower's industry. However, without having chosen to engage in repeated interactions with many borrowers in the industry, it may not benefit from information acquisition that results from a dedicated focus. We use the syndicated national credit registry (SNC) data to analyze the interplay between bank size and specialization.[20]

We explore this in detail in Appendix C. Arguably, none of the banks in our sample are small enough to be considered constrained. So, we take all SNC lenders, which include entities that are much smaller than the stress-tested banks in our Y14 sample, and split them according to size. We separate the LISCC banks as the largest bank category and otherwise

[20] Until 2018, the SNC tracked all syndicated loans held by at least two institutions supervised by the OCC, the FDIC, or the Federal Reserve that have a committed loan amount of at least 20 mil USD. It captures practically the entire syndicated loan market, especially for Term A and Credit Lines. See Blickle et al. (2020) for a discussion.

Table 5: ***Loan Performance – Different Possible Specialization Measures***

		(1)	(2)	(3)
		Loan ever becomes non-performing		
Excess Specialization	2-Digit Industry	-0.121***	-0.091***	-0.098***
		[0.014]	[0.013]	[0.013]
	4-Digit Industry	-0.098***	-0.049**	-0.065***
		[0.026]	[0.025]	[0.024]
Relative Specialization	2-Digit Industry	-0.005***	-0.004***	-0.005***
		[0.001]	[0.001]	[0.001]
	4-Digit Industry	-0.001**	-0.001*	-0.000*
		[0.000]	[0.000]	[0.000]
Portfolio Share	2-Digit Industry	-0.106***	-0.076***	-0.081***
		[0.012]	[0.012]	[0.012]
	4-Digit Industry	-0.070***	-0.038***	-0.046***
		[0.013]	[0.013]	[0.013]
Top Industry Dummy		-0.009***	-0.006***	-0.007***
		[0.002]	[0.001]	[0.001]
Baseline FE:		Bank*Time, Ind.*Time, Purpose		
Loan Rating at First Obs.		**No**	**No**	**Yes**
Interest Rate and Collateral Controls		**No**	**Yes**	**Yes**
Other Loan and Bank Controls		Yes	Yes	Yes
Loan Purpose and Type FE		Yes	Yes	Yes
N		298,043	296,951	296,951

Note: Each coefficient in this table is the result of a stand-alone regression. We vary our measure of specialization for equation:

$$Y_{l,i,b,s,T} = \beta_0 + \beta_l Specialization_{b,s,t} + \beta_2 \mathbf{X}_{l,b} + \beta_3 Relationship_{i,b} + \xi_{b,t} + \sigma_{s,t} + \phi_{loanriskrating} + \omega_{loanpurpose} + \epsilon_{l,i,b,s,t}$$

It regresses whether loan l to firm i in quarter t by bank b in sector/industry s ever becomes non-performing in future periods on the lending bank b's specialization in industry s. "Non-performing" is a dummy that takes the value of 1 if the loan falls in arrears or is otherwise in default. Specialization is defined as the degree to which a bank is over-invested in an industry, relative to a perfectly diversified portfolio. A diversified portfolio is one based solely on the size of an industry relative to all C&I lending. We use both relative (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}} / \frac{LoanAmount_{s,t}}{\sum_s LoanAmount_{s,t}}$) as well as excess (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}} - \frac{LoanAmount_{s,t}}{\sum_s LoanAmount_{s,t}}$) specialization at the 2-digit and 4-digit level. Further, we use the degree to which a bank has invested in a portfolio at the two and 4 digit levels as a further measure of specialization (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}}$) and whether a two-digit industry is a bank's favorite industry (i.e. a dummy). All specifications include bank*time, industry*time, and loan purpose fixed effects as well as loan size, bank-borrower relationship measures such as past and future bank-borrower interactions as well as the share of bank b's portfolio in borrower i's zip code. Columns (2) -(3) add collateral and interest rate controls and Column (3) adds loan rating at origination. Our sample includes only term loans when first observed in the sample. Standard errors are clustered at the bank-industry-year level while *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

split the remaining sample into 4 groups based on lender assets. We define our measure of 2-digit excess industry specialization for lenders in the SNC data in the same way we do for our Y14 lenders (see above) but use SIC codes reported by the arranging bank as opposed to NAICS codes.

Unfortunately, SNC data does not contain information on the rate paid by the borrower. Moreover, while some lenders report loan ratings, these are not standardized as in Y14 data. Nevertheless, we have several loan characteristics for which we can account. Moreover, we define a loan performance metric – based on loans becoming non-performing – that is based on whether a loan is ever in default or arrears. We relate this measure to lender specialization and other loan controls that include, type, purpose, size, and bank*time and industry*time fixed effects. As above, we focus on the period in which the loan was originated but use forward-looking measures of loan non-performance. We can see that the baseline effect of specialization on loan non-performance is positive, implying that specialized lenders are more likely to make loans that encounter issues.

However, if we interact our variable of interest with size-bucket dummies for our lenders, we see that the positive correlation between bad loan performance and lender specialization only holds for small lenders. The effect is reversed for larger – and arguably unconstrained – lenders. This is strong evidence in favor of the argument that specialization by unconstrained lenders can reflect informational advantages.

Our baseline specifications here use any loan held by a lender in SNC data – i.e. we use loans where the lender is a participant as well as loans where the lender is the lead arranger. The rationale is that an informed lender would only choose to participate in a loan that it considers safe. However, we can focus only on lead-arranging institutions as well. We see that the effects of specialization on loan performance are more pronounced for large banks that are also lead arrangers vs. large banks that are merely participants in syndicated loans. This result is more evidence in support of specialization leading reflecting informational advantages.

We corroborate that the specialization of the agent plays a role in Y14 data, where we can account for loan rating and loan interest rate (see Appendix Table A.7). If a loan is syndicated, the specialization of the loan's arranger matters. However, the effect is not completely explained by arranger specialization, implying that specialized banks still choose to participate in good loans. Moreover, the effect of lender specialization on loan performance is arguably still largest for small and un-syndicated loans, where there can be no concerns that uninformed lenders may try to free-ride on the information of an informed arranger.[21]

5.3 Robustness

5.3.1 Bank Deposits

In the above analyses, we briefly discuss the role of geography as a form of specialization, whereby banks could specialize in lending to certain regions. Banks may also naturally concentrate their lending in regions in which they obtain a large share of their deposits. We address this in Appendix D. We gather data on the share of deposits held at each branch of the US-based depository institutions in our sample. We can use this data to build a zip-code-based measure of deposit concentration.

We can see that deposit concentration and lending concentration are somewhat related. Banks tend to lend more in zip codes where they receive more deposits. However, if we relate deposit concentration to eventual loan performance, we see a positive relationship – i.e. loans are more likely to become non-performing. Banks are not automatically more informed in areas in which they receive more deposits. Lending where deposits are collected may be a passive/automated response to the bank's presence in a community and the economic environment of the area (see Nguyen (2019), Petersen and Rajan (2002)). As such, regional

[21] After all, if a loan is to remain on the balance sheet, the bank is most likely to invest in ex-ante screening and ex-post monitoring, given the exposure. Moreover, free-rider concerns (whereby other banks benefit from a specialized bank's knowledge) may be mitigated. In a similar vein, if a bank is syndicating a loan as the lead agent it has the greatest amount of control over loan terms. As such, we would expect some improved performance, relative to the terms for which we can account.

concentration may not yield the informational advantage we associate with a bank **choosing** to focus on a region or industry. We can define "excess" regional lending as the degree to which a bank has lent in a zip code in excess of the deposits it receives there. This measure is strongly negatively related to eventual loan performance issues. As such, it provides additional strong evidence in favor of the notion that bank specialization choices are associated with informational advantages. Naturally, our baseline measure of excess specialization remains negative and significant even if we account for deposit concentration or excess regional lending.

5.3.2 Alternative Definitions and Robustness

The above regressions make functional form and variable-definition assumptions that may impact our results. We wish to assuage such concerns somewhat in the extensive Appendix. We touch on some of the most important analyses briefly here, but refer the reader to the Appendix for a detailed discussion.

First, our definition of a borrower's industry does not impact the results. After all, we are relying on a single reporting source for a borrower's industry from the bank in Y14 data. Moreover, using SIC codes (instead of NAICS codes) and SNC data leads to very similar specialization results. We observe in Figure A.7 that the pattern of excess specialization in SNC data – for differently sized lenders – looks very similar to the pattern of specialization described in Figure 1 above. Moreover, in Appendix C, we use merged Y14 and COMPUSTAT data, whereby we then define specialization using the primary industry reported for a borrower in COMPUSTAT data. Even within this much smaller sample of merged data, we find that our specialization patterns described in Figure 1 still hold (see Appendix Figure A.8). Even using these alternative specialization definitions, we still see that specialization is correlated with positive future loan performance.

We further use the SNC data to test the idea that our results hold only in the recent period of low rates. Given that the SNC data starts in 1993, we can use the entire period in our tests, described above. Moreover, we can focus on the years of the recent financial crisis to test the idea that our results also hold in crises. Both conjectures are confirmed in Appendix C.

Next, we show that our functional form assumptions – and the fact that we do not observe all loans until maturity – do not impact our results. To this end, we run a dynamic panel regression in Section (E). The benefit of this regression is that it uses our entire data set and its panel structure, whereby each loan is observed multiple times. Some loans lose their non-performing status while in our sample, which we now exploit. Specifically, we relate a bank's specialization in the current and previous quarters to a loan's performance in the given quarter and include several controls.

In Appendix Table A.11, we see that specialization is still negatively related to the chance that a loan becomes non-performing in any given period. Given the high degree of auto-correlation within specialization measures, we also see that the effect of lagged specialization measures is insignificant after one quarter in the face of contemporaneous specialization. In a similar vein, we run a Cox-hazard model in Panel B. We again see that our results are confirmed and therefore not driven by the structure of our data.

Another concern may be that banks are picking safer borrowers in the industry in which they are specialized. This concern was addressed somewhat above when analyzing the relationship between specialization and loan rating and including rating-fixed effects in our analyses. However, the above tests rely on banks accurately reporting their ratings. In Appendix A.14, we use firm fixed effects to account for all unobservable firm-quality characteristics. While the relationship of interest can only be estimated among firms taking multiple loans from different banks within the short period of our sample, we still find that our coefficient of interest is negative and significant.

Finally, we analyze the impact of including lagged measures of specialization in Table A.15. Only contemporary specialization is relevant, even if we include lagged specialization. This makes sense, given the high degree of correlation within specialization across time.

As such, excluding contemporary specialization in favor of lagged specialization yields very similar results to our baseline specification.

6 Specialization and Loan Characteristics

Given the results discussed above, it is natural to ask how specialized banks attract and retain high-quality borrowers. We therefore relate loan terms to a bank's specialization. Specifically, we analyze whether specialized lenders grant larger loans, at favorable rates, with longer maturity, and with different collateral. We are not arguing for a causal link and acknowledge that loan terms are co-determined. We include loan terms that are not dependent variables as explanatory covariates. Results are depicted in Table 6.

From, Column (1), we can see that specialization is negatively correlated with interest rates. Moving from a bank's average industry its favorite industry would be related to a 45 basis point drop in rates – all else being equal. This is a sizeable effect, given the narrow estimation parameters. Similarly, a loan made to a borrower in a bank's preferred industry would be larger (9%) and of longer maturity (0.3 quarters). Finally, loans made by specialized banks are less likely to be unsecured.

In Appendix Table A.18, we explore the relationship between specialization and collateral further. We can distinguish different types of collateral, including real estate, marketable securities, accounts receivable, fixed assets blanket liens, and others. Specialization in an industry is associated with a higher likelihood that the lender accepts fixed assets, or "other assets". The latter is a group that is not further broken down but that does not fit into any of the other categories. Both are arguably assets that are more easily priced and, if the need arises, liquidated by banks with specialized knowledge.[22] Specialization is associated with a lower likelihood that the loan is secured by marketable securities, which include cash, or a blanket lien – both of which would require less specialized knowledge to liquidate and may

[22]The specialization banks can develop in valuing and dealing with certain collateral is explored in part in Gopal (2019).

Table 6: ***Specialization and Loan Terms***

Panel A

	(1) Interest rate	(2) Log loan amount	(3) Maturity remaining	(4) Unsecured
Excess Specialization	-0.569***	1.242***	3.803*	-0.044
	[0.134]	[0.114]	[2.008]	[0.043]
Fixed Effects	Bank*Time, Industry*Time, Loan Purpose, Loan Rating			
Controls	Loan Rate, Size, Maturity, Bank Industry Capture, Collateral			
Mean of dependent variable	3.6	8.2	22	0.14
R^2	0.36	0.25	0.28	0.38
N	296,951	296,951	296,951	296,951

Panel B

	(1) Interest rate	(2) Log loan amount	(3) Maturity remaining	(4) —
Excess Specialization	-0.193	1.220***	3.690*	–
	[0.154]	[0.107]	[2.047]	–
Int.: Unsecured * Specialization	-1.583***	0.294	15.267***	–
	[0.366]	[0.320]	[5.665]	–
Fixed Effects	Bank*Time, Industry*Time, Loan Purpose, Loan Rating			
Controls	Loan Rate, Size, Maturity, Bank Industry Capture, Collateral			
Mean of dependent variable	3.6	8.2	22	–
R^2	0.42	0 .26	0.28	–
N	296,951	296,951	296,951	–

Note: This table relates loan terms to a bank's specialization in that industry and a series of controls. We estimate the following regression:

$$Y_{l,i,b,s,T} = \beta_0 + \beta_l Specialization_{b,s,t} + \beta_2 \mathbf{X}_{l,b} + \beta_3 Relationship_{i,b} + \\ \xi_{b,t} + \sigma_{s,t} + \phi_{loanriskrating} + \omega_{loanpurpose} + \epsilon_{l,i,b,s,t}$$

It regresses the terms of loan l to firm i in quarter t by bank b in sector/industry s on the lending bank b's specialization in industry s. Specialization is defined as the degree to which a bank is over-invested in an industry, relative to a perfectly diversified portfolio. A diversified portfolio is one based solely on the size of an industry relative to all C&I lending. We use excess specialization (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}} - \frac{LoanAmount_{s,t}}{\sum_s LoanAmount_{s,t}}$) at the 2-digit level. All specifications include bank*time, industry*time, and loan purpose fixed effects. We include loan terms that are not the dependent variable as controls including loan size, loan rate, and loan maturity. We include dummies for different collateral types pledged, including real estate, marketable securities, accounts receivable, fixed assets, blanket lien, and "all other" in columns (1)-(3). Panel A makes use of the baseline specification. Panel B includes an interaction term for our variable of interest that takes the value of 1 if the loan is not secured by collateral. We exclude collateral fixed effects from these regressions. Our sample includes only term loans when first observed in the sample. Standard errors are clustered at the bank-industry-year level while *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

be preferred by unspecialized lenders.

In Panel B of Table 6 we interact our variable of interest with a dummy, denoting whether the loan is unsecured. We find that the interest rate effect is most pronounced among unsecured loans. In many cases, collateral may be a key determinant of rates. In unsecured loans, specialization is a key driver. Loan size is larger in all loans by specialized lenders. Maturity is larger in loans by specialized lenders, especially in the case of unsecured loans.

Our results suggest that a specialized bank has industry-specific knowledge which allows it to better evaluate potential borrowers. Better loan terms are reflective of borrowers being able to extract some of the rents banks derive from specialization. This is likely to be exacerbated by competition from non-banks or other sources of funding available to firms.

In Appendix Table A.19 we interact our variable of interest with a variable denoting whether the borrowing firm is borrowing from another Y14 bank in the same year. Firms that borrow from multiple lenders are typically large, rated, and established with known track records. As such, these results should be seen as indicative. We use both new originations and re-negotiations to have a larger selection of data. Nevertheless, we find that more specialized banks offer more attractive loan terms to those firms that have access to more than one lender in a given period. Larger borrowers with the option to borrow from many lender extract some of the rent banks obtain from specializing. In fact, given that firms may be averse to borrowing from a bank that might pass information to a competitor (see Asker and Ljungqvist (2010)) specialized banks may have to offer better loan terms to firms with better outside borrowing options.

7 Specialization and Bank Performance

In this section, we show that specialization can be associated with more stable performance but slightly lower bank profits, especially when a bank's favorite industry is performing badly. We thereby highlight some of the more negative consequences of specialization.

A concern may be that we have so far not made significant use of the time dimension of our data. It is possible that specialized banks are more likely to lend at times in which their preferred industries are doing well. While this may still imply that specialized banks are better at timing their loans and more selective about when they provide credit. However, we do not see that lending to a bank's favorite industry significantly changes over time (consider the examples discussed above and presented in Appendix Table A.4). Moreover, we can test the relative performance of a bank's loans to its specialized industry. In Table A.12 we first relate the chance that an individual loan becomes non-performing to the default rate of all *other* loans in its industry. Naturally, there is a strong positive correlation. However, this relationship does not hold for loans made by more specialized banks. Similarly, an individual loan's performance is related to the performance of the average loan in a bank's portfolio. A bank's average performance affects its entire loan portfolio, except in its specialized industry, where this relationship breaks down. As such, bank specialization counters bank-level and industry-level aggregate performance.

The above results show that banks charge slightly lower rates for larger loans in their industry of specialization. This is so that the bank may attract high-quality customers. It would naturally follow that a highly specialized bank may stand to earn less. We relate Y9C data, which tracks some aggregate bank balance sheet data at the quarterly level, to bank specialization. We discuss this in detail in Appendix F. We find that the degree to which a bank specializes in its top industry is somewhat negatively associated with profitability. This holds even though banks with higher specialization have somewhat more stable returns.

The interplay between return stability and profitability leads to aggregate dynamics we observed in Appendix A and discussed above. Banks with higher relative Tier-1 capital forego some of the stability offered by concentrating on the familiar to branch out. On the other hand, shocks to Tier 1 capital, or risk-weighted assets induce a bank to re-focus. In this case, a bank foregoes profitability in favor of stability (see Appendix Figure A.3 for details and discussion).

Furthermore, if we relate the average rate of non-performing loans in an industry to a bank's profitability, we find no relationship. However, if we additionally interact the share of non-performing loans in an industry with a measure of a bank's specialization in that industry, we find a negative relationship with income. More specialized banks are more exposed to downturns in their industry. We show above that banks are somewhat able to "counter" aggregate trends; loans in a bank's specialized industry outperform the industry average. Nevertheless, a specialized bank cannot avoid all performance issues in an industry. As such, aggregate downtrends in a bank's industry of preference translate to lower aggregate income. In aggregate, therefore, profitability and over-exposure may be considered the costs of specialization for a bank.

8 Extensions

8.1 SME Lending

In Appendix B we discuss the propensity of specialized lenders to lend to "SME"s. We see from Figure A.5 that the average assets of borrowers from specialized banks are smaller. Moreover, the propensity that an SME receives a loan from a specialized lender is much higher and directly related to the degree to which the lender is specialized. This result is corroborated when controlling for several loan characteristics in Table A.5. Unfortunately, data on firms is spotty. Especially smaller firms may have missing data, which means they cannot be identified as such. We therefore also make use of small loans as an indicator for small firm lending. Small loans – with less than 3 mil. USD – are more likely to be for small firms after all. The results are even stronger using this specification, indicating that large banks are still somewhat likely to lend to smaller firms in their industry of specialization.

8.2 Liquidity Provision and Monitoring

We show above that borrowers from specialized banks perform better, even in the face of aggregate industry downturns and after we account for a bank's loan ratings at origination. The natural question that arises is "why". We test two possibilities: better monitoring and better liquidity provision through crises.

First, using methodology developed by Gustafson et al. (2021), we use a subset of the SNC data to identify loans that are actively monitored by the bank. This includes loans for which the bank conducts repeated site visits or has repeated interactions with the borrower. We then relate the propensity to actively monitor a loan to loan characteristics and the specialization of the arranging lender. We find a strong positive correlation in Table A.9. Similarly, a specialized bank is more likely to review the collateral itself rather than hire an outside evaluator. This is indicative evidence of specialized lenders being more actively involved in the loans they make, which may help facilitate better performance.

Moreover, in Appendix G, we show that banks are more likely to provide credit to distressed borrowers in their industry of specialization. Borrowers are more likely to pull down credit from credit lines after they have defaulted on a loan. Similarly, borrowers draw down on credit lines after they have received any form of rating downgrade on any of their loans. This relationship is insignificant outside of a bank's specialized industry. This implies that only specialized banks are facilitating the draw-down.

In Figure A.10, we show the committed loan amount and the utilized loan share for credit lines. We split banks by whether they are specialized in an industry or not. We further focus only on firms that suffer at least a rating downgrade on any loan while in the data. The time of the downgrade is denoted as t=0. At the time of the rating downgrade, the amount of money committed to credit lines shrinks drastically. This is less true, however, for specialized lenders. Similarly, the amount a borrower can draw down is reduced after a downgrade, though this shift is only significant for borrowers from non-specialized banks.

Overall, we can argue that we have indicative evidence for the fact that specialized banks are better at providing liquidity during times of borrower distress, corroborating findings by Giannetti and Saidi (2019). This may follow from the fact that these lenders have better information on borrower and industry performance. The added liquidity may help borrowers overcome shortfalls, facilitating a more rapid and complete recovery.

8.3 COVID-19 Shock

At the outset of COVID-19, in March of 2020, aggregate deposits both grew (as purchases slowed and stimulus checks were saved) and were reshuffled across the financial system. Large banks – particularly those in our sample – were key beneficiaries of this reshuffling. As can be seen in Appendix Figure A.11, the total deposits of the banks in our sample increased 30% over a single quarter. This deposit growth was unsolicited. A smaller though similar deposit growth shock occurred around the SVB crash, where the four largest banks in our sample saw their deposits increase significantly again.

Banks invested these exogenous deposits in securities – such as treasuries. To the extent that the newly arrived funds were directed into C&I lending, they overwhelmingly flowed toward a bank's industry of specialization. We show that this holds in particular when we focus on the weeks with the largest deposit shocks, which were most likely to be exogenous to bank activities. Lending was encouraged during the COVID pandemic and the degree to which a bank was specialized influenced the direction of this lending.

The fact that banks invest these deposits in their industry of specialization, despite what aggregate loan demand or conditions may be, has implications for the firms in these sectors. We discuss this in detail in the Appendix I. We link firms in COMPUSTAT to our Y14 data. As such, we can determine which firms borrow from which lenders for a small subset of our largest firms. We relate firm growth and firm earnings in the COVID period to the degree to which their most specialized lender is specialized in the firm's industry and willing to lend. We find that the degree to which the firm's lender is specialized directly correlates with the firm's growth and earnings during COVID. After all, we can see that excess deposits were directed towards a bank's favorite industries as cheaper credit.

9 Conclusion

In this paper, we show that large stress-tested banks specialize their C&I loan portfolios in certain industries. Specialized banks direct a far larger share of their C&I lending towards their favorite industry, compared to what would be expected from a bank investing according to industry size. This specialization is consistent with superior information about borrowers in an industry – above and beyond the information provided by relationship lending. We show that specialization reduces borrower opacity and facilitates improved monitoring. As such we find that this type of specialization correlates with improved loan performance, even when controlling for borrower fixed effects and loan risk at origination. Banks attract high-quality borrowers, which specialized knowledge allows them to identify, with beneficial loan terms – such as larger amounts and higher rates.

Our results speak to the role of banks as designated intermediaries. We highlight that banks – even large banks – are not necessarily perfectly fungible in lending. During periods in which deposits are reshuffled in the banking system – such as during the COVID outbreak or possibly following the failure of SVB – borrowers of specialized banks may be differently affected than their competitors. This is a key insight for future research into banking as well as for policymakers.

We further show that banks trade off some profitability for stability by concentrating on their industry of specialization in times of lower Tier 1 capital. It is noteworthy that significant concentrations of lending into a single industry may pose a risk to banks during severe sector-specific downturns. Without data on events that cause disproportionate loan failures in a given sector, we are unable to complete the picture of the risks surrounding specialization.

References

Acharya, V., Hasan, I., Saunders, A., 2006. Should banks be diversified? evidence from individual bank loan portfolios. The Journal of Business 79, 1355–1412.

Acharya, V., Yorulmazer, T., 2007. Information contagion and bank herding. Journal of Money, Credit, and Banking 40.

Allen, F., Babus, A., Carletti, E., 2012. Asset commonality, debt maturity and systemic risk. Journal of Financial Economics 104.

Allen, F., Gale, D., 1998. Optimal financial crises. Journal of Finance 53, 1245–1284.

Allen, F., Gale, D., 2000. Financial contagion. Journal of Political Economy 108, 1–33. doi:10.1086/262109.

Asker, J., Ljungqvist, A., 2010. Competition and the structure of vertical relationships in capital markets. Journal of Political Economy 118.

Bander, J., Lewis, T., 1986. Oligopoly and financial structure: The limited liability effect. American Economic Review 76.

Barro, R., 1976. The loan market, collateral, and rates of interest. Journal of Money, Credit and Banking .

Battacharya, S., Thakor, A.V., 1993. Contemporary banking theory. Journal of Financial Intermediation 3, 2–50.

Beck, T., De Jonghe, O., Mulier, K., 2022. Bank sectoral concentration and risk: Evidence from a worldwide sample of banks. Journal of Money Credit and Banking 54.

Beck, T., Demirguc-Kunt, A., Laeven, L., Maksimovic, V., 2006a. The determinants of financing obstacles. Journal of International Money and Finance 25.

Beck, T., Demirguc-Kunt, A., Levine, R., 2006b. Bank concentration, competition, and crises: First results. Journal of Banking Finance 30.

Beck, T., Demirguc-Kunt, A., Levine, R., 2007. Bank concentration and fragility impact and mechanics, in: The Risks of Financial Institutions.

Behn, M., Haselmann, R., Vig, V., 2022. The limits of model-based regulation. The Journal of Finance 77, 1635–1684.

Berger, A., Miller, N., Petersen, M., Rajan, R., Stein, J., 2005. Does function follow organizational form? evidence from the lending practices of large and small banks. Journal of Financial Economics 76.

Berger, A., Saunders, A., Scalise, J., Udell, G., 1998. The effects of bank mergers and acquisitions on small business lending. Journal of Financial Economics 50.

Berger, A., Udell, G., 2008. Universal banking and the future of small business lending. NYU Working Paper No. FIN-95-009 .

Berger, A.B., Udell, G., 1995. Relationship lending and lines of credit in small firm finance. Journal of Business 68, 351–381.

Bernanke, B., 1983. Non monetary effects of the financial crisis in the propagation of the great depression. The American Economic Review 73.

Bikker, J., Haaf, K., 2002. Competition, concentration and their relationship: An empirical analysis of the banking industry. Journal of Banking and Finance 26.

Blickle, K., 2020. Local banks, credit supply, and house prices. Journal of Financial Economics 143, 876–896.

Blickle, K., Fleckenstein, Q., Hillenbrand, S., Saunders, A., 2020. The myth of the lead arranger's share. Federal Reserve Bank of New York Staff Report .

Blickle, K., He, Z., Huang, J., Parlatore, C., 2024. Information-based pricing in specialized lending. Working Paper .

Blickle, K., Luck, S., Brunnermeier, M., 2022. Who can tell which banks will fail? NBER Working Paper No. w29753 143.

Boyd, J., Prescott, E., 1986. Financial intermediary-coalitions. Journal of Economic Theory

38, 211–232.

Broecker, T., 1990. Credit-worthiness tests and interbank competition. Econometrica 58, 429–452.

Campbell, T.S., Kracaw, W.A., 1980. Information production, market signalling, and the theory of financial intermediation. The Journal of Finance 35, 863–882. URL: http://www.jstor.org/stable/2327206.

Carey, M., Post, M., Sharpe, S., 1998. Does corporate lending by banks and finance companies differ? evidence on specialization in private debt contracting. The Journal of Finance 53.

Cetorelli, N., Gambera, M., 2001. Banking market structure, financial dependence and growth: International evidence from industry data. Journal of Finance 56.

Cetorelli, N., Strahan, P., 2006. Finance as a barrier to entry: Bank competition and industry structure in local u.s. markets. Journal of Finance 61.

Degryse, H., Ongena, S., 2005. Distance, lending relationships, and competition. Journal of Finance 60, 231–266.

Diamond, D.W., 1984. Financial intermediation and delegated monitoring. The Review of Economic Studies 51, 393–414.

Diamond, D.W., Dybvig, P.H., 1983. Bank runs, deposit insurance, and liquidity. Journal of Political Economy 91, 401–419. doi:10.1086/261155.

Fernholz, R., Koch, C., 2016. Why are big banks getting bigger? Working Paper, Federal Reserve Bank of Dallas. .

Frattaroli, M., Herpfer, C., 2022. Information intermediaries: How commercial bankers facilitate strategic alliances. Journal of Financial and Quantitative Analysis 58.

Gertler, M., Gilchrist, S., 1994. Monetary policy, business cycles, and the behavior of small manufacturing firms. The Quarterly Journal of Economics .

Giannetti, M., Saidi, F., 2019. Shock propagation and banking structure. The Review of Financial Studies 32.

Giometti, M., Pietrosanti, S., 2022. Bank specialization and the design of loan contracts. FDIC Working Paper Series .

Goldstein, I., Huang, C., Yang, L., 2024. Open banking: Lending market competition and resource allocation efficiency. Working Paper .

Goldstein, I., Kopytov, A., Shen, L., Xiang, H., 2020. Bank heterogeneity and financial stability. working paper .

Gopal, M., 2019. How collateral affects small business lending: The role of lender specialization. NYU Working Paper .

Gorton, G., Winton, A., 2003. Chapter 8 - financial intermediation, in: Constantinides, G.M., Harris, M., Stulz, R.M. (Eds.), Corporate Finance. Elsevier. volume 1 of *Handbook of the Economics of Finance*, pp. 431–552. URL: https://www.sciencedirect.com/science/article/pii/S1574010203010124, doi:https://doi.org/10.1016/S1574-0102(03)01012-4.

Greenwood, R., Landier, A., Thesmar, D., 2015. Vulnerable banks. Journal of Financial Economics 115.

Gustafson, M.T., Ivanov, I.T., Meisenzahl, R.R., 2021. Bank monitoring: Evidence from syndicated loans. Journal of Financial Economics 139.

Haldane, A.G., May, R.M., 2011. Systemic risk in banking ecosystems. Nature , 351–355.

Hart, O., Moore, J., 1994. A theory of debt based on the inalienability of human capital. The Quarterly Journal of Economics .

He, Z., Huang, J., Parlatore, C., 2024. Information span in credit market competition. Working Paper .

He, Z., Huang, J., Zhou, J., 2023. Open banking: Credit market competition when borrowers own the data. **The Journal of Finance** , 863–882.

Heitz, A.R., Martin, C., Ufier, A., 2023. Bank monitoring with on-site inspection. FDIC working paper .

James, C., 1987. Some evidence on the uniqueness of bank loans. Journal of Financial Economics 19, 217–235.

Khwaja, A., Mian, A., 2008. Tracing the impact of bank liquidity shocks: Evidence from an emerging market. American Economic Review .

Laeven, L., Ratnovski, L., Tong, H., 2016. Bank size, capital, and systemic risk: Some international evidence. Journal of Banking Finance 69.

Leland, H.E., Pyle, D.H., 1977. Informational asymmetries, financial structure, and financial intermediation. The Journal of Finance 32, 371–387.

Loutskina, E., Strahan, P., 2011. Informed and uninformed investment in housing: The downside of diversification. The Review of Financial Studies 24.

Marquez, R., 2002. Competition, adverse selection, and information dispersion in the banking industry. The Review of Financial Studies 15.

Neuhann, D., Saidi, F., 2018. Do universal banks finance riskier but more productive firms? Journal of Financial Economics .

Nguyen, H., 2019. Are credit markets still local? evidence from bank branch closings. American Economic Journal: Applied Economics 2019 11.

Paravisini, D., Rappoport, V., Schnabl, P., 2020. Specialization in bank lending: Evidence from exporting firms. working paper .

Parlatore, C., Philippon, T., forthcoming. Designing stress scenarios. Journal of Finance .

Peek, J., Rosengren, E., 1998. Bank consolidation and small business lending: It's not just bank size that matters. Journal of Banking and Finance 22.

Petersen, M., Rajan, R., 2002. Does distance still matter? The Journal of Finance .

Petersen, M.A., Rajan, R.G., 1995. The effect of credit market competition on lending relationships. The Quarterly Journal of Economics 110, 407–443.

Rajan, R., 1992. Insiders and outsiders: The choice between informed and arm's-length debt. The Journal of Finance 47, 1367–1400.

Riordan, M.H., 1993. Chapter 11: Competition and bank performance: a theoretical perspective. C. Mayer, X. Vives (Eds.), Capital markets and financial intermediation, Cambridge University Press .

Ross, D., 2010. The "dominant bank effect:" how high lender reputation affects the information content and terms of bank loans. The Review of Financial Studies 23.

Saidi, F., Streitz, D., 2020. Bank concentration and product market competition. The Review of Financial Studies .

Sharpe, S., 1990. Asymmetric information, bank lending, and implicit contracts: A stylized model of customer relationships. The Journal of Finance 45, 1069–1087.

Stiglitz, J., Weiss, A., 1981. Credit rationing in markets with imperfect information. The American Economic Review .

Strahan, P., 2017. Are concentrated banks better informed than diversified ones? Journal of Accounting and Economics 64.

Strahan, P., Weston, J., 1998. Small business lending and the changing structure of the banking industry. Journal of Banking and Finance 22.

Tabak, B., Fazio, D., Cajueiro, D., 2011. The effects of loan portfolio concentration on brazilian banks' return and risk. Journal of Banking Finance 35, 3065–3076.

vanNieuwerburgh, S., Veldkamp, L., 2010. Information acquisition and under-diversification. The Review of Economic Studies 77, 779–805.

Wagner, W., 2010. Diversification at financial institutions and systemic crises. Journal of Financial Intermediation 30.

Westernhagen, N., Harada, E., Nagata, T., Vale, B., Ayuso, J., Saurina, J., Daltung, S., Ziegler, S., Kent, E., Reidhill, J., Peristiani, S., 2004. Bank failures in mature economies. BIS Working Paper Series 13 .

Yellen, J., 2013. Interconnectedness and Systemic Risk: Lessons from the Financial Crisis and Policy Implications: a speech at the American Economic Association/American Finance Association Joint Luncheon, San Diego, California, January 4, 2013. Speech 631. Board of Governors of the Federal Reserve System (U.S.). URL: https://EconPapers.repec.org/RePEc:fip:fedgsq:631.

Internet Appendix for "Specialization in Banking"

Kristian Blickle

Cecilia Parlatore

Anthony Saunders

A Persistence and Drivers of Specialization

In this section, we delve deeper into the mechanics of bank specialization. We discuss the persistence of specialization across time as well as the bank-specific characteristics that relate/ drive specialization into certain industries.

The banks that make up our sample are all large institutions. Specifically, our sample includes: Ally Financial, Bank of America, BBT, Bank of New York Mellon, Citigroup, Capital One, Fifth Third Bancorp, Goldman Sachs, JP Morgan Chase & Co, Keycorp, Morgan Stanley, PNC Financial Services Group, Regions Financial, Suntrust, State Street, U.S. Bancorp, Wells Fargo, Comerica, Huntington Bancshares, HSBC North America, M&T, Northern Trust, RBC USA, Santander Holdings USA, MUFG Holding, Zions Bancorporation.

We first show that there is a high degree of auto-correlation in specialization. The degree to which a bank has specialized in an industry in the past is predictive of the degree to which it will specialize in that industry in the future. We collapse our data to the bank*industry*time level to relate specialization to past specialization. Table A.1 shows that, in the absence of controls, the previous period's specialization in an industry explains 98% of the current period's specialization. Specialization from two periods before explains 96% of the current specialization. Even if we focus only on a bank's most favored industry, we still see that past specialization is highly predictive of current specialization.

Table A.2 shows a high degree of correlation between a bank's specialization calculated using its loan holdings and a bank's specialization calculated using only newly originated loans. Loan holdings may include loans (shares) the bank has purchased from other banks. It also explicitly excludes loans that it has chosen to divest. Throughout the paper, we use held loans, as these are arguably a better indicator of how a bank chooses to specialize. However, we can also calculate specialization by looking specifically at the loans originated by a bank in a given period. We find that the correlation between these two measures is almost 90%. If we focus on a bank's favorite industry, the relationship (that relates the two in a regression) rises above 100%. This is the result of increasing concentration in a bank's

Table A.1: *Lagged vs Current Specialization*

	(1)	(2)	(3)	(4)	(5)	(6)
			Excess Specialization			
Excess Specialization$_{t-1}$	0.975***		0.827***	0.953***		0.933***
	[0.000]		[0.000]	[0.000]		[0.001]
Excess Specialization$_{t-2}$		0.958***	0.153***		0.929***	0.022***
		[0.000]	[0.000]		[0.000]	[0.001]
Sample		All Loans			Top Industry	
Fixed Effects	No	No	Industry*Time	No	No	Industry*Time
Observations	20,828	20,828	20,828	872	872	872
R^2	0.95	0.92	0.96	0.93	0.88	0.96

Note: This Table relates excess specialization (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}} - \frac{LoanAmount_{s,t}}{\sum_s LoanAmount_{s,t}}$) to lagged excess specialization. Data is collapsed to the bank*industry*time level. Columns (1)-(3) use all industries and columns (4)-(6) use a bank's favorite industry. We include either no or industry*time fixed effects, as indicated. Standard errors are clustered at the bank-time level while *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

preferred industry over our sample period. Ultimately, we can see that specialization and concentration are highly stable over time and can be observed in newly originated and in held loans.

Next, we relate our primary measure of specialization (i.e. "excess" specialization, defined as $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}} - \frac{LoanAmount_{s,t}}{\sum_s LoanAmount_{s,t}}$) to other measures of loan specialization or concentration. We first show that "excess" and "relative" specialization (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}} / \frac{LoanAmount_{s,t}}{\sum_s LoanAmount_{s,t}}$) are highly correlated. This can be observed in panel (a) of Figure A.1 and mostly follows mechanically from the variables' construction. Similarly, we can see that industries in which a bank is more specialized are almost always also the industries in which a bank is most concentrated (i.e. where it has invested a larger share of its C&I loan portfolio) (panel (b)). However, at the extreme end of specialization, these measures diverge somewhat. This may follow from extreme specialization being more easily achieved in smaller industries. As such, in turn, it may be worthwhile looking at the impact of each measure of specialization on loan performance.

In panel (c) we relate our standard measure of "excess" specialization to the same measure that we would obtain by calculating specialization based solely on newly originated loans in that period (see above). We see a strong correlation between these variables. However, deviations between the two measures can emerge when banks make large loans to single industries that are not their historical "favorite" industry. The flow of new originations can

Table A.2: *New Loan Specialization and New Borrowers*

Panel A

	(1)	(2)	(3)	(4)	(5)	(6)
	Excess Specialization in New Loans					
Excess Specialization$_{t-1}$	0.888***	0.885***	0.888***	1.152***	1.029***	1.243***
	[0.012]	[0.012]	[0.012]	[0.088]	[0.101]	[0.126]
Constant	0.005***	0.005***	0.005***	-0.008	0.006	-0.012
	[0.000]	[0.000]	[0.000]	[0.011]	[0.012]	[0.015]
Sample	All Loans			Top Industry		
Fixed Effects	No	Industry	Industry* Time	No	Industry	Industry* Time
Observations	20,828	20,828	20,828	872	872	872
Adjusted R^2	0.213	0.220	0.220	0.162	0.245	0.220

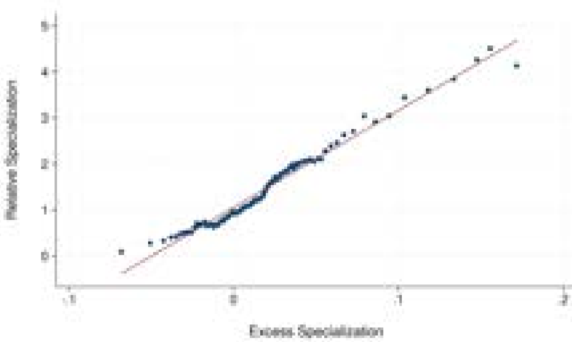
Panel B

	(1)	(2)	(3)	(4)
	New Borrower			
Excess Specialization$_{t-1}$	0.561***	0.321***	0.164***	0.382***
	[0.018]	[0.020]	[0.019]	[0.020]
Interest rate				0.007***
				[0.001]
Log loan amount				-0.053***
				[0.001]
Constant	0.398***	0.401***	0.404***	0.811***
	[0.001]	[0.001]	[0.001]	[0.006]
Fixed Effects	Time	Industry*Time, Rating	Bank*Time, Rating	Bank*Time, Rating
R^2	0.048	0.17	0.11	0.18
N	372,992	369,777	368,484	369,777

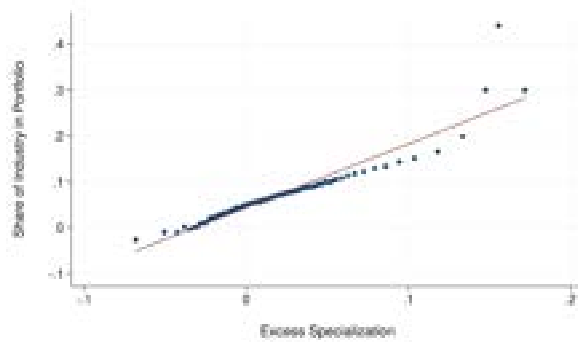
Note: In Panel A, this Table relates excess specialization as calculated with newly originated loans in a period to lagged excess specialization calculated with all loans on the bank's balance sheet. Data is collapsed to the bank*industry*time level. Columns (1)-(3) use all industries and columns (4)-(6) use a bank's favorite industry. We include either no, industry, or industry*time fixed effects, as indicated. Standard errors are clustered at the bank-time level. In Panel B, we relate whether a borrower appears in the data for the first time to whether it is borrowing from a specialized lender. Standard errors are clustered at the bank*industry*time level and *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

be impacted by outliers more easily than the stock of all held loans, especially in large banks. Finally, in panel (d), we relate specialization to the share of a bank's portfolio in a single ZIP code. There is very little correlation (which appears negative but insignificant) between regional concentration and specialization. We observe the same effect if we use county or state to denote "location" (not reported for brevity). Historical industrial concentrations are no longer as prevalent as they may have once been.

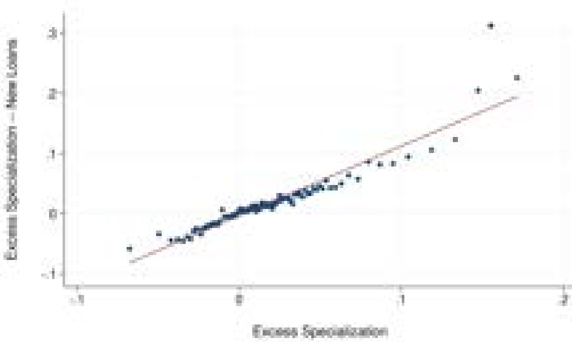
Figure A.1: ***Excess Specialization compared to other types of specialization***



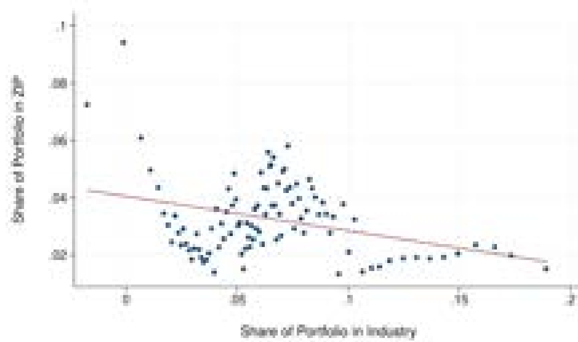
(a) Excess vs. Relative Specialization



(b) Excess Specialization vs. Portfolio Share



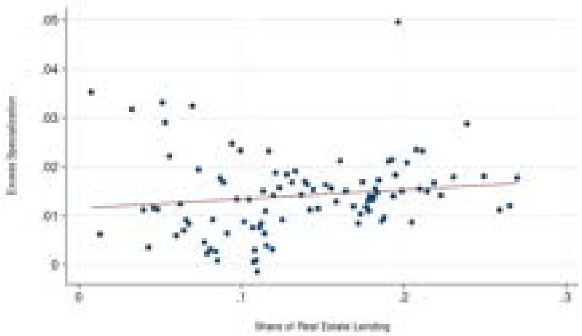
(c) Excess Specialization: New vs. Held Loans



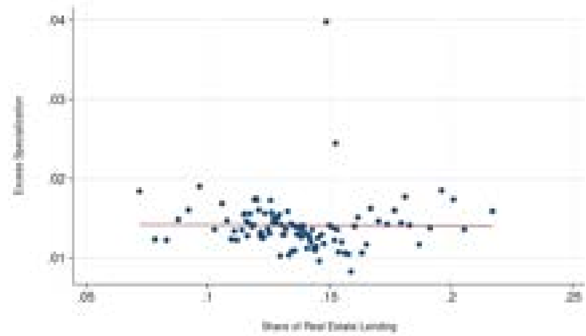
(d) Portfolio Share in ZIP vs. Industry

Note: This Figure relates our main variable of interest "Excess" specialization (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}} - \frac{LoanAmount_{s,t}}{\sum_s LoanAmount_{s,t}}$) to other forms of specialization in binned scatters, where each bin represents at least 5 observations. In Panel (a) we relate "excess" specialization to "relative" specialization (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}} / \frac{LoanAmount_{s,t}}{\sum_s LoanAmount_{s,t}}$). In Panel (b) we relate it to the share of a bank's portfolio in a single industry (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}}$) . In Panel (c) we compare "excess" specialization in a given period to the "excess" specialization that we would observe if we calculated it based solely on newly originated loans. In panel (d) we relate specialization to a bank's concentration in a single ZIP code.

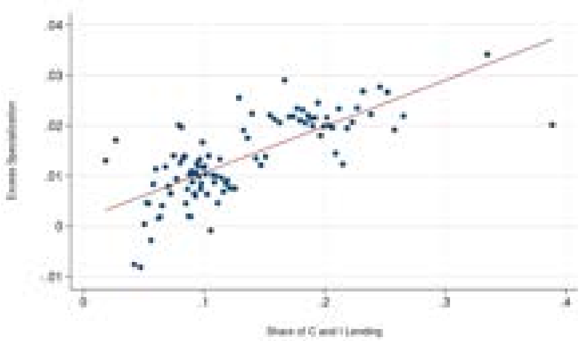
Figure A.2: ***Specialization and Bank Characteristics (I)***



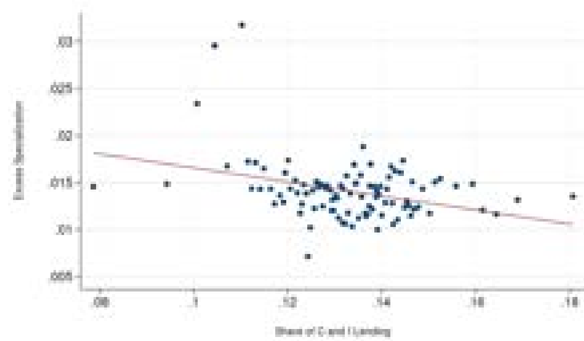
(a) Specialization vs. Share of Assets in R.E. – w/in Industry



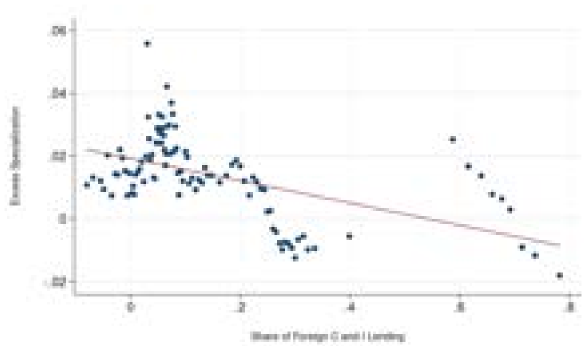
(b) Specialization vs. Share of Assets in R.E. – w/in Bank



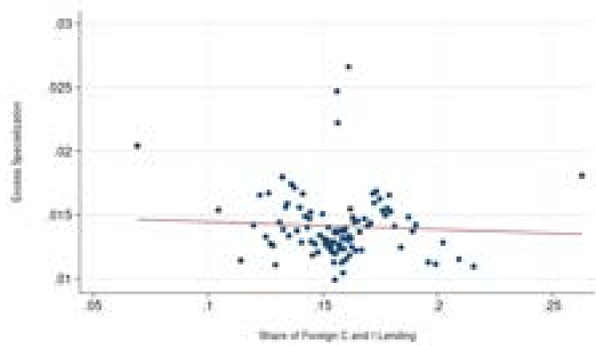
(c) Specialization vs. Share of Assets in C and I – w/in Industry



(d) Specialization vs. Share of Assets in C and I – w/in Bank



(e) Specialization vs. Share of Foreign – w/in Industry



(f) Specialization vs. Share of Foreign – w/in Bank

Note: This Figure relates a bank's "Excess" specialization (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}} - \frac{LoanAmount_{s,t}}{\sum_s LoanAmount_{s,t}}$) in its favorite industry to bank characteristics in binned scatters, where each bin represents at least 5 observations. In Panel (a) we relate specialization to the bank's share of real estate lending. We account for industry. In panel (b) we again relate specialization to real estate lending and account for banks. We follow the same pattern for all subsequent figures. In panels (c) and (d) we relate specialization to the share of C&I lending. Panels (e) and (f) relate specialization to the share of C&I lending that is foreign to the US.

In Figure A.2 we relate specialization to bank characteristics. In Panel (a) we relate our measure of excess specialization to the share of a bank's portfolio that is invested in real estate, absorbing industry fixed effects. We can see that the correlation between a bank's focus on real estate and specializations is almost zero. The same holds when we account for bank fixed effects in panel (b). An increase or decrease in a bank's focus on real estate – compared to its historical averages – does not change the degree to which it specializes in a single C&I industry. We next relate specialization to the share of a bank's portfolio invested in C&I lending. The greater the share of C&I lending, the greater a bank's specialization in its favorite industry. Banks focused on more C&I lending are liable to be more specialized. This pattern breaks down when we look within bank (panel (d)). The correlation becomes negative and insignificant, implying that banks that direct more of their assets into C&I, do not necessarily direct these into their most preferred industry of specialization.

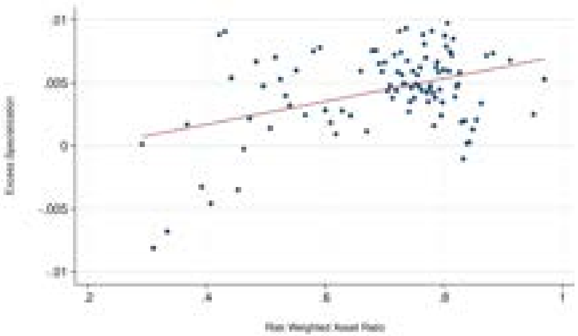
Finally, we calculate the share of C&I lending directed towards foreign entities as a share of total C&I lending. We find that it is slightly negatively correlated with favorite industry specialization. This may be because a foreign focus is its own form of specialization.

In Figure A.3 we relate a bank's excess specialization in its favorite industry to regulatory capital ratios. We account for bank fixed effects in each scatter plot. We first show there exists a negative relationship between the share of a bank's risk-weighted assets/total assets and the bank's specialization. Secondly, in panel (b) we find that banks with a greater than usual share of tier 1 capital to risk-weighted assets are less likely to focus on their industry of specialization. Similarly, banks with a higher (lower) share of Tier 1 capital relative to total assets are less (more) likely to focus on their industry of specialization.

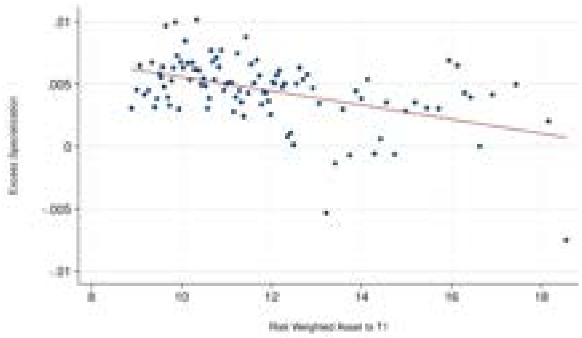
In Table A.3, we corroborate these results at the bank*industry*time level. We use all industries in columns (1)-(2) and a bank's favorite industry in columns (3)-(4). We relate a bank's excess specialization to the three regulatory capital ratios used above. When using all industries, we relate capital to the average degree to which a bank is specialized. When focusing on the top industry, we measure a bank's propensity to focus on its most favored industry when regulatory capital changes. Ultimately, we find that banks with greater regulatory capital are more likely to branch into other industries while banks with lower regulatory capital may re-focus on their industry of expertise.

Next, we test whether our observed measure of specialization can be the result of random lending patterns (i.e. the result of some banks randomly holding some large, long-maturity loans in certain industries). We could, after all, be measuring specialization which is the mechanical result of a bank's loan portfolio and loan maturity structure. To this end, we

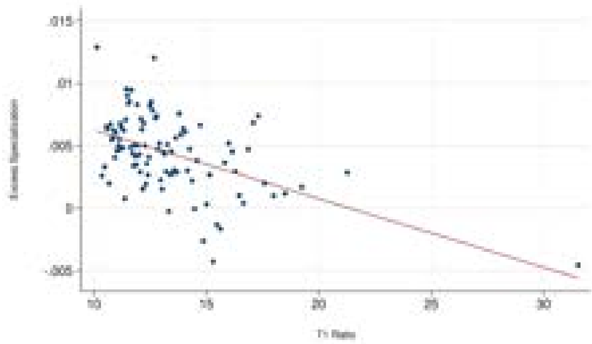
Figure A.3: ***Specialization and Bank Characteristics (II)***



(a) Share of Risk Weighted Assets



(b) Tier 1 to Risk-Weighted Assets



(c) Share of Tier 1 Capital

Note: This Figure relates a bank's "Excess" specialization (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}} - \frac{LoanAmount_{s,t}}{\sum_s LoanAmount_{s,t}}$) in its favorite industry to bank characteristics in binned scatters, where each bin represents at least 5 observations. In Panel (a) we relate specialization to the bank's share of risk-weighted assets. In panel (b) we relate specialization to the ratio of Tier 1 capital to risk-weighted assets and in panel (c) we use the share of tier 1 capital. We account for bank fixed effects in each Panel.

measure "specialization" in banks with randomly assigned loans from our sample. Specifically, we use two "simulation" techniques. In the first methodology, we randomly assign a newly originated loan to a bank in our sample. The bank then holds/renegotiates the loan in the same way that the original loan owner did. We limit a bank to holding the same number of loans they do in our original sample. This leads to some banks being larger and some smaller than in reality, as we do not place limits on loan size. In our second methodology, we again randomly assign loans at the first observation. However, we limit a bank to its approximate "true" size and assign loans randomly to banks that have lending capacity. This means some banks make fewer and some banks more loans than in our original sample. In neither case do we observe similarities with our true measures of specialization.

Figure A.4 shows some key graphical insights. Panel (a) depicts our traditional special-

Table A.3: ***Bank Characteristics and Specialization***

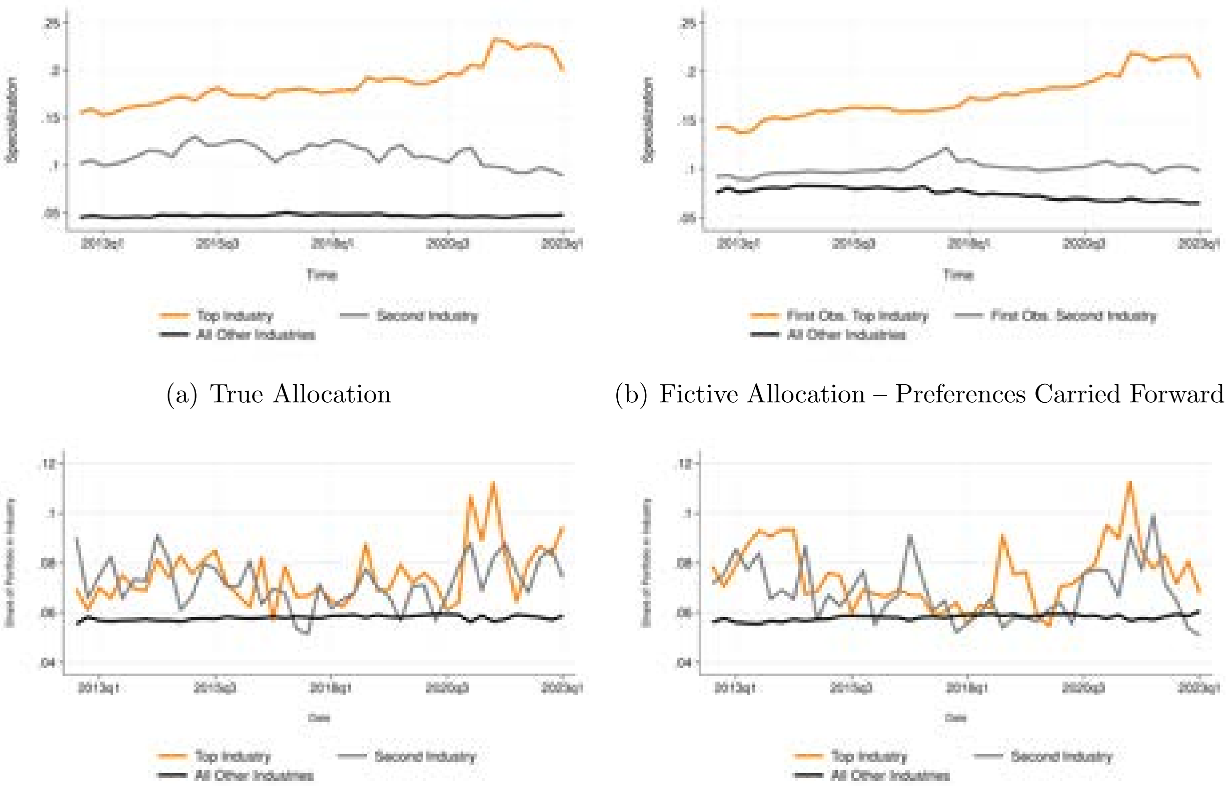
	(1)	(2)	(3)	(4)
	"Excess" Specialization			
Panel A				
Share of Risk-Weighted-Assets$_{t-1}$	0.009***	0.005	-0.016	0.058**
	[0.002]	[0.006]	[0.011]	[0.026]
Panel B				
Share of Tier 1 to Risk-Weighted-Assets$_{t-1}$	-0.047***	-0.064***	-0.037	-0.472***
	[0.008]	[0.022]	[0.047]	[0.079]
Panel C				
Share of Tier 1 Capital$_{t-1}$	-0.065***	-0.054**	-0.084**	-0.383***
	[0.008]	[0.025]	[0.043]	[0.089]
Fixed Effects	Industry	Bank	Industry	Bank
Controls	Bank Size, Leverage, and Time			
Sample	Full		Top Industry	
R^2	0.055	0.04	0.37	0.73
N	23,886	24,196	1,140	1,141

Note: This table makes use of our data at the bank*industry*time level. It relates a bank's excess specialization to relative capital measures such as the share of risk weighted assets in panel (a) the share of Tier 1 capital to risk weighted assets in panel (b) and the share of Tier 1 capital in Panel (c). We use all industries (columns (1)-(2)) and a bank's favorite industry (columns (3)-(4)). We include controls for bank size, leverage, and time. We alternating include industry or bank fixed effects. Standard errors are clustered at the firm-year level while *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

ization measure for reference. Panel (b) depicts a fictive measure of specialization, which would be obtained by keeping a bank's industry preference ranking fixed. We establish the ranking of industries when the lender first appears in our sample and calculate the portfolio shares in each group (top Industry, second favorite industry, all other industries) with actual lending in any given period. We can see that banks can change their top industry, though this is either temporary or a small reshuffling of preferences that does not change the patterns we observe in (a).

Panels (c) and (d) use of simulated data. A bank's preferred industry is the two-digit NAICS sector in which it happens to be relatively over-invested in a given period (as discussed above). In no instance do we observe the patterns we observed in our original data, i.e. that the bank's "preferred" industries consistently sees greater over-investment. In fact, random loan assignment appears to lead to very sporadic and inconsistent "specialization". As such, we conclude that, given the size of stress-tested banks, it is unlikely that our measure is the result of random chance.

Figure A.4: ***Specialization: Real vs Fictive***



(a) True Allocation

(b) Fictive Allocation – Preferences Carried Forward

(c) Random Allocation 1

(d) Random Allocation 2

Note: This Figure depicts the share of a bank's portfolio in its favorite, second favorite, and all other industries. Panel (a) replicates our original measure using the actual data, discussed above. Panel (b) fixes a bank's rankings of favorite, second favorite, and all other industries at first observation and carries the rankings forward. Portfolios are calculated with real loan data at any given time based on these historical industry rankings. Panels (c) and (d) assign loans to banks at random (based on bank size in Panel (d)) and calculate favorite, second favorite, and all other industries based on the degree of over-investment by that bank in that period.

We next analyze the degree to which specialization is consistent across time. We can distinguish two types of lenders: (i) those with very consistent industry rankings and (ii) those with less consistent industry rankings. In panel A we depict a masked industry identifier and show the number of banks from our sample that consider this industry their favorite – based on the degree of over-investment – in 2018q4. Naturally, there might be a degree of variance in which industries are considered a bank's favorite. In fact, 55% of banks in our sample keep the same favorite industry for over 75% of their time in the sample. As such, we show examples of banks with very consistent and examples of banks with somewhat inconsistent specialization. Panel B. of Table A.4, shows three examples of each type of lender. Specifically, the table shows a bank's favorite industry for the quarters between 2016 Q1 and 2018 Q4. We depict the industry of specialization and the degree of relative over-investment in brackets. We can see that even banks with lower consistency in their rankings are liable to keep a favorite industry for several periods. Moreover, they are likely to bounce between two or three preferred industries during their time in the sample. Banks with lower consistency in specialization tend to be slightly less specialized.

In Panel C we include a bank's favorite industry measured by held loans and by new loans in that period. We show two types of banks: Those with high overlap in their specialization in existing and new loans and those with less pronounced overlap. We can see some degree of overlap between existing and new specialization among all types of banks. Nevertheless, some banks with greater ex-ante specialization have a more singular focus.

Table A.4: *Consistency of Specialization*

Panel A: Number of Banks Specialized per Industry (2018q4)

Industry Code	1	2	3	4	5	5	7	8	9	10	11	12
Number of Specialized Banks	0	2	2	1	0	1	0	1	5	0	1	1
Industry Code (contd.)	13	15	16	17	18	19	20	21	22	23	24	
Number of Specialized Banks	2	5	0	1	0	1	6	0	2	3	3	

Panel B: Specialization in Top Industry – Examples

		2016q1	2016q2	2016q3	2016q4	2017q1	2017q2	2017q3	2017q4	2018q1	2018q2	2018q3	2018q4
Consistent Specialization	Example 1	9	9	9	9	9	9	9	9	9	9	9	9
		(5.1)	(5.1)	(5.1)	(5.1)	(5.1)	(5.3)	(5.2)	(5.2)	(5.1)	(5.1)	(5.1)	(5.1)
	Example 2	23	23	23	23	23	23	23	23	23	23	12	23
		(4.1)	(5.1)	(5.1)	(5.1)	(5.1)	(5.3)	(4.0)	(4.0)	(3.8)	(3.7)	(4.5)	(4.3)
	Example 3	22	22	22	22	22	22	22	22	22	22	22	22
		(5.1)	(5.1)	(5.1)	(5.1)	(5.1)	(5.3)	(5.2)	(5.2)	(5.1)	(4.9)	(4.8)	(4.7)
Inconsistent Specialization	Example 1	8	8	8	20	20	20	20	20	20	12	12	12
		(3.4)	(3.5)	(3.5)	(3.9)	(3.9)	(3.8)	(3.7)	(4.0)	(4.4)	(5.0)	(5.1)	(4.9)
	Example 2	17	17	1	1	17	17	17	17	17	17	15	15
		(2.2)	(1.8)	(1.9)	(2.2)	(1.9)	(1.9)	(1.8)	(1.7)	(1.9)	(1.9)	(1.6)	(1.5)
	Example 3	12	20	20	20	9	9	9	9	9	9	9	9
		(3.9)	(3.8)	(3.1)	(3.2)	(3.1)	(3.3)	(3.5)	(3.7)	(3.8)	(3.8)	(3.9)	(3.8)

Panel C: Existing vs. New Loan Specialization

			2016q1	2016q2	2016q3	2016q4	2017q1	2017q2	2017q3	2017q4	2018q1	2018q2	2018q3	2018q4
Consistent New and Existing Specialization	Example 1	Existing Specialization	9	9	9	9	9	9	9	9	9	9	9	9
			(5.1)	(5.1)	(5.1)	(5.1)	(5.1)	(5.3)	(5.2)	(5.2)	(5.1)	(5.1)	(5.1)	(5.1)
		Specialization New Loans	9	9	9	9	1	9	9	9	1	9	9	9
			(5.1)	(5.1)	(5.1)	(5.1)	(5.1)	(5.3)	(5.2)	(5.2)	(5.1)	(5.1)	(5.1)	(5.1)
	Example 2	Existing Specialization	24	24	24	24	24	24	24	17	17	17	17	17
			(5.1)	(5.1)	(5.1)	(5.1)	(5.1)	(5.3)	(5.2)	(5.2)	(5.1)	(5.1)	(5.1)	(5.1)
		Specialization New Loans	24	24	24	24	24	24	24	17	17	17	17	17
			(5.1)	(5.1)	(5.1)	(5.1)	(5.1)	(5.3)	(5.2)	(5.2)	(5.1)	(5.1)	(5.1)	(5.1)
Different New and Existing Specialization	Example 1	Existing Specialization	23	23	23	23	23	23	23	23	23	23	12	23
			(4.1)	(5.1)	(5.1)	(5.1)	(5.1)	(5.3)	(4.0)	(4.0)	(3.8)	(3.7)	(4.5)	(4.3)
		Specialization New Loans	19	23	23	23	23	23	21	7	21	17	23	17
			(4.1)	(5.1)	(5.1)	(5.1)	(5.1)	(5.3)	(4.0)	(4.0)	(3.8)	(3.7)	(4.5)	(4.3)
	Example 2	Existing Specialization	22	22	22	22	22	22	22	22	22	22	22	22
			(5.1)	(5.1)	(5.1)	(5.1)	(5.1)	(5.3)	(5.2)	(5.2)	(5.1)	(5.1)	(5.1)	(5.1)
		Specialization New Loans	22	22	22	21	22	7	12	2	21	22	21	22
			(5.1)	(5.1)	(5.1)	(1.4)	(5.1)	(1.8)	(1.1)	(1.2)	(2.8)	(4.9)	(2.5)	(4.7)

Note: This table depicts examples of banks and the industries in which they specialize. Panel A. shows industry codes (masked) and the number of banks that consider that industry their favorite in 2018q4. In Panel B we show three examples of banks with high consistency in terms of their industry preference and three with lower consistency. We depict the industry the bank considers its favorite in a given quarter, based on which industry a bank is most over-invested at that point in time. The relative degree of over-investment (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}} / \frac{LoanAmount_{s,t}}{\sum_s LoanAmount_{s,t}}$) is recorded in brackets. Panel Panel C makes use of specialization calculated in terms of newly originated loans as well as all outstanding loans (our preferred definition). We show two examples of banks with higher correlations between new and existing loans and two banks with lower correlations.

B Borrower Size and Specialization

In this section, we first relate a bank's specialization to some of its borrowers' characteristics. Second, we focus on a specialized lender's propensity to lend to small firms.

In Panel (a) of Figure A.5, we relate a bank's excess specialization in an industry to the average size of its borrowers in that industry. Specialization is negatively related to borrower size, as measured by assets. It should be noted that only around 60% of borrowers report assets, so this estimate may be biased. Given that smaller firms are less likely to report data, the bias is likely to run counter to our findings. Overall, this implies that the average borrower from a specialized bank is likely to be smaller than the average borrower from a less specialized bank. We can see from Panel (b) that a borrower from a more specialized bank is more likely to be an SME or small firm (less than 25 mil. USD in assets).

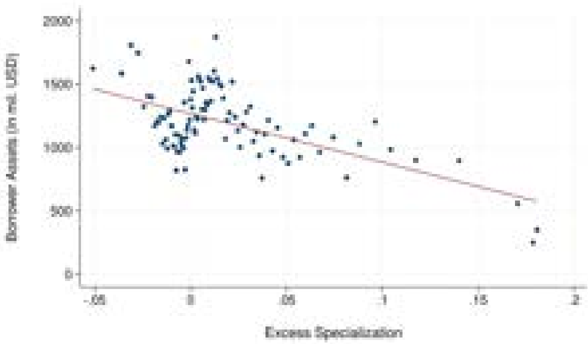
Although size does not necessarily imply lower profitability, we can see from Panel (c) that more specialized banks lend to less profitable firms (as measured by earnings to assets). This may imply that specialized lenders are more willing to take risks on growth firms that are not yet profitable or on firms suffering temporary reductions in profitability. The lender's expertise in an industry may help a borrower navigate possible issues around low profitability.

In Panel (d) of Figure A.5 we can see that lower profitability and smaller size do not necessarily impact firm risk ratings. In our data, firms with lower ratings are considered more creditworthy. More specialized lending is weakly associated with slightly better borrower ratings. However, it should be noted that this relationship is noisy. Unlike the other relationships, discussed above, it does not hold once we account for industry controls/fixed effects.

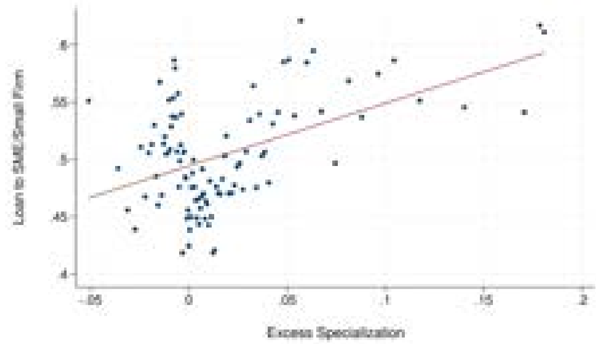
We next focus on lending to small firms (firms with less than 2 mil. USD in assets). Overall, the propensity of large banks in our sample to lend to small firms has decreased over time. This holds for a bank's most specialized (or favorite) industry as well as all others. However, as can be seen in Figure A.6, the propensity to lend to small firms has remained higher in a bank's favorite industry throughout our sample period, with only a small disruption during COVID.

Empirically, we can see that this effect is more pronounced when we account for borrower and loan characteristics. In Table A.5 we see that banks are generally more likely to lend to small firms. We can define small by asset size, EBIT, or the loan's size (between 1 and 2 mil USD). Under each definition, a small borrower is more likely to borrow from a specialized lender.

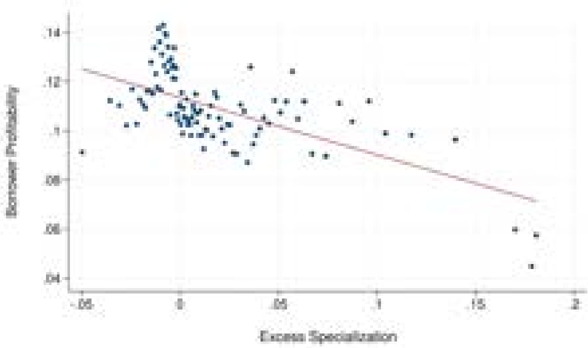
Figure A.5: *Specialization and Borrower Characteristics*



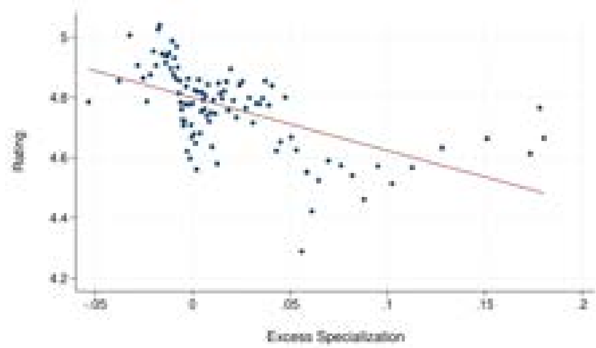
(a) Borrower Size vs. Excess Specialization



(b) Likelihood of Loans to SME and Small Firms vs. Excess Specialization



(c) Average Borrower Rating vs. Excess Specialization



(d) Borrower Earnings/Assets vs. Excess Specialization

Note: This Figure shows the relationship between borrower characteristics and a lender's excess specialization in binned scatters, where each bin represents at least 5 observations. We account for time and loan purpose fixed effects in all panels. Panel (a) depicts the relationship between lender specialization and borrower assets (for the 60% of borrowers who report assets). Panel (b) shows the propensity of an SME or small firm (below 25 mil. USD in assets) to borrow from a bank, given its level of specialization in the borrower's industry. Panel (c) relates borrower earnings (EBITDA)/assets (i.e. its profitability) to the lender's specialization. Panel (d) relates average firm ratings – as assigned by the bank's loan officers – to the bank's excess specialization.

Figure A.6: ***Small Firm Lending – Share of Portfolio***

Note: This figure depicts the share of a bank's lending that is directed towards small firms (less than 2 mil. USD in assets). We split the sample by a bank's favorite industry and all its other industries.

Table A.5: ***Small Firm Lending***

	(1)	(2)	(3)	(4)	(5)	(6)
	Small Loan		Small Firm (Asets)		Small Firm (EBIT)	
"Excess" Specialization	0.046*	0.079*	-0.276	0.589***	-0.010	0.472***
	[0.026]	[0.044]	[0.235]	[0.135]	[0.042]	[0.111]
Fixed Effects	Time, Rating, Purpose					
Bank*Time FE	Yes	No	Yes	No	Yes	No
Industry*Time FE	Yes	No	Yes	No	Yes	No
Frim FE	No	Yes	No	Yes	No	Yes
Controls	Loan Terms and Bank Characteristics					
R^2	0.49	0.81	0.39	0.89	0.32	0.87
N	296,951	296,951	174,108	174,108	151,895	151,895

Note: This table relates a dummy for whther a loan (Columns (1)-(2)) or a firm (Columns (3)-(6)) is small. We define small as 1-2mil. USD in loan size, 2 mil. USD or less in firm assets, or 2 mil USD or less in reported EBIT. We include loan controls such as size, rate, whether it is secured, and Purpose as well as borrower controls such as rating. We include either bank*quarter and industry*quarter or borrower fixed effects. Specialization is defined as the degree to which a bank is over-invested in an industry, relative to a perfectly diversified portfolio. A diversified portfolio is one based solely on the size of an industry relative to all C&I lending. We use "excess" (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}} - \frac{LoanAmount_{s,t}}{\sum_s LoanAmount_{s,t}}$) in Columns (3) and (4)) specialization at the two-digit industry level. Standard errors are clustered at the bank-industry-year level while *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

C Specialization calculated using SNC or COMPUSTAT Data

The fact that our measure of specialization is calculated using only Y14Q data raises two possible issues. Firstly, we cannot see smaller banks in the data. Our inferences are based on large banks. While smaller banks may also specialize, they may be driven to do so by size constraints as opposed to industry choices. Our hypotheses state that these smaller "specialized" banks should see fewer benefits from specialization. Testing this is an important part of our approach. Secondly, our definition of a borrower's industry is based on data reported by the bank. It may be inaccurate or too narrow for multi-industry borrowers.

To account for the above issues, we use two datasets: the Shared National Credit (SNC) and COMPUSTAT. SNC data is maintained by the Board of Governors of the Federal Reserve System, the Federal Deposit Insurance Corporation, and the Office of the Comptroller of the Currency. The dataset includes information for all syndicated loans with a minimum aggregate commitment of USD 20 million before 2018, where we end our data.[23] The loans must further be held by at least three supervised entities. The data and its benefits and shortcomings are discussed in detail in Blickle et al. (2020).

Critically, the SNC data includes information on entities that are much smaller than those stress tested using Y14Q data. It also contains information on the performance of loans. As such, we can calculate the degree of specialization for smaller lenders and relate this to loan performance. Unfortunately, the data does not include as much detail on the borrower, it contains no information on interest rates paid by borrowers, and it contains no information on loan-officer loan ratings for around 50% of loans.

We split our data into 5 bins based on the total size of a bank's SNC lending at the high-holder level. We assign the largest category to banks that fall under the Large Institution Supervision Coordinating Committee (LISCC) and quartile the rest of the bank lenders in our sample.[24] We exclude funds from this analysis, as they operate under a different business model.

As one can see from Figure A.7, small, medium, and large LISCC banks (we exclude medium-small and medium-large banks for ease of viewing) all have industries in which they predominantly specialize. In fact, they exhibit similar lending patterns – in terms of favoring one or two industries – to the banks in our Y14 data. Even though we use slightly different industry definitions, we find highly similar specialization rates. Smaller banks are more

[23]The threshold was raised to USD 100 million on January 1, 2018.

[24]Please see https://www.federalreserve.gov/supervisionreg/large-institution-supervision.htm for more information.

likely to be specialized. This may be a result of their smaller size. Being size-constrained, individual loans may make up a larger share of their portfolio leading to higher specialization values.

We next test whether this specialization relates to loan performance. We define non-performing loans similarly to the main text (including default, covenant breach, and amounts past due in the definition). We focus on the first observation of the loan and see whether the loan becomes non-performing at any point in which it appears in the sample. In Table A.6 one can see that, unconditionally, specialization is positively related to loans becoming non-performing. This holds no matter the controls we include (see column (2)). However, once we include interactions with lender size, the relationship inverts for larger banks. This perfectly matches our expectations. Large banks are specialized out of choice, as opposed to constraints. Their choice to focus on an industry correlates with better information. This, in turn, allows them to better perform their roles of reducing adverse selection and improving monitoring.

In columns (4) we include bank fixed effects. This specification is most closely related to the specifications in our main analysis. The results are confirmed for this specification. Thus far, regressions have made use of all banks in our sample. It is possible, however, that syndicate participants are less informed than syndicate arrangers. We thus re-run our regression of interest, using a sample of only lead arrangers. These lenders should have the most information on a borrower. We indeed find that our results hold. Coefficients are marginally larger for this sample. However, the effect is not solely driven by arrangers; specialized participants still appear to have some skill in selecting ideal loans in which to invest.

In Table A.7 we use Y14 data. Here we can include loan rating and interest rate controls that account for riskiness. We again focus on whether banks are more inclined to use their specialized knowledge if they are the syndication agent. This follows from the fact that the syndication agent is tasked with loan selection and monitoring. A specialized bank may be better able to conduct both. The analysis follows a similar pattern to the regression using SNC loans, discussed above. If we focus on all loans (column (1)), we do not see a difference between syndication agents and all other borrowers. Small loans are typically originated by a single bank that is liable to use its specialized knowledge to pick and monitor the best loan. As such, there is no difference between the effort expended for small loans or for large loans where a bank is the syndication agent.

If we focus only on syndicated loans (column (3)), however, we do notice that the specialization of the agent has an impact. It does not explain the entire effect that specialization has on loan performance, as specialized lenders can choose to invest in good loans only. Nev-

Table A.6: ***Loan Performance and Lender Size – SNC Data***

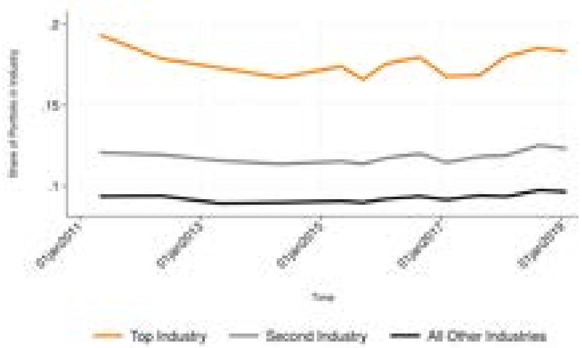
	(1)	(2)	(3)	(4)	(5)	(6)
	Indicator: Loan ever becomes non-performing					
Excess Specialization	0.098***	0.012***	0.066***	0.022***	-0.032	0.062*
	[0.002]	[0.002]	[0.002]	[0.003]	[0.022]	[0.035]
Medium Small × Excess Spec.			-0.011*	0.031***	-0.036	-0.002
			[0.006]	[0.006]	[0.022]	[0.055]
Medium × Excess Spec.			-0.173***	-0.090***	-0.289***	0.018
			[0.008]	[0.009]	[0.029]	[0.061]
Medium Large × Excess Spec.			-0.225***	-0.130***	-0.391***	-0.295***
			[0.010]	[0.010]	[0.039]	[0.066]
Large ×Excess Spec.			-0.223***	-0.160***	-0.378***	-0.207***
			[0.013]	[0.013]	[0.046]	[0.056]
Industry*Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Bank*Time FE	No	Yes	No	Yes	Yes	Yes
Controls	Loan and lender size					
Lender Sample	All					Arranger
Time Sample	1995-2018				2006-2009	1995-2018
R^2	0.2	0.23	0.2	0.23	0.24	0.14
N	2131559	2126159	2131559	2126159	274666	103798

Note: This table relates loan performance to excess specialization of banks in the borrower's industry. Specifically, we estimate:

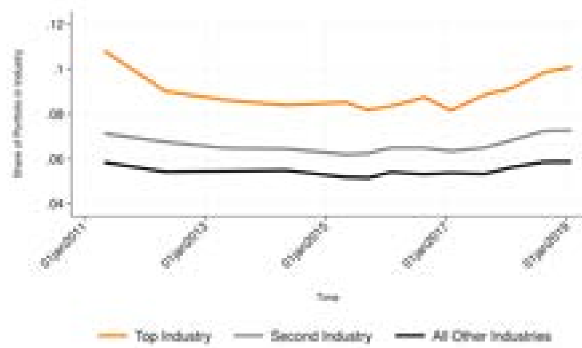
$$NonPerf_{l,i,b,s,T} = \beta_0 + \beta_l Specialization_{b,s,t} + \beta_2 Specialization_{b,s,t} * BankSize_{b,s,t} + \beta_3 \mathbf{X}_{l,b} + \xi_{b,t} + \sigma_{s,t} + \epsilon_{l,i,b,s,t,z}$$

We regress whether loan l to firm i in quarter t by bank b in sector/industry s ever becomes non-performing in future periods on the lending bank b's specialization in industry s. "Non performing" is a dummy that takes the value of 1 if the loan falls in arrears, breaches covenants, or is otherwise in default at any point in the future. We restrict our sample to when a loan is first observed. Specialization is defined as the degree to which a bank is over-invested in an industry, relative to a perfectly diversified portfolio. A diversified portfolio is one based solely on the size of an industry relative to all C&I lending. We use excess (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}} - \frac{LoanAmount_{s,t}}{\sum_s LoanAmount_{s,t}}$) specialization at the two-digit industry level. Columns (2) and (4) include bank-time fixed effects. Columns (3)-(6) interact our variable of interest with a dummy for a bank's size-category. Column (5) makes use of the period around the great recession. Column (6) makes use of only lead arrangers (i.e. the entity that SNC considers the reporting and originating agent). Data is collapsed to the bank high-holder level. Standard errors are clustered at the bank*time level while *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

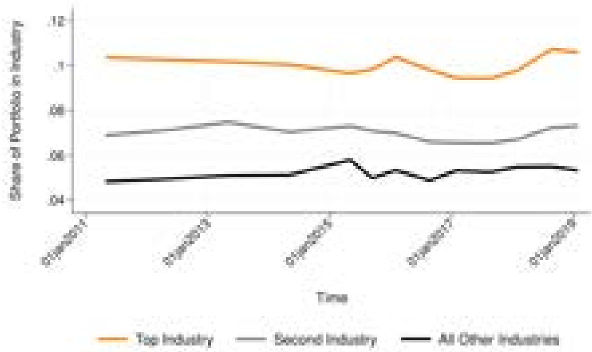
Figure A.7: ***Specialization in SNC***



(a) Small lenders



(b) Midsize lenders



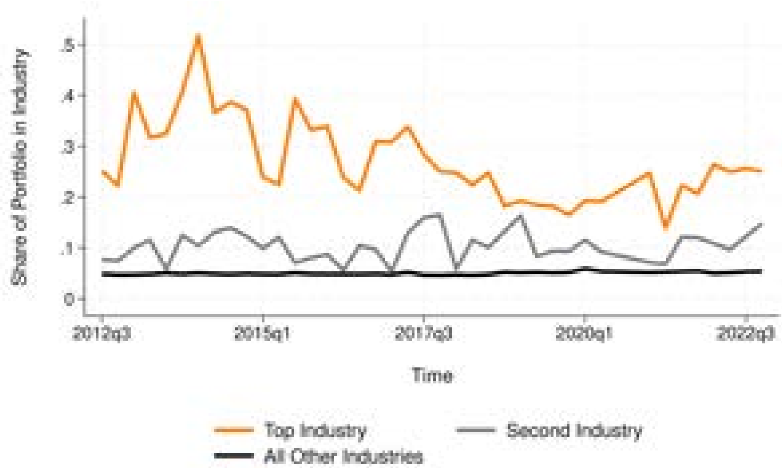
(c) LISCC lenders

Note: This Figure shows the share of loans banks in the SNC data have in single two digit industries (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}}$. We split our sample of banks into 5. The LISCC-banks are the largest and are their own category. The rest of the banks are quartiled based on their total SNC lending. We group all banks by their ultimate high-holder in a given period. We include only small, medium and LISCC banks for convenience (excluding medium-small and medium-large banks, who show very similar patterns).

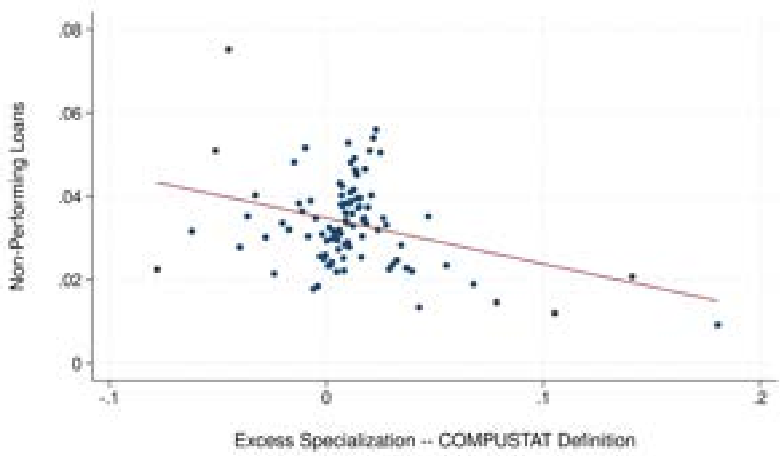
ertheless, specialized syndication agents are better able to perform their duties and select better loans. It is worth noting that the effect is still largest amongst small loans that are not sold or specialized, whereby the ender is most likely to use their specialized knowledge when they are the primary beneficiary.

COMPUSTAT Industries To address our second concern, that the industries of borrowers may be poorly recorded in Y14 data, we merge our primary data with COMPUSTAT information. COMPUSTAT records key characteristics of firms, including balance sheet characteristics, income statement details, and – importantly – information on a firm's primary industry. In this case we use two-digit SIC codes (which are distinct from NAICS codes) from COMPUSTAT to identify borrower industries. Unfortunately, we can only merge in

Figure A.8: ***Specialization using COMPUSTAT data***



(a) Specialization – COMPUSTAT Data



(b) Loan Performance – COMPUSTAT Data

Note: Panel (a) of this Figure shows the share of loans banks have in single two digit industries (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}}$. We calculate shares in an industry using the 78 two-digit sic codes. our data consists of merged Y14 and COMPUSTAT information. Panel (b) relates non-performance – i.e. whether the loan ever becomes non-performing during its time in the sample – to a bank's specialization in the industry of the borrower at the time the loan is first observed. We use in binned scatters, where each bin represents at least 5 observations.

Table A.7: ***Interactions with Syndication Agent***

	(1)	(2)	(3)
	Loan becomes non-performing		
"Excess" Specialization	-0.004***	-0.005***	-0.003***
	[0.001]	[0.001]	[0.001]
"Excess" Specialization *Agent	-0.001	0.002	-0.005**
	[0.002]	[0.002]	[0.002]
Agent	0.000	-0.006	0.007
	[0.004]	[0.004]	[0.005]
Fixed Effects	Bank*Time and Industry * Time, Loan Purpose, Loan Rating		
Controls	Collateral, Maturity, Rate, Relationship		
Sample	All Loans	Large Loans	Syndicated Loans
Mean of dependent variable	0.05	0.036	0.05
R^2	0.14	0.13	0.18
N	296,951	174,717	146,285

Note: This table shows the coefficients of interest for equation:

$$Y_{l,i,b,s,T} = \beta_0 + \beta_l Specialization_{b,s,t} + \beta_2 Specialization_{b,s,t} * SyndAgent$$
$$+\beta_3 \mathbf{X}_{l,b} + \beta_4 Relationship_{i,b} + \xi_{b,t} + \sigma_{s,t} + \phi_{loanriskrating} + \omega_{loanpurpose} + \epsilon_{l,i,b,s,t}$$

It regresses whether loan l to firm i in quarter t by bank b in sector/industry s ever becomes non-performing in future periods on the lending bank b's specialization in industry s. "Non-performing" is a dummy that takes the value of 1 if the loan falls in arrears or is otherwise in default. Specialization is defined as the degree to which a bank is over-invested in an industry, relative to a perfectly diversified portfolio. A diversified portfolio is one based solely on the size of an industry relative to all C&I lending. We use excess (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}} - \frac{LoanAmount_{s,t}}{\sum_s LoanAmount_{s,t}}$) specialization at the two-digit industry level. Syndication agent is a binary variable, it denotes the status of the bank in Y14Q, if it was identified as the lead arranger in a syndicated loan. All columns use bank*time and industry*time fixed effects. Other fixed effects and controls include loan purpose and type, bank size industry capture, loan size, whether a loan is secured by collateral, relationship to the borrower, and the interest rate. Relationship measures include past and future bank-borrower interactions as well as the specialization of bank b in borrower i's state. Our sample includes only term loans and focuses on the period a loan is first observed. Standard errors are clustered at the bank-industry-year level while *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

COMPUSTAT data along publicly identifiable characteristics – such as the ticker. We can merge in data for only around 9% of our total sample by loan*time. Companies for which we merge in data are larger than our average firm in the Y14. Nevertheless, we can compute a bank's specialization in industries based on COMPUSTAT data.

In Figure A.8, Panel (a), we show that we still observe similar patterns of specialization. Banks have a favorite, and some have a second favorite industry to which they devote a greater share of lending. Given that we have less data and loans to these large borrowers are inherently larger, our data shows more variability over time. Not nearly as much, however, as in figure A.4, discussed above. This implies that we are still observing specialization chosen by the banks.

In panel (b) we relate specialization to loan performance. We again find a negative correlation between specialization and a loan becoming non-performing. We corroborate that this relationship holds in the face of controls in Table A.8. All our measures of specialization remain negatively related to non-performance. In our most saturated regression, our coefficient of interest becomes insignificant. However, given the smaller set of borrowers, the full specification may be overly saturated with terms.

Finally, we use SNC data to relate monitoring activities to specialization. We do so in order to offer suggestive evidence that specialization is associated with monitoring effort. We relate lender specialization to whether a lender is an "active monitor". We define active monitoring as in Gustafson et al. (2021), whereby active monitoring is associated with the lead bank conducting "onsite inspections" or physical appraisals, etc. In Table A.9 we see that specialization is indeed associated with more active monitoring. Moreover, we can see from columns (4)-(6) that specialized lenders are slightly more likely to value collateral themselves. This implies that (as suggested in Gopal (2019)) bank knowledge of collateral is correlated with industry specialization. The caveats as to the granularity of SNC data, discussed above, apply. Nonetheless, we can see here indicative evidence for the fact that lender monitoring involvement correlates with specialization.

Table A.8: ***Loan Performance – COMPUSTAT Data***

	(1)	(2)	(3)
	Loan ever becomes non-performing		
Excess Specialization (COMPUSTAT)	-0.071***	-0.078***	-0.021
	[0.023]	[0.024]	[-0.036]
Interest rate		0.006***	0.007***
		[0.001]	[0.002]
Relationship		-0.000***	-0.000*
		[0.000]	[0.000]
Share in ZIP			-0.074***
			[0.012]
Bank-Time Fixed Effects	No	No	Yes
Industry-Time Fixed Effects	Yes	Yes	Yes
Rating and Purpose Fixed Effects	No	Yes	Yes
Controls	No	Loan and Bank	Loan and Bank
Mean of dependent variable	0.026	0.026	0.029
R^2	0.092	0.1	0.19
N	22,678	22,678	16,442

Note: This table relates loan performance to excess specialization of banks, calculated using COMPUSTAT SIC industry codes. Specifically, we estimate:

$$NonPerf_{l,i,b,s,T} = \beta_0 + \beta_l SpecializationCOMPUSTAT_{b,s,t} \\ +\beta_2 \mathbf{X}_{l,b} + \xi_{b,t} + \sigma_{s,t} + \epsilon_{l,i,b,s,t,z}$$

We regress whether loan l to firm i in quarter t by bank b in sector/industry s ever becomes non-performing in future periods on the lending bank b's specialization in industry s. "Non performing" is a dummy that takes the value of 1 if the loan falls in arrears, is past due, or is otherwise in default at any point in the future. We restrict our sample to when a loan is first observed. Specialization is defined as the degree to which a bank is over-invested in an industry, relative to a perfectly diversified portfolio. A diversified portfolio is one based solely on the size of an industry relative to all C&I lending. We use "excess" (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}} - \frac{LoanAmount_{s,t}}{\sum_s LoanAmount_{s,t}}$) specialization at the two-digit industry level. We add controls and fixed effects in each column. Data is collapsed to the bank high-holder level. Standard errors are clustered at the bank*industry*time level while *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table A.9: ***Monitoring Activity***

	(1)	(2)	(3)	(4)	(5)	(6)
		Active Monitoring			Bank Valued	
Specialization (share)	0.091***	0.025***	0.030***	0.037***	0.028***	0.003
	[0.002]	[0.002]	[0.003]	[0.002]	[0.002]	[0.002]
Date FE	X	X		X	X	
Agent FE		X			X	
Participant FE			X			X
Time*Industry FE			X			X
R^2	0.01	0.05	0.30	0.01	0.03	0.27
N	1,811,583	1,811,583	1,807,810	1,811,583	1,811,583	1,807,810

Note: This table relates monitoring activity, defined as in Gustafson et al. (2021) − using SNC data − , to lender specialization in the industry. In columns (4)-(6) we relate specialization to whether the bank assessed and valued the collateral posted for the loan itself. We include fixed effects as indicated. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

D Geographic Deposit Concentration

In this section, we use data on the geographic distribution of deposits and relate it to lending and loan performance. First, we use data on bank branches and calculate the share of deposits a bank has collected in a given zip code, using bank branch locations. In Figure A.9 one can see the distribution of deposits in zip codes for the fully US-based depository institutions in our data. In most cases, any individual zip code accounts for a very small share of a bank's deposits.

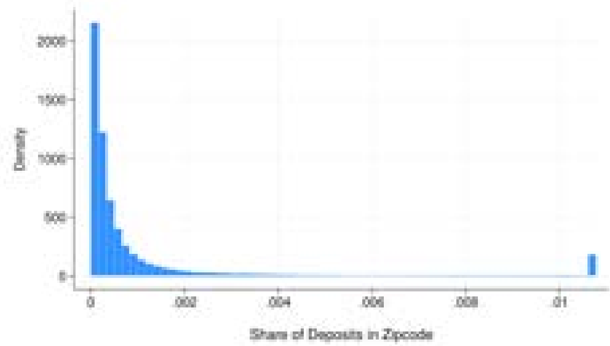
In Panel (b) we show that the relationship between the share of a bank's lending in a given ZIP code (as measured by the headquarters of the borrower) and a bank's deposit taking in a zip code are positively and significantly related. Even today, banks lend close to their branches (consider Petersen and Rajan (2002), Blickle (2020), or Nguyen (2019)). On average, zip codes in which a bank takes greater deposits are zip codes in which it will lend more. However, the relationship is not perfect. Many banks lend in zip codes in which they have no bank branch, for instance.

In Table A.10 we relate a bank's deposit concentration in a zip code at the time a loan is first observed in our data to eventual loan performance. We use our most saturated specification in each column (see column (4) of Table 3). We see that deposit concentration is positively related to loan default. This holds even as we include our key measure of "excess" specialization.

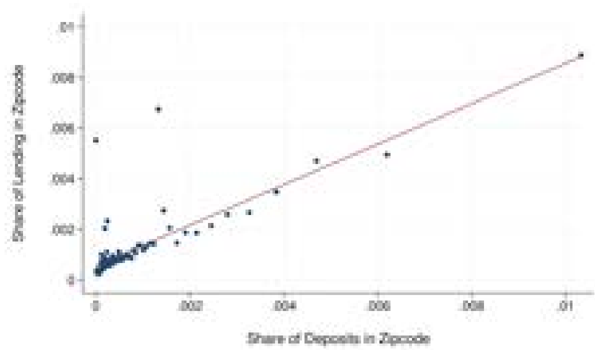
In column (3) we calculate the degree to which a bank is lending more in a zip code than it is taking in deposits. Loans in areas where the bank has a significantly higher share of lending than it takes in deposits are negatively – if weakly – related to loan non-performance. This may in itself be a reflection of the fact that the bank is making an active choice to lend in regions in which it takes fewer deposits. This active choice is likely associated with research in the region and the borrower, which lends itself to information asymmetry reductions.

In Column (4) we include both the share of lending in a zip code as well as the deposit share from that zip code. The coefficient on deposit concentration remains positive, implying that an over-reliance on a single area is negatively associated with loan performance. Overall, our results in this section indicate that our primary results, discussed above, are not affected by a bank's deposit-taking or other type of regional concentration.

Figure A.9: ***Geographic Deposit Concentration***



(a) Distribution of Deposit Concentration



(b) Deposits vs. Lending in ZIP

Note: This Figure shows the concentration of deposits. In Panel (a) we show the distribution of deposits at the zipcode level for the fully domestic banks in our sample. Panel (b) relates the share of a bank's deposits in a zipcode to the bank's lending in that same zipcode. We use in binned scatters, where each bin represents at least 5 observations.

Table A.10: ***Loan Performance – ZIP Concentration***

	(1)	(2)	(3)	(4)
	Loan ever becomes non-performing			
Deposits in ZIP	1.371***	1.337***		1.116**
	[0.461]	[0.462]		[0.458]
Excess Specialization		-0.084***	-0.083***	-0.080***
		[0.015]	[0.015]	[0.016]
Excess Lending in ZIP			-0.013*	
			[0.007]	
Share of Lending in ZIP				-0.259*
				[0.129]
Fixed Effects	Bank*Time, Industry*Time, Purpose, Rating			
Controls	Rate, Size, Maturity, Industry Capture, Collateral			
Mean of dependent variable	0.04	0.04	0.04	0.04
R^2	0.13	0.13	0.13	0.14
N	257,309	257,309	257,309	257,309

Note: This table shows the coefficients of interest for equation:

$$Y_{l,i,b,s,T,z} = \beta_0 + \beta_l Specialization_{b,s,t} + \beta_2 DepositConcentration_{b,z,t} + \beta_3 \mathbf{X}_{l,b} + \\ \beta_4 Relationship_{i,b} + \xi_{b,t} + \sigma_{s,t} + \phi_{loanriskrating} + \omega_{loanpurpose} + \epsilon_{l,i,b,s,t,z}$$

It regresses whether loan l to firm i in quarter t by bank b in sector/industry s, and zip code z ever becomes non-performing in future periods on the lending bank b's specialization in industry s. "Non-performing" is a dummy that takes the value of 1 if the loan falls in arrears or is otherwise in default. Specialization is defined as the degree to which a bank is over-invested in an industry, relative to a perfectly diversified portfolio. A diversified portfolio is one based solely on the size of an industry relative to all C&I lending. We use "excess" (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}} - \frac{LoanAmount_{s,t}}{\sum_s LoanAmount_{s,t}}$) specialization at the two-digit industry level. Depositor Concentration are share of a bank's deposits in a given zipcode z. Fixed effects include loan purpose and type as well as bank size. Relationship measures include past and future bank-borrower interactions as well as the specialization of bank b in borrower i's state. Regressions further include industry capture, loan size, whether a loan is secured by collateral, and the interest rate. Our sample includes only term loans and allows for re-negotiations. Standard errors are clustered at the bank-industry-year level while *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

E Loan Performance and Loan Terms – Additional Specifications

In this section, we include a large number of alternative specifications to showcase the validity of our baseline results under various assumptions.

First, we acknowledge that our measure of non-performing loans is only valid if we can observe the loan until maturity. Especially for loans that are negotiated towards the end of our sample, we are unable to verify whether these loans mature without becoming non-performing. While we do not expect this issue to introduce bias, as both specialized and non-specialized banks issue loans throughout the sample, we still wish to ensure our results are not driven by the structure of our data.

We can deal with this in one of three ways. Firstly, we can limit our sample to loans whose maturity we observe. This is a somewhat small sample, as loans are frequently renegotiated. Nevertheless, our baseline results discussed above hold for this sample (results not reported for brevity). Perhaps a more elegant solution would be to use a dynamic panel regression to better fit our panel data. This approach allows us to use the full panel structure of our data. Moreover, it allows a loan's non-performance status to vary with time. About half of loans in our sample recover from being labeled as non-performing.

In Table A.11, we relate whether a loan is non-performing in any given period to the degree to which the lender is specialized in the borrower's industry in the same period as well as lags of specialization and lags of whether the loan was non-performing in a previous period. Naturally, non-performance has a high degree of auto-correlation, implying that non-performance is predictive of future non-performance. Contemporaneous as well as lagged specialization are predictive of non-performance. However, given the high degree of auto-correlation present in the specialization measures, more than one lag is no longer predictive. These results hold if we use "excess" as well as "relative" specialization.

Thirdly, in Panel B we use a Cox-hazard regression. We define the "failure" event as a loan becoming non-performing, after which it is removed from the sample. Hazard regressions are among the most popular methods to deal with censored data. We find that, as before, loans by specialized banks are significantly less likely to become non-performing in any given period. The exact effect magnitude is larger in this specification than in our previous analyses, suggesting a more than 40% reduction in the likelihood of loans becoming non-performing.

Table A.11: ***Loan Performance – Dynamic Panel and COX Hazard Regressions***

Panel A

	(1)	(2)	(3)	(4)
	Non performing loan			
Non performing loan_{t-1}	0.168***	0.160***	0.168***	0.160***
	[0.002]	[0.002]	[0.002]	[0.002]
Excess Specialization	-0.017***	-0.018**		
	[0.006]	[0.007]		
Excess $\text{Specialization}_{t-1}$	-0.022***	-0.016**		
	[0.006]	[0.007]		
Excess $\text{Specialization}_{t-2}$		0.002		
		[0.007]		
Relative Specialization			-0.001***	-0.001***
			[0.000]	[0.000]
Relative $\text{Specialization}_{t-1}$			-0.001***	-0.001***
			[0.000]	[0.000]
Relative $\text{Specialization}_{t-2}$				-0.000
				[0.000]
Fixed Effects	Time,Purpose,Collateral			
Controls	Bank size, Industry Capture, Zip Share			
N	2,152,038	1,838,026	2,152,038	1,838,026

Panel B

	(1)	(2)	(3)
	Non-Performing Loan		
Excess Specialization	-1.890***	-1.442***	-1.535***
	[0.110]	[0.112]	[0.146]
Fixed Effects	No	No	Bank and Industry
Controls	No	Loan Charac.	Loan, Collateral, and Bank Charac.
N	3,820,052	3,820,052	3,808,508

Note: Panel A of this table shows coefficients for a dynamic panel regression. We relate whether a loan becomes non-performing in any given period to contemporaneous and lagged specialization measures as well as lagged measures of non-performance. We use "relative" (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}} / \frac{LoanAmount_{s,t}}{\sum_s LoanAmount_{s,t}}$) as well as "excess" (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}} - \frac{LoanAmount_{s,t}}{\sum_s LoanAmount_{s,t}}$) specialization at the two-digit industry level. All columns use time, loan purpose, and collateral-type fixed effects. Other controls include bank size, the degree to which a bank has captured an industry, and the share of the bank's lending in a given zipcode. Our sample includes only term loans and allows for re-negotiations. In Panel B we perform a standard cox-hazard analysis. The failure event is defined as loans becoming non-performing. We use "excess" specialization, as above. We include loan characteristics such as size, rate, purpose, etc. in column (2) and fixed effects for industry and bank in column (3). We display exponentiated coefficients and *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table A.12: ***Loan Performance Relative to Industry and Bank***

	(1)	(2)	(3)	(4)
	Non performing loan			
Avg. Defaults in Industry	0.236***	0.266***		0.256***
	[0.065]	[0.067]		[0.066]
Specialization * Ind. Default Rate		-2.275*		-2.187*
		[1.218]		[1.221]
Excess Specialization		0.003	0.022*	0.046***
		[0.016]	[0.012]	[0.018]
Avg. Default in Bank			0.588***	0.581***
			[0.061]	[0.061]
Specialization * Bank Default Rate			-1.915***	-1.847***
			[0.464]	[0.461]
Fixed Effects	Purpose, Rating, Bank, Time, Industry			
Controls	Loan size, Rate, Collateral, Industry capture			
R^2	0.29	0.3	0.3	0.3
N	296,951	296,951	296,951	296,951

Note: In this table, we relate the propensity of a loan becoming non-performing to the aggregate share of loans that are non-performing in the borrower's industry (columns (1) and (2)). We interact this variable with a measure of the lender's excess specialization in column (2). We relate the propensity of a loan becoming non-performing to the aggregate share of loans that are non-performing in the borrower's bank (columns (3) and (4)). We again interact our variable of interest with a bank's specialization in column (4). All columns use bank*time and industry*time fixed effects. Other fixed effects and controls include loan purpose and type, bank size industry capture, loan size, whether a loan is secured by collateral, relationship to the borrower, and the interest rate. Relationship measures include past and future bank-borrower interactions as well as the specialization of bank b in borrower i's state. Our sample includes only term loans and focuses on the period a loan is first observed. Standard errors are clustered at the bank-industry-year level while *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. include

From the above, we can conclude that the functional form assumptions made in the baseline analysis do not shape the results we present. In Table A.12 we look at the aggregate relationship between specialization and non-performing loans. Specifically, we relate the propensity that a loan to a firm becomes non-performing to the share of non-performing loans in the same industry at that time. We find a positive relationship. The more non-performing loans in an industry, the more likely a loan is to be non-performing. However, if we interact the share of non-performing loans with a bank's excess specialization in that industry, the relationship reverses. A more specialized bank can keep its loans from going bad, despite aggregate industry trends. We can see the same trend at the bank level (columns (3) and (4). Here we relate the share of non-performing loans in a bank's portfolio to the performance of individual loans. We again find a positive relationship that is reversed if we interact our variable of interest with a bank's specialization in the borrower's industry. Specialized industries perform better even when the bank as a whole is not.

Another potential issue could arise from the fact that bank size is negatively correlated with specialization. This implies that larger banks may be less specialized or otherwise less able to exploit specialization advantages. We relate specialization to non-performance – as in column (4) of Table 3 above – but now split our sample by bank size. It should be noted that all banks in the Y14 are large institutions. A split of banks "within" this group is extremely narrow. Table A.13 shows the results. The coefficients on specialization for each bank size group are extremely similar. We can focus on only the first observations or new origination loans, because in the case of small banks, individual large loans may disproportionately affect specialization calculated with the stock of loans. If we look only at the first observations of loans, the coefficient for the very largest group of banks is noisy and insignificant. However, if we look at newly originated loans or if we allow for renegotiations, then the coefficients are similar as well as statistically significant. We conclude that large banks are still able to exploit specialization. However, we concede that the results are slightly weaker if we focus on these banks. This finding corroborates results shown in A.6, above. Here we also split banks by size and find that the effect of specializing is more pronounced amongst larger institutions. Smaller banks (smaller than those found in our Y14 sample) are less able to specialize. Instead, their concentration in single industries is more likely to be the consequence of being size-constrained.

Table A.13: ***Specialization Split by Bank Size***

	(1)	(2)	(3)
	Loan ever becomes non-performing		
Panel A: First Loan Observation			
Excess Specialization	-0.050**	-0.085**	-0.068
	[0.021]	[0.037]	[0.051]
Panel B: New Loans			
Excess Specialization	-0.058**	-0.118**	-0.127*
	[0.026]	[0.046]	[0.040]
Panel C: All Loans incl. Renegotiations			
Excess Specialization	-0.116***	-0.072	-0.496***
	[0.032]	[0.050]	[0.077]
Sample	Small	Mid-Size	Largest
Fixed Effects	Bank*Time, Industry*Time, Loan Purpose, Loan Rating, Collateral		
Controls	Loan Rate, Size, Maturity, Bank Industry Capture		
R^2	0.17	0.2	0.094

Note: This table shows the coefficients of interest for equation:

$$Y_{l,i,b,s,T} = \beta_0 + \beta_l Specialization_{b,s,t} + \beta_2 \mathbf{X}_{l,b} + \beta_3 Relationship_{i,b}$$
$$+\xi_{b,t} + \sigma_{s,t} + \phi_{loanriskrating} + \omega_{loanpurpose} + \epsilon_{l,i,b,s,t}$$

It regresses whether loan l to firm i in quarter t by bank b in sector/industry s ever becomes non-performing in future periods on the lending bank b's specialization in industry s. "Non-performing" is a dummy that takes the value of 1 if the loan falls in arrears or is otherwise in default. Specialization is defined as the degree to which a bank is over-invested in an industry, relative to a perfectly diversified portfolio. A diversified portfolio is one based solely on the size of an industry relative to all C&I lending. We use excess (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}} - \frac{LoanAmount_{s,t}}{\sum_s LoanAmount_{s,t}}$) specialization at the two-digit industry level. We split our sample by bank size; we use terciles per quarter. Fixed effects include loan purpose, rating, bank*time and industry*time. Controls include industry capture, a bank's share in the zip code, and collateral types. Relationship measures include past and future bank-borrower interactions as well as the specialization of bank b in borrower i's state. Our sample includes only term loans, uses the first observations of loans in Panel A, purely new loans in Panel B, and all loans with re-negotiations in Panel C. Standard errors are clustered at the bank-industry-year level while *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

We next relate loan performance to lagged specialization measures in A.15. Here, we see that contemporaneous specialization has the largest negative impact on loan performance. The coefficient on specialization in past periods is negatively signed, though insignificant. This is in part due to the high degree of auto-correlation in specialization. If we exclude current specialization, we see that past specialization – measured as a continuous variable or a dummy for a bank's favorite/top industry (see column (3)) – is negative and significantly associated with loan non-performance.

In Table A.14 (Panel A), we show that including firm fixed effects does not fundamentally shape our results. The inclusion of firm fixed effects implies we are identifying our results off of a single firm that is borrowing from two differently specialized banks – all else held equal. Our above results, which show that loans by specialized banks perform better, may be picking up both the ability of specialized banks to select good borrowers as well as their ability to monitor said borrower after a loan has been originated. Including firm fixed effects removes the time-invariant borrower quality component. As such, we can see the degree to which specialization by the bank can help it monitor. The firm fixed effect absorbs between 50% and 60% of the baseline effect, estimated above. The residual effect may be in large part due to the bank's ability to monitor the borrower. It should be noted, however, that our identification rests on the somewhat smaller sample of larger borrowers that borrow from more than one bank. Moreover, the identification of individual firms is not perfect. Especially smaller firms are not tracked well across different banks, making these results only indicative.

We additionally test our baseline proposition using specialization calculated using new loans when they are first observed in our sample (see Panel B). We again find that our results are confirmed. Specialization calculated using only loans originated or purchased in the same period that a given loan is made does not depend on legacy loans – where single large loans could influence our results. The results are slightly smaller (and the specialization measure slightly noisier) than for our baseline results, but the overall direction and economic magnitude are preserved.

In the regressions above, we use lending at the zip-level as a measure of regional concentration/specialization. This may arguably be too granular a level, as similar borrowers may be spread across several zip-codes in a single state. We therefore additionally run a series of regressions that measure concentration at the state level. For simplicity, we include the "horse-race" specification here in Table A.17. We can see that lending concentration at the state level does not significantly impact loan performance. A ZIPcode or county (not reported for brevity) is arguably a better geographic distinction for the effect in question.

Additionally, we analyze loan terms in Table A.16. In this specification, we allow for re-negotiations. It is possible, that a bank lures borrowers with attractive initial rates only to change the terms once a borrower is locked into a relationship. We see that, if anything, the opposite holds. Allowing for re-negotiations shows a similar relationship between specialization and loan terms established above. Borrowers from specialized firms pay lower rates and obtain larger loans, at longer maturity. The coefficients on these regressions are larger than those that ignore renegotiation. Lastly, the loans from specialized banks are still less likely to be unsecured.

We can see in Table A.18 that specialized lenders are less likely to take common, easily pledged, collateral. These types typically include real estate, blanket liens on the business, or marketable securities. They do not require specialized lender knowledge to liquidate in the case of borrower distress. Specialized lenders, on the other hand, take fixed assets, or "other" more obscure forms of collateral. These are harder to value and harder to liquidate in the case of borrower distress. Specialized knowledge can be valuable for these forms of collateral (as discussed in Gopal (2019)).

We show that competition can affect the terms offered to borrowers in Table A.19. Here, we interact our variable of interest – specialization – with a variable that counts the number of other banks that a borrower has access to in a given period. We see that the loan terms are more generous for those borrowers who have alternatives. Similarly, the loan performance for borrowers attracted in times of higher competition is slightly worse (not reported for brevity). While the latter effect is small, it helps solidify the point that specialized banks are attracting desirable borrowers with good loan terms and that competition induces them to be slightly more aggressive and less discerning.

Table A.14: ***Firm Fixed Effects and Specialization at Origination***

Panel A: Firm Fixed Effects

	(1)	(2)	(3)	(4)
	Loan ever becomes non-performing			
Excess Specialization	-0.054**	-0.057**	-0.055**	-0.068**
	[0.027]	[0.027]	[0.026]	[0.028]
General Fixed Effects	Industry*Time, Bank*Time, Purpose, Firm			
Interest Rate	No	Yes	Yes	Yes
Loan Rating Fixed Effects	No	No	Yes	Yes
Collateral Fixed Effects	No	No	No	Yes
Mean of dependent variable	0.05	0.05	0.05	0.05
R^2	0.7	0.7	0.71	0.71
N	296,951	296,951	296,951	296,951

Panel B: Specialization at Origination

	(1)	(2)	(3)	(4)
	Loan ever becomes non-performing			
Excess Specialization	-0.079***	-0.071***	-0.050***	-0.052***
	[0.007]	[0.007]	[0.006]	[0.006]
General Fixed Effects	Industry*Time, Bank*Time, Purpose			
Interest Rate	No	Yes	Yes	Yes
Loan Rating Fixed Effects	No	No	Yes	Yes
Collateral Fixed Effects	No	No	No	Yes
Mean of dependent variable	0.05	0.05	0.05	0.05
R^2	0.02	0.03	0.15	0.16
N	298,951	298,951	298,951	298,951

Note: This table shows the coefficients of interest for equation:

$$Y_{l,i,b,s,T} = \beta_0 + \beta_l Specialization_{b,s,t} + \beta_2 \mathbf{X}_{l,b} + \beta_3 Relationship_{i,b} + \\ \xi_{b,t} + \sigma_{s,t} + \gamma_i + \phi_{loanriskrating} + \omega_{loanpurpose} + \epsilon_{l,i,b,s,t}$$

It regresses whether loan l to firm i in quarter t by bank b in sector/industry s ever becomes non-performing in future periods on the lending bank b's specialization in industry s. "Non performing" is a dummy that takes the value of 1 if the loan falls in arrears or is otherwise in default. Specialization is defined as the degree to which a bank is over-invested in an industry, relative to a perfectly diversified portfolio. A diversified portfolio is one based solely on the size of an industry relative to all C&I lending. We use excess (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}} - \frac{LoanAmount_{s,t}}{\sum_s LoanAmount_{s,t}}$) specialization at the two-digit industry level. Other fixed effects and controls include loan purpose and type, bank size, industry capture, loan size, whether a loan is secured by collateral, relationship to the borrower, and the interest rate. Relationship measures include past and future bank-borrower interactions as well as the specialization of bank b in borrower i's state. Our sample includes only term loans and focuses on the period a loan is first observed. All columns use bank*time, industry*time, and loan purpose fixed effects. Panel A makes use of firm fixed effects. In Panel B, our specialization measure is calculated with only loans first observed in the given quarter as opposed to the stock of all loans held in that quarter. Standard errors are clustered at the bank-industry-year level while *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table A.15: *Specialization Lags*

	(1)	(2)	(3)
	Loan becomes non-performing		
Excess Specialization	-0.175***		
	[0.055]		
Excess Specialization$_{t-1}$	0.040	-0.099*	
	[0.072]	[0.054]	
Excess Specialization$_{t-2}$	0.048	0.020	
	[0.052]	[0.054]	
Top Industry Dummy$_{t-1}$			-0.004***
			[0.001]
Fixed Effects	Bank*Time and Industry * Time, Loan Purpose, Loan Rating		
Controls	Collateral, Maturity, Rate, Relationship		
R^2	0.13	0.13	0.14
N	258,557	258,557	296,951

Note: This table shows the coefficients of interest for equation:

$$Y_{l,i,b,s,T} = \beta_0 + \beta_l Specialization_{b,s,t} + \beta_2 Specialization_{b,s,t-x}$$
$$+\beta_3 \mathbf{X}_{l,b} + \beta_4 Relationship_{i,b} + \xi_{b,t} + \sigma_{s,t} + \phi_{loanriskrating} + \omega_{loanpurpose} + \epsilon_{l,i,b,s,t}$$

It regresses whether loan l to firm i in quarter t by bank b in sector/industry s ever becomes non-performing in future periods on the lending bank b's specialization in industry s. "Non-performing" is a dummy that takes the value of 1 if the loan falls in arrears or is otherwise in default. Specialization is defined as the degree to which a bank is over-invested in an industry, relative to a perfectly diversified portfolio. A diversified portfolio is one based solely on the size of an industry relative to all C&I lending. We use excess (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}} - \frac{LoanAmount_{s,t}}{\sum_s LoanAmount_{s,t}}$) specialization at the two-digit industry level. We include lags of specialization as well as lags for whether a two-digit industry is a lender's favorite (i.e. has the highest specialization). All columns use bank*time and industry*time fixed effects. Other fixed effects and controls include loan purpose and type, bank size, industry capture, loan size, whether a loan is secured by collateral, relationship to the borrower, and the interest rate. Relationship measures include past and future bank-borrower interactions as well as the specialization of bank b in borrower i's state. Our sample includes only term loans and focuses on the period a loan is first observed. Standard errors are clustered at the bank-industry-year level while *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table A.16: *Specialization and Loan Terms – All Observations*

	(1) Interest rate	(2) Log loan amount	(3) Maturity remaining	(4) Unsecured
Specialization	-0.534***	1.357***	5.738***	-0.159***
	[0.074]	[0.064]	[1.123]	[0.028]
Fixed Effects	Bank*Time, Industry*Time, Loan Purpose, Loan Rating			
Controls	Loan Rate, Size, Maturity, Bank Industry Capture, Collateral			
Mean of dependent variable	3.6	8.2	22	0.14
R^2	0.28	0.24	0.24	0.31
N	2,874,888	2,874,888	2,823,618	2,874,888

Note: This table shows the coefficients of interest for equation:

$$Y_{l,i,b,s,T} = \beta_0 + \beta_l Specialization_{b,s,t} + \beta_2 \mathbf{X}_{l,b} + \beta_3 Relationship_{i,b} + \xi_{b,t} + \sigma_{s,t} + \phi_{loanriskrating} + \omega_{loanpurpose} + \epsilon_{l,i,b,s,t}$$

It regresses different characteristics of loan l to firm i in quarter t by bank b in sector/industry s on the lending bank b's specialization in industry s. We look at the rate paid by the borrower, the loan amount (logged), the maturity, and whether the loan is unsecured. Specialization is defined as the degree to which a bank is over-invested in an industry, relative to a perfectly diversified portfolio. A diversified portfolio is one based solely on the size of an industry relative to all C&I lending. We use excess (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}} - \frac{LoanAmount_{s,t}}{\sum_s LoanAmount_{s,t}}$) specialization at the two-digit industry level. All columns use bank*time and industry*time fixed effects. Other fixed effects/controls include loan purpose and type, bank size, industry capture, loan size, whether a loan is secured by collateral, relationship to the borrower, and the interest rate (unless the outcome in question is the specified variable). Relationship measures include past and future bank-borrower interactions as well as the specialization of bank b in borrower i's state. Our sample includes only term loans and allows for re-negotiations. Standard errors are clustered at the bank-industry-year level while *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table A.17: ***Alternative Specialization Measures and Loan Performance – State Level***

	(1)	(2)	(3)	(4)	(5)	(6)
	Loan ever becomes non-performing					
Excess Specialization	-0.097***				-0.201***	-0.011***
	[0.013]				[0.045]	[0.002]
Share of Portfolio in industry		-0.081***			0.097**	0.008**
		[0.012]			[0.039]	[0.004]
Share of Portfolio in State			-0.003		-0.004	-0.00
			[0.023]		[0.007]	[0.00]
Borrower Relationship				-0.001***	-0.001***	-0.003***
				[0.000]	[0.000]	[0.001]
Fixed Effects	Bank*Time, Industry*Time, Loan Purpose, Loan Rating					
Controls	Loan Rate, Size, Maturity, Bank Industry Capture, Collateral					
Standardized Coefficients	No	No	No	No	No	Yes
Mean of dependent variable	0.05	0.04	0.04	0.04	0.04	0.04
R^2	0.16	0.16	0.16	0.16	0.17	0.17
N	296,951	296,951	296,951	296,951	296,951	296,951

Note: This table shows the coefficients of interest for equation:

$$NonPerformance_{l,i,b,s,T} = \beta_0 + \beta_l Specialization_{b,s,t} + \beta_2 \mathbf{X}_{l,b} + \xi_{b,t} + \sigma_{s,t} + \phi_{loanriskrating} + \omega_{loanpurpose} + \epsilon_{l,i,b,s,t}$$

It regresses whether loan l to firm i in quarter t operating in sector/industry s and located in a given State, which is made by bank b ever becomes non-performing in future periods on bank b's specialization in industry s. "Non performing" is a dummy that takes the value of 1 if the loan ever falls in arrears, has negative maturity, or is otherwise in default after it is first observed in our data. Specialization is defined in a number of different ways. First, we use the degree to which a bank is over-invested in an industry, relative to a perfectly diversified portfolio. A diversified portfolio is one based solely on the size of an industry relative to all C&I lending. We use excess (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}} - \frac{LoanAmount_{s,t}}{\sum_s LoanAmount_{s,t}}$)) specialization at the two-digit industry level. Second, we use the share of a bank's portfolio invested in a single industry (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}}$). Third, we use the share of a bank's portfolio invested in a single state (i.e. $\frac{LoanAmount_{b,state,t}}{\sum_s LoanAmount_{b,state,t}}$). Finally, we use the relationship between a bank and a borrower, measured as the number of interactions the two have over our entire dataset. Column (6) makes use of standardized coefficients. All columns contain bank*time, industry*time, collateral type and loan purpose fixed effects. We additionally account for the interest rate paid, the loan's rating at first observation, and the size of the loan. Our data is focused on the first observation of a loan and contains only term loans. Standard errors are clustered at the bank-industry-year level while *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table A.18: *Specialization and Collateral – All Observations*

	(1) Real estate	(2) Marketable securities	(3) AR	(4) Fixed assets	(5) Blanket lien	(6) Other
Specialization	-0.386***	-0.098***	0.079***	0.125***	-0.186***	0.582***
	[0.030]	[0.016]	[0.017]	[0.030]	[0.027]	[0.046]
Fixed Effects	Bank*Time, Industry*Time, Loan Purpose, Loan Rating					
Controls	Loan Rate, Size, Maturity, Bank Industry Capture, Collateral					
Mean of dep. var.	0.34	0.034	0.077	0.16	0.14	0.1
R^2	0.63	0.19	0.29	0.43	0.37	0.25
N	2,471,640	2,471,640	2,471,640	2,471,640	2,471,640	2,471,640

Note: This table shows the coefficients of interest for equation:

$$Y_{l,i,b,s,T} = \beta_0 + \beta_l Specialization_{b,s,t} + \beta_2 \mathbf{X}_{l,b} + \beta_3 Relationship_{i,b} + \xi_{b,t} + \sigma_{s,t} + \phi_{loanriskrating} + \omega_{loanpurpose} + \epsilon_{l,i,b,s,t}$$

It regresses the propensity that loan l to firm i in quarter t by bank b in sector/industry s is secured by collateral of type Y on the lending bank b's specialization in industry s. We look at whether a loan is secured by real estate, marketable securities, Accounts receivable. Specialization is defined as the degree to which a bank is over-invested in an industry, relative to a perfectly diversified portfolio. A diversified portfolio is one based solely on the size of an industry relative to all C&I lending. We use excess (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}} - \frac{LoanAmount_{s,t}}{\sum_s LoanAmount_{s,t}}$) specialization at the two-digit industry level. All columns use bank*time and industry*time fixed effects. Other fixed effects/controls include loan purpose and type, bank size, industry capture, loan size, whether a loan is secured by collateral, relationship to the borrower, and the interest rate (unless the outcome in question is the specified variable). Relationship measures include past and future bank-borrower interactions as well as the specialization of bank b in borrower i's state. Our sample includes only term loans and allows for re-negotiations. Standard errors are clustered at the firm-year level while *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table A.19: *Specialization and Loan Terms – Competition*

	(1) Interest rate	(2) Log loan amount	(3) Maturity remaining	(4) Unsecured
Specialization	-0.417***	1.208***	4.700***	-0.168***
	[0.079]	[0.066]	[1.184]	[0.025]
Interaction: Numb. Lenders	-0.076**	0.022	1.649***	-0.035**
	[0.032]	[0.024]	[0.227]	[0.013]
Numb. Lenders	-0.013***	0.109***	-0.261***	0.014***
	[0.001]	[0.001]	[0.010]	[0.000]
Fixed Effects	Bank*Time, Industry*Time, Loan Purpose, Loan Rating			
Controls	Loan Rate, Size, Maturity, Bank Industry Capture, Collateral			
Mean of dependent variable	3.6	8.2	22	0.14
R^2	0.32	0.27	0.23	0.31
N	2,874,888	2,874,888	2,823,618	2,874,888

Note: This table shows the coefficients of interest for equation:

$$Y_{l,i,b,s,T} = \beta_0 + \beta_l Specialization_{b,s,t} + \beta_2 Specialization_{b,s,t} * NumbLenders + \beta_3 \mathbf{X}_{l,b} + \\ \beta_4 Relationship_{i,b} + \xi_{b,t} + \sigma_{s,t} + \phi_{loanriskrating} + \omega_{loanpurpose} + \epsilon_{l,i,b,s,t}$$

It regresses different characteristics of loan l to firm i in quarter t by bank b in sector/industry s on the lending bank b's specialization in industry s. We look at the rate paid by the borrower, the loan amount (logged), the maturity, and whether the loan is unsecured. Specialization is defined as the degree to which a bank is over-invested in an industry, relative to a perfectly diversified portfolio. A diversified portfolio is one based solely on the size of an industry relative to all C&I lending. We use excess (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}} - \frac{LoanAmount_{s,t}}{\sum_s LoanAmount_{s,t}}$) specialization at the two-digit industry level. We include an interaction variable which counts the number of lenders a borrower has access to in a given quarter. All columns use bank*time and industry*time fixed effects. Other fixed effects/controls include loan purpose and type, bank size, industry capture, loan size, whether a loan is secured by collateral, relationship to the borrower, and the interest rate (unless the outcome in question is the specified variable). Relationship measures include past and future bank-borrower interactions as well as the specialization of bank b in borrower i's state. Our sample includes only term loans and allows for re-negotiations. Standard errors are clustered at the firm-year level while *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

F Aggregate Results

In this section we show that specialization is negatively associated with bank income. Intuitively, this can be the natural consequence of the fact that specialized banks provide larger loans at lower rates to their preferred borrowers. From Table A.20 we can see that the degree to which a bank is specialized in its top industry as well as the degree to which a bank is specialized on average are negatively associated with aggregate income as well as profitability (loosely defined as income to assets here). A one standard deviation increase in a bank's specialization in its top industry is associated with around 3 mil. USD in lower income, relative to a bank's typical earnings. The effect is statistically significant, though economically small.

Secondly, concentration in an industry can be somewhat unprofitable for banks during a downturn in that industry. We can see from Table A.21 that an increase in non-performing loans in an industry is related to lower income for banks. This holds especially if the bank in question is heavily specialized in this industry. Since the borrowers of specialized lenders are less likely to become non-performing during downturns, the magnitude of the effect is economically small. Nevertheless, this clearly depicts a shortcoming of aggressive specialization. At the aggregate level, banks may be overly exposed to negative performance issues in an industry if they are overly focused on said industry.

Table A.20: ***Bank Income to Specialization***

	(1)	(2)	(3)	(4)
	Income/Assets ratio		Net Income	
Specialization in Top Industry	-0.007***		-2.598***	
	[0.001]		[0.542]	
Average Specialization		-0.001		-0.102
		[0.001]		[0.476]
Fixed Effects	Industry	Industry	Bank	Bank
Controls	Time, Bank size, Leverage			
R^2	0.36	0.36	0.75	0.76
N	26,040	25,158	27,235	25,491

Note: This table shows the relationship between bank income and the degree to which the bank is specialized. We collapse our data to the bank*industry*time level and relate specialization in a bank's top industry (columns (1) and (3) and specialization in general (columns (2) and (4) to various measures of bank income. These measures include income to assets (columns (1) and (2)) as well as net income scaled by millions of dollars (columns (3) and (4)). We include time fixed effects as well as bank size and leverage controls in all specifications. We either include industry or bank fixed effects. Standard errors are heteroskedasticity robust while *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table A.21: ***Bank Income to Specialization***

	(1) Income/Assets	(2) Int. Income/Assets	(3) Income	(4) Interest Income
Share of non-performing loans	0.001	0.003	0.367	0.709
	[0.003]	[0.005]	[1.871]	[3.879]
Interaction: Share Non-Perf * Specialization	-0.206***	-0.475***	-14.943	-128.478**
	[0.052]	[0.083]	[28.967]	[52.033]
Excess Specialization	0.002*	0.011***	-0.834	2.795***
	[0.001]	[0.002]	[0.511]	[1.085]
Fixed Effects	Time and Industry			
Controls	Assets and Leverage			
R^2	0.36	0.52	0.71	0.74
N	25,158	25,158	25,158	25,158

Note: This table shows the relationship between bank income and the degree to which the bank is specialized interacted with the share of non-performing loans in an industry. We collapse our data to the bank*industry*time level and relate specialization in a bank's top industry, the share of non-performing loans in an industry, and the interaction of these variables to various measures of bank income. These measures include income or interest income to assets (columns (1) and (2)) as well as net income or net interest income scaled by millions of dollars (columns (3) and (4)). We include time as well as industry fixed effects as well as bank size and leverage controls in all specifications. Standard errors are heteroskedasticity robust while *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

G Liquidity Provision by Specialized Banks after Borrower Default

In this section, we show that specialized banks provide more liquidity to their borrowers. This holds, most crucially, after a borrower has previously had downgrade or non-performing loan issues. We thus corroborate findings by Giannetti and Saidi (2019), who show that concentrated lenders are more likely to provide liquidity in times of distress.

In Table A.22, column (1), we can first see that specialized lenders grant larger credit lines to their borrowers. This result was already observed and discussed above for term loans, but we can confirm that it holds for credit lines. From column (2) we can additionally see that borrowers are more likely to utilize a larger share of the credit lines provided by specialized lenders.

Some borrowers run into issues of loan non-performance or downgrades at some point during their time in the sample. We can analyze credit provided by specialized and non-specialized lenders around these issues. We first take loan amounts, de-trend them by the interest rate charged as well as the rating and purpose of the loan, and then group these by whether the loan is made by a specialized bank in their favorite (i.e. most specialized) industry or not. In Figure A.10, Panel (a), we show the average log loan amount (de-trended) provided by specialized lenders and non-specialized lenders. We keep only those borrowers who suffer issues with loan downgrades at some point during their time in the sample. This means that the rating of any of the borrower's loans was dropped by one or more lenders at some point. Time 0 denotes the period the downgrade issue occurred.

We can see that the loan amount provided by banks drops around the time that the issue occurs. However, it drops less for specialized lenders. Moreover, specialized lenders are more likely to provide more credit again after the downgrade issue. We repeat the above exercise in panel (b), using the share of utilized credit lines as opposed to loan size as our variable of interest. We again de-trend the share utilized by the interest rate, the loan purpose, and the rating before grouping them by whether the lender is lending to their most favored industry or not. As indicated in the Table A.22, we see that specialized lenders allow borrowers to draw down greater shares of the credit line. Moreover, we again find that the share of utilized credit lines drops less and rises more quickly for loans provided by specialized lenders after a borrower faces downgrade issues.

These results hold if we use " any non-performing loans" as an indicator of borrower issues (not reported for brevity). These observations are confirmed in columns (3) and (4) of Table A.22. Borrowers who have previously experienced issues can use a larger share of their credit lines if their lender is more specialized in their industry.

Taken together, the above is indicative evidence of two possible mechanics at play in our data. First, it seems that specialized lenders are more quickly able to judge the quality of a borrower after a non-performance issues has occurred[25]. Second, specialized lenders may help avoid more serious negative consequences or repeated non-performance issues by extending credit to their borrowers during times of distress. This may make the relationship to specialized lenders especially critical for opaque borrowers experiencing turbulence.

Table A.22: ***Specialization and Lending after Non-Performance***

	(1)	(2)	(3)	(4)
	(log) Loan Size	Utilized Share		
Specialization	1.275***	0.142***	0.141***	0.131***
	[0.095]	[0.023]	[0.023]	[0.023]
Interaction: Spec. * Past Non-Performance			0.209**	
			[0.083]	
Interaction: Spec. * Past Rating Downdrade				0.083**
				[0.041]
Fixed Effects	Bank*Time and Industry * Time, Loan Purpose, Loan Rating			
Controls	Collateral, Maturity, Rate, Relationship			
R^2	0.31	0.36	0.36	0.36
N	3,089,745	3,089,745	3,089,745	3,089,745

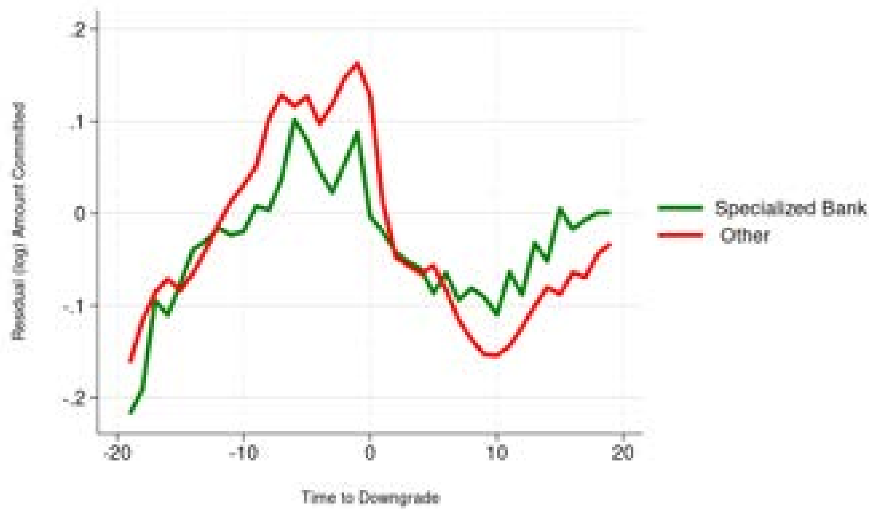
Note: This table shows the coefficients of interest for equation:

$$Y_{l,i,b,s,T} = \beta_0 + \beta_l Specialization_{b,s,t} + \beta_3 \mathbf{X}_{l,b} + \beta_4 Relationship_{i,b} + \\ \xi_{b,t} + \sigma_{s,t} + \phi_{loanriskrating} + \omega_{loanpurpose} + \epsilon_{l,i,b,s,t}$$

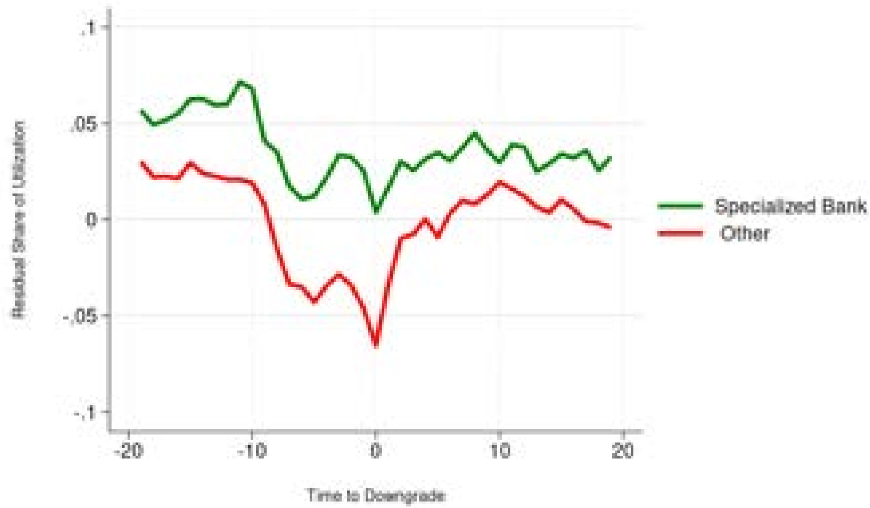
It regresses the size of credit line l (or the share of the credit line that is drawn) to firm i in quarter t by bank b in sector/industry s is secured by collateral of type Y on the lending bank b's specialization in industry s. A diversified portfolio is one based solely on the size of an industry relative to all C&I lending, where specialization is over-investment in an industry. We use "excess" (i.e. $\frac{LoanAmount_{b,s,t}}{\sum_s LoanAmount_{b,s,t}} - \frac{LoanAmount_{s,t}}{\sum_s LoanAmount_{s,t}}$) specialization at the two-digit industry level. All columns use bank*time and industry*time fixed effects. Other fixed effects/controls include loan purpose and type, bank size, industry capture, loan size, a bank's relationship to the borrower, and the interest rate. Relationship measures include past and future bank-borrower interactions as well as the specialization of bank b in borrower i's state. Our sample includes only credit lines and allows for re-negotiations. Standard errors are clustered at the bank-industry-year level while *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

[25] Borrowers who never borrow again leave our sample. Borrowers who remain eventually recover. The speed at which specialized lenders increase credit provision again following an issue shows the speed at which they can evaluate these borrowers.

Figure A.10: ***Credit Provision around Borrower Non-Performance***



(a) Committed Amount



(b) Share of Credit Line Utilized

Note: This figure depicts log loan amounts (panel (a)) and the share of utilized credit lines (panel (b)) around the time a borrower has any non-performing loan or loan downgrade issues. We depict 20 quarters before and 20 quarters after the event, though this is an unbalanced panel. Borrowers with more than one non-performing loan/downgrade event in succession do not appear for the full period, but would appear multiple times around the event. We use downgrades as a measure of performance issues. We de-trend (residualize) both the loan amount and the share utilized by the interest rate paid, the loan purpose as well as the loan rating at first observation before averaging the data according to whether the loan is made by a bank to its favorite (i.e. most specialized) industry or all other industries.

H Deposit Changes and Specialization

In this section, we show that sudden exogenous shocks to deposits are redirected towards a bank's favorite industry of specialization. As can be seen from Figure A.11, deposits in the largest banks have been rising steadily over our sample period. However, the sudden jump in deposits during the onset of COVID as well as the jump in deposits around the failure of SVB were exogenous to the bank's business models. These depository shocks were driven by changes in depositor sentiment and behavior. Consequently, the new deposits were unsolicited.

We can see from Table A.23 that unsolicited deposit growth is most likely to be directed into a bank's favorite industry while general deposit growth is not or negatively associated with top-industry specialization. We define the "COVID period" as the first quarter of 2020 and the post COVID period as 2020 and 2021. The period around the SVB shock is coded as the first quarter of 2023. We either average deposit changes over the quarter (columns (1) -(3)) or take the maximum value of weekly deposit changes in the quarter (columns (4)-(6)). The latter option is particularly useful as most banks saw the most significant deposit changes occur over the course of 1-2 weeks. These rapid changes in deposits are most likely to be unsolicited.

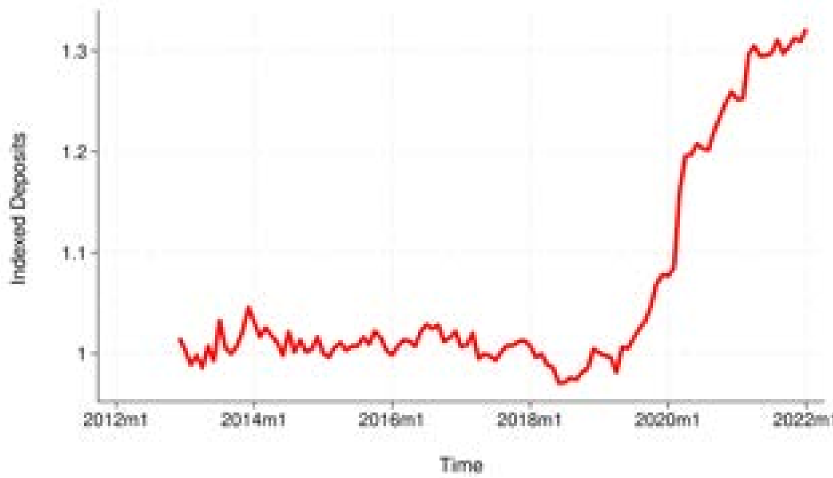
As indicated, we find that the COVID shock as well as the SVB shock are positively associated with growth of "excess" specialization in a bank's favorite industry. This is particularly true if we use the maximum, deposit change in a given quarter as the explanatory variable of choice. In our most saturated regression, a 10% increase in deposits would result in a 60 basis points increase in excess specialization in the following period. This is a sizeable effect, given that most loan contracts are not negotiated with great speed.

Table A.23: *Deposit Shock and Specialization*

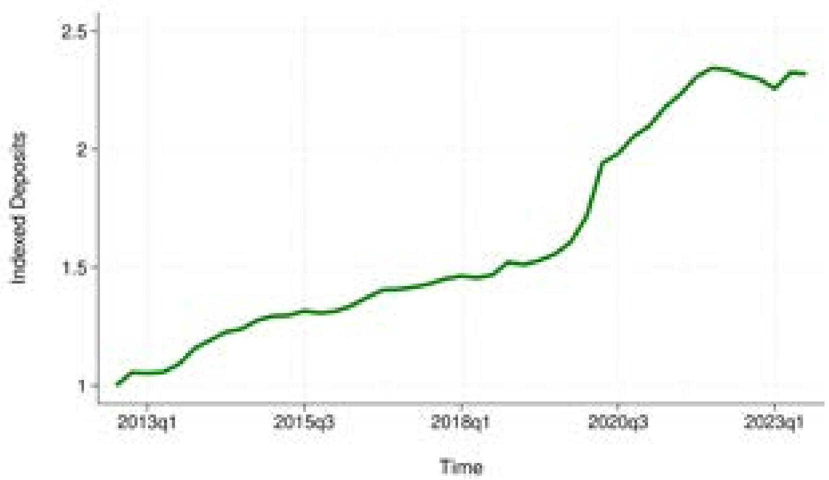
	(1)	(2)	(3)	(4)	(5)	(6)
			Excess Specialization			
Deposit Change$_{t-1}$	-0.026	-0.017	-0.064**			
	[0.020]	[0.016]	[0.030]			
Covid-Period × Deposit Change$_{t-1}$	0.040	0.040	0.084**			
	[0.034]	[0.029]	[0.040]			
Post Covid-Period × Deposit Change$_{t-1}$	0.048	0.044	0.090*			
	[0.047]	[0.038]	[0.053]			
SVB-Period × Deposit Change$_{t-1}$	0.009	-0.006	0.097			
	[0.059]	[0.059]	[0.094]			
Maximum Deposit Change$_{t-1}$				-0.024***	-0.009	-0.055***
				[0.008]	[0.007]	[0.015]
Covid-Period × Max. Deposit Change$_{t-1}$				0.030**	0.018*	0.060***
				[0.012]	[0.011]	[0.017]
Post Covid-Period × Max. Deposit Change$_{t-1}$				0.040**	0.031**	0.074***
				[0.019]	[0.016]	[0.022]
SVB-Period × Max. Deposit Change$_{t-1}$				0.022	-0.000	0.057*
				[0.023]	[0.023]	[0.033]
Fixed Effects	Time	Bank	Bank, Ind.*Time	Time	Bank	Bank, Ind.*Time
Controls			Bank size, leverage			
R^2	0.34	0.5	0.66	0.35	0.5	0.66
N	981	981	729	981	981	729

Note: This table relates the excess specialization of banks in their top/favorite industry to deposit growth in the previous quarter. We average deposit changes over the quarter (columns (1) -(3)) or take the maximum value of weekly deposit changes in the quarter (columns (4)-(6)). We include either quarter, bank, or quarter*industry and bank fixed effects. We account for bank size and leverage. Data is collapsed to the industry*quarter*bank level and we focus only on a bank's favorite industry. Standard errors are clustered at the bank*time level while *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Figure A.11: ***Deposits in Large Banks***



(a) Weekly Deposits in Large Banks (2013-2022)



(b) Monthly Deposits in Y14 Banks (2013-2023)

Note: This Figure's panel (a) shows indexed deposits all large banks (¿50 bn. USD) at the weekly level. It focuses on the period between 2012 and 2022 and includes the COVID-19 crisis. Deposits are indexed to January 2015. Panel (b) makes use of monthly data for the Y-14 banks in our sample. It includes the COVID and the more recent SVB shock (after March 2023).

I Real Effects: Firm Growth during COVID

As shown in the paper above, the COVID pandemic saw many large banks experience a sharp increase in deposits. These banks purchased securities and lent to firms – primarily in their preferred sector of specialization. This growth in firm-level funding was largely exogenous to banks and certainly exogenous to the firms that borrowed from these banks. The excess availability of credit for some firms may well have affected the ability of these firms to grow. To test this proposition, we therefore merge COMPUSTAT with Y14 data using a complicated hand-matching procedure based on the names of firms, or the ticker symbols where these are available. Our combined sample accounts for 60,000 individually verified firm-year observations between 2012 and 2022. We relate the growth in firms during the COVID 19 pandemic (fiscal years 2019, 2020, 2021, and 2022) to the degree of specialization of the firm's primary lender over this same time period[26]. We limit our sample to those firms that have borrowed from a Y14 lender within five years (2018 to 2022) at least once.

Mechanically, we are interested in determining the degree to which lender specialization affects year-on-year firm growth of its borrowers during a period of exogenous growth in deposits. Our primary variable of interest is the degree to which a firm's lender is specialized in its industry. For many firms, this measure is simply the degree to which its sole lender is specialized. For firms that borrow from multiple lenders during the period, we rank lenders by their specialization in the borrower's industry and use the most specialized lender. We count the number of lenders a firm engages within a given period and control for this number separately. However, it should be noted that the intuition of our results, found in table A.24 does not change if we use the average specialization of all lenders.

As can be seen, the more specialized a firm's lender is – measured with our relative specialization measure –, the more likely it is that the firm grows its liabilities during COVID. A highly specialized lender would facilitate a 5% greater growth in borrowing, all else equal. This ability to borrow more than competitors during a key period facilitates a similarly sized growth in EBITDA. In columns (4) - (6) we account for lagged growth in EBITDA, to avoid accidentally capturing the fact that firms growing profitably borrow from specialized lenders.

Overall, our results are significant and sizeable, given that firms in our sample see an annualized liabilities growth-rate of 7% and an EBITDA growth-rate of 4% during the period in question. Our results are unaffected by whether we use our adjusted measure of specialization (not shown for brevity). Finally, given that we are much more likely to be able to merge large firms to Y14, due to the nature of COMPUSTAT data and the likelihood that larger borrower information is better recorded in Y14, our results are liable to be a lower-bound

[26]We include 2019 in the sample as we measure year-on year growth rates

Table A.24: ***Firm Growth During COVID***

	(1)	(2)	(3)	(4)	(5)	(6)
	Change in Liabilities			Change in EBITDA		
Most Specialized Lender	0.015***	0.010***	0.018***	0.014***	0.018***	0.034***
	[0.003]	[0.003]	[0.005]	[0.005]	[0.005]	[0.009]
Liabilities to EBITDA_{t-1}	-0.001***	-0.001***	-0.002***	0.020***	0.019***	0.021***
	[0.000]	[0.000]	[0.000]	[0.001]	[0.001]	[0.001]
Leverage_{t-1}	-0.116***	-0.127***	-0.114***	-0.167***	-0.155***	-0.134***
	[0.008]	[0.008]	[0.010]	[0.022]	[0.023]	[0.033]
Controls	Firm Size, count of other lenders					
Industry and Time FE	No	Yes	Yes	No	Yes	Yes
R^2	0.036	0.067	0.096	0.13	0.19	0.19
N	8,844	8,844	4,506	7,164	7,164	3,397

Note: This table relates a firm's year-on-year growth (in total liabilities or EBITDA) to the specialization of its *most specialized* lender (using our relative measure). We include a number of firm-specific characteristics such as lagged assets, lagged leverage (debt/assets), lagged, profitability (debt/EBITDA). We account for industry (2 digit) and year fixed effects in columns (2), (3), (5) and (6). Columns (3) and (6) are focused on smaller firms. Columns (4)-(6) include lagged changes in EBITDA. Standard errors are clustered at the firm-year level while *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

estimate. After all, small firms with less access to outside funding will be more reliant on their banking relationship, especially in times of turmoil. We show this by highlighting (in columns (3) and (6)) that smaller firms that lie below our median asset-value experience a more pronounced growth due to lender specialization.

Naturally, given that better firms may have stronger relationships with more specialized lenders, these results should be viewed with caution. However, these findings still speak to the real effects of an exogenous reshuffling of deposits in the banking sector. Not all banks are fungible, given that they specialize in different industries. This specialization has very real consequences for the firms that borrow from them.