The Industry Expertise Channel of Mortgage Lending*

Yongqiang Chu Zhanbing Xiao Yuxiang Zheng

September 2023

Abstract

We show that banks use industry knowledge acquired through corporate lending in mortgage lending, a phenomenon we refer to as the "industry expertise channel." Specifically, we show that banks with specialization in particular industries increase mortgage lending in regions with significant industry concentrations. The impact of industry expertise increases with information asymmetry and elevated borrower risk. Additionally, mortgages originated from this channel contain more soft information and perform better. The importance of the channel also increases after unexpected industry distress and the 2008 financial crisis, suggesting that the effects are likely causal. Our findings underscore how insights gleaned from the corporate loan market contribute to enhancing banks' screening and monitoring capabilities within the mortgage market.

Keywords: Lending Specialization; Industry Expertise; Information Asymmetry; Income Risk; Mortgage Lending; Syndicated Loans.

JEL Codes: G21, G30, D82.

---

*Yongqiang Chu: Belk College of Business and Childress Klein Center for Real Estate, University of North Carolina at Charlotte. Email:yonqiang.chu@uncc.edu. Zhanbing Xiao: The Salata Institute for Climate and Sustainability, Harvard University. Email:zhanbingxiao@fas.harvard.edu. Yuxiang Zheng: College of Business, University of Akron, Email:yzheng2@uakron.edu. We thank Thomas Davidoff, Cameron LaPoint (discussant), Mikhail Mamonov (discussant), David Martinez-Miera, Klaas Mulier (discussant), Evren Ors, José Luis Peydro, Simon Rother (discussant) and seminar participants at the MFA 2023, ASSA-IBEFA Meetings 2022, AFA Ph.D. Student Poster Session 2022, FMA Annual Meetings 2021, AAA Annual Meetings 2021, 28*th* AEFIN Ph.D. Mentoring Day, 28*th* AEFIN Finance Forum, and SWFA Annual Meetings 2021 for helpful comments. All errors are our own.

1 Introduction

Banks acquire information through their interactions with borrowers. Research in lending relationships shows that banks use borrower-specific information to screen and monitor future borrowers (Berger and Udell, 1995; Petersen and Rajan, 1995). The recent literature shows that, in addition to borrower-specific information, banks cultivate specialized knowledge within specific industries by concentrating their lending activities in those sectors (Acharya, Hasan, and Saunders, 2006; Berger, Minnis, and Sutherland, 2017; Blickle, Parlatore, and Saunders, 2021). In particular, Blickle et al. (2021) show that banks apply industry-specific knowledge when lending to opaque firms operating within those industries. We ask whether the impact of industry-specific knowledge transcends the realm of commercial lending, as we delve into how banks' industry proficiency influences their residential mortgage lending.

Specifically, we investigate the impact of banks' industry expertise on their mortgage lending in areas where those industries are concentrated. We hypothesize that banks' industry expertise could mitigate the information asymmetry between borrowers and lenders, and thereby alleviating credit rationing. This hypothesis is built on two conceptual lenses. One is that household income growth positively correlates with the performance of leading industries in a county. Such correlation holds for both households working in the leading industries and those in non-leading industries due to a spillover effect.[1] The other is that industry expertise helps banks gain a deeper understanding of the local economy in areas concentrated with these industries. Considering the importance of regular income in mortgage repayment (Elul, Souleles, Chomsisengphet, Glennon, and Hunt, 2010), the industry-specific knowledge possessed by banks could reduce the information asymmetry between lenders and borrowers, allowing them to better assess a borrower's income risk and, hence, mortgage affordability. As articulated by Stiglitz

[1] For example, a collapse of the auto industry in Detroit negatively affects both auto workers and non-auto workers (e.g., workers in the service industry like restaurants or the retail industry like shopping malls).

and Weiss (1981), the reduction in information asymmetry could curtail credit rationing, thereby leading to increases in credit supply.

To empirically test the effects of industry expertise on mortgage lending, we construct a measure of industry specialization using the DealScan syndicated loan data. Specifically, we classify a bank as specialized in an industry if the bank's loan share in that industry is above the 75^{th} percentile of the distribution of all banks' portfolio shares in the industry plus the 1.5 inter-quartile range of the distribution (Paravisini, Rappoport, and Schnabl, 2023). This classification method controls for heterogeneity in the sizes of different banks and different industries. We then define that a bank and a county are connected through the industry expertise channel if the bank has one or more specialized industries that provide at least 5% jobs in the county.

We compare mortgage credits to borrowers in a county by banks connected to the county through the industry expertise channel relative to those not. We find that industry expertise significantly increases banks' mortgage lending. The results hold after adding county-by-year fixed effects to control for county-specific time-varying trends and bank-by-state fixed effects to control for any links between banks and states. The results also remain robust after adding bank-by-year fixed effects to control for time-varying heterogeneities across banks or bank-by-county fixed effects to control for time-invariant links between banks and counties. The economic magnitude is also significant. The channel increases banks' mortgage lending by 6.3% in the number of mortgages and 6.5% in dollar volumes. The findings highlight the importance of information embedded in the industry expertise channel in banks' mortgage decisions.

We also examine the effects on banks' mortgage approval rates. The approval rate reflects banks' lending decisions conditional on received mortgage applications, therefore isolating demand-side factors from contaminating our estimations. We find that industry expertise increases banks' number- and volume-based approval rates by 40 basis points. The evidence suggests that the demand-side forces do not likely drive our findings. Instead, it is the banks' supply decisions that matter.

Next, we provide six sets of evidence supporting the information mechanism of the industry expertise channel. First, a prerequisite for the industry expertise channel is that household income growth and mortgage affordability positively correlate with the performance of leading industries in a county. Therefore, industry expertise allows banks to assess local borrowers' income dynamics and mortgage default risks after origination. Consistent with this conjecture, we find that sales growth of a county's key industries positively affects the county's household income growth and negatively affects the county's mortgage delinquency rates. The economic effect is large - a one standard deviation increase in sales growth is associated with a 14.9% increase in household income growth.

Second, we examine the information asymmetry between banks and mortgage borrowers. We find that banks' use of industry expertise increases with the distance between banks' headquarters and borrowers' home counties, suggesting that banks' industry expertise can mitigate the distance-generated information friction. A one standard deviation increase in the distance doubles the effects of industry expertise on mortgage lending. In addition, we find that social connections between banks and borrowers decrease banks' reliance on industry expertise, suggesting that the soft information from industry expertise can substitute for the soft information from social connections.

Third, banks' information needs in mortgage origination should be larger for high-risk borrowers because they are more likely to miss their mortgage payments and default. Our first proxy for borrower risk is a county's house price volatility. High volatility increases the downside risk of house prices and could lead to negative home equity. We find that banks use the channel more when local house prices become more volatile. The effect of industry expertise on mortgage lending increases from 6.7% to 11.4% for a one standard deviation increase in house price volatility. In addition, we also use borrowers' loan-to-income ratios (LTI) as another proxy for borrower risk. We find that banks rely more on industry expertise when lending to high-LTI borrowers. The effect increases from 5.3% to 11.1% for a one standard deviation increase in the LTI.

Fourth, we examine the soft information contained in mortgage contracts to provide more direct evidence of the information mechanism. The screening model in Cornell and Welch (1996) shows that lower information frictions lead to larger loan term dispersion, as better information allows banks to better discriminate between “good" and “bad" borrowers. As such, banks can grant mortgages with favorable terms to “good" borrowers and mortgages with strict terms to “bad" borrowers (Fisman, Paravisini, and Vig, 2017; Lim and Nguyen, 2020). Otherwise, banks can only give loans with similar terms based on the average quality of all borrowers. We find industry expertise significantly increases the dispersion in mortgage amounts, loan-to-income ratios, interest rates, and loan-to-value ratios. In particular, the standard deviations of mortgage amounts, loan-to-income ratios, interest rates, and loan-to-value ratios are 0.6%, 0.5%, 2.1% and 2.2% higher for mortgages originated through the channel.

Fifth, we examine the differential impact of industry expertise on conventional versus government-insured mortgages. Government insurance provided by the Federal Housing Administration (FHA) and Veterans Affairs (VA) makes lenders’ mortgage exposure less information sensitive, and hence underwriting government-insured mortgages is less subject to information asymmetry and credit rationing. To this end, we find that banks with industry expertise originate more conventional mortgages relative to government-insured mortgages.

Finally, we test the performance implications of the industry expertise channel. If the channel indeed improves banks’ screening and monitoring efficiencies in mortgage decisions, we expect a positive effect on banks’ mortgage performance. Using the HMDA data matched with the Fannie Mae, Freddie Mac, and McDash loan performance data, we find that mortgages originated through the industry expertise channel experience lower delinquency and foreclosure rates.

The results could be driven by omitted variables at the bank-county level or by reverse causalities. For example, banks could strategically choose which industries to lend to according to their mortgage business. We address these concerns using a difference-in-differences setting around unexpected industry-wide distress. Specifically, we compare the effects of industry distress on mortgage lending of banks with differential ex-ante industry specializations. This setting helps test the hypothesis that the industry expertise channel is most useful in industries fraught with uncertainties, which generate significant downside risks to household income. This is because industry expertise allows banks to better price the borrowers' income risk and to avoid large-scale default by timely selling mortgages to third parties like Fannie Mae and Freddie Mac. The industry-level shocks are plausibly exogenous to any given bank, county, or mortgage borrower, mitigating the omitted variables and reverse causality concerns. Our empirical results show that the industry expertise channel becomes much more important in industries and periods of distress. The effect of the channel on mortgage lending increases from 2% in non-distress periods to 6.4% in distress periods. Furthermore, we also use the 2008 financial crisis as another shock for difference-in-differences analyses. We find that industry expertise becomes more valuable in mortgage underwriting during and after the crisis.

Our paper contributes to the growing literature on banks' lending specialization. This literature shows that lending concentration allows banks to develop relevant industry expertise, which improves efficiencies in information collection and monitoring of corporate borrowers, leading to lower risks and higher bank values (e.g., Acharya, Hasan, and Saunders, 2006; Loutskina and Strahan, 2011; Berger, Minnis, and Sutherland, 2017; Gopal, 2019; Giometti and Pietrosanti, 2020; Beck, De Jonghe, and Mulier, 2021; Blickle, Parlatore, and Saunders, 2021; Paravisini, Rappoport, and Schnabl, 2023). Our paper extends the discussion by investigating the implications of industry expertise on banks' mortgage lending. Specifically, we show that banks use industry knowledge acquired through corporate lending in screening and monitoring mortgage borrowers. Hence, our paper reveals an information spillover from the corporate loan market to the mortgage market.

Our paper also contributes to the literature on information asymmetry and credit access in the mortgage market. Even though hard information like credit reports and employment records significantly alleviate information frictions in mortgage origination

(Ergungor, 2010; Gilje, Loutskina, and Strahan, 2016), widespread mortgage fraud exists (Garmaise, 2015; Mian and Sufi, 2017). Recent studies document the importance of social capital and financial technology in information collection and mortgage lending (Fuster, Plosser, Schnabl, and Vickery, 2019; Rehbein and Rother, 2020). Our paper uncovers a new soft information channel, the industry expertise channel, that banks use to overcome information frictions. We show that this channel contains credible information regarding borrowers' future income dynamics, therefore improving banks' screening and monitoring efficiencies in the mortgage market.

Some studies investigate banks' allocation of mortgage credits across areas. Cortés and Strahan (2017) find that banks reallocate mortgage capital from non-core markets to disaster-hit areas to meet local demand increases. Chavaz and Rose (2019) and Chu and Zhang (2022) examine the influence of political forces on banks' lending decisions. In addition, social connectedness between banks and mortgage borrowers also plays an important role (Lim and Nguyen, 2020; Rehbein and Rother, 2020). Our paper complements these studies by showing that banks allocate more mortgage credits to counties with shared industry concentrations.

Last, our paper adds to the literature on household income risk and mortgage default (e.g., Elul, Souleles, Chomsisengphet, Glennon, and Hunt, 2010; Campbell and Cocco, 2015; Gerardi, Herkenhoff, Ohanian, and Willen, 2018). Even though household income is a critical factor in standard models of mortgage defaults, empirical estimates of its effects are small. For example, Foote, Gerardi, Goette, and Willen (2010) find that debt-to-income ratio (DTI) is not a strong predictor of future mortgage defaults, particularly as time passes since origination.[2] We show that the industry expertise channel complements the hard information on income collected at origination, enabling banks to predict borrowers' future income dynamics. Therefore, this channel improves banks' ability to assess borrowers' income risk.

The remainder of this paper is organized as follows. Section 2 describes data and

[2]They argue that it's because "*income today is an imperfect predictor of income tomorrow*" and "*a mortgage that is affordable at origination may be substantially less so later on, and vice versa*".

measures used in empirical analyses. Section 3 shows the baseline results of the effect of bank industry expertise on mortgage lending. Section 4 presents evidence supporting the information mechanism of the channel. Section 5 addresses endogeneity issues using unexpected industry distress and the 2008 financial crisis. Section 6 concludes.

2 Data and Measures

2.1 Sample Construction

We use the LPC DealScan database to measure how banks specialize in the corporate loan market. This database covers the syndicated loan market extensively since the mid-1980s and provides detailed information on each loan, such as the lender, the borrower, loan amount, starting date, ending date, interest rate, etc. Syndicated loans account for almost half of all commercial and industrial loans in the United States. They are commonly used to study banks' lending policies and real impacts, although they do not capture all lending activities (e.g., Bharath, Dahiya, Saunders, and Srinivasan, 2007; Chodorow-Reich, 2014; Chakraborty, Goldstein, and MacKinlay, 2018).

We use the link tables provided by Schwert (2018) and Gomez, Landier, Sraer, and Thesmar (2020) to merge lenders in Dealscan with bank call report data.[3] We obtain information on banks' branch characteristics (e.g., branch name, geographic coordinates, address, the BHC, deposits, etc.) from the Summary of Deposits (SOD) data, which covers the universe of banks' depository branches annually from 1994. We also use the link table provided by Chava and Roberts (2008) to link borrowers with their accounting and industry information in Compustat.

We obtain data on banks' small business lending from the Community and Reinvestment Act (CRA) small business loans database provided by the Federal Financial Institutions Examination Council (FFIEC). This data set contains information on the total num-

[3]Banks are aggregated at the bank holding company (BHC) level in the link tables. Throughout the paper, we use the term "bank" to refer to BHCs.

ber and volume of small business loans originated by each reporting bank in each county starting in 1996.

We obtain mortgage data from the Home Mortgage Loan Disclosure Act (HMDA) database, which covers more than 90% of all mortgages originated in the U.S. We follow the prior literature and drop non-conventional loans and loans for manufactured housing and multifamily dwellings to remove the impact of government subsidies on banks' lending decisions.[4] We also exclude other on-standard mortgages, such as mortgages for home improvement and non-owner-occupied dwellings. We exclude counties in which a bank has fewer than five mortgage applications per year to ensure that our results are not driven by outliers.[5] We follow Dagher and Kazimov (2015) to merge the HMDA data with banks in the call reports by matching agency-specific IDs in HMDA (e.g., Federal Reserve RSSD-ID, FDIC Certificate Number, and OCC Charter Number) to RSSD IDs.

We complement the HMDA data with information on monthly loan-level performance from three sources: Fannie Mae and Freddie Mac single-family loan-level data sets and McDash loan-level data. The Fannie Mae data cover the fixed-rate single-family mortgage loans acquired by Fannie Mae from January 2000 to December 2022, with the origination year starting from 1999. The Freddie Mac data cover approximately 52.2 million fixed-rate single-family mortgage loans originated between January 1, 1999 and September 30, 2022 that are acquired by Freddie Mac. The McDash data are a proprietary database compiled by Black Knight. The data track the dynamic performance of both agency and non-agency loans and is widely used in the literature (e.g., Fuster, Goldsmith-Pinkham, Ramadorai, and Walther, 2022; Gerardi, Willen, and Zhang, 2023). Depending on the years, the McDash data cover 60% to 80% of the US mortgage market. Important to our study, all three datasets include a rich set of information not available in HMDA, including the borrowers' credit scores, loan-to-value (LTV) ratios, interest rates, and ex-post monthly loan performance (e.g., delinquency status and foreclosure). We follow Chu, Ma, and Zhang

[4]Non-conventional loans include the Federal Housing Administration (FHA)-insured loans, Veterans Affairs (VA)-guaranteed loans, Farm Service Agency (FSA) loans, and Rural Housing Service (RHS) loans.

[5]Our results are robust if we require at least ten or twenty mortgage applications or remove this requirement.

(2022) and match the three datasets to HMDA.[6] The matched government-sponsored enterprise (GSE) mortgage sample is based on Fannie Mae and Freddie Mac data, and the non-GSE mortgage sample is based on the McDash data. We combine the GSE and non-GSE mortgages and focus on the period from 1999 to 2017.

We obtain data on county-level employment and wages from the Bureau of Labor Statistics (BLS) Quarterly Census of Employment and Wages (QCEW) database. The QCEW program publishes a quarterly count of employment and wages reported by employers covering more than 95 percent of all jobs in the U.S., available at the county, MSA, state, and national levels by industry. For this study, we use the QCEW data that cover all six-digit NAICS industries for more than 3,000 counties in the U.S. at the annual frequency from 1990 to 2018.

We obtain data on county-to-county distances and data on county-level characteristics (e.g., income, housing price index, population, race, age, etc.) from the Bureau of Economic Analysis (BEA), the Federal Housing Finance Agency (FHFA), and the NBER database. We obtain county-level mortgage delinquency rates from the Consumer Financial Protection Bureau (CFPB) and the county-to-county social connectedness index (SCI) from Bailey, Cao, Kuchler, Stroebel, and Wong (2018). The SCI is constructed using Facebook friendship links in the year 2016.[7]

2.2 Measuring a Bank's Industry Specialization

We first measure the industry specialization for each bank using the DealScan data. The borrowers in DealScan are relatively large, and interacting with these large borrowers helps banks access the most advanced and comprehensive industry knowledge. We use the deal origination dates and maturities to create a panel that dynamically tracks each

[6]See Appendix B in Chu et al. (2022) for details of the matching process.

[7]$Social\ Connectedness\ Index_{i,j} = \frac{FB\_Connections_{i,j}}{FB\_Users_i * FB\_Users_j}$, where $FB\_Connections_{i,j}$ is the total number of Facebook friendship connections between individuals in the two counties i and j, and FB_Users_i and FB_Users_j are the number of Facebook users in the two counties separately. See Bailey et al. (2018) and Bailey, Dávila, Kuchler, and Stroebel (2019) for more details.

bank's lending portfolio at any given time based on the starting and ending dates.

Most loans in Dealscan are syndicated and thus have multiple lenders. However, only lead lenders bear the monitoring responsibilities (Sufi, 2007; Gustafson, Ivanov, and Meisenzahl, 2021). Lead lenders have stronger incentives and better opportunities than participating lenders to acquire information about borrowers and accumulate industry expertise. As a result, lending specialization matters more for lead lenders than for participating lenders (Blickle, Parlatore, and Saunders, 2021). In addition, lead lenders are less likely to sell all of their loan shares in the secondary market (Irani, Iyer, Meisenzahl, and Peydro, 2021). We therefore focus on lead lenders of syndicated loans.[8]

We assume that lead lenders commit all capital in a loan because the allocation of loan shares is missing for most loans in DealScan, and the lead lenders obtain industry knowledge by monitoring the total loan amount rather than their own capital (e.g., Bharath, Dahiya, Saunders, and Srinivasan, 2007; Giannetti and Saidi, 2019; Saidi and Streitz, 2021). For loans with multiple lead lenders, we split the loan amount equally among the lead lenders.[9]

We aggregate banks' outstanding loans at the three-digit NAICS industry level each year. Our choice of the three-digit NAICS code level ensures sufficient precision of industry breakdowns and a reasonable number of firms and loans in each industry. We exclude firms in the financial industry.

Following Paravisini et al. (2023), we classify a bank as being specialized in an indus-

[8]We define lead lenders in each syndicated loan following the procedure outlined in Chakraborty et al. (2018). Specifically, lead lenders are identified under the following ranking hierarchy: 1) a lender is denoted as "Admin Agent", 2) a lender is denoted as "Lead bank", 3) a lender is denoted as "Lead arranger", 4) a lender is denoted as "Mandated lead arranger", 5) a lender is denoted as "Mandated arranger", 6) a lender is denoted as either "Arranger" or "Agent" and has a "yes" for the lead arranger credit, 7) a lender is denoted as either "Arranger" or "Agent" and has a "no" for the lead arranger credit, 8) a lender has a "yes" for the lead arranger credit but has a role other than those previously listed ("Participant" and "Secondary investor" are also excluded), 9) a lender has a "no" for the lead arranger credit but has a role other than those previously listed ("Participant" and "Secondary investor" are also excluded), and 10) a lender is denoted as a "Participant" or "Secondary investor". For a given loan package, the lender with the highest title (following the ten-part hierarchy) is considered the lead agent.

[9]We get similar results if we set loan shares retained by lead lenders equal to the median of the sample with non-missing information on the syndicate allocation (Chodorow-Reich, 2014; Giannetti and Saidi, 2019).

try if the bank's loan share in that industry is above the 75^{th} percentile plus one and a half times the inter-quartile range of the distribution of all banks' portfolio shares in the industry (Hodge and Austin, 2004).

$$Specialization^b_{i,t} = \begin{cases} 1 & L^b_{i,t} \geqq L^*_{i,t} \\ 0 & otherwise \end{cases} \tag{1}$$

where b denotes bank, i denotes industry, and t denotes year. $L^b_{i,t} = \frac{Loan^b_{i,t}}{\sum_{i=1}^{I} Loan^b_{i,t}}$ is bank b's portfolio share of syndicated loans towards industry i in the list of industries from 1 to I, at time t. $L^*_{i,t}$ is the threshold to identify the outlier in the distribution of $L^b_{i,t}$ among all banks in industry i. For each industry, the threshold is the 75^{th} percentile plus one and a half times the inter-quartile range of the distribution of all banks' portfolio shares in the industry (Hodge and Austin, 2004).

There are at least two advantages to measuring lending specialization in a relative way. First, this method accounts for the heterogeneity in the sizes of different banks and industries. Specifically, scaling a bank's loans to a given industry using the bank's total loans makes the measure impervious to bank sizes. Comparing different banks' loan shares within the same industry makes the measure impervious to industry sizes. Second, as we will discuss later, we include county-by-year fixed effects in our main empirical specifications to compare different banks' mortgage lending in the same county. A relative measure enables us to focus on banks' relative industry advantages in a county.

2.3 Measuring a County's Industry Specialization

We use the employment information provided by the QCEW to identify key industries in a county. We exclude employment by government-owned entities and the financial industry and aggregate employment at the three-digit NAICS level. An average county has 59 three-digit NAICS industries.[10] Figure 1 presents the employment shares by the top-20 industries in a county. The percentages of jobs provided by the top-ten industries in a county are: 19.35%, 11.90%, 8.85%, 7.14%, 5.98%, 5.09%, 4.41%, 3.85%, 3.43%, and 3.04%. We classify industries that provide at least 5% jobs in a county as the county's specialized industries. Our choice of 5% ensures that an industry has material impact on local economy and household income. In total, these industries provide about 58% jobs in an average county.

2.4 Measuring the Industry Expertise Channel

Using industry specialization measures for each bank and county, we classify that a bank and a county are connected through the industry expertise channel if the bank has one or more specialized industries that provide at least 5% jobs in the county. Banks can use their industry expertise to better screen eligible mortgage borrowers and monitor their income risks, allowing them to extend more mortgage credits to local residents.

Figure 2 presents the geographic distribution of counties in the contiguous U.S. that are connected with at least one bank in our sample through the industry expertise channel in years 1999, 2004, 2009, and 2014 (in orange). The figures suggest that connected counties are evenly distributed across the U.S. over the sample period.

2.5 Variable Construction

We aggregate mortgage applications at the bank-county-year level. The sample consists of 78 unique banks with mortgage businesses in 3,165 counties from 1999 to 2017. For each bank-county-year observation, we calculate the natural logarithm of the number and the dollar volume of approved mortgages to measure a bank's mortgage originations in a county. We also calculate approval rates based on both the number and the dollar volume of mortgages. The number-based (volume-based) approval rate is the number

[10] The 25^{th} percentile, the median, and the 75^{th} percentile are 50, 62 and 72, respectively.

(dollar volume) of mortgages a bank approves scaled by the number (dollar volume) of mortgage applications a bank receives in a county. To measure the dispersion in the terms of approved mortgage contracts, we estimate the standard deviations of loan amounts, loan-to-income (LTI) ratios, interest rates, and loan-to-value (LTV) ratios for mortgages a bank approves in a county (Fisman et al., 2017; Lim and Nguyen, 2020). We measure mortgage performance using mortgage delinquency and foreclosure rates: *Delinquency 60 Days*, *Delinquency 90 Days*, and *Foreclosure*. *Delinquency 60 Days* is the percentage of mortgages that are more than 60 days past due on monthly payments. *Delinquency 90 Days* is the percentage of mortgages that are more than 90 days past due on monthly payments. *Foreclosure* is the percentage of mortgages that have gone through a foreclosure. We also calculate the average interest rate of a bank's approved mortgages in a county each year. To further compare banks' originations of conventional mortgages with government-insured mortgages (i.e., FHA and VA loans), we extend the HMDA-based mortgage sample to include government-insured mortgages and calculate the percentage of conventional mortgages a bank approves in a county.

We construct three mortgage-related control variables: the average loan-to-income ratio of all mortgage applicants (*LTI*), the percentage of male applicants (*Male*), and the percentage of minority applicants (*Minority*). We further control for the average credit score (*Credit Score*), the average loan-to-value ratio (*LTV*), the average debt-to-income ratio (*DTI*), and the average interest rate (*Interest Rate*) when using the matched sample between HMDA and Fannie Mae, Freddie Mac, and McDash datasets. We also add controls of geographic distance and existing business relations for each bank-county pair in a year: the natural logarithm of one plus the number of branches a bank has in the county (*Branch*), the natural logarithm of the geographic distance between the headquarters county of a bank and the borrower's home county (*Distance*), the natural logarithm of one plus the number of small business loans a bank originates in the borrower's home county (*SBL*), and the average percentage of mortgages retained on balance sheets in the borrower's home county in the past three years (*Mortgage Exposure*). Bank-level control variables include the natural logarithm of bank assets (*Log(Assets*), total loans scaled by assets (*Total Loans/Assets*), deposits scaled by assets (*Deposits/Assets*), commercial and industrial (C&I) loans scaled by total loans (*C&I Loans/Total Loans*), real estate loans scaled by total loans (*RE Loans/Total Loans*), return on assets (*ROA*), and total liquidity scaled by assets (*Liquidity/Assets*).

2.6 Summary Statistics

Table 1 reports the summary statistics of the variables used in our empirical analyses. Panel A presents the county-level statistics; panel B presents the bank-level statistics; panel C presents the HMDA-based main sample at the bank-county level; and panel D presents the matched bank-county-level sample between HMDA and monthly loan-level performance from the Fannie Mae, the Freddie Mac and the McDash datasets. The sample period is from 1999 to 2017, except that the county-level mortgage delinquency in panel A is only available from 2008 to 2017.

The number of mortgages a bank approves in a county has a mean of 88.0 and a median of 19.0. The standard deviation is 193.8, suggesting large variations across bank-county pairs. The mean dollar volume (in millions) of approved mortgages is 14.4, and the median is 2.3. The average number-based mortgage approval rate is 74.6%, and the average volume-based mortgage approval rate is 75.6%. 16.5% of the 316,552 bank-county pairs are connected through the industry expertise channel.

3 The Industry Expertise Channel and Mortgage Lending

3.1 The Number and Volume of Approved Mortgages

We conjecture that industry expertise enhances banks' abilities to assess household income risk and therefore reduces information frictions in the mortgage lending process. Lower information asymmetry thus mitigates credit rationing, leading to more credit supply. We test this conjecture using the following empirical specification:

$$Y_{bct} = \pi_{ct} + \mu_{bs} + \beta \, Industry \; Expertise_{bct} + \delta \boldsymbol{X}_{bct} + \varepsilon_{bct} \tag{2}$$

where b denotes bank, c denotes home county of the borrower, s denotes home state of the borrower, t denotes year. Y_{bct} is the natural logarithm of the number or the dollar volume (in millions) of mortgages bank b approves to borrowers in county c in year t. $Industry\ Expertise_{bct}$ is a dummy equal to one for a bank-county pair if there exists at least one industry in which the bank b specializes and provides at least 5% jobs in county c in year t. X_{bct} is a vector of controls, including the average loan-to-income ratio of all mortgage applicants, the percentage of male applicants, the percentage of minority applicants, the natural logarithm of one plus the number of branches a bank has in the county, the natural logarithm of the geographic distance between the headquarters county of a bank and the borrower's home county, the natural logarithm of one plus the number of small business loans that a bank originates in the borrower's home county, the average fraction of mortgages retained in the balance sheets in the borrower's home county in the past three years, the natural logarithm of bank assets, total loans scaled by assets, deposits scaled by assets, commercial and industrial (C&I) loans scaled by total loans, real estate loans scaled by total loans, return on assets, and total liquidity scaled by assets. π_{ct} is county-by-year fixed effects, which allows us to compare different banks' mortgage lending in the same county. μ_{bs} is bank-by-state fixed effects, which controls for hidden links between banks and states, such as political rent-seeking (Chu and Zhang, 2022).

We present the results of estimating Equation (2) in Table 2. In column (1), the coefficient estimate on *Industry Expertise* is positive and statistically significant, indicating that industry expertise increases banks' mortgage lending. The significance remains after adding mortgage-level or bank-level controls in columns (2) and (3). We use county-by-year fixed effects to replace borrower home county and year fixed effects in column (4) and further use bank-by-state fixed effects to replace bank fixed effects in column (5). The results continue to hold. The economic effect is also significant. The estimate in column (5) suggests that industry expertise increases banks' mortgage lending by 6.3%. Columns (6) - (10) repeat the analyses using the dollar volume of approved mortgages as the dependent variable. The results are consistent with those in columns (1) - (5).

3.2 Robustness

In Internet Appendix A, we conduct a series of additional tests to show that our baseline results are robust when using alternative measures, more fixed effects, and alternative model specifications. Specifically, we first reconstruct our measure of the industry expertise channel by taking into account each borrowing firm's market position, assuming that lending to industry leaders allows banks to better acquire industry expertise. Second, we construct two continuous measures that capture the intensity of the connections between banks and counties through the channel. One is based on the fraction of a county's residents working in industries that a bank specializes in and provides at least 5% jobs in the county. The other is based on the fraction of a county's residents working in any industry in which a bank specializes, regardless of the number of jobs provided in the county. Third, we reconstruct the measure to quantify the level of a bank's industry expertise, which is calculated as the difference between a bank's loan share in an industry minus the threshold $L^*_{i,t}$ to identify an outlier loan share in equation (1). We also control for time-varying bank-level characteristics with bank-by-year fixed effects and time-invariant links between banks and counties using bank-by-county fixed effects. The results hold. Last, we use both linear regression models and the fixed effects Poisson model (Cohn, Liu, and Wardlaw, 2022) to address concerns using the log form of the dependent variables. The results remain statistically significant and the economic effects are even larger.

In Internet Appendix B, we exclude an alternative channel - banks prioritizing lending mortgages to employees of their corporate borrowers. Specifically, we drop a bank-state pair from the HMDA mortgage sample if a bank's syndicated loan borrower is located in the state or has a major establishment/subsidiary reported in the state in 10-K filings. Our results hold.

3.3 Mortgage Approval Rates

Although the results hold after controlling for county-by-year and bank-by-state fixed effects, the results could still be driven by demand-side factors. For example, certain households may prefer to borrow from a bank for reasons such as brand preferences or the availability and usability of mobile phone apps. These factors are either not observable or not directly measurable and therefore are not controlled in our baseline model specification. To alleviate this concern, we examine banks' mortgage approval decisions conditional on received applications. Specifically, we use a bank's mortgage approval rate as the dependent variable, defined as the number (dollar volume) of mortgages approved scaled by the number (dollar volume) of mortgage applications received. The results are reported in Table 3. We find that, conditional on received applications, industry expertise significantly increases the number- and volume-based approval rates by 40 basis points. The evidence suggests that the findings in Table 2 are unlikely to be driven by demand-side factors.

Overall, the results support the conjecture that the industry expertise channel mitigates information asymmetry, and hence credit rationing.

4 The Information Mechanism

We hypothesize that banks use the industry expertise channel in mortgage lending because it provides credible soft information for banks to better assess income risk at origination. In this section, we provide six pieces of consistent evidence to support the information mechanism of the channel.

4.1 Industry Growth and Household Income and Mortgage Delinquency

A prerequisite for the industry expertise channel is that the conditions of the key industries in a county are useful in assessing borrower credit quality; that is, banks can use their industry expertise to assess local borrowers' income dynamics and mortgage default probabilities. We therefore test whether this is true using the following empirical model:

$$Y_{ct} = \theta_c + \tau_t + \beta \, Sales\ Growth_{ct} + \delta \boldsymbol{X}_{ct} + \varepsilon_{ct} \tag{3}$$

where c denotes county, t denotes year. Y_{ct} is the dependent variable, the income growth rate or the annual change in the mortgage delinquency rate. $Sales\ Growth_{ct}$ is the standardized employment-weighted industry sales growth rate in a county. The weights are the fractions of local residents working in a given industry. The sales growth rates for each industry are estimated using the sales of all U.S. public firms in the industry. $\boldsymbol{X}_{ct}$ is a vector of county-level controls, including the natural logarithm of the population, the percentage of the population over 65, the percentage of the male population, the percentage of the minority population, and the percentage of the population with a bachelor's degree or above. θ_c is county fixed effects and τ_t is year fixed effects.

The results are presented in Table 4. The dependent variable in columns (1) - (3) is the income growth rate. The coefficient estimate on sales growth is positive and statistically significant, suggesting that faster industry growth is associated with greater growth in household income. The correlation is also economically significant. In column (3), a one standard deviation increase in sales growth is associated with a 14.9% increase in household income growth. In columns (4) - (6), we examine the mortgage delinquency rate, an indicator of mortgage performance.[11] Consistent with our expectation, the coefficient estimate on sales growth is negative and statistically significant, suggesting that faster industry growth is associated with lower mortgage delinquency rates.

[11] The sample is much smaller because the data on mortgage delinquency rate from the CFPB only covers 470 counties per year from 2008. The data is based on a nationally representative five percent sample of closed-end, first-lien, 1–4 family residential mortgages.

Overall, the findings suggest that the growth of a county's key industries positively correlates with the county's household income growth and negatively correlates with the county's mortgage delinquency rates. The evidence builds the foundation for the key argument in this paper: industry expertise enables banks to predict local household income dynamics and, therefore, mortgage default risks after origination.

4.2 Information Asymmetry

We then investigate the effect of information asymmetry on banks' use of the industry expertise channel in mortgage lending. We start with the geographic distance between banks' headquarters and mortgage borrowers. Previous studies show that long geographical distance erodes banks' abilities in information acquisition, creating significant barriers for banks to reach distant borrowers (Agarwal and Hauswald, 2010; Hollander and Verriest, 2016). We expect that industry expertise mitigates information barriers and enables banks to extend mortgage credits to distant borrowers. We test this prediction in columns (1) and (3) of Table 5. Consistent with the literature, mortgage credits decrease significantly with the distance between banks' headquarters and borrowers. More importantly, the effect of industry expertise increases with distance. The economic magnitude is also large. The effects of the industry expertise channel on mortgage lending more than double for a one standard deviation increase in distance. The evidence suggests that industry expertise reduces distance-induced information frictions between banks and mortgage borrowers.

We also examine how soft information embedded in industry expertise interacts with soft information banks collect from other sources. To this end, we use social networks as a proxy for alternative soft information in columns (2) and (4) of Table 5. Consistent with Rehbein and Rother (2020), social connections between a bank's headquarters county and a borrower's home county significantly increase banks' mortgage lending. However, social connections decrease banks' reliance on industry expertise. The estimate in column (2) suggests that a one standard deviation increase in SCI is associated with a 31.3% decrease in the effect of industry expertise. The evidence further suggests that industry expertise provides additional soft information, which could substitute for soft information from social connections.

4.3 Borrower Risk

Credit rationing caused by information asymmetry should be more severe for ex ante riskier borrowers. We therefore expect that the impact of industry expertise should be stronger for riskier borrowers. Our first proxy for borrower risk is the local house price volatility. Large volatility increases the downside risk of house prices and hence mortgage default risk (Gerardi et al., 2018). We report the results in columns (1) and (3) of Table 6. Consistent with our expectation, the coefficient estimates on the interaction term between *Industry Expertise* and *HP Volatility*, the standardized county-level housing price volatility, are positive and statistically significant. This result suggests that banks use their industry expertise more when local house prices become more volatile. The estimate in column (1) suggests that the effect of industry expertise on mortgage lending increases from 6.7% to 11.4% for a one standard deviation increase in local house price volatility.

Our second proxy for borrower risk is the LTI ratio. A higher LTI ratio indicates higher mortgage leverage and is associated with lower mortgage affordability and higher borrowing constraints (Campbell and Cocco, 2015). The results in columns (2) and (4) of Table 6 show that the effect of industry expertise is stronger for borrowers with higher LTI ratios. The estimate in column (2) suggests that the effect of industry expertise increases from 5.3% to 11.1% for a one standard deviation increase in the LTI ratio.

4.4 Soft Information in Mortgage Contracts

To provide more direct evidence on the information mechanism, we test soft information contained in mortgage contracts by examining whether mortgages originated through the industry expertise channel are less standardized, that is, greater dispersion in contractual terms. This is because better information allows banks to better distinguish between “good" and “bad" borrowers (Cornell and Welch, 1996; Athreya, Tam, and Young, 2012; Rajan, Seru, and Vig, 2015). As a result, banks can grant mortgages with favorable terms to “good" borrowers and mortgages with strict terms to “bad" borrowers. In contrast, if banks do not have much information to evaluate borrowers, they can only design mortgage terms based on the average quality of all mortgage borrowers and thus originate loans with similar terms.

We follow the literature and construct four variables to capture the dispersion in the terms of approved mortgage contracts: the natural logarithm of the standard deviations of the loan amounts, loan-to-income ratios, interest rates, and loan-to-value ratios (Fisman et al., 2017; Lim and Nguyen, 2020). We present the results in Table 7. Consistent with our prediction, mortgages originated through the industry expertise channel have less standardized contractual terms, i.e., more dispersed loan amounts, LTI ratios, interest rates, and LTV ratios. In particular, the standard deviations of loan amounts, LTI ratios, interest rates, and LTV ratios are 0.6%, 0.5%, 2.1% and 2.2% higher for mortgages originated through the channel.[12]

4.5 Conventional and Government-Insured Mortgages

Government-insured mortgages, i.e., FHA and VA loans, are less subject to credit rationing (Duca and Rosenthal, 1991; Gabriel and Rosenthal, 1991; Ambrose, Pennington-Cross, and Yezer, 2002). Banks should therefore originate more conventional mortgages relative to government-insured mortgages in counties connected by the industry expertise channel if industry expertise truly alleviates credit rationing. To test, we re-estimate Equation (2) by extending the HMDA-based mortgage sample to include government-insured mortgages and replacing the dependent variable with the percentage of conventional loans a bank originates in a county.

[12]We obtain similar results (untabulated) using the natural logarithm of the interquartile ranges of the four contractual terms.

The results are presented in Table 8, with columns (1) - (3) for the number-based percentage of conventional mortgages and columns (4) - (6) for the volume-based percentage of conventional mortgages. The coefficient estimates on *Industry Expertise* are all positive and statistically significant, suggesting that banks increase conventional mortgage lending relative to government-insured mortgages in counties connected by the industry expertise channel, consistent with the argument that banks' industry expertise mitigates credit rationing.

4.6 Mortgage Performance

Lastly, we examine how industry expertise affects *ex post* mortgage performance. If industry expertise provides useful soft information for banks to better screen mortgage applicants and monitor their income risks, we expect industry expertise to positively affect mortgage performance.

To test the performance implications, we focus on mortgage delinquency and foreclosure rates using the matched sample between HMDA and monthly loan-level performance from the Fannie Mae, Freddie Mac, and McDash datasets. Specifically, we track each mortgage's monthly payment records to identify whether a mortgage ever had a 60-day-plus delinquency, a 90-day-plus delinquency, or a foreclosure. We aggregate the loan-level records at the bank-county-year level and construct three variables: *Delinquency 60 Days*, *Delinquency 90 Days*, and *Foreclosure*. *Delinquency 60 Days* is the percentage of mortgages that are more than 60 days past due on monthly payments. *Delinquency 90 Days* is the percentage of mortgages that are more than 90 days past due on monthly payments. *Foreclosure* is the fraction of mortgages that have gone through a foreclosure. We further match the measures to the HMDA-based sample of banks and counties in our baseline analyses.

The results are presented in Table 9. The coefficient estimates on the *Industry Expertise* are all negative and statistically significant, suggesting a negative effect of industry expertise on subsequent mortgage delinquency and foreclosure rates. The economic magnitudes are also large. On average, mortgages originated by banks with industry expertise have 4.1% lower 60-day-plus delinquency rates, 4.0% lower 90-day-plus delinquency rates, and 4.8% lower foreclosure rates, respectively.

In summary, the analyses in this section show that banks rely more on the industry expertise channel in mortgage lending when they face significant barriers in acquiring borrowers' information. The channel becomes more important when the borrowers are riskier. Additionally, mortgages originated through the channel contain more soft information, i.e., more dispersed mortgage terms. Furthermore, banks originate more conventional mortgages relative to government-insured mortgages in counties connected by the industry expertise channel to alleviate credit rationing. Finally, the channel reduces mortgage delinquency and foreclosure rates. Together, these findings provide strong support for the information mechanism of the industry expertise channel.

5 Addressing Endogeneity using Two Types of Shocks

The results above provide consistent evidence that industry expertise provides credible soft information that facilitates bank mortgage lending. The results hold after including county-by-year, bank-by-state, and bank-by-year fixed effects, indicating that the findings are not driven by county-level or bank-level heterogeneities. However, the results could still be biased by omitted variables at the bank-by-county level. In addition, results could be due to reverse causalities - banks may strategically choose which industries to specialize in according to the expansions of their mortgage business. To alleviate the endogeneity concerns, we design two empirical tests using two shocks that are plausibly exogenous to a given bank's use of industry expertise in mortgage lending.

5.1 Industry Distress

We first design a difference-in-differences test using unexpected industry-wide distress. Industry distress is a serious decline in the industry accompanied by a great deal of uncertainty, leading to considerable strain on company performance. The industry-level and firm-level negative performance can be transmitted to the income of households working in distressed industries or in counties where these distressed industries concentrate. In the worst case, households may get fired and lose all income during industry distress.

Relevant industry expertise allows banks to better evaluate the duration and severity of distress and its impact on mortgage risk. Therefore, banks with industry expertise can better price the income risks of affected mortgage borrowers. These banks can avoid large-scale defaults by timely selling mortgages to third parties such as Fannie Mae and Freddie Mac. The positive effects of industry expertise on mortgage lending should thus be more pronounced for distressed industries. More importantly, industry-wide shocks are plausibly exogenous for any given bank, county, or mortgage borrower, mitigating the issues of omitted variables and reverse causality (Giannetti and Saidi, 2019; Babina, 2020).

We measure industry distress following previous studies (Opler and Titman, 1994; Babina, 2020). Specifically, we classify a three-digit NAICS industry as distressed in a year if the industry-level two-year sales growth is negative and the industry-level two-year stock return is less than –10% from the beginning of that year. For robustness checks, we also use two additional stock return thresholds: –20% and –30%. We then compare the effects of industry distress on mortgage lending with differential ex-ante industry specializations using the following model:

$$\begin{aligned} Y_{bct} = \pi_{ct} + \mu_{bs} + \tau_{bt} + \beta_1\, \textit{Industry Expertise}_{bct-2} \times \textit{Distress}_{bct-1} \\ + \beta_2 \textit{Industry Expertise}_{bct-2} + \delta \boldsymbol{X}_{bct-2} + \varepsilon_{bct} \end{aligned} \tag{4}$$

where b denotes bank, c denotes borrower home county, s denotes borrower home state, t denotes year. Y_{bct} is the dependent variable, the natural logarithm of the number or the dollar volume of mortgages (in millions) bank b approves to borrowers in county c in year t. $\textit{Industry Expertise}_{bct-2}$ is a dummy variable equal to one for a bank-county pair if there exists at least one industry in which bank b specializes and provides at least 5%

jobs in county c, measured at $t-2$. $Distress_{bct-1}$ is a dummy that equals one for a bank-county pair if distress happens in any of the industries in which bank b specializes and provides at least 5% jobs in county j, measured at $t-1$.[13] In addition to county-by-year fixed effects π_{ct} and bank-by-state fixed effects μ_{bs}, we also add bank-by-year fixed effects τ_{bt} to account for potential negative effects of industry distress on bank capital.

The results are reported in Table 10. Columns (1) & (4) are based on the return threshold of -10%, columns (2) & (5) are based on the return threshold of -20%, and columns (3) & (6) are based on the return threshold of -30%. The coefficient estimates on the interaction term between *Industry Expertise* and *Distress* are positive and statistically significant, suggesting that banks rely more on their industry expertise in distress periods. In column (1), the effect of industry expertise on mortgage lending increases from 2% in non-distress periods to 6.4% in distress periods. In addition, the incremental effect is larger in more distressed scenarios. Moving from the -10% to the -30% return threshold, the incremental effect changes from 4.4% to 5.6%, a 27% increase. Similar patterns hold for the volume-based measure of mortgage lending.

5.2 The 2008 Financial Crisis

We also use the 2008 financial crisis as another shock. Mortgages and housing markets were at the center of the recession. National house prices dropped by more than 10% from 2007 to 2009. The average delinquency rate on single-family residential mortgages rose from 1.84% in the pre-crisis period (2004 - 2007) to 7.04% during the crisis period (2008 - 2009).[14] These widespread mortgage defaults eventually led to significant losses for banks. Importantly, fraudulently overstated income in mortgage applications contributed significantly to the oversupply of mortgages to marginal borrowers (Mian and Sufi, 2017). We expect banks to be more cautious in screening mortgage borrowers and

[13] We intentionally measure the industry expertise channel at $t-2$ and industry distress at $t-1$ to avoid the concern that industry distress may affect banks' loan originations and thus choices of industry specialization.

[14] Estimated using data on housing price indexes and mortgage delinquency rates from the website of the Federal Reserve Bank of St. Louis.

extending credit during and after the crisis. Therefore, industry expertise should become more valuable in mortgage underwriting.

We design a difference-in-differences test to examine how the crisis affects banks' use of industry expertise in mortgage lending around the crisis period. The event window is from 2004 to 2010[15] and we estimate the following specification:

$$\begin{aligned} Y_{bct} = \pi_{ct} + \mu_{bs} + \tau_{bt} + \beta_1 \, \textit{Industry Expertise}_{bc2003} \times \textit{Crisis}_t \\ + \beta_2 \, \textit{Industry Expertise}_{bc2003} + \delta \mathbf{X}_{bct} + \varepsilon_{bct} \end{aligned} \tag{5}$$

where b denotes bank, c denotes borrower home county, s denotes borrower home state, t denotes year. Y_{bct} is the dependent variable, the natural logarithm of the number or the dollar volume of mortgages (in millions) bank b approves to borrowers in county c in year t. $\textit{Industry Expertise}_{bc2003}$ is an indicator variable equal to one for a bank-county pair if there exists at least one industry in which bank b specializes and provides at least 5% jobs in county c, measured in 2003. *Crisis* is an indicator variable equal to zero for the period 2004-2007 and one for the period 2008-2010. π_{ct} denotes county-by-year fixed effects, μ_{bs} denotes bank-by-state fixed effects, and τ_{bt} denotes bank-by-year fixed effects.

Table 11 presents the results. The estimates in columns (1) and (3) show that industry expertise positively affects mortgage lending in the pre-crisis period. More importantly, banks' reliance on the channel significantly increases during the crisis period. The economic magnitude increases from 3.8% to 12.3% in column (1), and from 4.7% to 12% in column (3). In columns (2) and (4), we further break down the *Crisis* dummy into year dummies. Year 2007 is the base year and thus omitted. The coefficient estimates on the interaction terms *Industry Expertise* × *Year 2004*, *Industry Expertise* × *Year 2005*, and *Industry Expertise* × *Year 2006* are not statistically significant, suggesting that the effect of industry expertise on mortgage lending are relatively stable during the pre-crisis period. Since 2009, the effect of industry expertise increases significantly by 14.2% and 12.6% in

[15] The crisis ended in 2009. We include 2010 in the sample because our goal is to assess banks' use of the channel before, during, and after the crisis. For simplicity, we use "crisis" to represent the period 2008 - 2010.

columns (2) and (4). The effect slightly decreases in 2010 after the worst time, but is positive and statistically significant. We also plot the dynamics of the coefficient estimates in Figure 3.

To summarize, these findings suggest that the effect of industry expertise becomes more important in times of chaos, during which household income risk becomes much more salient. More importantly, these tests alleviate the endogeneity concerns outlined above, implying that the effects of industry expertise on mortgage lending are likely to be causal.

6 Conclusion

In this paper, we show that industry knowledge banks gain from corporate lending help them overcome informational frictions in mortgage markets. In particular, we show that banks specialized in certain industries increase mortgage lending in areas concentrated with those industries, which we call the industry expertise channel. We find that the effect of the channel is more pronounced when information asymmetry is more severe or borrowers are riskier. We also find that mortgages originated through the channel contain more soft information and perform better. Further analyses based on unexpected industry distress suggest that the effects are likely causal. Overall, our work demonstrates a broader impact of banks' lending concentration at the industry level. The industry expertise developed through lending concentration benefits banks in corporate lending and mortgage lending. Our paper also shows that information could flow from the corporate lending division to the mortgage lending division within a bank.

References

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

Agarwal, S. and R. Hauswald (2010). Distance and private information in lending. *Review of Financial Studies 23*(7), 2757–2788.

Ambrose, B. W., A. Pennington-Cross, and A. M. Yezer (2002). Credit rationing in the us mortgage market: Evidence from variation in fha market shares. *Journal of Urban Economics 51*(2), 272–294.

Athreya, K., X. S. Tam, and E. R. Young (2012). A quantitative theory of information and unsecured credit. *American Economic Journal: Macroeconomics 4*(3), 153–183.

Babina, T. (2020). Destructive creation at work: How financial distress spurs entrepreneurship. *Review of Financial Studies 33*(9), 4061–4101.

Bailey, M., R. Cao, T. Kuchler, J. Stroebel, and A. Wong (2018). Social connectedness: Measurement, determinants, and effects. *Journal of Economic Perspectives 32*(3), 259–80.

Bailey, M., E. Dávila, T. Kuchler, and J. Stroebel (2019). House price beliefs and mortgage leverage choice. *Review of Economic Studies 86*(6), 2403–2452.

Beck, T., O. De Jonghe, and K. Mulier (2021). Bank sectoral concentration and risk: Evidence from a worldwide sample of banks. *Journal of Money Credit and Banking*.

Berger, A. N. and G. F. Udell (1995). Relationship lending and lines of credit in small firm finance. *Journal of Business 110*(2), 351–381.

Berger, P. G., M. Minnis, and A. Sutherland (2017). Commercial lending concentration and bank expertise: Evidence from borrower financial statements. *Journal of Accounting and Economics 64*(2-3), 253–277.

Bharath, S., S. Dahiya, A. Saunders, and A. Srinivasan (2007). So what do I get? The bank's view of lending relationships. *Journal of Financial Economics 85*(2), 368–419.

Blickle, K., C. Parlatore, and A. Saunders (2021). Specialization in banking. *NBER Working Paper 31077*.

Campbell, J. Y. and J. F. Cocco (2015). A model of mortgage default. *Journal of Finance 70*(4), 1495–1554.

Chakraborty, I., I. Goldstein, and A. MacKinlay (2018). Housing price booms and crowding-out effects in bank lending. *Review of Financial Studies 31*(7), 2806–2853.

Chava, S. and M. R. Roberts (2008). How does financing impact investment? The role of debt covenants. *Journal of Finance 63*(5), 2085–2121.

Chavaz, M. and A. K. Rose (2019). Political borders and bank lending in post-crisis america. *Review of Finance 23*(5), 935–959.

Chodorow-Reich, G. (2014). The employment effects of credit market disruptions: Firm-level evidence from the 2008–9 financial crisis. *Quarterly Journal of Economics 129*(1), 1–59.

Chu, Y., X. F. Ma, and T. Zhang (2022). Bank public status and the racial gap in mortgage pricing. *Available at SSRN 3461061*.

Chu, Y. and T. Zhang (2022). Political influence and banks: Evidence from mortgage lending. *Journal of Financial Intermediation 52*, 100982.

Cohn, J. B., Z. Liu, and M. I. Wardlaw (2022). Count (and count-like) data in finance. *Journal of Financial Economics 146*(2), 529–551.

Cornell, B. and I. Welch (1996). Culture, information, and screening discrimination. *Journal of Political Economy 104*(3), 542–571.

Cortés, K. R. and P. E. Strahan (2017). Tracing out capital flows: How financially integrated banks respond to natural disasters. *Journal of Financial Economics 125*(1), 182–199.

Dagher, J. and K. Kazimov (2015). Banks' liability structure and mortgage lending during the financial crisis. *Journal of Financial Economics 116*(3), 565–582.

Duca, J. V. and S. S. Rosenthal (1991). An empirical test of credit rationing in the mortgage market. *Journal of Urban Economics 29*(2), 218–234.

Elul, R., N. S. Souleles, S. Chomsisengphet, D. Glennon, and R. Hunt (2010). What "triggers" mortgage default? *American Economic Review 100*(2), 490–94.

Ergungor, O. E. (2010). Bank branch presence and access to credit in low-to moderate-income neighborhoods. *Journal of Money, Credit and Banking 42*(7), 1321–1349.

Fisman, R., D. Paravisini, and V. Vig (2017). Cultural proximity and loan outcomes. *American Economic Review 107*(2), 457–92.

Foote, C., K. Gerardi, L. Goette, and P. Willen (2010). Reducing foreclosures: No easy answers. *NBER Macroeconomics Annual 24*(1), 89–138.

Fuster, A., P. Goldsmith-Pinkham, T. Ramadorai, and A. Walther (2022). Predictably unequal? The effects of machine learning on credit markets. *Journal of Finance 77*(1), 5–47.

Fuster, A., M. Plosser, P. Schnabl, and J. Vickery (2019). The role of technology in mortgage lending. *Review of Financial Studies 32*(5), 1854–1899.

Gabriel, S. A. and S. S. Rosenthal (1991). Credit rationing, race, and the mortgage market. *Journal of Urban Economics 29*(3), 371–379.

Garmaise, M. J. (2015). Borrower misreporting and loan performance. *Journal of Finance 70*(1), 449–484.

Gerardi, K., K. F. Herkenhoff, L. E. Ohanian, and P. S. Willen (2018). Can't pay or won't pay? Unemployment, negative equity, and strategic default. *Review of Financial Studies 31*(3), 1098–1131.

Gerardi, K., P. S. Willen, and D. H. Zhang (2023). Mortgage prepayment, race, and monetary policy. *Journal of Financial Economics 147*(3), 498–524.

Giannetti, M. and F. Saidi (2019). Shock propagation and banking structure. *Review of Financial Studies 32*(7), 2499–2540.

Gilje, E. P., E. Loutskina, and P. E. Strahan (2016). Exporting liquidity: Branch banking and financial integration. *Journal of Finance 71*(3), 1159–1184.

Giometti, M. and S. Pietrosanti (2020). Bank specialization and the design of loan contracts. *Unpublished working paper*.

Gomez, M., A. Landier, D. Sraer, and D. Thesmar (2020). Banks' exposure to interest rate risk and the transmission of monetary policy. *Journal of Monetary Economics*.

Gopal, M. (2019). How collateral affects small business lending: The role of lender specialization. *Unpublished working paper*.

Gustafson, M. T., I. T. Ivanov, and R. R. Meisenzahl (2021). Bank monitoring: Evidence from syndicated loans. *Journal of Financial Economics 139*(2), 452–477.

Hodge, V. and J. Austin (2004). A survey of outlier detection methodologies. *Artificial Intelligence Review 22*, 85–126.

Hollander, S. and A. Verriest (2016). Bridging the gap: the design of bank loan contracts and distance. *Journal of Financial Economics 119*(2), 399–419.

Irani, R. M., R. Iyer, R. R. Meisenzahl, and J.-L. Peydro (2021). The rise of shadow banking: Evidence from capital regulation. *Review of Financial Studies 34*(5), 2181–2235.

Lim, I. and D. D. Nguyen (2020). Hometown lending. *Journal of Financial and Quantitative Analysis*, 1–65.

Loutskina, E. and P. E. Strahan (2011). Informed and uninformed investment in housing: The downside of diversification. *Review of Financial Studies 24*(5), 1447–1480.

Mian, A. and A. Sufi (2017). Fraudulent income overstatement on mortgage applications during the credit expansion of 2002 to 2005. *Review of Financial Studies 30*(6), 1832–1864.

Opler, T. C. and S. Titman (1994). Financial distress and corporate performance. *Journal of Finance 49*(3), 1015–1040.

Paravisini, D., V. Rappoport, and P. Schnabl (2023). Specialization in bank lending: Evidence from exporting firms. *Journal of Finance 78*(4), 2049–2085.

Petersen, M. A. and R. G. Rajan (1995). The effect of credit market competition on lending relationships. *Quarterly Journal of Economics 110*(2), 407–443.

Rajan, U., A. Seru, and V. Vig (2015). The failure of models that predict failure: Distance, incentives, and defaults. *Journal of Financial Economics 115*(2), 237–260.

Rehbein, O. and S. Rother (2020). Social connectedness (and distance) in bank lending. *Unpublished working paper*.

Saidi, F. and D. Streitz (2021). Bank concentration and product market competition. *Review of Financial Studies 34*(10), 4999–5035.

Schwert, M. (2018). Bank capital and lending relationships. *Journal of Finance 73*(2), 787–830.

Stiglitz, J. E. and A. Weiss (1981). Credit rationing in markets with imperfect information. *American Economic Review 71*(3), 393–410.

Sufi, A. (2007). Information asymmetry and financing arrangements: Evidence from syndicated loans. *Journal of Finance 62*(2), 629–668.

Appendix A. Variable Definitions

Variables	Description
Dependent Variables	
Log(Number of Approved Mortgages)	The natural logarithm of the number of mortgages a bank approves in a county.
Log(Volume of Approved Mortgages)	The natural logarithm of the dollar volume of mortgages (in millions) a bank approves in a county.
Approval Rate - Number	The number of mortgages a bank approves scaled by the number of mortgage applications a bank receives in a county.
Approval Rate - Volume	The dollar volume (in millions) of mortgages a bank approves scaled by the dollar volume of mortgage applications a bank receives in a county.
Income Growth (%)	A county's household income growth rate (%).
Delta Delinquency Rate (%)	The annual change in a county's 1-4 family residential mortgage delinquency rate (%).
Log(STD. Mortgage Size)	The natural logarithm of the standard deviation of the amounts of approved mortgages.
Log(STD. LTI)	The natural logarithm of the standard deviation of the loan-to-income (LTI) ratios of approved mortgages.
Log(STD. Interest Rates)	The natural logarithm of the standard deviation of the interest rates of approved mortgages.
Log(STD. LTV)	The natural logarithm of the standard deviation of the loan-to-value (LTV) ratios of approved mortgages.
Delinquency 60 Days	The percentage of mortgages that are more than 60 days past due on monthly payments.
Delinquency 90 Days	The percentage of mortgages that are more than 90 days past due on monthly payments.
Foreclosure	The percentage of mortgages that hat have gone through a foreclosure.
% Conventional Mortgages	The number-based (or volume-based) percentage of conventional mortgages a bank approves in a county.
Key Independent Variables	
Industry Expertise	A dummy that equals one for a bank-county pair if there exists at least one industry which a bank specializes in and provides at least 5% jobs in a county.
	The rank is from 1 to 10, with 10 indicating the highest ratio.
Sales Growth	The standardized employment-weighted industry-level sales growth rate in a county. The sales growth rate for each industry is calculated as the average sales growth rate of all public U.S firms in the industry.
Distress	A dummy that equals one for a bank-county pair if distress happens in any of the industries that a bank specializes in and provide at least 5% jobs in a county.
Crisis	A dummy that equals one for the period 2008 - 2010 and zero for the period 2004 - 2007.

Variables	Description
Other Independent Variables	
LTI	The average of the loan-to-income (LTI) ratios of mortgage applicants.
Male	The fraction of mortgage applicants that are male.
Minority	The fraction of mortgage applicants that are minorities.
Credit Score	The average credit score of approved mortgages.
LTV	The average loan-to-value (LTV) ratio of approved mortgages.
DTI	The average debt-to-income ratio (DTI) ratio of approved mortgages.
Interest Rate	The average interest rate of approved mortgages.
Branch	The logarithm of one plus the number of branches a bank has in a county.
Distance	The natural logarithm of one plus the geographic distance between a mortgage borrower's home county and a bank's headquarter county.
SBL	The natural logarithm of one plus the number of small business loans a bank lends out in a county.
Mortgage Exposure	The average fraction of mortgages retained on balance sheets in the borrower's home county in the past. three years.
SCI	The standardized social connectedness index between a mortgage borrower's home county and a bank's headquarter county.
HP Volatility	The standardized county-level house price volatility, based on a county's housing prices in the past five years.
Log(Assets)	The natural logarithm of bank assets.
Total Loans/Assets	Total loans scaled by assets.
Deposits/Assets	Total deposits scaled by assets.
C&I Loans/Total Loans	Commercial & industrial (C&I) loans scaled by total loans.
RE Loans/Total Loans	Real estate loans scaled by total loans.
ROA	Total income scaled by assets.
Liquidity/Assets	The sum of total investment securities, total assets held in trading accounts, and federal funds sold and securities purchased under agreements to resell scaled by assets.
Population	The natural logarithm of the population in a county.
Above 65	The fraction of the population above 65 in a county.
Male	The fraction of the male population in a county.
Minority	The fraction of the minority population in a county.
Bachelor	The fraction of the population with a bachelor's degree or above in a county.

Figures

Figure 1. Average Employment Share by Top-20 Industries in a County

The figure presents the average employment share by top-20 industries in a county.

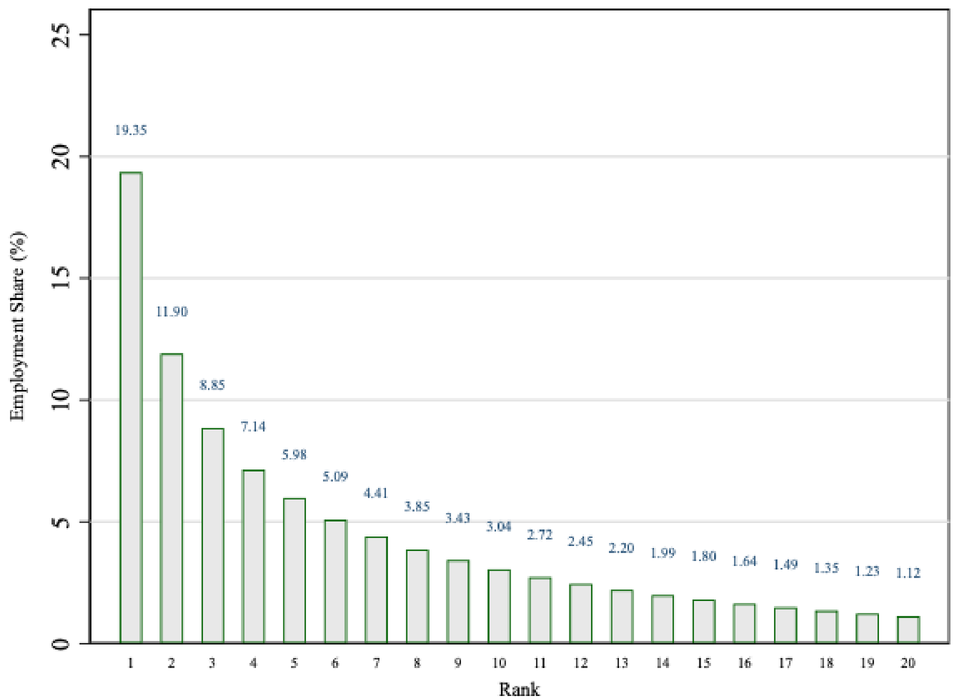
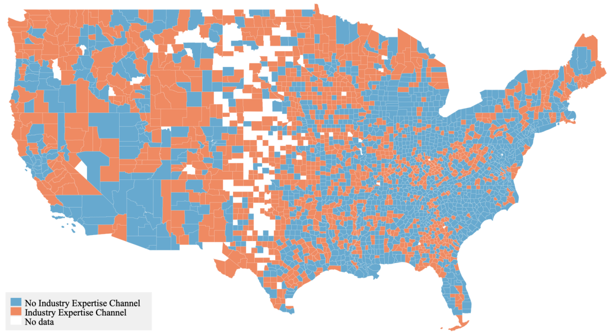


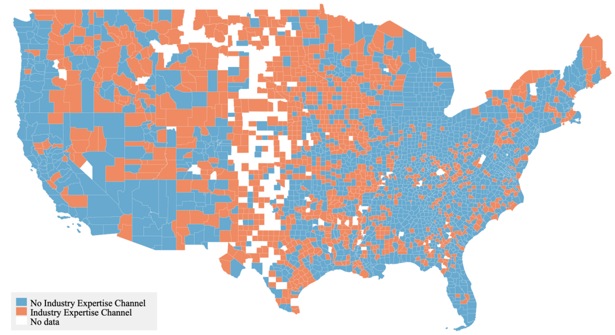
Figure 2. The Distribution of Counties Connected with Banks through the Industry Expertise Channel

The figures present the geographic distribution of counties in the contiguous U.S. that are connected with at least one bank in our sample through the industry expertise channel in the years 1999, 2004, 2009, and 2014 (in orange). Counties in blue denote those without such connections. Counties in white denote those where banks in our sample do not have mortgage businesses. A bank and a county are connected if there exists at least one industry which a bank specializes in and provides at least 5% jobs in a county.

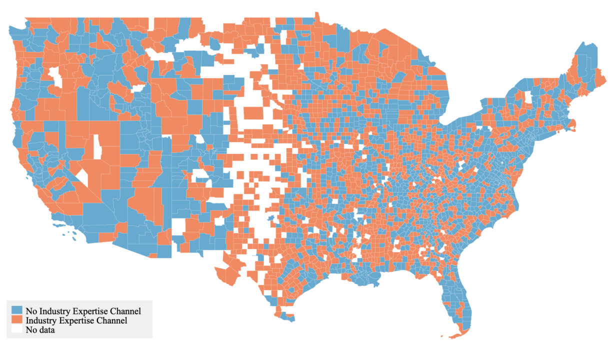
(A) Year 1999



(B) Year 2004



(C) Year 2009



(D) Year 2014

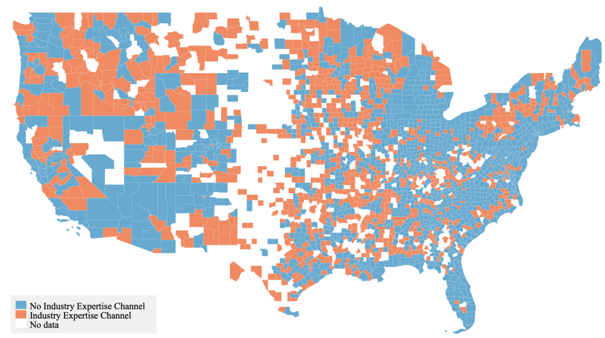
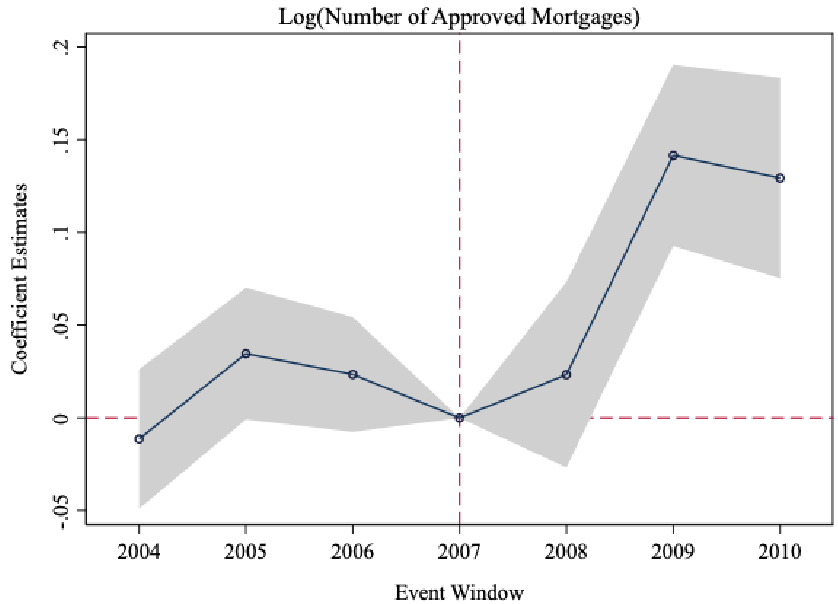


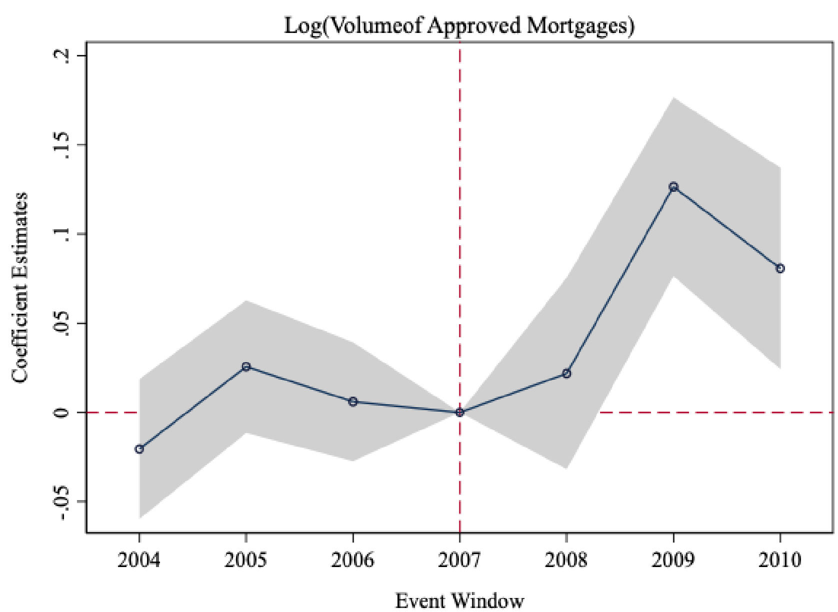
Figure 3. The 2008 Financial Crisis and the Industry Expertise Channel

The figures present the dynamic treatment effects of the 2008 financial crisis on banks' use of the industry expertise channel in mortgage lending. Figures (A) and (B) present the effects on the number and volume of approved mortgages, respectively. The regression results behind the figures are reported in columns (2) and (4) of Table 11.

(A) Number of Approved Mortgages



(B) Volume of Approved Mortgages



Tables

Table 1. Summary Statistics

This table presents the summary statistics of variables used in empirical analyses. Panel A presents the county-level statistics. Panel B presents the bank-level statistics. Panel C presents the HMDA-based main sample at the bank by county level. Panel D presents the matched bank-county-level sample between HMDA and monthly loan-level performance from the Fannie Mae, Freddie Mac, and McDash datasets. The sample period is 1999 to 2017, except that the data on county-level mortgage delinquency is from 2008 to 2017. See Appendix A for variable definitions.

	N	Mean	SD	P25	P50	P75
Panel A. County Level						
Income Growth (%)	58,394	3.926	4.809	1.521	3.848	6.205
Delta Mortgage Delinquency (%)	4,230	-0.209	1.088	-0.800	-0.383	0.117
Sale Growth	60,857	-0.003	0.998	-0.496	-0.079	0.449
Population	58,395	10.267	1.381	9.321	10.151	11.097
Above 65	59,322	0.112	0.031	0.090	0.109	0.131
Male	59,322	0.498	0.017	0.488	0.495	0.503
Minority	59,297	0.128	0.152	0.023	0.059	0.173
Bachelor	57,607	0.171	0.077	0.116	0.151	0.205
Panel B. Bank Level						
Log(Assets)	592	11.020	1.333	10.055	10.832	11.838
Total Loans/Assets	592	0.631	0.114	0.562	0.659	0.719
Deposits/Assets	592	0.719	0.076	0.667	0.724	0.772
C&I Loans/Total Loans	592	0.246	0.080	0.188	0.239	0.292
RE Loans/Total Loans	592	0.521	0.141	0.420	0.515	0.649
ROA	592	0.010	0.007	0.008	0.011	0.013
Liquidity/Assets	592	0.230	0.095	0.161	0.211	0.292

	N	Mean	SD	P25	P50	P75
Panel C. Bank-County Level (HMDA)						
Number of Approved Mortgages	316,552	87.987	193.768	7.000	19.000	66.000
Log(Number of Approved Mortgages)	316,552	3.196	1.503	1.946	2.944	4.190
Volume of Approved Mortgages	316,552	14.407	34.847	0.820	2.330	9.142
Log(Volume of Approved Mortgages)	316,552	1.102	1.699	-0.198	0.846	2.213
Approval Rate-Number	316,552	0.746	0.167	0.638	0.775	0.867
Approval Rate-Volume	316,552	0.756	0.174	0.649	0.786	0.889
Log(STD. LTI)	316,551	4.283	0.643	3.859	4.266	4.678
Log(STD. Mortgage Size)	314,818	-0.049	0.377	-0.246	-0.016	0.189
Industry Expertise	316,552	0.165	0.371	0.000	0.000	0.000
LTI	316,552	2.042	0.553	1.653	1.991	2.377
Male	315,054	0.731	0.136	0.653	0.734	0.812
Minority	316,552	0.096	0.128	0.000	0.048	0.143
Branch	316,552	0.400	0.749	0.000	0.000	0.693
Distance	316,552	6.285	0.956	5.659	6.352	7.003
SBL	316,552	2.477	2.096	0.000	2.303	4.094
Mortgage Exposure	270,672	0.390	0.257	0.185	0.348	0.556
Panel D. Bank-County Level (Matched - HMDA, Fannie Mae, Freddie Mac, and McDash)						
Delinquency 60 Days	73,732	0.121	0.154	0.000	0.067	0.194
Delinquency 90 Days	73,732	0.100	0.140	0.000	0.034	0.164
Foreclosure	73,732	0.063	0.107	0.000	0.000	0.100
Log(STD. Interest Rates)	73,726	-0.405	0.597	-0.845	-0.505	0.066
Log(STD. LTV)	73,712	2.845	0.416	2.589	2.897	3.156
Industry Expertise	73,732	0.180	0.385	0.000	0.000	0.000
Credit Score	73,732	729.481	33.914	707.750	733.527	756.667
LTV	73,732	70.848	9.554	64.888	72.044	77.900
DTI	73,354	32.833	5.464	29.545	32.924	36.355
Interest Rate	73,732	5.624	1.432	4.232	5.822	6.744
LTI	73,732	2.221	0.530	1.840	2.182	2.559
Male	73,706	0.736	0.148	0.643	0.750	0.833
Minority	73,686	0.089	0.124	0.000	0.026	0.143
Branch	73,732	1.010	1.041	0.000	0.693	1.792
Distance	73,732	6.326	1.079	5.577	6.407	7.305
SBL	73,732	4.119	2.074	2.944	4.431	5.665
Mortgage Exposure	70,524	0.365	0.200	0.222	0.342	0.477

Table 2. Mortgage Lending Through the Industry Expertise Channel: Number and Volume

This table presents the effects of the industry expertise channel on banks' mortgage lending across counties. The dependent variables are the natural logarithm of the number of mortgages a bank approves in a county in columns (1) - (5) and the natural logarithm of the dollar volume (in millions) of mortgages a bank approves in a county in columns (6) - (10). The key independent variable is *Industry Expertise*, a dummy that equals one for a bank-county pair if there exists at least one industry in which a bank specializes and provides at least 5% jobs in a county. Controls include the average loan-to-income ratio of all mortgage applicants, the percentage of male applicants, the percentage of minority applicants, the natural logarithm of one plus the number of branches a bank has in the county, the natural logarithm of the geographic distance between the headquarters county of a bank and the borrower's home county, the natural logarithm of one plus the number of small business loans a bank originates in the borrower's home county, the average percentage of mortgages retained on balance sheets in the borrower's home county in the past three years, the natural logarithm of bank assets, total loans scaled by assets, deposits scaled by assets, commercial and industrial (C&I) loans scaled by total loans, real estate loans scaled by total loans, return on assets, and total liquidity scaled by assets. The sample period is 1999 to 2017. Standard errors clustered by county are reported in parentheses below the coefficient estimates. ***, **, and * indicate p-values of 1%, 5%, and 10%, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Log(Number of Approved Mortgages)					Log(Volume of Approved Mortgages)				
Industry Expertise	0.024***	0.025***	0.056***	0.067***	0.063***	0.028***	0.028***	0.060***	0.068***	0.065***
	(0.008)	(0.008)	(0.006)	(0.007)	(0.006)	(0.008)	(0.008)	(0.006)	(0.007)	(0.006)
LTI		-0.331***	-0.062***	-0.056***	-0.004		-0.046***	0.202***	0.200***	0.235***
		(0.010)	(0.007)	(0.008)	(0.007)		(0.010)	(0.007)	(0.008)	(0.008)
Male		-0.020	-0.012	-0.011	0.031*		0.315***	0.320***	0.322***	0.322***
		(0.020)	(0.017)	(0.020)	(0.017)		(0.020)	(0.017)	(0.019)	(0.017)
Minority		0.913***	0.392***	0.410***	0.285***		0.669***	0.154***	0.127***	0.064*
		(0.055)	(0.037)	(0.042)	(0.034)		(0.054)	(0.036)	(0.041)	(0.035)
Branch			0.540***	0.548***	0.458***			0.507***	0.513***	0.430***
			(0.010)	(0.010)	(0.010)			(0.010)	(0.010)	(0.010)
Distance			-0.157***	-0.152***	-0.190***			-0.159***	-0.155***	-0.200***
			(0.007)	(0.007)	(0.018)			(0.008)	(0.008)	(0.018)
SBL			0.246***	0.251***	0.206***			0.255***	0.261***	0.213***
			(0.003)	(0.003)	(0.003)			(0.003)	(0.003)	(0.003)
Mortgage Exposure			0.172***	0.139***	0.200***			0.082***	0.039**	0.107***
			(0.016)	(0.017)	(0.016)			(0.015)	(0.017)	(0.016)
Log(Assets)			0.138***	0.110***	0.218***			0.151***	0.123***	0.241***
			(0.015)	(0.016)	(0.017)			(0.016)	(0.017)	(0.018)
Total Loans/Assets			1.000***	1.039***	1.080***			1.489***	1.508***	1.547***
			(0.074)	(0.079)	(0.072)			(0.074)	(0.078)	(0.073)
Deposits/Assets			-1.309***	-1.486***	-1.367***			-1.478***	-1.632***	-1.530***
			(0.049)	(0.052)	(0.052)			(0.051)	(0.054)	(0.053)
C&I Loans/Total Loans			3.975***	3.999***	4.090***			3.880***	3.902***	4.000***
			(0.097)	(0.105)	(0.098)			(0.099)	(0.107)	(0.101)
RE Loans/Total Loans			2.270***	2.103***	2.334***			2.346***	2.151***	2.419***
			(0.068)	(0.074)	(0.069)			(0.072)	(0.077)	(0.073)
ROA			-0.187	-1.663***	-2.223***			1.007**	-0.620	-1.045**
			(0.458)	(0.487)	(0.449)			(0.477)	(0.505)	(0.467)
Liquidity/Assets			-1.480***	-1.793***	-1.872***			-0.951***	-1.291***	-1.377***
			(0.076)	(0.082)	(0.086)			(0.079)	(0.084)	(0.089)
Observations	316,524	315,026	265,134	257,492	257,382	316,524	315,026	265,134	257,492	257,382
Year FE	Yes	Yes	Yes	No	No	Yes	Yes	Yes	No	No
Bank FE	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	No
County FE	Yes	Yes	Yes	No	No	Yes	Yes	Yes	No	No
County×Year FE	No	No	No	Yes	Yes	No	No	No	Yes	Yes
Bank×State FE	No	No	No	No	Yes	No	No	No	No	Yes
Adjusted R^2	0.472	0.482	0.718	0.704	0.773	0.580	0.583	0.762	0.750	0.804

Table 3. Mortgage Lending Through the Industry Expertise Channel: Approval Rates

This table presents the effects of the industry expertise channel on banks' mortgage approval rates across counties. The dependent variables are the number-based approval rate in columns (1) - (3) and the volume-based approval rate in columns (4) - (6). The key independent variable is *Industry Expertise*, a dummy that equals one for a bank-county pair if there exists at least one industry in which a bank specializes and provides at least 5% jobs in a county. Controls include the average loan-to-income ratio of all mortgage applicants, the percentage of male applicants, the percentage of minority applicants, the natural logarithm of one plus the number of branches a bank has in the county, the natural logarithm of the geographic distance between the headquarters county of a bank and the borrower's home county, the natural logarithm of one plus the number of small business loans a bank originates in the borrower's home county, the average percentage of mortgages retained on balance sheets in the borrower's home county in the past three years, the natural logarithm of bank assets, total loans scaled by assets, deposits scaled by assets, commercial and industrial (C&I) loans scaled by total loans, real estate loans scaled by total loans, return on assets, and total liquidity scaled by assets. The sample period is 1999 to 2017. Standard errors clustered by county are reported in parentheses below the coefficient estimates. ***, **, and * indicate p-values of 1%, 5%, and 10%, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	Approval Rate-Number			Approval Rate-Volume		
Industry Expertise	0.003***	0.004***	0.004***	0.004***	0.005***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	265,134	257,492	257,382	265,134	257,492	257,382
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	No	No	Yes	No	No
Bank FE	Yes	Yes	No	Yes	Yes	No
County FE	Yes	No	No	Yes	No	No
County×Year FE	No	Yes	Yes	No	Yes	Yes
Bank×State FE	No	No	Yes	No	No	Yes
Adjusted R^2	0.378	0.386	0.436	0.330	0.343	0.389

Table 4. Industry Growth and Household Income and Mortgage Delinquency

This table presents the relation between the growth of a county's key industries and the county's household income growth and mortgage delinquency rates. The dependent variable in columns (1) - (3) is a county's average income growth rate (%). The dependent variable in columns (4) - (6) is the annual change in a county's mortgage delinquency rate (%). The key independent variable is a county's employment-weighted industry-level sales growth rate. The weight is the fraction of local residents working in a given industry. The sales growth rate for each industry is estimated using the sales of all U.S. public firms in the industry. Controls include the natural logarithm of the population, the percentage of the population above 65, the percentage of the male population, the percentage of the minority population, the percentage of the population with a bachelor's degree or above. The sample that examines the income growth rate is from 1999 to 2017, and the sample that examines the mortgage delinquency rate is from 2008 to 2017. Standard errors clustered by county are reported in parentheses below the coefficient estimates. ***, **, and * indicate p-values of 1%, 5%, and 10%, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	Income Growth (%)			Delta Mortgage Delinquency (%)		
Sales Growth	1.257***	1.339***	0.585***	-1.577***	-1.700***	-0.169***
	(0.033)	(0.035)	(0.044)	(0.054)	(0.050)	(0.052)
Population		0.100***	-2.771***		-0.082***	1.840***
		(0.023)	(0.286)		(0.028)	(0.543)
Above 65		-16.116***	23.559***		-2.747***	-9.766**
		(0.999)	(2.801)		(0.668)	(4.778)
Male		-6.025***	-11.616**		-10.751***	68.652***
		(1.268)	(5.023)		(2.927)	(12.013)
Minority		-1.999***	-0.242		0.001	0.253
		(0.125)	(1.597)		(0.167)	(2.373)
Bachelor		3.214***	9.738***		-0.623***	-2.526***
		(0.307)	(1.784)		(0.187)	(0.778)
Observations	58,394	55,212	55,212	4,230	3,728	3,728
County FE	No	No	Yes	No	No	Yes
Year FE	No	No	Yes	No	No	Yes
Adjusted R^2	0.0681	0.0865	0.222	0.355	0.405	0.714

Table 5. Information Asymmetry and the Industry Expertise Channel

This table presents the effects of information asymmetry on banks' use of the industry expertise channel in mortgage lending. The dependent variables are the natural logarithm of the number of mortgages a bank approves in a county in columns (1) - (2) and the natural logarithm of the dollar volume (in millions) of mortgages a bank approves in a county in columns (3) - (4). The key independent variable is the interaction term between the *Industry Expertise* and the partition variables. *Industry Expertise* is a dummy that equals one for a bank-county pair if there exists at least one industry in which a bank specializes and provides at least 5% jobs in a county. *Distance* is the standardized distance between a bank's headquarters county and a borrower's home county. *SCI* is the standardized social connectedness index between a bank's headquarters county and a borrower's home county. Controls include the average loan-to-income ratio of all mortgage applicants, the percentage of male applicants, the percentage of minority applicants, the natural logarithm of one plus the number of branches a bank has in the county, the natural logarithm of the geographic distance between the headquarters county of a bank and the borrower's home county, the natural logarithm of one plus the number of small business loans a bank originates in the borrower's home county, the average percentage of mortgages retained on balance sheets in the borrower's home county in the past three years, the natural logarithm of bank assets, total loans scaled by assets, deposits scaled by assets, commercial and industrial (C&I) loans scaled by total loans, real estate loans scaled by total loans, return on assets, and total liquidity scaled by assets. The sample period is 1999 to 2017. Standard errors clustered by county are reported in parentheses below the coefficient estimates. ***, **, and * indicate p-values of 1%, 5%, and 10%, respectively.

	(1)	(2)	(3)	(4)
	Log(Number of Approved Mortgages)		Log(Volume of Approved Mortgages)	
Industry Expertise	0.070***	0.064***	0.072***	0.066***
	(0.006)	(0.006)	(0.006)	(0.006)
Industry Expertise × Distance	0.083***		0.085***	
	(0.006)		(0.006)	
Distance	-0.302***		-0.307***	
	(0.035)		(0.034)	
Industry Expertise × SCI		-0.020***		-0.020***
		(0.006)		(0.006)
SCI		0.045***		0.048***
		(0.009)		(0.009)
Observations	257,382	257,302	257,382	257,302
Controls	Yes	Yes	Yes	Yes
County×Year FE	Yes	Yes	Yes	Yes
Bank×State FE	Yes	Yes	Yes	Yes
Adjusted R^2	0.773	0.773	0.804	0.804

Table 6. Borrower Risk and the Industry Expertise Channel

This table presents the effects of borrower risk on banks' use of the industry expertise channel in mortgage lending. The dependent variables are the natural logarithm of the number of mortgages a bank approves in a county in columns (1) - (2) and the natural logarithm of the dollar volume (in millions) of mortgages a bank approves in a county in columns (3) - (4). The key independent variable is the interaction term between *Industry Expertise* and the partition variables. *Industry Expertise* is a dummy that equals one for a bank-county pair if there exists at least one industry in which a bank specializes and provides at least 5% jobs in a county. *HP Volatility* is the standardized county-level house price volatility. *LTI* is the average LTI ratio for all mortgage applicants in a county. Controls include the average loan-to-income ratio of all mortgage applicants, the percentage of male applicants, the percentage of minority applicants, the natural logarithm of one plus the number of branches a bank has in the county, the natural logarithm of the geographic distance between the headquarters county of a bank and the borrower's home county, the natural logarithm of one plus the number of small business loans a bank originates in the borrower's home county, the average percentage of mortgages retained on balance sheets in the borrower's home county in the past three years, the natural logarithm of bank assets, total loans scaled by assets, deposits scaled by assets, commercial and industrial (C&I) loans scaled by total loans, real estate loans scaled by total loans, return on assets, and total liquidity scaled by assets. The sample period is 1999 to 2017. Standard errors clustered by county are reported in parentheses below the coefficient estimates. ***, **, and * indicate p-values of 1%, 5%, and 10%, respectively.

	(1)	(2)	(3)	(4)
	Log(Number of Approved Mortgages)		Log(Volume of Approved Mortgages)	
Industry Expertise	0.067***	0.053***	0.067***	0.055***
	(0.007)	(0.006)	(0.007)	(0.006)
Industry Expertise × HP Volatility	0.047***		0.050***	
	(0.005)		(0.006)	
Industry Expertise × LTI		0.058***		0.058***
		(0.006)		(0.006)
Observations	188,511	257,382	188,511	257,382
Controls	Yes	Yes	Yes	Yes
County×Year FE	Yes	Yes	Yes	Yes
Bank×State FE	Yes	Yes	Yes	Yes
Adjusted R^2	0.775	0.773	0.796	0.804

Table 7. Dispersion in Mortgage Contractual Terms

This table presents the effects of the industry expertise channel on the dispersion of mortgage contractual terms. The dependent variables are the natural logarithm of the standard deviations of loan amounts, loan-to-income ratios, interest rates and loan-to-value ratios. The key independent variable is *Industry Expertise*, a dummy that equals one for a bank-county pair if there exists at least one industry in which a bank specializes and provides at least 5% jobs in a county. Estimations in columns (1) and (2) use the HMDA sample. Estimations in columns (3) and (4) use the matched sample between HMDA and the Fannie Mae, Freddie Mac, and McDash datasets. Common controls in all four columns include the natural logarithm of one plus the number of branches a bank has in the county, the natural logarithm of the geographic distance between the headquarters county of a bank and the borrower's home county, the natural logarithm of one plus the number of small business loans a bank originates in the borrower's home county, the average percentage of mortgages retained on balance sheets in the borrower's home county in the past three years, the natural logarithm of bank assets, total loans scaled by assets, deposits scaled by assets, commercial and industrial (C&I) loans scaled by total loans, real estate loans scaled by total loans, return on assets, and total liquidity scaled by assets. In addition, columns (1) and (2) control for the average loan-to-income ratio of all mortgage applicants, the percentage of male applicants, the percentage of minority applicants. Columns (3) and (4) control for the average loan-to-value ratio, the percentage of male, the percentage of minority, the average credit score, the average loan-to-value ratio, and the average interest rate of approved mortgages. The sample period is 1999 to 2017. Standard errors clustered by county are reported in parentheses below the coefficient estimates. ***, **, and * indicate p-values of 1%, 5%, and 10%, respectively.

	(1)	(2)	(3)	(4)
	Log(STD. Mortgage Size)	Log(STD. LTI)	Log(STD. Interest Rates)	Log(STD. LTV)
Industry Expertise	0.006***	0.005**	0.021***	0.022***
	(0.002)	(0.002)	(0.006)	(0.004)
Observations	256,055	257,382	58,772	58,764
Controls	Yes	Yes	Yes	Yes
County×Year FE	Yes	Yes	Yes	Yes
Bank×State FE	Yes	Yes	Yes	Yes
Adjusted R^2	0.450	0.688	0.595	0.568

Table 8. The Percentage of Conventional Mortgages

This table presents the effects of the industry expertise channel on banks' originations of conventional versus government-insured mortgages. The dependent variables are the number-based percentage of conventional mortgages a bank approves in a county in columns (1) - (3) and the volume-based percentage of conventional mortgages a bank approves in a county in columns (4) - (6). The key independent variable is *Industry Expertise*, a dummy that equals one for a bank-county pair if there exists at least one industry in which a bank specializes and provides at least 5% jobs in a county. Controls include the average loan-to-income of mortgage borrowers, the percentage of male borrowers, the percentage of minority borrowers, the natural logarithm of one plus the number of branches a bank has in the county, the natural logarithm of the geographic distance between the headquarters county of a bank and the borrower's home county, the natural logarithm of one plus the number of small business loans a bank originates in the borrower's home county, the average percentage of mortgages retained on balance sheets in the borrower's home county in the past three years, the natural logarithm of bank assets, total loans scaled by assets, deposits scaled by assets, commercial and industrial (C&I) loans scaled by total loans, real estate loans scaled by total loans, return on assets, and total liquidity scaled by assets. The sample period is 1999 to 2017. Standard errors clustered by county are reported in parentheses below the coefficient estimates. ***, **, and * indicate p-values of 1%, 5%, and 10%, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	% Conventional Mortgages					
	Number			Volume		
Industry Expertise	0.004***	0.003***	0.002**	0.004***	0.003***	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	318,940	267,037	259,510	318,940	267,037	259,510
Controls	No	Yes	Yes	No	Yes	Yes
Year FE	Yes	Yes	No	Yes	Yes	No
Bank FE	Yes	Yes	No	Yes	Yes	No
County FE	Yes	Yes	No	Yes	Yes	No
County×Year FE	No	No	Yes	No	No	Yes
Bank×State FE	No	No	Yes	No	No	Yes
Adjusted R^2	0.265	0.308	0.459	0.255	0.295	0.452

Table 9. Mortgage Delinquency and Foreclosure

This table presents the effects of the industry expertise channel on banks' mortgage delinquency and foreclosure rates. The dependent variables are *Delinquency 60 Days*, *Delinquency 90 Days*, and *Foreclosure*. *Delinquency 60 Days* is the percentage of mortgages that are more than 60 days past due on monthly payments. *Delinquency 90 Days* is the percentage of mortgages that are more than 90 days past due on monthly payments. *Foreclosure* is the percentage of mortgages that have gone through a foreclosure. The key independent variable is *Industry Expertise*, a dummy that equals one for a bank-county pair if there exists at least one industry in which a bank specializes and provides at least 5% jobs in a county. Controls include the average credit score of approved mortgages, the average loan-to-value ratio of approved mortgages, the average debt-to-income ratio of approved mortgages, the average interest rate of approved mortgages, the average loan-to-income ratio of approved mortgages, the percentage of male borrowers, the percentage of minority borrowers, the natural logarithm of one plus the number of branches a bank has in the county, the natural logarithm of the geographic distance between a bank's headquarters county and the borrower's home county, the natural logarithm of one plus the number of small business loans a bank originates in the borrower's home county, the average percentage of mortgages retained on balance sheets in the borrower's home county in the past three years, the natural logarithm of bank assets, total loans scaled by assets, deposits scaled by assets, commercial and industrial (C&I) loans scaled by total loans, real estate loans scaled by total loans, return on assets, and total liquidity scaled by assets. The sample period is 1999 to 2017. Standard errors clustered by county are reported in parentheses below the coefficient estimates. ***, **, and * indicate p-values of 1%, 5%, and 10%, respectively.

	(1)	(2)	(3)
	Delinquency 60 Days	Delinquency 90 Days	Foreclosure
Industry Expertise	-0.005***	-0.004***	-0.003**
	(0.001)	(0.001)	(0.001)
Credit Score	-0.002***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)
LTV	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)
DTI	0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
Interest Rate	0.045***	0.041***	0.023***
	(0.002)	(0.002)	(0.002)
LTI	0.008***	0.008***	0.008***
	(0.002)	(0.002)	(0.001)
Male Applicants	-0.015***	-0.015***	-0.007**
	(0.004)	(0.003)	(0.003)
Minority Applicants	0.044***	0.038***	0.018***
	(0.006)	(0.005)	(0.004)
Branch	-0.002**	-0.003***	-0.001*
	(0.001)	(0.001)	(0.001)
Distance	0.000	0.000	0.002
	(0.002)	(0.002)	(0.001)
SBL	0.004***	0.005***	0.004***
	(0.001)	(0.001)	(0.000)
Mortgage Exposure	-0.024***	-0.018***	-0.033***
	(0.004)	(0.004)	(0.003)
Log(Assets)	0.014***	0.016***	-0.007***
	(0.003)	(0.003)	(0.002)
Total Loans/Assets	0.110***	0.086***	0.082***
	(0.019)	(0.018)	(0.014)
Deposits/Assets	0.144***	0.147***	0.186***
	(0.016)	(0.014)	(0.012)
C&I Loans/Total Loans	0.028	0.016	0.003
	(0.027)	(0.025)	(0.019)
RE Loans/Total Loans	0.033*	0.009	-0.015
	(0.019)	(0.018)	(0.014)
ROA	0.139	0.156	-0.435***
	(0.122)	(0.116)	(0.096)
Liquidity/Assets	0.157***	0.159***	0.110***
	(0.022)	(0.020)	(0.015)
Observations	58,779	58,779	58,779
County×Year FE	Yes	Yes	Yes
Bank×State FE	Yes	Yes	Yes
Adjusted R^2	0.649	0.634	0.578

Table 10. Industry Distress and the Industry Expertise Channel

This table presents the effects of industry distress on banks' use of the industry expertise channel in mortgage lending. The dependent variables are the logarithm of the number of mortgages a bank approves in a county in columns (1) - (3) and the logarithm of the dollar volume (in millions) of mortgages a bank approves in a county in columns (4) - (6). The key independent variable is the interaction term between *Industry Expertise* and *Distress*. *Industry Expertise* is a dummy that equals one for a bank-county pair if there exists at least one industry in which a bank specializes and provides at least 5% jobs in a county, measured at t-2. *Distress* is a dummy that equals one for a bank-county pair if distress happens in any of the industries that a bank specializes in and provides at least 5% jobs in a county, measured at t-1. A three-digit NAICS industry is classified as distressed in a year if, from the beginning of that year, the industry-level two-year sales growth is negative and the industry-level two-year stock return is less than –10% (columns (1) & (4)), -20% (columns (2) & (5)), or -30% (columns (3) & (6)). Controls include the average loan-to-income ratio of all mortgage applicants, the percentage of male applicants, the percentage of minority applicants, the natural logarithm of one plus the number of branches a bank has in the county, the natural logarithm of the geographic distance between the headquarters county of a bank and the borrower's home county, the natural logarithm of one plus the number of small business loans a bank originates in the borrower's home county, the average percentage of mortgages retained on balance sheets in the borrower's home county in the past three years, the natural logarithm of bank assets, total loans scaled by assets, deposits scaled by assets, commercial and industrial (C&I) loans scaled by total loans, real estate loans scaled by total loans, return on assets, and total liquidity scaled by assets. The sample period is 1999 to 2017. Standard errors clustered by county are reported in parentheses below the coefficient estimates. ***, **, and * indicate p-values of 1%, 5%, and 10%, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	Log(Number of Approved Mortgages)			Log(Volume of Approved Mortgages)		
	Return<-10%	Return<-20%	Return<-30%	Return<-10%	Return<-20%	Return<-30%
Industry Expertise	0.020**	0.020**	0.020**	0.025***	0.025***	0.025***
	(0.009)	(0.009)	(0.009)	(0.010)	(0.010)	(0.010)
Industry Expertise $\times$ Distress	0.044*	0.052**	0.056**	0.041	0.046*	0.049*
	(0.025)	(0.026)	(0.028)	(0.026)	(0.027)	(0.028)
Observations	165,306	165,306	165,306	165,306	165,306	165,306
Controls	Yes	Yes	Yes	Yes	Yes	Yes
County$\times$Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Bank$\times$State FE	Yes	Yes	Yes	Yes	Yes	Yes
Bank$\times$Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.820	0.820	0.820	0.840	0.840	0.840

Table 11. The 2008 Financial Crisis and the Industry Expertise Channel

This table presents the effects of the 2008 financial crisis on banks' use of the industry expertise channel in mortgage lending. The dependent variables are the logarithm of the number of mortgages a bank approves in a county in columns (1) - (2) and the logarithm of the dollar volume (in millions) of mortgages a bank approves in a county in columns (3) - (4). The key independent variable is the interaction term between *Industry Expertise* and *Crisis* in columns (1) & (3), the interaction terms between *Industry Expertise* and year dummies in columns (2) & (4). *Industry Expertise* is a dummy that equals one for a bank-county pair if there exists at least one industry in which a bank specializes and provides at least 5% jobs in a county, measured at the year 2003. *Crisis* is a dummy that equals one for the period 2008 - 2010 and zero for the period 2004 - 2007. *Year 2004*, *Year 2005*, *Year 2006*, *Year 2008*, *Year 2009* and *Year 2010* are year dummies. Year 2007 is the base year and thus omitted. Controls include the average loan-to-income ratio of all mortgage applicants, the percentage of male applicants, the percentage of minority applicants, the natural logarithm of one plus the number of branches a bank has in the county, the natural logarithm of the geographic distance between the headquarters county of a bank and the borrower's home county, the natural logarithm of one plus the number of small business loans a bank originates in the borrower's home county, the average percentage of mortgages retained on balance sheets in the borrower's home county in the past three years, the natural logarithm of bank assets, total loans scaled by assets, deposits scaled by assets, commercial and industrial (C&I) loans scaled by total loans, real estate loans scaled by total loans, return on assets, and total liquidity scaled by assets. The sample period is 2004 to 2010. Standard errors clustered by county are reported in parentheses below the coefficient estimates. ***, **, and * indicate p-values of 1%, 5%, and 10%, respectively.

	(1)	(2)	(3)	(4)
	Log(Number of Approved Mortgages)		Log(Volume of Approved Mortgages)	
Industry Expertise	0.038**	0.026	0.047***	0.045**
	(0.018)	(0.022)	(0.018)	(0.022)
Industry Expertise × Crisis	0.085***		0.073***	
	(0.023)		(0.024)	
Industry Expertise × Year 2004		-0.011		-0.021
		(0.023)		(0.024)
Industry Expertise × Year 2005		0.035		0.026
		(0.022)		(0.023)
Industry Expertise × Year 2006		0.023		0.006
		(0.019)		(0.020)
Industry Expertise × Year 2008		0.023		0.022
		(0.030)		(0.033)
Industry Expertise × Year 2009		0.142***		0.126***
		(0.030)		(0.030)
Industry Expertise × Year 2010		0.129***		0.081**
		(0.033)		(0.034)
Observations	87,166	87,166	87,166	87,166
Controls	Yes	Yes	Yes	Yes
County×Year FE	Yes	Yes	Yes	Yes
Bank×State FE	Yes	Yes	Yes	Yes
Bank×Year FE	Yes	Yes	Yes	Yes
Adjusted R^2	0.829	0.829	0.853	0.853

Internet Appendix to

“The Industry Expertise Channel of Mortgage Lending"

Yongqiang Chu   Zhanbing Xiao   Yuxiang Zheng

September 2023

A Robustness Checks of Baseline Results in Table 2

In this section, we conduct several additional tests to show the robustness of our baseline results in Table 2 using alternative measures, more fixed effects, and alternative model specifications.

A.1 Alternative Measures of the Industry Expertise Channel

In general, industry leaders have more advanced technologies and are more closely linked to the latest industry dynamics relative to followers. Thus, lending to industry leaders allows banks to accumulate industry expertise faster and better, relative to lending to followers. To capture the knowledge gap between industry leaders and followers, we construct a new measure of banks' lending specialization. Specifically, we use a firm's size to proxy for its position in an industry. We first classify all firms in an industry into ten groups based on total assets, with group 10 including firms with the largest assets.[1] We then use the rank as the weight to calculate a bank's total lending to a given industry in the following way:

$$L_{i,t}^{b} = \frac{\sum_{j=1}^{K} Loan_{i,j,t}^{b} * Rank_{i,j,t}}{\sum_{i=1}^{I} \sum_{j=1}^{K} Loan_{i,j,t}^{b} * Rank_{i,j,t}} \tag{IA.1}$$

where b denotes bank, i denotes industry, j denotes firm. $Rank_{i,j,t}$ is the rank of a firm's assets in its industry i. Then we reconstruct the dummy *Industry Expertise* using the same method in equation (1). Columns (1) and (5) in Table IA.1 present the results.

In addition, we construct two continuous measures that capture the intensity of the connections between banks and counties through the industry expertise channel. The first measure, *Industry Expertise (Fraction, 5%)*, is the fraction of a county's residents working in any industries that a bank specializes in and provide at least 5% jobs in the county. The second measure, *Industry Expertise (Fraction, All)*, is the fraction of a county's residents

[1] Using a rank variable rather than the assets avoids high skewness in the distribution of firm assets in an industry and the uneven distribution of firm assets across industries.

working in any industry in which a bank specializes, regardless of the number of jobs provided in the county. The results using the two measures are reported in columns (2), (3), (6), and (7) of Table IA.1 and are consistent with Table 2.

Lastly, we construct a measure that reflects the level of a bank's industry expertise. This measure is calculated as the difference between a bank's loan share in an industry minus the threshold used to identify an outlier loan share in equation (1). Results using this measure are consistent and are reported in columns (4) and (8) of Table IA.1.

A.2 More Fixed Effects

Even though we add bank and bank-by-state fixed effects and bank-level variables to control for heterogeneities across banks in Table 2 & Table 3, the concern over omitted time-varying bank-level characteristics remains. To better address the concern, we add bank-by-year fixed effects in columns (1) and (3) of Table IA.2. In addition, we use bank-by-county fixed effects to replace bank-by-state fixed effects to control for time-invariant links between banks and counties in columns (2) & (4) of Table IA.2. Our results hold.

A.3 Alternative Empirical Specifications

Another concern is that using the log form of the dependent variables may produce biased estimations. To address this issue, we redo the tests in Table 2 using alternative empirical specifications. Specifically, in Table IA.3 columns (1) and (4), we use a linear regression model to estimate the effect of industry expertise on the raw number and the raw dollar volume of a bank's mortgage originations in a county. The coefficient estimates are statistically significant and the economic effects are important - industry expertise increases banks' mortgage lending by 7.2% in numbers and 9.2% in dollar volumes. We also use the population-scaled raw number and dollar volume of approved mortgages as the dependent variables and get consistent results in columns (2) and (5). In columns (3) and (6), we follow Cohn, Liu, and Wardlaw (2022) and use the fixed effects Poisson model to redo the estimation. Our results still hold. Economic effects are even larger - industry expertise increases banks' mortgage lending by 10.6% in numbers and 12.3% in dollar volumes.

Table IA.1. Robustness Checks - Alternative Measures of the Industry Expertise Channel

This table presents robustness checks of baseline results in Table 2 using alternative measures of the industry expertise channel. The dependent variables are the natural logarithm of the number of mortgages a bank approves in a county in columns (1) - (4) and the natural logarithm of the dollar volume (in millions) of mortgages a bank approves in a county in columns (5) - (8). The independent variable in columns (1) & (5), *Industry Expertise (Weighted)*, is adjusted by each corporate borrower's market position (see equation (IA.1)). The independent variable in columns (2) & (6), *Industry Expertise (Fraction, 5%)*, is the fraction of a county's residents working in industries that a bank specializes in and provide at least 5% jobs in the county. The independent variable in columns (3) & (6), *Industry Expertise (Fraction, All)*, is the fraction of a county's residents that work in any industry in which a bank specializes, regardless of the number of jobs provided in the county. The independent variable in columns (4) & (8), *Industry Expertise (Level)*, is the level of a bank's industry expertise, measured as the the of difference between the bank's loan share in an industry minus the threshold used to identify an outlier loan share. Controls include the average loan-to-income ratio of all mortgage applicants, the percentage of male applicants, the percentage of minority applicants, the natural logarithm of one plus the number of branches a bank has in the county, the natural logarithm of the geographic distance between the headquarters county of a bank and the borrower's home county, the natural logarithm of one plus the number of small business loans a bank originates in the borrower's home county, the average percentage of mortgages retained on balance sheets in the borrower's home county in the past three years, the natural logarithm of bank assets, total loans scaled by assets, deposits scaled by assets, commercial and industrial (C&I) loans scaled by total loans, real estate loans scaled by total loans, return on assets, and total liquidity scaled by assets. The sample period is from 1999 to 2017. Standard errors clustered by county are reported in parentheses below the coefficient estimates. ***, **, and * indicate p-values of 1%, 5%, and 10%, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Log(Number of Approved Mortgages)				Log(Volume of Approved Mortgages)			
Industry Expertise (Weighted)	0.063***				0.066***			
	(0.006)				(0.006)			
Industry Expertise (Fraction, 5%)		0.537***				0.569***		
		(0.047)				(0.048)		
Industry Expertise (Fraction, All)			0.357***				0.381***	
			(0.046)				(0.046)	
Industry Expertise (Level)				0.138***				0.143***
				(0.022)				(0.022)
Observations	265,664	257,382	257,382	257,382	265,664	257,382	257,382	257,382
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County×Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bank×State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.774	0.773	0.773	0.773	0.807	0.804	0.804	0.804

Table IA.2. Robustness Checks - More Fixed Effects

This table presents robustness checks of baseline results in Table 2 with additional fixed effects. The dependent variables are the natural logarithm of the number of mortgages a bank approves in a county in columns (1) - (2) and the natural logarithm of the dollar volume (in millions) of mortgages a bank approves in a county in columns (3) - (4). Columns (1) and (3) present the results with bank-by-year fixed effects. Columns (2) and (4) present the results using bank-by-county fixed effects to replace bank-by-state fixed effects. The key independent variable is *Industry Expertise*, a dummy that equals one for a bank-county pair if there exists at least one industry which a bank specializes in and provides at least 5% jobs in a county. Controls include the average loan-to-income ratio of all mortgage applicants, the percentage of male applicants, the percentage of minority applicants, the natural logarithm of one plus the number of branches a bank has in the county, the natural logarithm of the geographic distance between the headquarters county of a bank and the borrower's home county, the natural logarithm of one plus the number of small business loans a bank originates in the borrower's home county, the average percentage of mortgages retained on balance sheets in the borrower's home county in the past three years, the natural logarithm of bank assets, total loans scaled by assets, deposits scaled by assets, commercial and industrial (C&I) loans scaled by total loans, real estate loans scaled by total loans, return on assets, and total liquidity scaled by assets. The sample period is from 1999 to 2017. Standard errors clustered by county are reported in parentheses below the coefficient estimates. ***, **, and * indicate p-values of 1%, 5%, and 10%, respectively.

	(1)	(2)	(3)	(4)
	Log(Number of Approved Mortgages)		Log(Volume of Approved Mortgages)	
Industry Expertise	0.028***	0.051***	0.029***	0.052***
	(0.007)	(0.005)	(0.008)	(0.005)
Observations	257,378	248,844	257,378	248,844
Controls	Yes	Yes	Yes	Yes
County×Year FE	Yes	Yes	Yes	Yes
Bank×State FE	Yes	No	Yes	No
Bank×Year FE	Yes	No	Yes	No
Bank×County FE	No	Yes	No	Yes
Adjusted R^2	0.807	0.852	0.832	0.868

Table IA.3. Robustness Checks - Alternative Empirical Specifications

This table presents robustness checks of baseline results in Table 2 using alternative empirical specifications. We use the linear regression model to estimate equation (2) in columns (1), (2), (4) and (5), and the fixed effects Poisson model in columns (3) and (6). In columns (1) and (3), the dependent variable is the number of mortgages a bank approves in a county. In columns (2), the dependent variable is the number of mortgages a bank approves in a county scaled by the county's population and multiplied by 1000. In columns (4) and (6), the dependent variable is the dollar volume (in millions) of mortgages a bank approves in a county. In column (5), the dependent variable is the dollar volume (in thousands) of mortgages a bank approves in a county scaled by the county's population. The key independent variable is *Industry Expertise*, a dummy that equals one for a bank-county pair if there exists at least one industry which a bank specializes in and provides at least 5% jobs in a county. Controls include the average loan-to-income ratio of all mortgage applicants, the percentage of male applicants, the percentage of minority applicants, the natural logarithm of one plus the number of branches a bank has in the county, the natural logarithm of the geographic distance between the headquarters county of a bank and the borrower's home county, the natural logarithm of one plus the number of small business loans a bank originates in the borrower's home county, the average percentage of mortgages retained on balance sheets in the borrower's home county in the past three years, the natural logarithm of bank assets, total loans scaled by assets, deposits scaled by assets, commercial and industrial (C&I) loans scaled by total loans, real estate loans scaled by total loans, return on assets, and total liquidity scaled by assets. The sample period is from 1999 to 2017. Numbers in parentheses are standard errors. Standard errors clustered by county are reported in parentheses below the coefficient estimates. ***, **, and * indicate p-values of 1%, 5%, and 10%, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	Number of Approved Mortgages			Volume of Approved Mortgages		
	Linear	Linear Scaled by Population	Poisson	Linear	Linear Scaled by Population	Poisson
Industry Expertise	6.377***	0.074***	0.106***	1.330***	0.013***	0.123***
	(1.076)	(0.007)	(0.007)	(0.185)	(0.001)	(0.007)
Observations	257,382	251,718	257,382	257,382	251,718	257,382
Controls	Yes	Yes	Yes	Yes	Yes	Yes
County×Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Bank×State FE	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.631	0.525	0.864	0.645	0.514	0.836

B Excluding an Alternative Channel

An alternative story to the industry expertise channel we propose in this paper is that a bank may prioritize lending mortgages to employees of its corporate borrowers. Undoubtedly, a bank knows the income streams of a mortgage borrower better if she is an employee of the bank's corporate borrowers, as the bank has access to these firms' private information through corporate lending. Some firms even have joint programs with their relationship banks to help employees get mortgages with favourable terms. Therefore, this alternative channel coincides with the industry expertise channel when a bank has corporate borrowers in industries in which the bank specializes and provide significant jobs in a county. Controlling for a bank's small business lending in a county in our regressions alleviates the concern to some extent, assuming that small business lending is positively correlated with small firms' employment and thus the importance of this alternative channel. However, this does not capture employees of banks' big borrowers in the syndicated loan market. To better address the concern, we conduct additional tests in Appendix Table IB.1.

Specifically, we first obtain firms' historical headquarter states from the Compustat. Then, each year, we drop a bank-state pair from the HMDA mortgage sample if the bank has a syndicated loan borrower located in the state. This helps exclude mortgage borrowers that could be syndicated loan borrowers' employees. The underlying assumption is that most of a firm's workers are employed in its headquarter state, which is a common assumption in the literature (e.g., Serfling, 2016; Bena, Ortiz-Molina, and Simintzi, 2022). Table IB.1 columns (1) and (3) present the results. After applying the exclusion, the sample size is smaller. However, the results hold and the economic magnitude is comparable to that in Table 2.

Some may still worry the validity of the assumption that most of a firm's workers are employed in its headquarter state. To further address the concern, we obtain the geographic dispersion of a firm's business operations at the state level from Garcia and Norli (2012), which is based on firms' 10-K filings. However, the data covering most firms is only available until 2007. To better use the data, we assume that a firm's geographic dispersion in 2008 - 2017 is the same as 2007. Then, each year, we drop a bank-state pair from the HMDA mortgage sample if the bank has a syndicated loan borrower located in the state or the borrower has a reported establishment/subsidiary in the state. For unmatched syndicated borrowers, we still drop their headquarter states. Results are reported in columns (2) and (4) of Table IB.1 and are consistent.

Combined together, the evidence in Table IB.1 helps exclude the alternative story that banks allocate more mortgage credit to a county simply because they prioritize lending mortgages to employees of their corporate borrowers in the county.

Table IB.1. Excluding an Alternative Channel

This table presents robustness checks of the baseline results in Table 2 when excluding an alternative channel. The sample in columns (1) - (3) drops a bank's mortgage lending in a state if it has a syndicated loan borrower located in the state. The sample in columns (2) - (4) drops a bank's mortgage lending in a state if the bank has a syndicated loan borrower located in the state or the borrower has a reported establishment/subsidiary in the state. The dependent variables are the natural logarithm of the number of mortgages a bank approves in a county in columns (1) and (2) and the natural logarithm of the dollar volume (in millions) of mortgages a bank approves in a county in columns (3) and (4). The key independent variable is *Industry Expertise*, a dummy that equals one for a bank-county pair if there exists at least one industry which a bank specializes in and provides at least 5% jobs in a county. Controls include the average loan-to-income ratio of all mortgage applicants, the percentage of male applicants, the percentage of minority applicants, the natural logarithm of one plus the number of branches a bank has in the county, the natural logarithm of the geographic distance between the headquarters county of a bank and the borrower's home county, the natural logarithm of one plus the number of small business loans a bank originates in the borrower's home county, the average percentage of mortgages retained on balance sheets in the borrower's home county in the past three years, the natural logarithm of bank assets, total loans scaled by assets, deposits scaled by assets, commercial and industrial (C&I) loans scaled by total loans, real estate loans scaled by total loans, return on assets, and total liquidity scaled by assets. The sample period is from 1999 to 2017. Standard errors clustered by county are reported in parentheses below the coefficient estimates. ***, **, and * indicate p-values of 1%, 5%, and 10%, respectively.

	(1)	(2)	(3)	(4)
	Log(Number of Approved Mortgages)		Log(Volume of Approved Mortgages)	
	HQ States	Reported States	HQ States	Reported States
Industry Expertise	0.066***	0.087***	0.067***	0.090***
	(0.007)	(0.013)	(0.007)	(0.013)
Observations	211,293	79,871	211,293	79,871
Controls	Yes	Yes	Yes	Yes
County×Year FE	Yes	Yes	Yes	Yes
Bank×State FE	Yes	Yes	Yes	Yes
Adjusted R^2	0.764	0.753	0.796	0.782

References

Bena, J., H. Ortiz-Molina, and E. Simintzi (2022). Shielding firm value: Employment protection and process innovation. *Journal of Financial Economics 146*(2), 637–664.

Cohn, J. B., Z. Liu, and M. I. Wardlaw (2022). Count (and count-like) data in finance. *Journal of Financial Economics 146*(2), 529–551.

Garcia, D. and Ø. Norli (2012). Geographic dispersion and stock returns. *Journal of Financial Economics 106*(3), 547–565.

Serfling, M. (2016). Firing costs and capital structure decisions. *Journal of Finance 71*(5), 2239–2286.