Banks' Geographic Expansion: New Location, Same Old Neighbours

Ka-leung Lam *

Princeton University

November 3, 2023

Abstract

This paper explores how information asymmetry shapes the patterns of banks' geographic expansion. Theory suggests that, despite the removal of legal entry barriers, informational frictions can still hinder bank entry into the deregulated markets. However, the presence of "familiar" firms can alleviate this issue. Leveraging the US interstate banking deregulation as a natural experiment and comprehensive data on locations of bank branches and firm establishments, I find that banks are indeed more likely to expand to new locations with a stronger presence of "familiar" firms. These firms are "familiar" as they already operate in the bank's original neighbourhoods. And I confirm that banks likely have financial interactions with their neighbouring firms, using a novel dataset combining corporate loans with information on borrower and lender locations. I further document that banks' lending patterns in the newly deregulated markets naturally mirror their entry patterns, as credit provision is closely associated with the location of branches. Due to the informational barriers to entry, the statewide deregulation may not benefit all regions and firms equally. Areas where more firms are "known" to the banks from other states experience more entries and higher employment growth. Furthermore, small businesses may not benefit as much as large firms. These findings highlight the potential trade-offs of banking deregulation, and offer guidance for more effective and equitable financial reforms.

---

*Contact: jl88[at]princeton.edu. I am grateful for continuous support and guidance from my advisor Atif Mian. This project benefited greatly from many helpful suggestions from and inspiring discussions with Ernest Liu, Wei Xiong, Motohiro Yogo, as well as seminar participants at Princeton Civitas Finance Seminar and finance student workshop.

1 Introduction

1.1 Motivation

Extensive research has established a strong positive link between financial development, particularly banking deregulation, and economic growth.[1] A fundamental rationale behind banking deregulation is the removal of entry barriers, which encourages foreign banks to enter and expand local credit availability. The physical presence of banking institutions, represented by brick-and-mortar branches, is crucial for enabling credit provision, a prerequisite for economic growth. This significance arises from the information-intensive nature of financing services, where asymmetric information between lenders and borrowers requires monitoring and due diligence by the lenders following credit provision, and geographic proximity can reduce such costs.[2]

However, information asymmetry can also pose challenges for banks seeking to enter new markets in the first place, potentially undermining the effectiveness of deregulation policies. This notion can be traced back to the winner's curse hypothesis within the banking sector, as formulated by Broecker (1990). In this model, banks independently assess the creditworthiness of potential borrowers, but are faced with adversely selected customers who may have been previously rejected by other banks. As the number of banks in the market increases, the probability of at least one bank approving the borrower's financing application increases, leading to a decline in the average quality of successful applicants. Building on this framework, Dell'Ariccia et al. (1999) and Marquez (2002) explicitly demonstrate that adverse selection can limit the entry of new banks into the market. The key assumption is that incumbent banks possess superior information about their local clients while the potential entrants do not. The incumbents can reject refinancing when they identify poor-quality projects, thereby leaving the riskier borrowers to new entrants who can no longer profit. As a consequence, informational barriers discourage market entry.

Despite the compelling theoretical framework, the empirical relevance of the informational barriers to entry remains largely untested. However, given that

[1] See, for instance, King and Levine (1993), Jayaratne and Strahan (1996), Rajan and Zingales (1998), and Levine and Zervos (1998).

[2] Previous studies (e.g., Petersen and Rajan 2002; Nguyen 2019) have highlighted the critical role of geographic proximity in lending relationships.

information frictions influence many aspects of banking operations, their impact on banks' expansion destinations is a critical question. The ramifications of statewide or nationwide banking deregulation might be constrained if information asymmetry affects banks' location choices. Consequently, assessing whether information frictions affect banks' geographic expansion can provide insights into the dynamics of financial development post-deregulation and the mechanisms underlying the effects of financial liberalisation. This, in turn, will inform the adaptation and the design of future policy interventions.

1.2 Preview of the analysis

In this paper, I present evidence that information asymmetry does affect the patterns of banks' geographic expansion. A natural prediction of the theory is that the presence of "familiar" firms in the newly deregulated market can mitigate the issue of information asymmetry, thus lowering the informational barriers to entry. The first part of the analysis tests this hypothesis. I exploit a quasi-experimental design facilitated by US interstate banking deregulation initiated in the 1980s. I start by assuming that banks possess better information about firms located in close proximity—they share operational areas (in this case counties)—due to possible previous financial interactions. This premise is supported by previous studies on the relation between geographic proximity and credit provision, and will be directly tested with loan level data in the second part of the analysis. I then leverage detailed information on firm establishment locations and employment from Dun & Bradstreet database to calculate the employment shares in the newly deregulated counties contributed by each potential entrant bank's original neighbouring firms. This measure represents the bank's familiarity with the new locations and is used to predict entry likelihood. A more substantial presence of these old neighbouring firms should limit the exposure to unfamiliar and potentially adversely selected borrowers, improving the likelihood of bank entry. Consistent with this prediction, I find that banks are indeed more likely to enter a new location where more firms are familiar old neighbours. The specification includes a full set of bank and county fixed effects, absorbing all bank or county specific characteristics that drive expansion. I also rule out the results being driven purely by geographic distance between the bank's headquarters and the destination county. Moreover, the effects are stronger for banks specialised in commercial and industrial lendings, for whom firm information is more relevant.

In the second part of the analysis, I verify the premise that banks have superior information about their neighbouring firms. This can be reflected by the existence of their lending relationships. I construct a novel dataset merging corporate loan data from DealScan with data on firm establishment location from Dun & Bradstreet and data on bank branch locations from Summary of Deposits. First, I compare the list of large multistate firms used to measure informational entry barriers in the first part of the analysis against those appear in the corporate loan data. I find that a significant fraction of these firms accessed banking credit, particularly larger firms. Second, I show that, conditional on borrowing, firms are more likely to borrow from neighbouring banks. These findings add to the previous studies on geographic proximity and lending relationships, and support the argument that banks have better knowledge about their neighbouring firms.

In the third part of the analysis, I examine banks' lending to both businesses and households in the deregulated states. Specifically, I ask how informational barriers of entry affect the geographic distribution of credits following the deregulation. In terms of business lendings, I continue to exploit the merged DealScan-D&B data. To account for the fact that firms operate in multiple regions, I allocate loans across borrowers' establishment locations to measure a bank's lending to each deregulated county. Although this does not necessarily capture where the funds are spent, it does capture where the firms (and their establishments) being financed are located, which is of primary interest for this part of the analysis. First, I document that average lending increases and grows faster in the deregulated regions after the removal of legal restrictions. Second, the growth of credits is more prominent in counties with higher concentration of the bank's old neighbouring firms, i.e., with lower informational barriers to entry. This is partly due to the growth of lending to the bank's original neighbouring firms, and partly to the lending to firms outside the original neighbourhoods but are located in counties where banks are likely to enter due to informational advantages. In terms of household lendings, I use mortgage origination data from Home Mortgage Disclosure Act (HMDA) and find similar results: lending increases in the deregulated markets since deregulation, and mainly in counties with lower informational barriers to entry. These results are consistent with the previously results that information frictions affect entry, and the fact that geographic proximity is strongly associated with credit provision.

In the last part of the analysis, I investigate the aggregate real effects of the information channel of banks' geographic expansion. The previous results suggest that different areas and firms may have been affected differently by the statewide deregulation due to information frictions. Employing a stacked difference-in-differences empirical design and exploiting variation in the aggregate informational barriers to entry across counties within the same state, I find that, counties with higher concentration of banks' familiar firms exhibit higher employment growth in the long run. However, in the short and medium run, these counties experience a lower employment growth, and the results are mainly driven by small businesses. These findings are in line with idea that large multi-regional firms who may have had connections with out-of-state banks elsewhere facilitate entry, while the informationally opaque local small businesses constitute the main source of information asymmetries in the new market. In the long run, however, as the entrant banks settle, and become familiar with the local clients, small businesses can also benefit from the increased banking services brought by the entrants. The results also support the theoretical argument by Petersen and Rajan (1995) that increased banking competition harms small businesses who rely more on local long-term banking relationships. Fortunately, the effects may be short-lived.

1.3 Related literature

This paper relates to several strands of the literature.

First, the paper contributes to the extensive literature on information frictions in the credit markets. Banking theories have long been emphasising the effects of information asymmetry between lenders and borrowers on financial and real outcomes (Leland and Pyle, 1977; Diamond, 1991; Holmström, 1979; Holmstrom and Tirole, 1997). Empirically, Dennis and Mullineaux (2000), Sufi (2007) and Ivashina (2009) investigate how information asymmetry affects the structure and terms of loan contracts; Petersen and Rajan (2002), Agarwal and Hauswald (2010), and Nguyen (2019) show that geographic proximity can reduce information asymmetry, especially for small businesses; Bharath et al. (2007)

and Chodorow-Reich (2014) document the persistence in banking relationships; and finally Chodorow-Reich (2014), Huber (2018), and Greenstone et al. (2020) examine the real effects of the credit market frictions during financial crisis.

The contributions of the current paper to these previous studies are thus twofold. First, I provide new evidence that geographic proximity matters for financing outcomes, even for large firms. Most previous studies have focused on small business lendings, or usually measure distance between lenders and borrowers based on locations of their headquarters, whereas in this paper, I show that whether multi-regional banks and firms operate in the same neighbourhoods also plays a role in forming lending relationships. Second, I document that information frictions also influence banks' geographic expansion. The fact that banks are more likely to enter a new location where firms are old neighbours partly contribute to the persistence of lending relationships shown by the previous studies. The results in this paper can be viewed through the models by Dell'Ariccia et al. (1999) and Marquez (2002) who demonstrate that adverse selection prevents entry. Closely related empirical studies include Mian (2006) and Bofondi and Gobbi (2006). Mian (2006) shows that foreign banks still lend to informationally transparent firms in Pakistan following the banking deregulation, even after establishing local branches. Bofondi and Gobbi (2006) evaluate banks' lending performance following Italian banking deregulation, and find that loans originated by entrant banks have higher default rates than incumbents, which suggests that entrant banks do face adversely selected borrowers.

Second, this article contributes to the large literature on banking deregulation, most of which focus on the effects of deregulation on the banking sector or the real economy. For example, Jayaratne and Strahan (1996) show that relaxations of intrastate branch restrictions leads to economic growth; Chava et al. (2013) and Cornaggia et al. (2015) investigate how banking deregulation spurs innovative activities; Goetz et al. (2013, 2016) and Chu et al. (2020) document how deregulation induced geographic diversification affects the efficiency and riskiness of the banking sector. Findings on the effects of banking deregulation on small businesses are mixed. Theoretically, Petersen and Rajan (1995) argue that increased competition in the banking sector makes it harder for lenders to extract rents from long term lending relationship, which harms small businesses who rely more on relationship financing. Boot and Thakor (2000), on the other hand, argue that competition could encourage more relationship lending since relationship lending insulates banks from competition relative to transaction lending. Empirically, Black and Strahan (2002) find that banking deregulation helps new business formation; Rice and Strahan (2010) find that credits become cheaper for small businesses after the deregulation, but there is no effect on the amount of borrowing; Cannon and Lynch (2023), however, show that deregulation reduces credit availability to small businesses.

The results in this paper complements these previous findings in various aspects. First, limited attention has been paid to what factors are driving banks' geographic expansion. In this paper, by taking the banks' location choice as an endogenous outcome, I identify the role of information frictions in shaping the pattern of banks' geographic expansion. A similar attempt is by Gropp et al. (2019) who show that banks expand to areas where local natural disasters are more correlated with their home territories. They conclude that banks may have been leveraging their expertise in dealing with natural disasters when choosing new location. This result is similar in spirit to this paper where I find that the presence of familiar firms from the original neighbourhoods mitigates the problem of information asymmetry. Second, the findings that banks are more likely to enter a location where more firms are old neighbours cautions that the benefits of geographic diversification, which has been the focus of most previous research, may be limited. Third, I find that variation in informational barriers result in different regions and firms being affected differently by the deregulation. In particular, counties with lower informational barriers of entry (in the aggregate) experienced a larger number of bank entries and higher employment growth in the long run. However, in the short and medium run, there is some adverse effects on employment in these counties, primarily on small businesses. These results are in line with the theory by Petersen and Rajan (1995) who argue that banking competition diminish the value of relationship lending, and particularly harms small businesses who rely more on this type of financing due to their informational opaqueness. However, in the long run, they also enjoy the benefits from deregulation possibly because new entrant banks become familiar with the local businesses after entry.

Third, the paper is broadly related to research on the interaction between the financial sector and the real economy. Apart from the above-mentioned studies on how financial crises affect domestic economy, and how banking deregulation causes economic growth, a series of paper by Schularick and Taylor (2012), Jordà et al. (2013), Mian et al. (2017), and Mian et al. (2020) document how credit booms and busts drive business cycles; Peek and Rosengren (1997, 2000), Chava and Purnanandam (2011) and Schnabl (2012) explore the international transmission of financial shocks; Federico et al. (2023) investigate how trade shocks propagate through the banking sector. This paper shows that the landscape of the real economy can shape the development of the financial sector, which in turn affects the way financial liberalisation influences the real economy.

1.4 Structure of the paper

The rest of the paper is organised as follows. Section 2 lays out the conceptual framework. Section 3 describes the data sources. Section 4 introduces relevant institutional background on the US interstate banking deregulation. Section 5 discusses the empirical strategy and report the main results on entry outcomes. Section 6 provides evidence on lending relationship between firms and neighbouring banks. Section 7 examines lending outcomes in the deregulated regions. Section 8 investigate the heterogeneous employment effects of banking deregulation due to informational barriers of entry. Finally, Section 9 concludes.

2 Conceptual Framework

The empirical inquiry of this paper can be motivated by the theoretical models presented in Broecker (1990), Dell'Ariccia et al. (1999) and Marquez (2002). Broecker's model illustrates a competitive credit market in which the problem of asymmetric information leads to adverse selection and the winner's curse. In this model, banks conduct independent tests to access potential borrowers' creditworthiness before granting financing. However, they are unaware of whether the applicant has been rejected by other banks. As the number of banks increases, the fraction of borrowers passing at least one test and obtaining financing increases, but the average quality of the successful applicants deteriorates.

Dell'Ariccia et al. (1999) and Marquez (2002) explicitly demonstrate how adverse selection constrains the number of banks in the market. In their models, incumbent banks possess superior information about some local clients due to previous interactions, allowing them to reject refinancing once they identify poor quality projects. Potential entrant banks, on the other hand, cannot distinguish between new borrowers due to turnover and the old ones who may have been rejected by the incumbents. As a result, they cannot profit from the adversely selected pool of borrowers, and entry is blocked in equilibrium.

A natural implication of the theory is that if entrant banks also possess information about the borrowers in the new market, then the adverse selection problem would be less severe. How could a bank have gained knowledge about the borrowers in the new location? Bank clients are typically households or firms. While households are mostly unfamiliar in the new location, firms may operate in multiple regions, and some of them may have been operating in the bank's original neighbourhoods. Given that previous studies have documented that geographic proximity is strongly related to lending relationships, it is plausible that banks have better knowledge about these neighbouring firms than those farther away through financial interactions. The presence of these familiar firms in the new location can thus reduce the informational barriers to entry. The upcoming analysis aims to test this prediction.

A general identification concern is reverse causality. Bank locations and firm locations may be jointly endogenously determined, making it difficult to establish the direction of causality. To address this concern, I leverage the US interstate banking deregulation as a natural experiment. The interstate banking regulation limits banking activities across state borders, while firms were not subject to similar restrictions. Additionally, as will be demonstrated shortly, many firms were already operating in multiple regions before banks. Therefore, I am able to use the pre-deregulation distribution of firm locations as the key explanatory variable to predict bank entry outcomes.

3 Data

The analysis utilises data from multiple sources at the bank, firm, and loan levels.

The timing of interstate banking deregulation is documented by Amel (1993). Based on this documentation, I have compiled a chronology of newly deregulated state pairs, which can be found in Appendix A.

The data regarding the locations of bank branches are sourced from the Summary of Deposits (SoD). The data for the later period (1987–) were obtained from the FDIC website.[3] For the earlier period (1981–1986), this information was acquired from the work of Christa H.S. Bouwman.[4] SoD includes details about the county locations of bank branches, the amount of deposits as of June 30 each year, and the identification of banks' ultimate parent holding companies. In the case of SoD from 1987 onwards, it also provides information about the headquarter counties of the parent BHCs, which I use to determine their home states. For the earlier period (1981–1986), I gathered information about BHCs' headquarter locations from FFIEC National Information Center online data repository.[5] Finally, bank balance sheet data comes from Call Reports.

Data related to firm location and employment are extracted from Dun & Bradstreet (D&B) database. This comprehensive database includes information on the location, employment figures, and industry classification (SIC) of all business establishments across the United States. It also offers insights into firm ownership structure as of the reported year.

For the analysis of lending outcomes, I utilise data on syndicated loans from DealScan with deal active date spanning from 1987 to 2005. Borrowers are manually matched to D&B database based on company name, location, and industry information provided by both sources. Borrowers are then assigned identifiers (DUNSNO) from D&B. Lenders are linked to the ultimate parent BHCs in SoD using names and locations, and they are assigned identifiers (RSSDID) from SoD.

For the analysis on employment outcomes, I acquire data from Census Business Dynamics Statistics (BDS) spanning from 1978 to 2008. This database provides information on employment categorised by firm size at the county level.

[3] Data from 1987–1993 is available at https://www.fdic.gov/foia/sod/, and data from 1994 onwards can be found at https://www.fdic.gov/bank/statistical/index.html.

[4] Data for 1981–1993 is available in part D of the data page on https://sites.google.com/a/tamu.edu/bouwman/home.

[5] Data can be accessed at https://www.ffiec.gov/NPW

4 Institutional Background

The empirical analysis of this paper leverages the natural experiment from interstate banking deregulation in the United States. During this period, the removal of legal entry barriers allowed bank holding companies (BHC) to operate across state borders. In this section, I provide an overview of the relevant institutional background regarding the interstate banking deregulation. I will focus particular features of the deregulation that facilitate my empirical design.

According to Amel (1993), the state of Maine was the first to eliminate its interstate banking restrictions and open its banking sector to all other states in 1978. However, Maine's approach included a reciprocity requirement, stating that out-of-state bank holding companies could only enter if the foreign state in which the BHC was headquartered also permitted entry of banks from Maine. No no other state took actions until 1982 when New York also passed its deregulation legislation. Afterwards, many states adopted similar measures and deregulated their banking industries. Ultimately, in 1994, a federal law called Riegle-Neal Interstate Banking and Branching Efficiency Act went into effect, essentially removing the remaining restrictions on banking activities across state borders. The deregulation process is often characterised as chaotic for two primary reasons. First, deregulation legislation often contained reciprocity provisions, similar to Maine's, making the ability of a bank to enter a deregulated state dependent on the banking regulations of its home state as well. Second, not all states opened their banking sectors to the entire country in a single legislation move; instead, they often began by opening to some specific designated states and gradually expanded the list.

Figure 1 visualises the evolution of the interstate banking deregulation, and a detailed chronology can be found in Appendix A. Panel A displays the number of newly deregulated state pairs each year, while Panel B shows the cumulative fraction of deregulated state pairs. The year of deregulation (for this figure and the rest of the paper) is the one in which one state effectively permitted entry of BHCs from another state, accounting for the reciprocity requirement in the relevant legislations. Each state pair is classified according to their deregulation status: "Unilateral" indicates that only one of the states allowed entry of BHCs from the other, while "Bilateral" indicates that BHCs headquartered in both states were allowed to enter each other's market. In Panel A, the dereg-

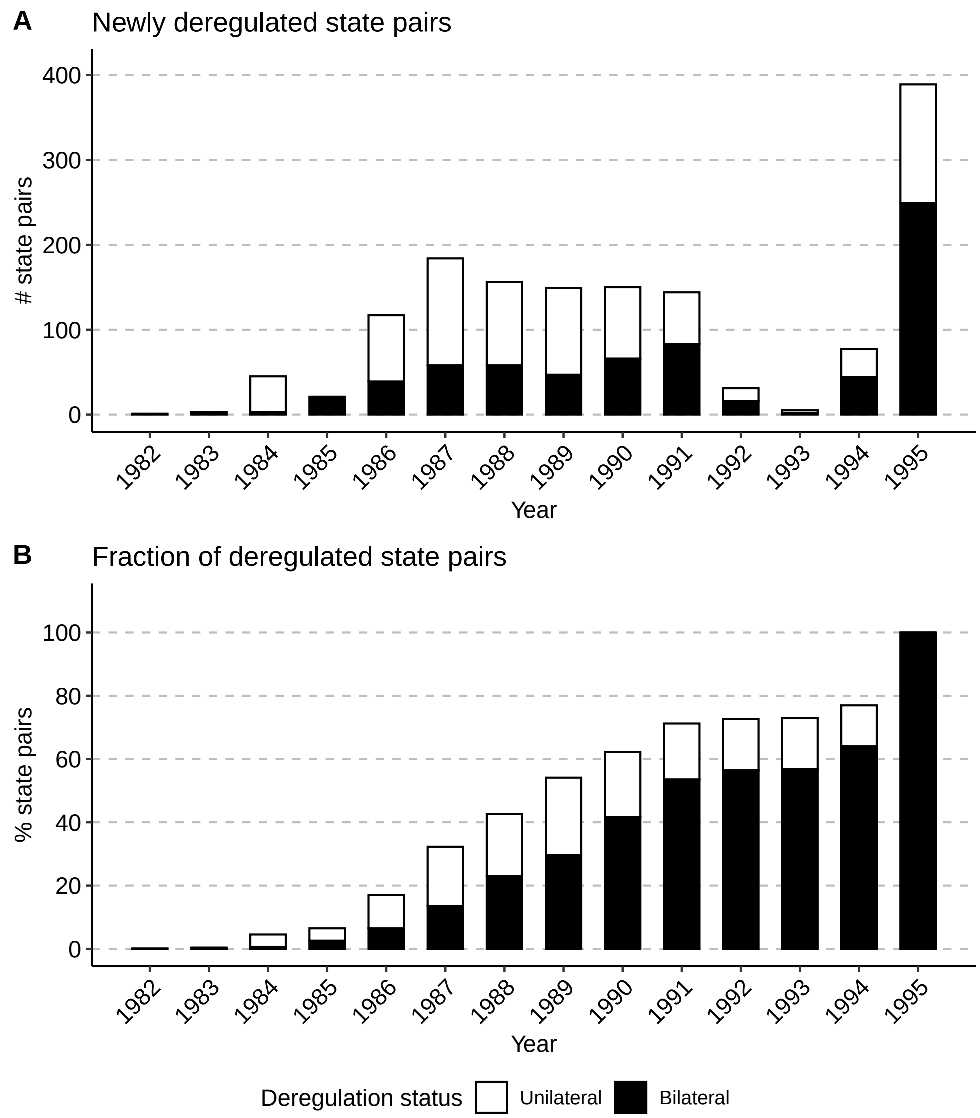

Figure 1: Evolution of Interstate Banking Deregulation

Notes: This figure plots the evolution of interstate banking deregulation. Panel A shows the number of newly deregulated state pairs in each year. The deregulation status of each pair is classified as unilateral if only one state permitted entry of BHCs from the other in that year, and as bilateral if both states allowed entry of BHCs from each other. Panel B shows the cumulative fraction of deregulated state pairs. The deregulation status of each pair is classified as unilateral if only one state has permitted entry of BHCs from the other by the end of the year, and as bilateral if both states have permitted entry from each other. The sample includes 47 contiguous states (i.e., Alaska, Delaware, Hawaii and South Dakota are excluded.) and thus 1081 (= 47 × 46/2) state pairs in total.

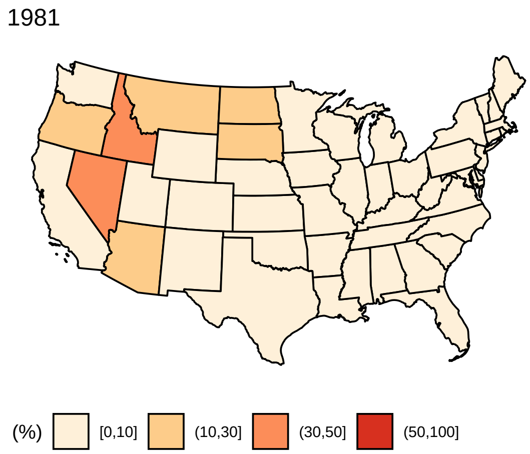

Figure 2: Share of Branches by Out-of-State BHCs

Notes: This figure plots the share of branches controlled by out-of-state BHCs in 1981 versus 2006. Source: Summary of Deposits.

ulation status is classified based on the new legislation, while in Panel B, it is classified based on the end-of-year status of the state pair. For example, Texas deregulated in 1987 to allow entry of BHCs from Alabama, and Alabama also deregulated in 1988. Therefore, in Panel A, the state pair AL-TX is counted as one of the unilaterally deregulated pairs in 1987 and again as one of the unilaterally deregulated pairs in 1988. In Panel B, however, AL-TX is counted as one of the unilaterally deregulated pairs in 1987 and as one of the bilaterally deregulated pairs in 1988. The sample includes 47 contiguous states, with Hawaii excluded due to its remote geographic location, Alaska excluded due to its lack of regular county subdivisions, and Delaware and South Dakota excluded due to their special arrangements for credit card businesses. Therefore, there are 1,081 state pairs in total ($= 47 \times 46/2$). Panel A shows that during the most active deregulation years (1986–1991), around 150 state pairs that were involved in deregulation events, with the majority being unilateral deregulations.

Figure 2 illustrates the impact of interstate banking deregulation on the entry of out-of-state bank holding companies. It compares the situation in 1981, one year before the interstate banking deregulation began, to 2006, over a decade after the passage of IBBE. Prior to deregulation, almost three-quarters of states had virtually no presence of out-of-state BHCs. However, in 2006, out-of-state BHCs had become prevalent. Nearly all states had at least 10% of branches

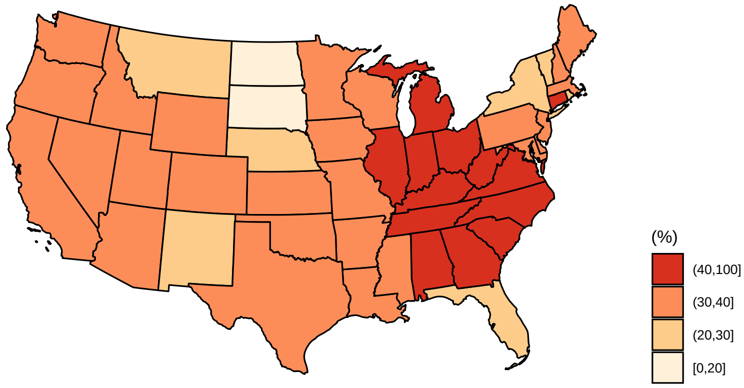

Figure 3: Employment Share of Multistate Firms

Notes: This figure plots the employment share in each state and county by large multistate firms in 1980. Source: Dun & Bradstreet.

controlled by out-of-state BHCs, with some states exceeding 50%.

On the other hand, many firms were already operating across state lines before banks did. Panel A of Figure 3 displays the employment share of large multistate firms (those with more than 500 employees) in each state in 1980. Most states had at least 10% of the labour force employed by these multistate firms, with some states exceeding 40%. This makes it possible to examine the role of multistate firms in reducing the informational barriers to entry and shaping the pattern of banks' geographic expansion.

5 Old Neighbours and Bank Entry

5.1 Specification

I estimate the following linear probability model for bank entry outcomes:

$$Entry_{bc,t(d)+h} = \beta[Old\ Neighbours]_{bc,t(d)-1} + \phi_{b,t(d)} + \phi_{c,t(d)} + \gamma X_{bc,t(d)-1} + \varepsilon_{bc,t(d)}. \tag{1}$$

The dependent variable $Entry_{bc,t(d)+h}$ is an indicator of whether bank b enters county c, h years after the deregulation. It assumes a value of one if, in year $t(d)+h$, a branch in county c is controlled by bank b. The index d represents the event in which the state of county c deregulated its banking sector to the state where bank b is headquartered, and $t(d)$ is the year in which deregulation took effect.

The key explanatory variable $[Old\ Neighbours]_{bc,t(d)-1}$ measures the employment share in the destination county c by the original neighbouring firms of bank b (to be defined shortly). It is computed in the year prior to deregulation. This variable reflects the potential entrant bank's familiarity with the new location, and consequently, the degree of information asymmetry. Specifically, a higher share of old neighbouring firms reduces information asymmetry.

The variables $\phi_{b,t(d)}$ and $\phi_{c,t(d)}$ represent bank-year and county-year fixed effects, which absorb bank- or county- specific characteristics that may influence bank entry outcomes and correlated with employment share of banks' neighbouring firms. For example, larger banks may have greater capacity for expansion and may also be connected to a larger number of firms.

Finally, the variables X include controls that vary only across bank-county pairs. In particular, I include the geodesic distance between bank's headquarters and the destination county as a control. The distance captures the administrative costs of operating a new branch, and it is typically less costly to manage a branch nearby. However, it is likely that the same firms are also operating in those areas. Including distance as a control makes sure that the variable *Old Neighbours* is not simply capturing the effects of distance and management costs.

5.2 Sample and variables

To conduct the test described above, I constructed a dataset with bank-county pairs as the unit of observation. For each deregulation event in which a state opened its banking sector to banks from other states, I designated the state that initiated the deregulation as the "home state" and the states whose banks were permitted entry as the "foreign states." This made all counties in the home state as potential destinations for entry. As for potential entrants, I compiled a list of BHCs headquartered in the foreign states one year before the deregulation, excluding those had already entered the home state. To ensure the sample of banks was most relevant for studying of geographic expansion, I only included BHCs that were reasonably sizable, those that had ever achieved more than one billion deposits, and had expanded across state borders during the sample period

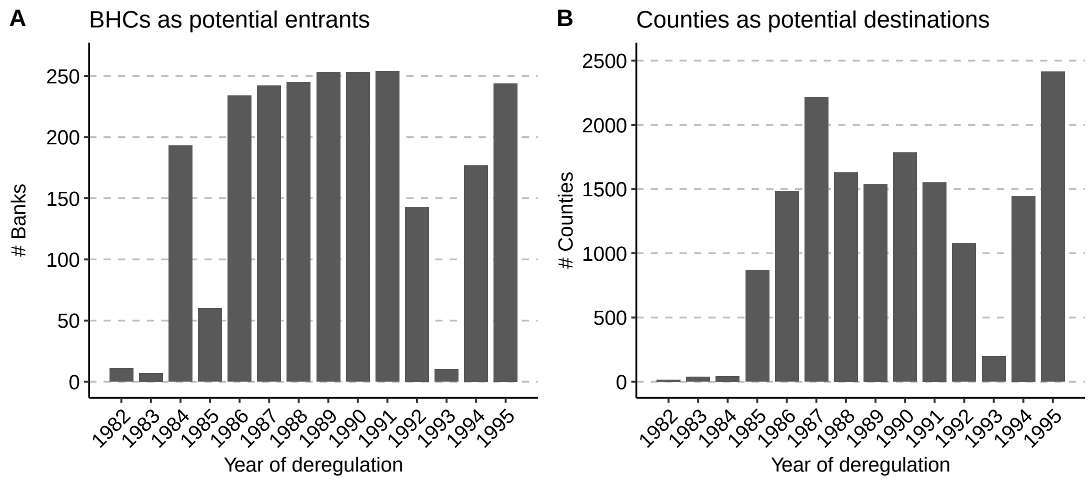

Figure 4: Bank Entry Sample

Notes: This figure plots the number of potential entrant banks and potential desitnation counties that are affected by the deregulation in each year. Potential entrant banks are banks that have not entered the deregulated states at the time of deregulation, that have ever had at least one billion deposits and that have ever expanded across state borders during the sample period (1982–2005). Potential destination counties are all counties in the deregulated states. Note that not all banks are paired with all counties displayed in this figure for the entry analysis, only those affected by the deregulation are included.

(1982–2005). Subsequently, each potential entrant bank was associated with each potential destination county it could enter. Then I stacked bank-county pairs from all deregulation events to create a large cross-section of bank-county pairs.

In Figure 4, Panel A illustrates the total number of BHCs as potential entrants for each year of deregulation, while Panel B shows the total number of counties within the deregulated states. Notably, during the most active years of deregulation, there were more than 200 banks permitted to enter new markets, with approximately 1500 counties becoming potential destinations.

To calculate the employment share of a bank's neighbouring firms, I first define a bank b and a firm f being neighbours in year t if they operate in the same county:

$$Neighbour_{fbt} = \mathbb{1}\{C_{ft} \cap C_{bt} \neq \emptyset\}, \tag{2}$$

where C_b represents the set of counties where bank b has a branch, and C_f the set of counties where firm f has an establishment. Using this indicator

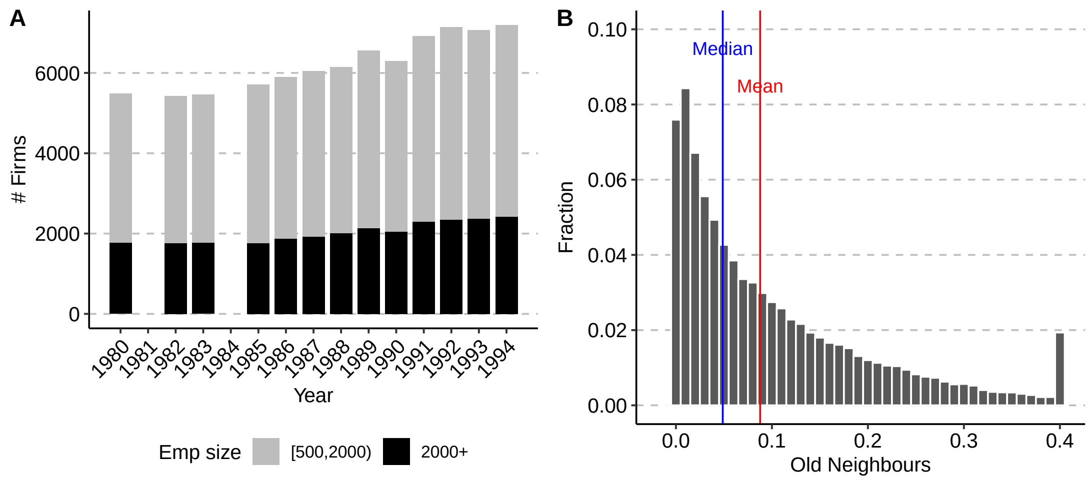

Figure 5: Multistate Firms and Distribution of Old Neighbours

Notes: Panel A plots the number of large multistate nonfinancial firms used to construct variable *Old Neighbours*. Panel B plots the distribution of variable *Old Neighbours* across bank-county pairs. For illustrative purpose, values of zero are omitted from the figure (around 15% of the observations), but are included in the calculation of the mean and median). The last bin contains all values greater than 0.4.

variable, I then calculate the employment share of a bank' neighbouring firms in the destination as follows:

$$[Old\ Neighbours]_{bct} = \frac{\sum_f Employment_{fct} \times Neighbour_{fbt}}{\sum_f Employment_{fct}}, \tag{3}$$

where the numerator represents the employment by the bank's neighbouring firms and the denominator is the total employment of the county.

Panel A of Figure 5 shows the number of firms used to construct the variable *Old Neighbours*. These firms are non-financial non-governmental entities with more than 500 employees, operating across multiple states. There are approximately 6,000 of them, with about one-third having more than 2,000 employees. Note that D&B data were not produced for years 1981 and 1984. Therefore, when data from these two years are needed, I use the data from the previous year. For example, if data from 1984 is required, I use the data from 1983.

Panel B of Figure 5 presents the distribution of the variable *Old Neighbours* across bank-county pairs. For illustrative purpose, approximately 15% of the observations with zero values are omitted from the figure. The last bin of the histogram includes all values greater than 0.4. The figure reveals substantial

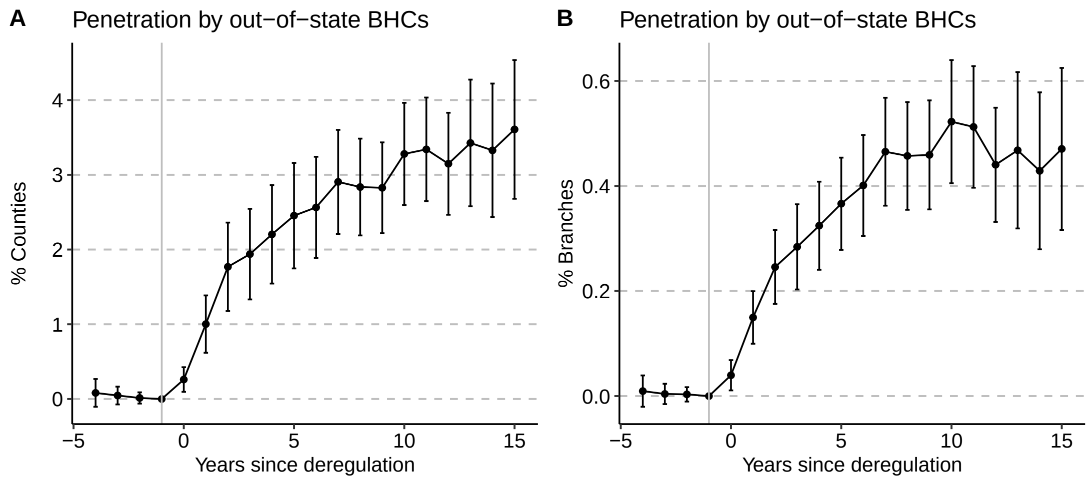

Figure 6: Speed of Entry

Notes: This figure plots the coefficients from Equation 4. The dependent variable $FrgnBHCShare_{ijh}$ is the penetration rate of home state i by banks in foreign state j, h years since the deregulations. Panel A measures the penetration rate by the share of counties entered by foreign banks, while Panel B measures the penetration rate by the share of branches owned by foreign banks. The error bars indicate 95% confidence intervals, with standard errors clustered at home-foreign state pair level.

variation in the variable of interest. The sample mean (including zeros) is 8.8%, while the median is 4.9%. The standard deviation is 10.6%.

5.3 Timing

Before presenting the regression results, I assess the speed of entry by conducting the following event study regression using a state-pair panel dataset:

$$FrgnBHCShare_{ijh} = \phi_{ij} + \sum_{r \neq -1} \beta_r \mathbb{I}_{r=h}, \tag{4}$$

where the dependent variable $\text{FrgnBHCShare}_{ijh}$ measures the extent of penetration in the home state i by banks from foreign state j, h years since their deregulation. I create two measures of penetration rate: the fraction of counties entered by out-of-state BHCs and the share of branches owned by those banks. ϕ_{ij} represents state pair fixed effects. Therefore, the coefficients of interest, β_r's, indicate the average change in penetration rate over time, relative to the year before deregulation.

Table 1: Old Neighbours and Bank Entry

	(1) Entry$_{+2}$	(2) Entry$_{+2}$	(3) Entry$_{+10}$	(4) Entry$_{+10}$
Old Neighbours	1.15***	0.926***	3.39***	2.72***
	(0.28)	(0.266)	(0.71)	(0.668)
LogDist		−0.748***		−2.02***
		(0.168)		(0.329)
Bank-Year FE	Yes	Yes	Yes	Yes
County-Year FE	Yes	Yes	Yes	Yes
Mean of Dep Var (%)	0.19	0.19	0.75	0.75
R^2 Adj	0.144	0.146	0.237	0.242
Observations	653,257	653,257	442,097	442,097

Notes: This table reports regression results from Equation 1. The dependent variable $Entry_{+h}$ is an indicator for whether bank controls a branch in the destination county, h years after deregulation. Variable *Old Neighbours* is the employment share in the destination county of a bank's original neighbouring firms, divided by 100. Variable *LogDist* is the log geodesic distance between the bank's headquarters and the destination county (in miles), divided by 100. Standard errors are clustered at bank and county levels. Significance levels: ***1%, **5%, *10%.

Figure 6 plots the estimates from four years prior to deregulation to fifteen years after. The coefficient for the year immdediately before deregulation is normalised to zero. The figure suggests that entries occur rapidly within the first three years following deregulation, with penetration rates reaching their peak at approximately ten years before stabilising. Both penetration measures exhibit similar patterns.

In the following section, I report results on entry outcomes in both the short term (two years since deregulation) and the long term (ten years since deregulation). Additional results can be found in Appendix C.

5.4 Entry outcomes

Table 1 reports results from Equation 1. The coefficients are scaled up by 100. In column (1), where I estimate the likelihood of entry two years after deregulation, the estimated coefficient on *Old Neighbours* is 1.15, indicating that a one-standard-deviation increase in the employment share by the bank's original neighbouring firms (10.5%) increases the probability of entry by 11.5 basis points. This magnitude is nearly two thirds of the unconditional mean (19 bp).

Column (2) introduces control for geodesic distance between bank headquarters and destination counties. The negative coefficient on distance suggests that a longer distance leads to a reduced likelihood of entry. While the coefficient on *Old Neighbours* decreases slightly in magnitude, reflecting its partial correlation with geographic distance, it remains statistically significant. This implies that familiar firms play an additional role alongside geographic distance in influencing banks' entry. In particular, I attribute the effects to the reduction in information frictions.

Columns (3) and (4) present results on entry likelihood ten years after deregulation. The coefficients for *Old Neighbours* remain positive and statistically significant. In column (4), for instance, the estimated value is 2.72, which suggests that a one-standard-deviation increase in the neighbouring firm share (10.1%) increases the probability of entry in ten years by 27.2 basis point, more than one third of the unconditional mean (75 bp).

Moreover, the inclusion of bank-industry fixed effects, where the industry is the 2-digit SIC with the largest employment share in the county, has minimal impact on the results (see Table C.2), suggesting that the findings are not solely driven by banks' industry specialisation.

5.5 Heterogeneous effects

To provide further evidence in support of the hypothesis that the presence of familiar firms reduces information asymmetry, I explore heterogeneous effects of old neighbours on bank entry across banks with different business models. Theory implies that information asymmetry deriving from the composition of firms in the new markets should be more relevant to potential entrant banks who have a greater focus on corporate lending businesses. The results, as presented in Table 2, substantiate this hypothesis.

The coefficients associated with the interaction between *Old Neighbours* and *C&I Loan*, which measures the proportion of the bank's assets allocated to commercial and industrial loans, are both positive and statistically significant. This indicates that the effects of neighbouring firm share are stronger for banks specialising in corporate lending.

Furthermore, the coefficients on the non-interaction term *Old Neighbours* are now close to zero and statistically insignificant. This suggests that if a bank does not have corporate lending businesses (i.e., *C&I Loan* = 0), then the presence of

Table 2: Old Neighbours and Bank Entry: Heterogeneity

	(1) Entry_{+2}	(2) Entry_{+10}
Old Neighbours	0.00326	−1.06
	(0.298)	(0.883)
Old Neighbours × C&I Loan	6.05^{**}	25.1^{***}
	(2.59)	(8.62)
LogDist	-0.75^{***}	-2.03^{***}
	(0.168)	(0.329)
Bank-Year FE	Yes	Yes
County-Year FE	Yes	Yes
Mean of Dep Var (%)	0.19	0.75
R^2 Adj	0.146	0.242
Observations	653,241	442,097

Notes: This table reports regression results from Equation 1, allowing effect heterogeneity across bank characteristics. The dependent variable $Entry_{+h}$ is an indicator for whether the bank controls a branch in the destination county, h years after deregulation. Variable *Old Neighbours* is the employment share in the destination county by the bank's original neighbouring firms. Variable *C&I Loan* is the share of assets invested in commercial and industrial loans. Standard errors are clustered at bank and county levels. Significance levels: ***1%, **5%, *10%.

old neighbouring firms is irrelevant in reducing the informational barriers to entry.

5.6 Placebo tests

To ensure that geographic expansion was genuinely constrained before deregulation and that the presence of old neighbours should have insignificant effects on expansion, I conduct placebo tests predicting banks' entry into states that were still under banking regulation, using the same specification as in Equation 1. For example, to construct the dataset for a placebo test with a two-year horizon, I select state pairs from years 1982 to 1992, where the home state remains regulated two years later. Specifically, in year 1982, state pairs (ordered as home-foreign) such as Arizona-California are included in the sample, since banks in California were still prohibited from entering Arizona in 1984. However, state pairs like California-Arizona are excluded, as they would be deregu-

Table 3: Old Neighbours and Bank Entry: Placebo Tests

	(1)	(2)
	Entry_{+2}	Entry_{+10}
Old Neighbours	2.81	−4.50
	(2.04)	(3.32)
LogDist	−3.93	−9.26
	(2.70)	(6.81)
Bank-Year FE	Yes	Yes
County-Year FE	Yes	Yes
Mean of Dep Var (bp)	1.3	3.1
R^2 Adj	0.011	0.037
Observations	4,235,481	461,116

Notes: This table reports results from placebo tests of bank entry model in Equation 1. The dependent variable $Entry_{+h}$ is an indicator for whether bank controls a branch in the destination county, h years after deregulation. Variable *Old Neighbours* is the employment share in the destination county of a bank's original neighbouring firms, divided by 10^4. Variable *LogDist* is the log geodesic distance between the bank's headquarters and the destination county, divided by 10^4. Standard errors are clustered at bank and county levels. Significant level: ***1%, **5%, *10%.

lated in 1984 and banks in Arizona were permitted entry to California. I then associate banks in California to counties in Arizona, and predict their entry as if they were deregulated, using banks' old neighbouring firm shares calculated in 1980. For the ten-year horizon, it is only possible to designate placebo deregulation in years 1982–1984 because starting in 1985, all state pairs would have undergone deregulation within the next ten years.

The results of placebo regressions are presented in Table 3. The reported coefficients on *Old Neighbours* are scaled up by 10^4, so the effects are close to zero and statistically insignificant. This suggests that bank expansion is indeed a result of deregulation.

6 Lending Relationship

The validity of the previous empirical approach relies on the premise that banks possess superior knowledge about firms in their local area compared to those further away. This premise can be supported by the fact that neighbouring firms and banks have financial interactions. In this section, I provide evidence in sup-

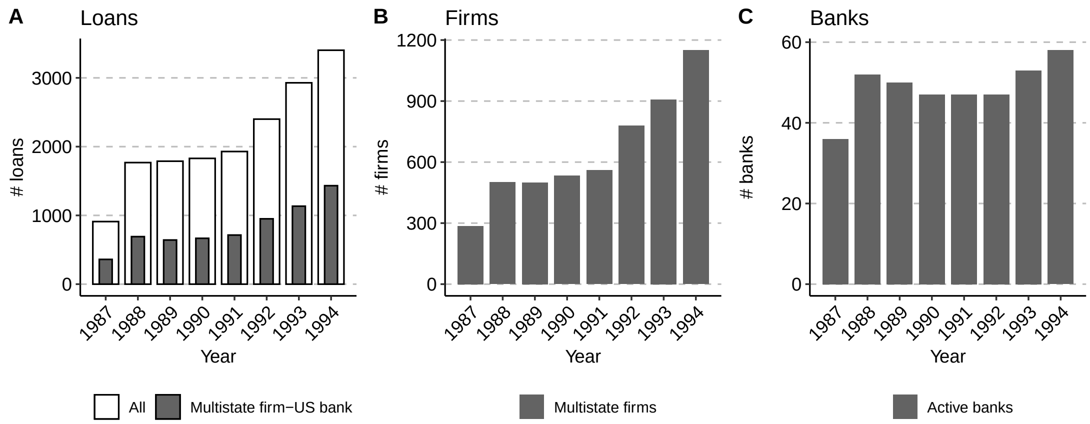

Figure 7: Lending Relationship Sample

Notes: This figure plots the number of loans and unique borrowers and lenders used to investigate lending relationships. Panel A plots the total number of loans in DealScan to US borrowers. The grey portion represents the fraction of loans for which the borrower can be matched to Dun & Bradstreet and its ultimate parent is a non-financial or non-government firm operating in multiple states. Panel B plots the number of unique firms, and Panel C the number of active lenders (appear on more than 10 loans).

port of this assumption, using a novel dataset that combines corporate loan data from DealScan with data on firm establishment locations from Dun & Bradstreet and data on bank branch locations from Summary of Deposits.

Although DealScan covers mostly syndicated loans, a specific type of corporate loans, and has limited observations in the earliest part of the database, I find that a significant fraction of multistate firms, especially larger firms, appear in the database. This suggests that these firms do rely on banking credits. More importantly, conditional on borrowing, firms are more likely to borrow from neighbouring banks, confirming the conjecture that banks have better knowledge about neighbouring firms.

The sample of loans for this part of the analysis is derived from DealScan database, with deal active dates between 1987 and 1994, as this period is most relevant, due to ongoing deregulation, for validating the link between geographic proximity and lending relationships.[6] Panel A of Figure 7 illustrates the total number of loans designated to US borrowers by DealScan in each year, as represented by the white bars. I manually match the borrowers to D&B based on information such as company name, industry and location provided by both data

[6] DealScan has very few observations before 1987.

sources, excluding financial and government entities. I aggregate these loans to the firm's ultimate parent, and retain only firms that operate across multiple states. I also match the lenders to banks in the Summary of Deposits, excluding loans in which no US banks are involved. This matching process reduces the sample size by approximately two-thirds. The remaining loans, as represented by the grey bars in Panel A, are thus those borrowed by multistate firms and lent by at least one US bank.

Panel B of Figure 7 displays the number of unique multistate firms, with roughly 300 to over 1000 firms borrowing in any given year. Panel C shows the number of active active lenders (BHCs) that appeared in at least 10 loans in a given year. There are approximately 50 active banks in each year.

First, I assess the coverage of multistate firms by DealScan borrowers. Panel A of Figure 8 shows the total number of multistate firms used to construct the variable *Old Neighbours*. These firms are non-financial, non-governmental entities with more than 500 employees that operate in multiple states. There are approximately 6,000 such firms, with one-third of them having more than 2,000 employees. Panel B reproduces Panel B of Figure 7, representing the borrowing firms in DealScan, with more than half of them having over 2,000 employees.

Panel C of Figure 8 indicates that around 10% of all multistate firms (with more than 500 employees) carried out new loans in a given year. This figure is calculated by dividing the grey bars in Panel B with grey bars in Panel A. And these borrowing firms account for 20% to 50% of total employment by all multistate firms. Panel D of Figure 8 shows that approximately 35% of the multistate firms had borrowed at least once during the sample period, and they account for more than 70% of total employment.

The coverage of firms becomes more significant for larger firms, as shown in Panels E and F of Figure 8. Among firms with more than 2,000 employees, 10% to 30% of them borrowed each year, accounting for 20% to 50% of the employment of these firms. Furthermore, 60% of these firms borrowed at some point during the sample period, accounting for around 80% of the total employment.

Next, to demonstrate that firms are more likely to borrow from neighbouring banks, I estimate the following linear probability model:

$$Loan_{fbt} = \phi_{ft} + \phi_{bt} + \beta Neighbour_{fbt} + \varepsilon_{fbt}, \tag{5}$$

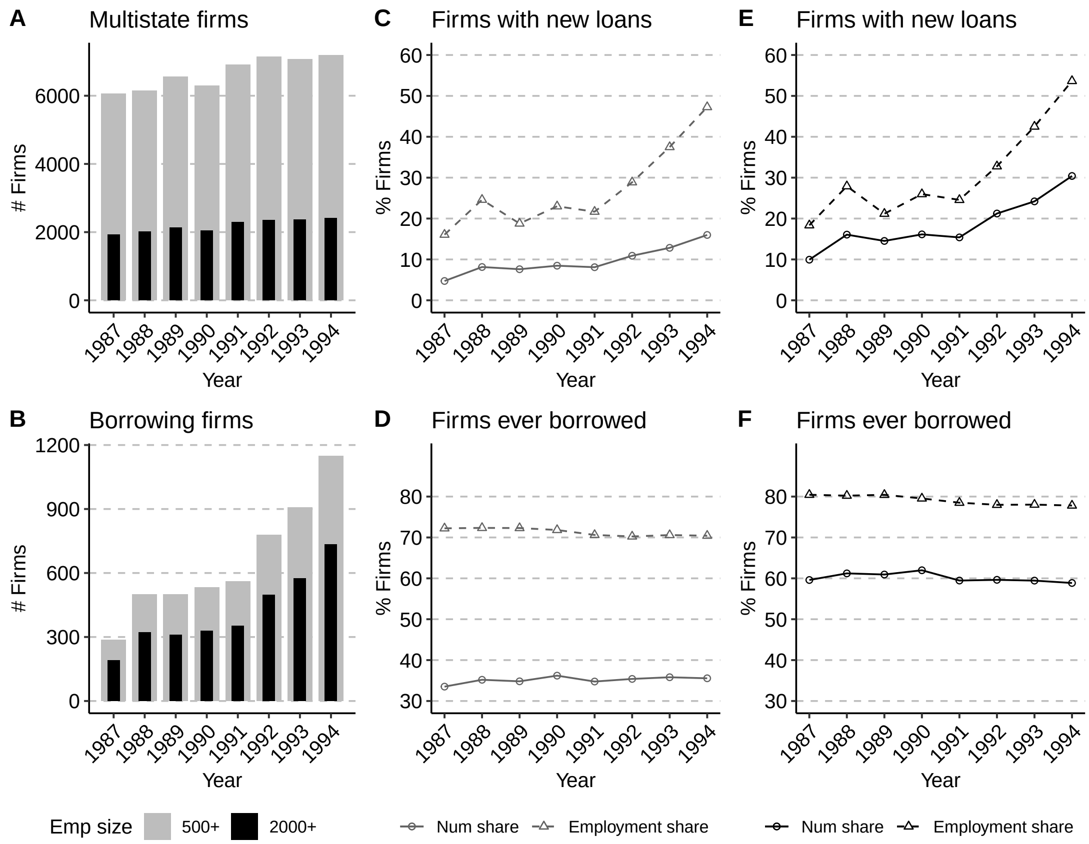

Figure 8: Bank Dependence of Multistate Firms

Notes: This figure examines the coverage of multistate firms by DealScan borrowers. Panel A plots the number of total number of multistate firms categorised by employment size. Panel B plots the number of firms in DealScan. Panels C and D show the fraction among all multisate firms with more than 500 employees that borrowed in any given year, or ever borrowed during the sample period. Panels E and F show the fraction among all multisate firms with more than 2000 employees that borrowed in any given year, or ever borrowed during the sample period.

where the dependent variable $Loan_{fbt}$ is an indicator for whether firm f borrows from bank b in year t, and the primary explanatory variable is the indicator $Neighbour_{fbt}$ constructed in section 5.2. Additionally, firm-year and bank-year fixed effects, ϕ_{ft} and ϕ_{bt}, are included to account for firm or bank-specific factors that affect loan demand or supply. The sample include all multistate firms in DealScan as shown in Panel B of Figure Figure 7, each associated with all active bank as shown in Panel C of 7.

Table 4 reports the results from the above regression. The coefficients on $Neighbour_{fbt}$ are both positive and statistically significant, which suggests that firms are more likely to borrow from neighbouring banks rather than those lo-

Table 4: Lending Relationships between Neighbours

	(1)	(2)
	Loan	Loan
Neighbour	8.14***	4.02***
	(1.61)	(0.417)
Firm-Year FE		Yes
Bank-Year FE		Yes
Mean of Dep Var	7.97	7.97
R^2 Adj	0.022	0.227
Observations	268,017	268,017

Notes: This table reports regression results from Equation 5. The dependent variable *Loan* is an indicator for whether a firm borrows from a bank, multiplied by 100. The explanatory variable *Neighbour* is an indicator for whether the firm and bank operate in the same county. Standard errors are double clustered at firm and bank levels. Significance levels: ***1%, **5%, *10%.

cated farther away. Specifically, the estimated value in column (2) is 4.02, suggesting that being neighbours increases the likelihood of lending by 4.02 percentage points. For comparison, the unconditional mean any bank being a lender is 7.79 percentage points.

In summary, these findings confirm that banks have better knowledge about neighbouring firms than those farther away.

7 Lending in Deregulated Regions

I now turn to investing the lending outcomes within the deregulated regions. An important objective for banking deregulation is to encourage foreign banks to enter these markets and provide financial support for local economic activities. The physical proximity between bank branches to borrowers can reduce monitoring costs and facilitate credit provision. However, as shown previously that informational frictions influences entry decisions, it is reasonable to expect that lending patterns within the deregulated markets would mirror entry patterns. In this section, I demonstrate that banks increase their lending in the deregulated regions after deregulation. However, the credit growth is more pronounced in areas where banks have a greater likelihood of entering, thanks to the informational advantage derived from the presence of old neighbouring firms.

In the following subsections, I will investigate banks' lending activities to

Table 5: Allocation of Loans to Bank-County Level

Bank	County	# Loans	$M Loans
B_1	C_1	0.6	$5.4(= 10 \times .6 \times .9)$
B_1	C_2	0.6	$0.6(= 10 \times .6 \times .1)$
B_2	C_1	0.4	$0.4(= 10 \times .4 \times .1)$
B_2	C_2	0.4	$0.4(= 10 \times .4 \times .1)$

Notes: This table demonstrates how each loan in DealScan is allocated to the bank-county level. The calculation is based on the following numerical example: a firm borrows $10M from banks B_1 and B_2 with 60% and 40% lender shares respectively, and the firm operates establishments in counties C_1 and C_2 with 90% and 10% of its total employment respectively.

both businesses and households. For business lending, I will utilise an extended sample of corporate loans from DealScan spanning the period from 1987 to 2005. For household credits, I will employ mortgage data recorded by Home Mortgage Disclosure Act (HMDA) for years from 1990 to 2005.

7.1 Business lending

7.1.1 Measure of lending volumes

I begin by examining business lending outcomes using data on corporate loans from DealScan. I focus on a sample of loans by US non-financial borrowers during 1987–2005. To obtain relevant information on firm locations, I again match borrowers to firms in Dun & Bradstreet.

To evaluate the geographic distribution of credit provision, it is necessary to measure lending volumes (the dollar amount or the number of loans) at the bank-county level. However, two complexities arise when working with syndicated loan data: (1) a single loan deal often involves multiple lenders, and (2) the borrowing firm may have operations in multiple counties. To address these issues, I take the following approach.

First, for calculating the dollar amount of loans at the bank-county level, I first allocate the loan deal amount across lenders based on (imputed) lender shares, and then further distribute each lender's share across counties according to the employment shares of the borrower in each county.[7] For a practical

[7] Following Chodorow-Reich (2014), missing values of lender shares are imputed based on the average of other loans with similar syndicate structure (same number of lead and participant lenders) but with non-missing values of lender shares.

illustration, consider the following example: Suppose a firm borrows \$10M from banks B_1 and B_2 with lender shares of 60% and 40%, respectively. The firm operates establishments in counties C_1 and C_2 where 90% and 10% of its employment is located, respectively. In this scenario, the dollar amount of lending by bank B_1 to county C_1 is $5.4(= 10 \times .6 \times .9)$, and similar calculations can be made for other allocations, as exemplified in Table 5.

Second, for calculating the number of loans at the bank-county level, I simply assign the lender's share of the loan to each county that the firm has establishments. In the previous example (Table 5), the number of loans made by bank B_1 to county C_1 is equal to 0.6, representing the lender share of bank B_1 in this loan contract. Compared to the dollar amount measure of loan volumes, this measure of loan volume is immune to the measurement errors of employment at the firm's establishments, but is proportionally scaled with the firm's geographic operations in the aggregate. Importantly, the two measures yield similar qualitative conclusions.

It is important to note that the proposed measures of lending volumes above do not capture where the funds are spent. While investigating fund allocation could be an interesting avenue for investigation, unfortunately, I do not obverse how firms internally allocate these funds. Nonetheless, the proposed measures provide insights into the locations of those firms and their establishments being financed by out-of-state banking organisations, specifically whether they are in areas where banks are likely to enter due to lower informational barriers. This aspect aligns with the primary focus of this part of the analysis, which is to examine credit provision by entrant banks to firms in the deregulated regions.

7.1.2 Lending outcomes

I start by examining the average lending growth in the deregulated regions. I estimate the following specification:

$$Loan_{bct} = \phi_{bt} + \phi_{ct} + \sum_{k \neq -1} \beta_k \times \mathbb{1}\{YsD_{bct} = k\} + \varepsilon_{bct}, \tag{6}$$

where the dependent variable $Loan_{bct}$ is either the logarithm of one plus the dollar amount of loans or the logarithm of one plus the number of loans by bank b in county c in calendar year t; the main explanatory variable of interest is

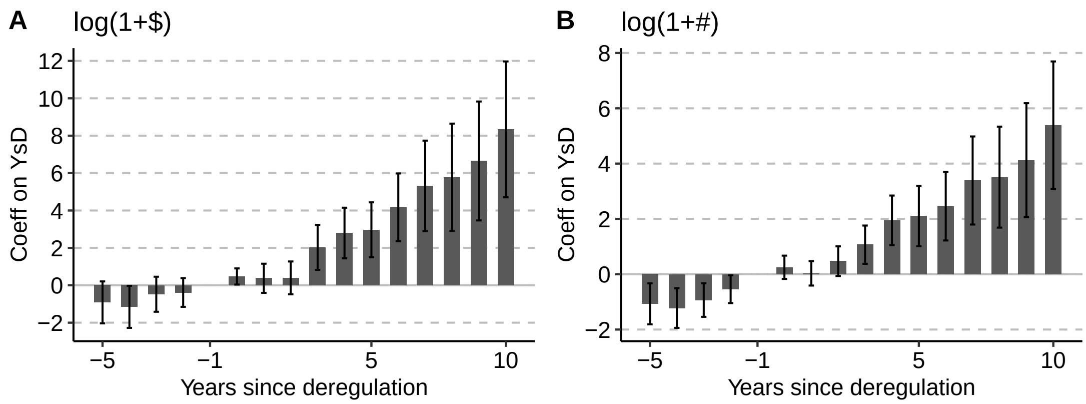

Figure 9: Corporate Lending Growth in Deregulated Counties

Notes: This figure plots the coefficient estimates from Equation 6. The dependent varible in Panel A is the logarithm of one plus the dollar amount of loans. The dependent varible in Panel B is the logarithm of one plus the number of loans. Standard errors are clustered at the bank and county levels. The error bars represent 95% confidence intervals.

the dummy variables $\mathbb{1}\{YsD_{bct} = k\}$ where YsD stands for number of years since deregulation, and thus the coefficients β_k's measure the change in lending volumes relative to the year prior to the deregulation; finally, ϕ_{bt} and ϕ_{ct} are bank-year and county-year fixed effects that absorb credit supply and demand shocks at the bank and county levels. The sample include all bank-county pairs that would be affected by the interstate banking deregulation during 1982–1995. And I trace the lending outcomes from 5 years prior to the deregulation to 10 years after, whenever possible.

Figure 9 plots the coefficient estimates β_k's. The figure shows that there is only mild growth before deregulation, but substantial growth after deregulation. Table 6 performs a formal test on the change in credit growth rate by fitting different linear time trends before and after the deregulation:

$$Loan_{bct} = \phi_{bt} + \phi_{ct} + \beta_1 YsD_{bct} + \beta_2 YsD_{bct} \times Post_{bct} + \varepsilon_{bct}, \tag{7}$$

where $Post_{bct}$ is an indicator for years after deregulation, i.e., when $YsD_{bct} \geq 0$. Thus, the coefficient β_1 estimates the average annual growth rate during the five years before deregulation, and β_2 estimates the change in growth rate after deregulation. The results in Table 6 show that the coefficient on the interaction term $YsD_{bct} \times Post_{bct}$ is positive and statistically significant, which suggests that credit growth accelerates since deregulation, consistent with the idea that entry

Table 6: Deregulation and Corporate Lending Growth

	(1) Log(1+$)	(2) Log(1+#)
YsD	0.11	0.20**
	(0.12)	(0.087)
YsD×Post	0.66***	0.29**
	(0.20)	(0.120)
Bank-Year FE	Yes	Yes
County-Year FE	Yes	Yes
R^2 Adj	0.45	0.55
Observations	2,526,679	2,526,679

Notes: This table reports the regression results from Equation 7. The sample includes obvervations at bank-county-year level from 5 years before deregulation to 10 years after. The dependent variables are lending volumes measured by either the logarithm of one plus the dollar amount of loans or logarithm of one plus the number of loans. Variable *YsD* is the number of years since deregulation. Variable *Post* is an indicator for years after deregulation, i.e., when $YsD \geq 0$. Standard errors are clustered at bank and county levels. Significance levels: ***1%, **5%, *10%.

in physical forms is essential for credit provision.

Next, I ask whether lending growth differs across areas with different informational barriers of entry. To answer this question, I include the interaction between the year-since-deregulation dummies with the measure of information barrier *Old Neighbours* in Equation 6 as follows:

$$\begin{aligned} Loan_{bct} = \sum_k \beta_k \left[Old\ Neighbours\right]_{bc} \times \mathbb{I}\{YsD_{bct} = k\} \\ + \gamma' X_{bct} + \phi_{YsD(bct)} + \phi_{bt} + \phi_{ct} + \varepsilon_{bct}, \end{aligned} \tag{8}$$

where the dependant variable $Loan_{bct}$ is loan volume by bank b to county c in calendar year t; X_{bct} include interactions of the *YsD* dummies with the geographic distance between the bank and the county to control for the time varying effects of distance on lending outcomes; $\phi_{YsD(bct)}$ are dummy variables representing the number of years since deregulation; ϕ_{bt} and ϕ_{ct} represent bank-year and county-year fixed effects. The coefficients of interests are β_k's. They capture the evolution of the difference in lending volumes across counties with different informational barriers of entry before and after deregulation.

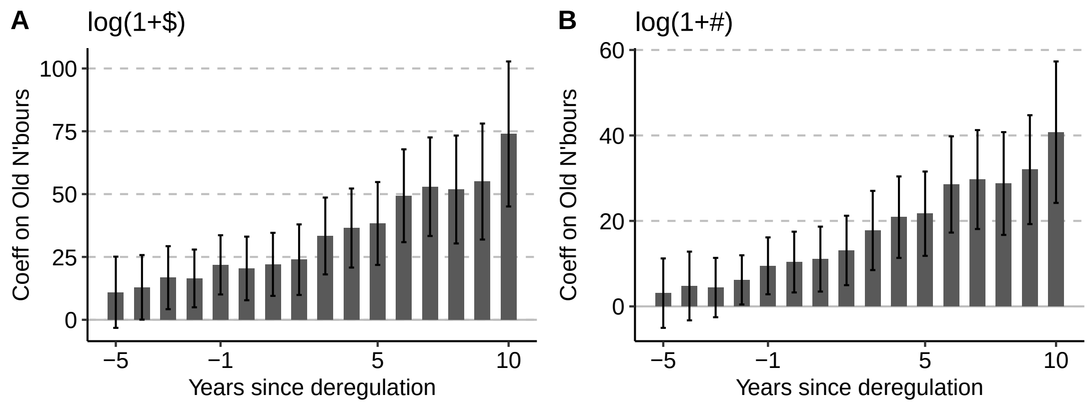

Figure 10: Old Neighbours and Corporate Lending Growth

Notes: This figure plots the coefficient estimates from Equation 8. The dependent variable in Panel A is the logarithm of one plus the dollar amount of loans. The dependent variable in Panel B is the logarithm of one plus the number of loans. Standard errors are clustered at the bank and county levels. The error bars represent 95% confidence intervals.

Figure 10 plots the coefficient estimates β_k. The coefficients are positive prior to the deregulation, mainly because banks are lending to neighbouring firms that also operate in the newly deregulated regions, and thus counties with higher concentration of old neighbouring firms should have higher lending volumes even prior to the deregulation. What's more important, the coefficient is higher and increasing faster ever since the deregulation, which suggests that banks lending increases by more in counties with lower informational barriers of entry. I formally test these two statements using the following specifications:

$$\begin{aligned} Loan_{bct} = \beta_1 [Old\ Neighbours]_{bc} + \beta_2 [Old\ Neighbours]_{bc} \times Post_{bct} \\ + \gamma' X_{bct} + \phi_{YsD(bct)} + \phi_{bt} + \phi_{ct} + \varepsilon_{bct}, \end{aligned} \tag{9}$$

and

$$\begin{aligned} Loan_{bct} = & \beta_1 [Old\ Neighbours]_{bc} \times YsD_{bct} \\ & + \beta_2 [Old\ Neighbours]_{bc} \times YsD_{bct} \times Post_{bct} \\ & + \gamma' X_{bct} + \phi_{YsD(bct)} + \phi_{bt} + \phi_{ct} + \varepsilon_{bct}, \end{aligned} \tag{10}$$

Table 7 reports the results from the above two regressions. Columns (1) and (3) show that lending is higher in areas with more familiar firms, and the differences increases after the deregulation. Columns (2) and (4) show that credit growth

Table 7: Old Neighbours and Corporate Lending Growth

	(1) Log(1+$)	(2) Log(1+$)	(3) Log(1+#)	(4) Log(1+#)
Old N'bours	20.47***	18.30***	8.37**	9.04**
	(5.78)	(6.37)	(3.23)	(3.85)
Old N'bours×Post	19.30***		13.75***	
	(5.00)		(3.07)	
Old N'bours×YsD		1.08		1.20*
		(1.22)		(0.70)
Old N'bours×YsD×Post		3.62**		1.62
		(1.75)		(0.99)
Bank-Year FE	Yes	Yes	Yes	Yes
County-Year FE	Yes	Yes	Yes	Yes
YsD FE	Yes	Yes	Yes	Yes
Dist Cntrl	Yes	Yes	Yes	Yes
R^2 Adj	0.46	0.46	0.56	0.56
Observations	2,526,679	2,526,679	2,526,679	2,526,679

Notes: This table reports the regression results from Equations 9 and 10. The sample includes obvervations at bank-county-year level from 5 years before deregulation to 10 years after. The dependent variables are lending volumes measured by either the logarithm of one plus the dollar amount of loans or logarithm of one plus the number of loans. Variable *Old N'bours* is the employment share in the destination county of a bank's old neighbouring firms. Variable *YsD* is the number of years since deregulation. Variable *Post* is an indicator for years after deregulation, i.e., when $YsD \geq 0$. Standard errors are clustered at bank and county levels. Significance levels: ***1%, **5%, *10%.

rate after deregulation is higher in counties with lower informational barriers to entry, and is significant for the dollar amount of loans.

Finally, I ask whether lending growth are driven by firms in the old neighbourhoods, or those outside the old neighbourhoods. To answer this question, I estimate Equation 8 separately for lending to firms in and outside the old neighbourhood. Panels A and B in Figure 11 plots the difference in lending volumes to firms in the banks' old neighbourhood between counties with low and high informational barriers to entry. The coefficients on the variable *Old Neighbours* are naturally positive prior to the deregulation, because places where more old neighbouring firms are concentrated will be allocated more loans. More importantly, the coefficient keeps increasing after the deregulation, which suggests that the growth in lending to firms in the old neighbourhoods contribute to the

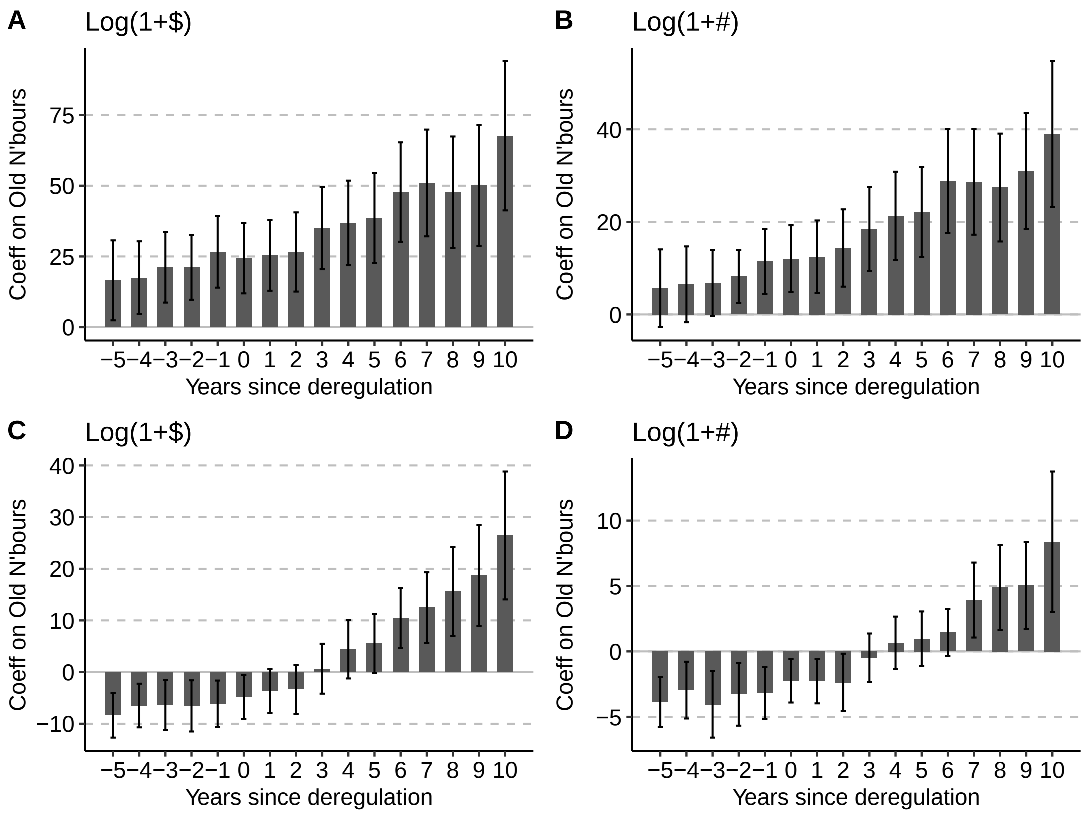

Figure 11: Lending to Firms In or Outside Old Neighbourhoods

Notes: This figure plots the coefficient estimates from Equation 8. Panels A and B plot the effects on lending to firms in the old neighbourhood. Panels C and D plot the effects on lending to firms outside the old neighbourhood. The dependent variables in Panels A and C are the logarithm of one plus the dollar amount of loans. The dependent variables in Panels B and D are the logarithm of one plus the number of loans. Standard errors are clustered at the bank and county levels. The error bars represent 95% confidence intervals.

increasing gap in lending volumes across deregulated counties with high and low informational entry barriers. Panel A of Table 8 performs a formal tests on the change in both the level and growth rate of lending volumes to firms in the old neighbourhoods before and after the deregulation, based on Equations 9 and 10. The positive and statistically significant coefficients on *Old Neighbours* in columns (1) and (3) suggest that a bank's lending to firms in the old neighbourhoods are higher in counties with higher concentration of old neighbours prior to the deregulation. The positive and statistically significant coefficients on the variable *Old Neighbours*×*Post* suggest that this lending gap becomes even larger after the deregulation. Columns (2) and (4) show that the credit growth rate is

Table 8: Lending to Firms In or Outside the Old Neighbourhoods

	(1)	(2)	(3)	(4)
Panel A: Firms in the old neighbourhoods				
	Log(1+$)	Log(1+$)	Log(1+#)	Log(1+#)
Old N'bours	24.38^{***}	22.83^{***}	10.16^{***}	10.88^{***}
	(5.90)	(6.34)	(3.32)	(3.93)
Old N'bours×Post	15.12^{***}		12.16^{***}	
	(4.49)		(2.92)	
Old N'bours×YsD		0.98		1.11^{*}
		(1.11)		(0.66)
Old N'bours×YsD×Post		2.69^{*}		1.37
		(1.52)		(0.90)
Panel B: Firms outside the old neighbourhoods				
	(1)	(2)	(3)	(4)
	Log(1+$)	Log(1+$)	Log(1+#)	Log(1+#)
Old N'bours	-4.17^{**}	-7.22^{***}	-2.57^{***}	-3.47^{***}
	(1.94)	(2.43)	(0.91)	(1.07)
Old N'bours×Post	10.64^{***}		3.87^{***}	
	(2.51)		(1.09)	
Old N'bours×YsD		−0.07		0.02
		(0.45)		(0.16)
Old N'bours×YsD×Post		3.02^{***}		1.00^{***}
		(0.87)		(0.34)
Bank-Year FE	Yes	Yes	Yes	Yes
County-Year FE	Yes	Yes	Yes	Yes
YsD FE	Yes	Yes	Yes	Yes
Dist Cntrl	Yes	Yes	Yes	Yes
Observations	2,526,679	2,526,679	2,526,679	2,526,679

Notes: This table reports the regression results from Equations 9 and 10 separately for lendings to firms in and outside the banks' old neighbourhoods. The sample includes obvervations at bank-county-year level from 5 years before deregulation to 10 years after. The dependent variables are lending volumes measured by either the logarithm of one plus the dollar amount of loans or logarithm of one plus the number of loans. Variable *Old N'bours* is the employment share in the destination county of a bank's old neighbouring firms. Variable *YsD* is the number of years since deregulation. Variable *Post* is an indicator for years after deregulation, i.e., when $YsD \geq 0$. Standard errors are clustered at bank and county levels. Significance levels: ***1%, **5%, *10%.

(weakly) higher after the deregulation than before.

Panels C and D in Figure 11 shows the difference in lending volumes to firms outside a bank's old neighbourhoods between counties with different informational barriers of entry. Prior to the deregulation, the lending gap is negative, meaning that firms being financed but outside the banks' old neighbourhoods are mostly in counties with lower concentration of old neighbouring firms, and the gap remains stable during the five years prior to the deregulation. However, this gap starts to shrink after deregulation and eventually becomes positive, meaning that firms being financed but outside the bank's old neighbourhoods are now located mostly in counties where informational entry barriers are low. This is consistent with the story that the presence of old neighbours lowers informational barriers of entry, and once legal restrictions are removed, banks are more likely to enter locations where more firms are familiar, following which new relations can then be established with local businesses. Panel B of Table 8 performs the tests based on Equations 9 and 10 where the outcomes are now lending volumes to firms outside the banks' old neighbourhoods. The negative coefficients on *Old Neighbours* in columns (1) and (3) indicates that lendings to firms outside the old neighbourhoods is lower in counties with higher concentration of old neighbours prior to the deregulation. And the positive and statistically significant coefficients on the interaction term *Old Neighbours* × *Post* suggest that the lending gap is reversed after the deregulation. Columns (2) and (4) show that the coefficients on *Old Neighbours* × *YsD* is close to zero and statistically insignificant, indicating that there is no significant change in the lending growth between counties with high or low informational barriers of entry before deregulation. The coefficient on *Old Neighbours* × *YsD* × *Post* is positive and statistically significant, which suggests that, after the deregulation, lending grows faster in counties with lower informational barriers of entry.

7.2 Household lending

I now turn to mortgage lending outcomes. I obtain mortgage data from Home Mortgage Disclosure Act (HMDA) 1990–2005. Using the concordance table prepared by Robert B. Avery, I link HMDA reporters to BHCs. Unlike syndicated loan data from DealScan, it is straightforward to measure mortgage lending volumes at the bank-county level, because each mortgage has a unique originator and the borrower has a unique location. As in the case of business lendings, I also construct two measures of lending volumes: the dollar amount and the number of mortgage originations.

Panels A and B of Figure 12 show the average lending volumes in deregu-

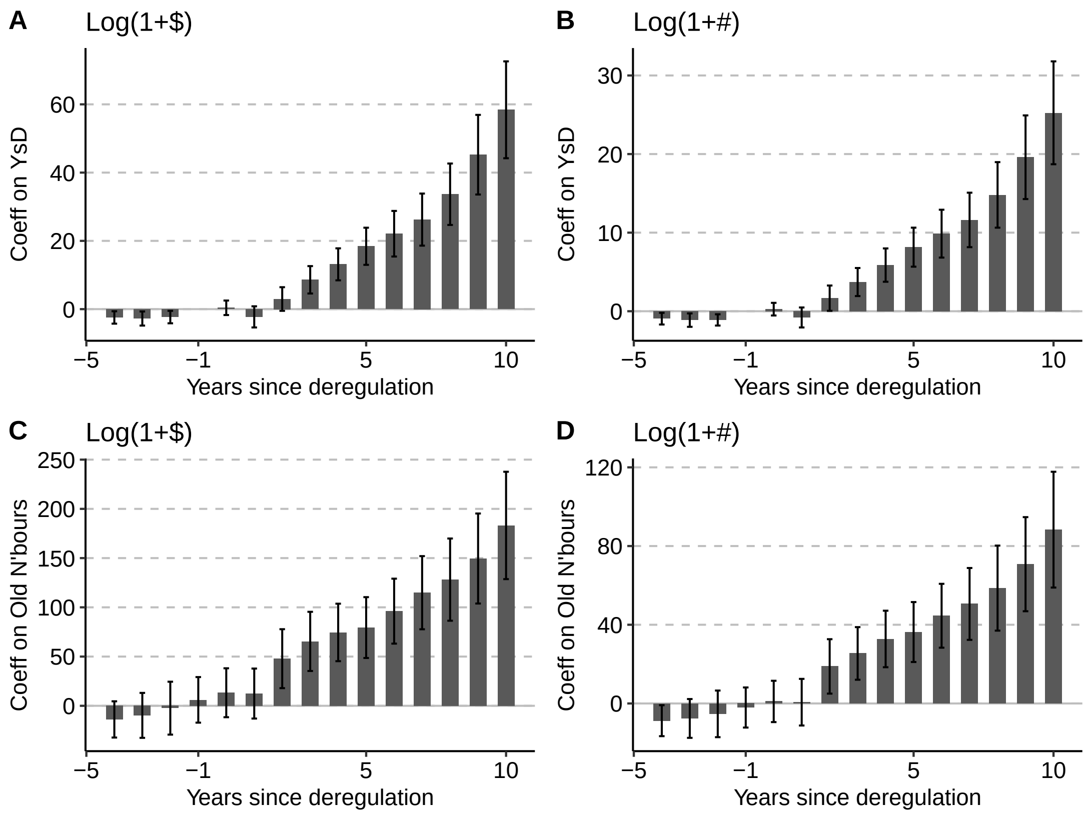

Figure 12: Old Neighbours and Home Mortgage Lending Growth

Notes: This figure plots the coefficient estimates from Equation 8. The dependent variable in Panel A is the logarithm of one plus the dollar amount of mortgages. The dependent variable in Panel B is the logarithm of one plus the number of mortgages. Standard errors are clustered at the bank and county levels. The error bars represent 95% confidence intervals.

lated counties relative to the year prior to the deregulation. These estimates are obtained from regressing lending volumes on the dummy variables representing the number of years since deregulation, as in Equation 6. The sample include observations of bank-county mortgage outcomes from 4 years before deregulation to 10 years after. The figures suggest that mortgage lending growth is substantially higher after the deregulation.

Panels C and D plot the coefficients from Equation 8. They suggest that the lending growth is faster in counties with higher concentration of familiar firms, i.e., lower informational barriers to entry.

Table 9 presents results from Equations 9 and 10. Columns (1) and (4) again show that credit growth is faster following the deregulation. Columns (2) and (5) show that lending concentrates in counties with greater presence of old neigh-

Table 9: Old Neighbours and Mortgage Lending Growth

	(1) Log(1+$)	(2) Log(1+$)	(3) Log(1+$)	(4) Log(1+#)	(5) Log(1+#)	(6) Log(1+#)
YsD	1.31***			0.48***		
	(0.28)			(0.11)		
YsD×Post	4.11***			1.85***		
	(0.64)			(0.29)		
Old N'bours		4.62	10.39		−0.65	−0.71
		(10.07)	(14.19)		(4.46)	(6.43)
Old N'bours×Post		81.90***			39.10***	
		(11.66)			(5.88)	
Old N'bours×YsD			6.13**			2.01
			(2.94)			(1.35)
Old N'bours×YsD×Post			9.43*			5.96**
			(4.87)			(2.40)
Bank-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
County-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
YsD FE		Yes	Yes		Yes	Yes
Dist Cntrl		Yes	Yes		Yes	Yes
R^2 Adj	0.46	0.47	0.47	0.41	0.41	0.42
Observations	5,274,575	5,274,575	5,274,575	5,274,575	5,274,575	5,274,575

Notes: This table reports regression results on mortgage lending outcomes. Standard errors are clustered at bank and county levels. Significance levels: ***1%, **5%, *10%.

bouring firms, where banks are more likely to enter. Columns (3) and (6) show that the credit growth is faster in counties with lower informational barriers to entry following the deregulation.

8 Employment Effects

In this section, I examine how informational entry barriers shape the real economic outcomes of banking deregulation. I aggregate the measure of information advantage across banks that a county is exposed to, and compare employment outcomes for counties within each deregulated state using a stacked difference-in-differences design. I first show that counties experience more entries after the deregulation that have more firms with which the out-of-state banks are familiar. I then show that employment grows faster in these counties in the long run. However, in the short term, employment is lower in those counties, mainly because small business growth slows down. The data for this part of the analysis comes from Census Business Dynamics Statistics (BDS, 1978–2008). It contains information on county employment categorised by firm size.

Since the states may experience multiple times of deregulation, I designate the first deregulation year y_0 as the treatment year for each state. Then I calculate each county's exposure to banks over subsequent deregulations as a weighted average of *Old Neighbours* across all banks. The weights depend on the time since the state's first deregulation when the county was exposed to the bank. Specifically, let y be any deregulation year of county c (note that $\min(y) = y_0$), then the weights assigned to the candidate banks in that year is:

$$w(y) = \frac{1}{y - y_0 + 1}, \tag{11}$$

and the aggregate bank exposure for county c is defined as

$$AggOldNbour_c = \sum_b w(y_b) \times [Old\ Neighbours]_{bc}, \tag{12}$$

where y_b denotes the year in which the state of county c deregulates to the state of bank b. For example, NY first deregulated to ME and then to AZ and so on. Then banks in ME will all receive weights 1 and banks in AZ will all receive weights $\frac{1}{3}$. I similarly aggregate distance to all banks for each county as controls.

Table 10 reports the effects of aggregate share of old neighbours on the total number of entries within 10 years since first deregulation using the following cross-sectional regression:

$$\#\ Entry_c = \phi_s + \beta AggOldNbour_c + \gamma X_c + \varepsilon_c, \tag{13}$$

where the dependent variable $\#\ Entry_c$ is the total number of entries within ten years since the first deregulation; ϕ_s represent state fixed effects; X_c include the aggregate distance and pre-deregulation employment levels as controls. The coefficients on *AggOldNbours* are both positive and statistically significant, suggesting that lower aggregate informational barriers predict more entries. The estimated value in column (3) is 4.48, which suggests that a 10% increase in aggregate share of old neighbours increases the number of entries by around 0.45.

I then trace the employment outcomes from 5 years before the first deregulation to 15 years after for each county, and stack all observations using this

Table 10: Aggregate Old Neighbours and Total Entries

	(1) # Entry	(2) # Entry	(3) # Entry
AggOldNbours	5.71^{***}	5.51^{***}	4.48^{***}
	(0.602)	(0.588)	(0.551)
AggLogDist		-1.08^{***}	-0.994^{***}
		(0.254)	(0.239)
Emp_{-1}			5.07^{***}
			(1.68)
State FE	Yes	Yes	Yes
Mean of Dep Var	1.23	1.23	1.23
R^2 Adj	0.37	0.37	0.46
Observations	3,030	3,030	3,030

Notes: This table reports the regression results from Equation 13. The dependent variable *# Entry* is the total number of entries within ten years since first deregulation. Variable *AggOldNbours* is the aggregate measure of *Old Neighbours*. Variable *AggLogDist* is the aggregate measure of distance between the county and all potential entrant banks. Variable Emp_{-1} is the the county employment level one year prior to first deregulation. Standard errors are clustered at the state level. Significance levels: ***1%, **5%, *10%.

common time window. I estimate the following stacked dynamic diff-in-diff specification:

$$Emp_{ct} = \alpha_{st} + \alpha_c + \sum_k \beta_k \left[AggOldNbours\right]_c \times \mathbb{1}\{t = k\} + \gamma X_{ct} + \varepsilon_{ct}, \tag{14}$$

where the time index t represents the number of years since first deregulation; the dependent variable Emp_{ct} is the employment level in county c, t years since first deregulation; α_{st} represent state-time fixed effects; α_c are county fixed effects; X_{ct} include interactions of time dummies with aggregate distance to banks, and those with pre-deregulation employment level. Figure 13 plots the coefficients estimates β_k. Panel A plots the effects on total employment. Note that there is no significant differences in the employment growth path between low exposure counties and high exposure counties before deregulation. After the deregulation, employment is higher in counties that have lower informational barriers of entry. A 10% increase in the aggregate share of old neighbouring firms implies around 3000 more employees for the county, 15 years after first deregulation. However, in the short and medium run, employment is lower in

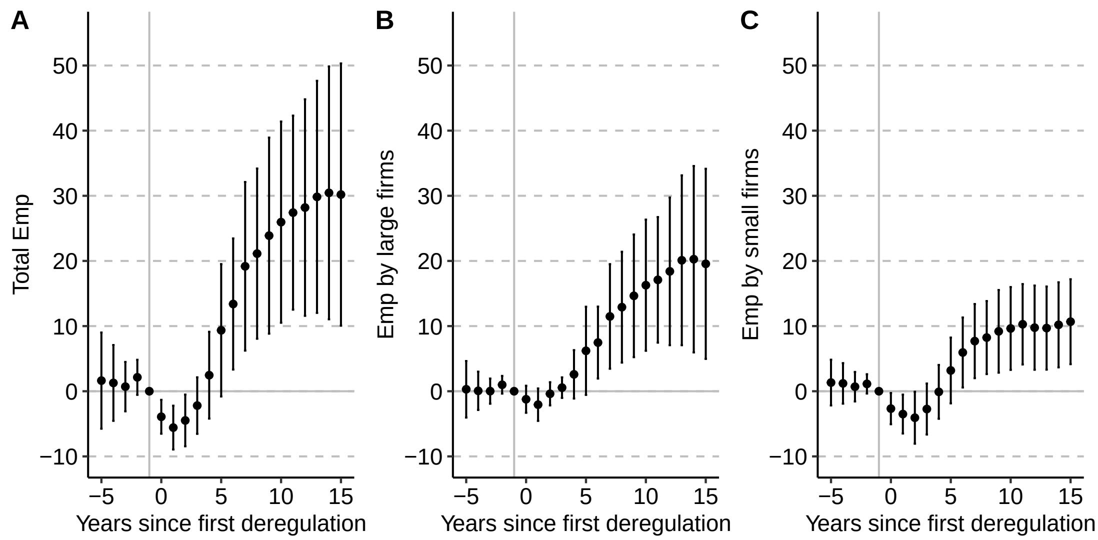

Figure 13: Employment Effects

Notes: Panels A–C plots the effects on levels of employment in thousands. All specifications control for the time varying effects of aggregate distance to banks and pre-deregulation employment level.

counties where informational barriers are low. I separately investigate the employment outcomes by large firms with at least 500 employees, and those by small firms with less than 500 employees. Panels B and C plot the effects on employment by large and small firms respectively. Note that the negative effects are mainly driven by small businesses.

Table 13 presents results from the following stacked diff-in-diff regression, where I divide post-deregulation period into three episode:

$$Emp_{ct} = \alpha_{st} + \alpha_c + \sum_{k=1}^{3} \beta_k \left[AggOldNbours\right]_c \times Post_{kct} + \gamma X_{ct} + \varepsilon_{ct}. \tag{15}$$

In this specification, the variables $Post_{kct}$ are indicators for three different periods after deregulation. Specifically, $Post_{1,ct} = 1$ for $t \in [0,2]$, $Post_{2,ct} = 1$ for $t \in [3,5]$, and $Post_{3,ct} = 1$ for $t \in [5,10]$. Therefore, β_1 estimate the effects of *AggOldNbours* on employment in the first , β_2 the medium term, and β_3 the long term. Control variables X_{ct} include interactions of $Post_{kct}$ with aggregate distance and pre-deregulation employment level. α_{st} and α_c are state-time and county fixed effects, respectively. Column (1) presents the estimates of the effects on total employment, column (2) the effects on employment by large firms,

Table 11: Employment Effects

	(1) Emp	(2) EmpLarge	(3) EmpSmall
AggOldNbours×$Post_1$	-5.81^{*}	-1.5	-4.28^{*}
	(2.98)	(1.1)	(2.33)
AggOldNbours×$Post_2$	2.05	2.84	-0.752
	(3.79)	(1.75)	(2.63)
AggOldNbours×$Post_3$	19.6^{***}	12.3^{***}	7.27^{**}
	(6.76)	(3.91)	(3.35)
County FE	Yes	Yes	Yes
State-Time FE	Yes	Yes	Yes
Mean of Dep Var	28.5	13.3	15.3
R^2 Adj	0.99	0.99	1.00
Observations	48,464	48,464	48,464

Notes: This table reports regression results on mortgage lending outcomes. Standard errors are clustered at bank and county levels. Significance levels: ***1%, **5%, *10%.

and column (3) the effects on employment by small firms. The coefficients on $AggOldNbours \times Post_1$ is negative and statistically significant, which suggests that the lower informational barriers to entry reduces employment growth in the short run, and these effects are primarily from small businesses. In the medium term, there is no significant difference in employment between counties with high and low share of old neighbouring firms. In the long run, there is strong and positive effects on employment—counties with lower informational barriers to entry exhibit higher growth—and the results are mainly driven by large firms.

9 Conclusions

In this paper, I document that information asymmetry limits banks' geographic expansions. The presence of familiar firms from the bank's original neighbourhoods serves to alleviate issues related to adverse selection in new markets, thereby improving the likelihood of bank entry. I also document that banks' lending patterns mirror entry outcomes, as geographic proximity facilitates credit provision. In terms of real economic impacts, I find that the information frictions can generate differential growth effects of the statewide deregulation across regions and firms within states. Areas with lower informational barriers to entry experience stronger long-term employment growth, primarily driven by larger firms. However, in the initial years following deregulation, employment growth temporarily decelerates, mainly because of hampered growth of small businesses.

The findings presented in this paper carry significant policy implications. First, given that information frictions can discourage entry of banks, despite the removal of legal restrictions, a more effective financial liberalisation could involve simultaneous reforms in the real sector. Strengthening accounting regulations to enhance informational transparency of domestic firms or encouraging entry of multinational firms may be instrumental in this regard. This approach may be particularly relevant for emerging markets and underdeveloped economies, where merely opening up the financial sector may not be sufficient to foster economic growth. Second, it is important to recognise the impact of banking deregulation on small businesses, particularly in the initial phase. Small enterprises may not experience the same level of benefits as larger firms. And they could be a major source of information asymmetry that prevents entry in the first place. However, a key objective of banking deregulation is to enhance credit availability, especially for local businesses. The results in this paper suggest that the removal of legal entry barriers may not be sufficient, or could even be harmful. Therefore, it is advisable to implement additional measures such as financial incentives for entrant banks to serve local businesses. Third, in spite of many previous studies documenting the benefits of diversification from geographic expansion, the evidence presented in this paper indicate that such benefits may be limited as banks tend to stay close to their original neighbouring firms, even when they are allowed to expand. This may lead to concerns regarding customer concentration, posing potential threats to financial stability.

References

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

Amel, D. F. (1993). State laws affecting the geographic expansion of commercial banks. *Federal Reserve Board*.

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

Black, S. E. and Strahan, P. E. (2002). Entrepreneurship and bank credit availability. *The Journal of Finance*, 57(6):2807–2833.

Bofondi, M. and Gobbi, G. (2006). Informational barriers to entry into credit markets. *Review of Finance*, 10(1):39–67.

Boot, A. W. and Thakor, A. V. (2000). Can relationship banking survive competition? *The Journal of Finance*, 55(2):679–713.

Broecker, T. (1990). Credit-worthiness tests and interbank competition. *Econometrica: Journal of the Econometric Society*, pages 429–452.

Cannon, B. and Lynch, J. (2023). Does Banking Competition Really Increase Credit for All? The Effect of Bank Branching Deregulation on Small Business Credit.

Chava, S., Oettl, A., Subramanian, A., and Subramanian, K. V. (2013). Banking deregulation and innovation. *Journal of Financial Economics*, 109(3):759–774.

Chava, S. and Purnanandam, A. (2011). The effect of banking crisis on bank-dependent borrowers. *Journal of Financial Economics*, 99(1):116–135.

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

Chu, Y., Deng, S., and Xia, C. (2020). Bank geographic diversification and systemic risk. *The Review of Financial Studies*, 33(10):4811–4838.

Cornaggia, J., Mao, Y., Tian, X., and Wolfe, B. (2015). Does banking competition affect innovation? *Journal of Financial Economics*, 115(1):189–209.

Dell'Ariccia, G., Friedman, E., and Marquez, R. (1999). Adverse selection as a barrier to entry in the banking industry. *The RAND Journal of Economics*, pages 515–534.

Dennis, S. A. and Mullineaux, D. J. (2000). Syndicated loans. *Journal of Financial Intermediation*, 9(4):404–426.

Diamond, D. W. (1991). Monitoring and reputation: The choice between bank loans and directly placed debt. *Journal of Political Economy*, 99(4):689–721.

Federico, S., Hassan, F., and Rappoport, V. (2023). Trade shocks and credit reallocation. Technical report, National Bureau of Economic Research.

Goetz, M. R., Laeven, L., and Levine, R. (2013). Identifying the valuation effects and agency costs of corporate diversification: Evidence from the geographic diversification of US banks. *The Review of Financial Studies*, 26(7):1787–1823.

Goetz, M. R., Laeven, L., and Levine, R. (2016). Does the geographic expansion of banks reduce risk? *Journal of Financial Economics*, 120(2):346–362.

Greenstone, M., Mas, A., and Nguyen, H.-L. (2020). Do credit market shocks affect the real economy? Quasi-experimental evidence from the great recession and "normal" economic times. *American Economic Journal: Economic Policy*, 12(1):200–225.

Gropp, R., Noth, F., and Schüwer, U. (2019). What drives banks' geographic expansion? The role of locally non-diversifiable risk. *SAFE Working Paper No.246*.

Holmström, B. (1979). Moral hazard and observability. *The Bell Journal of Economics*, pages 74–91.

Holmstrom, B. and Tirole, J. (1997). Financial intermediation, loanable funds, and the real sector. *the Quarterly Journal of Economics*, 112(3):663–691.

Huber, K. (2018). Disentangling the effects of a banking crisis: Evidence from german firms and counties. *American Economic Review*, 108(3):868–898.

Ivashina, V. (2009). Asymmetric information effects on loan spreads. *Journal of Financial Economics*, 92(2):300–319.

Jayaratne, J. and Strahan, P. E. (1996). The finance-growth nexus: Evidence from bank branch deregulation. *The Quarterly Journal of Economics*, 111(3):639–670.

Jordà, Ò., Schularick, M., and Taylor, A. M. (2013). When credit bites back. *Journal of Money, Credit and Banking*, 45(s2):3–28.

King, R. G. and Levine, R. (1993). Finance and growth: Schumpeter might be right. *The Quarterly Journal of Economics*, 108(3):717–737.

Leland, H. E. and Pyle, D. H. (1977). Informational asymmetries, financial structure, and financial intermediation. *The Journal of Finance*, 32(2):371–387.

Levine, R. and Zervos, S. (1998). Stock markets, banks, and economic growth. *American Economic Review*, pages 537–558.

Marquez, R. (2002). Competition, adverse selection, and information dispersion in the banking industry. *The Review of Financial Studies*, 15(3):901–926.

Mian, A. (2006). Distance constraints: The limits of foreign lending in poor economies. *The Journal of Finance*, 61(3):1465–1505.

Mian, A., Sufi, A., and Verner, E. (2017). Household debt and business cycles worldwide. *The Quarterly Journal of Economics*, 132(4):1755–1817.

Mian, A., Sufi, A., and Verner, E. (2020). How does credit supply expansion affect the real economy? The productive capacity and household demand channels. *The Journal of Finance*, 75(2):949–994.

Nguyen, H.-L. Q. (2019). Are credit markets still local? Evidence from bank branch closings. *American Economic Journal: Applied Economics*, 11(1):1–32.

Peek, J. and Rosengren, E. S. (1997). The international transmission of financial shocks: The case of japan. *The American Economic Review*, pages 495–505.

Peek, J. and Rosengren, E. S. (2000). Collateral damage: Effects of the japanese bank crisis on real activity in the united states. *American Economic Review*, 91(1):30–45.

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

Petersen, M. A. and Rajan, R. G. (2002). Does distance still matter? The information revolution in small business lending. *The Journal of Finance*, 57(6):2533–2570.

Rajan, R. and Zingales, L. (1998). Financial development and growth. *American Economic Review*, 88(3):559–586.

Rice, T. and Strahan, P. E. (2010). Does credit competition affect small-firm finance? *The Journal of Finance*, 65(3):861–889.

Schnabl, P. (2012). The international transmission of bank liquidity shocks: Evidence from an emerging market. *The Journal of Finance*, 67(3):897–932.

Schularick, M. and Taylor, A. M. (2012). Credit booms gone bust: monetary policy, leverage cycles, and financial crises, 1870–2008. *American Economic Review*, 102(2):1029–1061.

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

Appendix

A Chronology of Interstate Banking Deregulation

In this section, I present the chronology of interstate banking deregulation constructed based on Amel (1993)[8]. In Table A.1 below, each record contains information on the year of deregulation, the home state which opens the banking sector, and the list of foreign states whose BHCs are allowed to expand. The year of deregulation is the year in which the entry of out-of-state BHCs became effective, taking into account the reciprocity requirements set forth by the relevant legislations. For example, the state of Maine deregulated in 1978 with reciprocity requirement, while New York, the second deregulated state, only passed a similar law in 1982, which made Maine's deregulation effective in 1982, as shown by the first record in the table below. I include only 47 contiguous states (excluding Alaska, Hawaii, Delaware and South Dakota). A value "All" for FOREIGN column indicates that the home state is open to all other states.

Table A.1: Chronology of Interstate Banking Deregulation

NO	YEAR	HOME	FOREIGN
1	1982	ME	NY
2	1982	NY	ME
3	1983	CT	MA, ME
4	1983	MA	CT, ME
5	1983	ME	CT, MA
6	1984	CT	RI
7	1984	MA	RI
8	1984	ME	AL, AR, AZ, CA, CO, DC, FL, GA, IA, ID, IL, IN, KS, KY, LA, MD, MI, MN, MO, MS, MT, NC, ND, NE, NH, NJ, NM, NV, OH, OK, OR, PA, RI, SC, TN, TX, UT, VA, VT, WA, WI, WV, WY
9	1984	RI	CT, MA, ME
10	1985	DC	FL, MD, NC, VA
11	1985	FL	DC, GA, NC, TN, VA
12	1985	GA	FL, NC, TN, VA
13	1985	ID	NV, UT
14	1985	KY	OH, TN, VA
15	1985	MD	DC, VA
16	1985	NC	DC, FL, GA, TN, VA
17	1985	NV	ID, UT
18	1985	OH	KY
19	1985	TN	FL, GA, KY, NC, VA
20	1985	UT	ID, NV
21	1985	VA	DC, FL, GA, KY, MD, NC, TN
22	1986	AZ	All
23	1986	DC	AL, AR, AZ, CA, CO, CT, GA, IA, ID, IL, IN, KS, KY, LA, MA, ME, MI, MN, MO, MS, MT, ND, NE, NH, NJ, NM, NV, NY, OH, OK, OR, PA, RI, SC, TN, TX, UT, VT, WA, WI, WV, WY
24	1986	FL	SC

continued on next page

[8] Utah's national nonreciprocal law effective on Dec 31, 1987 is considered effective in 1988.

Table A.1 – *Continued from previous page*

NO	YEAR	HOME	FOREIGN
25	1986	GA	KY, SC
26	1986	ID	OR
27	1986	IL	IN, KY, MI, MO
28	1986	IN	IL, KY, MI, OH
29	1986	KY	AZ, DC, GA, IL, IN, ME, MO, NC, NJ, NY, PA, SC
30	1986	MI	IL, IN, OH
31	1986	MO	IL, KY, TN
32	1986	NC	KY, SC
33	1986	NJ	DC, KY, OH, PA
34	1986	NV	AZ, OR
35	1986	NY	AZ, DC, KY
36	1986	OH	DC, IN, MI, NJ, PA
37	1986	OR	AZ, CA, ID, NV, UT, WA
38	1986	PA	DC, KY, NJ, OH
39	1986	SC	DC, FL, GA, KY, NC, TN, VA
40	1986	TN	MO, SC
41	1986	UT	AZ, OR
42	1986	VA	SC
43	1987	AL	DC, FL, GA, KY, LA, MD, NC, SC, TN, VA
44	1987	CA	AZ, OR, TX, WA
45	1987	CT	NH
46	1987	FL	AL, LA, MD
47	1987	GA	AL, DC, LA, MD
48	1987	ID	WA, WY
49	1987	IL	WI
50	1987	IN	TN, WI
51	1987	KY	AL, LA, MD, OK, TX, WA, WI, WY
52	1987	LA	AL, DC, FL, GA, KY, MD, NC, OK, SC, TN, TX, VA
53	1987	MA	NH
54	1987	MD	AL, FL, GA, KY, LA, NC, PA, SC
55	1987	MI	WI
56	1987	MN	WI
57	1987	MO	OK
58	1987	NC	AL, LA, MD
59	1987	NH	CT, MA, ME, RI
60	1987	NV	WA, WY
61	1987	NY	OK, TX, WA, WY
62	1987	OH	WI
63	1987	OK	All
64	1987	PA	MD
65	1987	RI	NH
66	1987	SC	AL, LA, MD
67	1987	TN	AL, IN, LA
68	1987	TX	All
69	1987	UT	WA, WY
70	1987	VA	AL, LA
71	1987	WA	AZ, CA, DC, ID, KY, ME, NV, NY, OK, OR, TX, UT, WY
72	1987	WI	IL, IN, KY, MI, MN, OH
73	1987	WY	All
74	1988	AL	MS, TX, WV
75	1988	CA	ID, UT
76	1988	CO	AZ, OK, UT, WY

continued on next page

Table A.1 – *Continued from previous page*

NO	YEAR	HOME	FOREIGN
77	1988	CT	VT
78	1988	FL	WV
79	1988	ID	AL, AR, AZ, CA, CO, CT, DC, FL, GA, IA, IL, IN, KS, KY, LA, MA, MD, ME, MI, MN, MO, MS, MT, NC, ND, NE, NH, NJ, NM, NY, OH, OK, PA, RI, SC, TN, TX, VA, VT, WI, WV
80	1988	IN	WV
81	1988	KY	ID, MI, RI, UT, WV
82	1988	LA	MS, WV
83	1988	MA	VT
84	1988	MD	TN, WV
85	1988	MI	AZ, DC, ID, KY, ME, NJ, NY, OK, RI, TX, UT, WA, WV, WY
86	1988	MN	ID, WA, WY
87	1988	MS	AL, LA, TN
88	1988	NC	TX, WV
89	1988	NH	VT
90	1988	NJ	AZ, ID, ME, MI, NY, OK, RI, TX, UT, WA, WV, WY
91	1988	NY	ID, MI, NJ, OH, RI, UT, WV
92	1988	OH	AZ, ID, ME, NY, OK, RI, TX, UT, WA, WV, WY
93	1988	PA	WV
94	1988	RI	AZ, DC, ID, KY, MI, NJ, NY, OH, OK, TX, UT, VT, WA, WV, WY
95	1988	SC	WV
96	1988	TN	DC, MD, MS, WV
97	1988	UT	AL, AR, CA, CO, CT, DC, FL, GA, IA, IL, IN, KS, KY, LA, MA, MD, ME, MI, MN, MO, MS, MT, NC, ND, NE, NH, NJ, NM, NY, OH, OK, PA, RI, SC, TN, TX, VA, VT, WI, WV
98	1988	VA	WV
99	1988	VT	CT, MA, ME, NH, RI
100	1988	WA	MI, MN, NJ, OH, RI, WV
101	1988	WV	AL, AZ, DC, FL, ID, IN, KY, LA, MD, ME, MI, NC, NJ, NY, OH, OK, PA, RI, SC, TN, TX, UT, VA, WA, WY
102	1989	AL	AR
103	1989	AR	AL, DC, FL, LA, MD, MO, MS, NC, OK, SC, TN, TX, VA, WV
104	1989	CA	NM, NV
105	1989	CO	NM
106	1989	FL	AR
107	1989	KY	NM, NV, OR
108	1989	LA	AR, AZ, ID, ME, MI, NJ, NM, NV, NY, OH, OR, RI, UT, WA, WY
109	1989	MD	AR
110	1989	MI	LA, NM, NV, OR
111	1989	MO	AR
112	1989	MS	AR
113	1989	NC	AR
114	1989	NJ	LA, NM, NV, OR
115	1989	NM	All
116	1989	NV	AL, AR, CA, CO, CT, DC, FL, GA, IA, IL, IN, KS, KY, LA, MA, MD, ME, MI, MN, MO, MS, MT, NC, ND, NE, NH, NJ, NM, NY, OH, OK, PA, RI, SC, TN, TX, VA, VT, WI, WV
117	1989	NY	LA, NM, NV, OR
118	1989	OH	LA, NM, NV, OR
119	1989	OR	AL, AR, CO, CT, DC, FL, GA, IA, IL, IN, KS, KY, LA, MA, MD, ME, MI, MN, MO, MS, MT, NC, ND, NE, NH, NJ, NM, NY, OH, OK, PA, RI, SC, TN, TX, VA, VT, WI, WV, WY
120	1989	RI	LA, NM, NV, OR
121	1989	SC	AR
122	1989	TN	AR
123	1989	VA	AR

continued on next page

Table A.1 – *Continued from previous page*

NO	YEAR	HOME	FOREIGN
124	1989	WA	LA, NM
125	1989	WV	AR, NM, NV, OR
126	1990	CO	NE
127	1990	CT	AZ, DC, ID, IL, KY, LA, MI, NJ, NM, NV, NY, OH, OK, OR, PA, TX, UT, WA, WV, WY
128	1990	FL	MS
129	1990	GA	MS
130	1990	IL	AZ, CT, DC, ID, LA, MA, ME, MN, NH, NJ, NM, NV, NY, OH, OK, OR, PA, RI, TX, UT, VT, WA, WV, WY
131	1990	IN	MN, PA
132	1990	KY	CT, MA, MS, NH, VT
133	1990	LA	CT, IL, MA, NH, PA, VT
134	1990	MA	AZ, DC, ID, IL, KY, LA, MI, NJ, NM, NV, NY, OH, OK, OR, PA, TX, UT, WA, WV, WY
135	1990	MI	CT, MA, NH, PA, VT
136	1990	MN	IL, IN, NE
137	1990	MO	NE
138	1990	MS	FL, GA, KY, NC, SC, TX, VA, WV
139	1990	NC	MS
140	1990	NE	CO, MN, MO, WY
141	1990	NH	AL, AR, AZ, CA, CO, DC, FL, GA, IA, ID, IL, IN, KS, KY, LA, MD, MI, MN, MO, MS, MT, NC, ND, NE, NJ, NM, NV, NY, OH, OK, OR, PA, SC, TN, TX, UT, VA, WA, WI, WV, WY
142	1990	NJ	CT, IL, MA, NH, VT
143	1990	NY	CT, IL, MA, NH, PA, VT
144	1990	OH	CT, IL, MA, NH, VT
145	1990	PA	AZ, CT, ID, IL, IN, LA, MA, ME, MI, NH, NM, NV, NY, OK, OR, RI, TX, UT, VT, WA, WY
146	1990	RI	IL, PA
147	1990	SC	MS
148	1990	VA	MS
149	1990	VT	AZ, DC, ID, IL, KY, LA, MI, NJ, NM, NV, NY, OH, OK, OR, PA, TX, UT, WA, WV, WY
150	1990	WA	CT, IL, MA, NH, PA, VT
151	1990	WV	CT, IL, MA, MS, NH, VT
152	1991	AR	NE
153	1991	CA	CO, CT, DC, IL, KY, LA, MA, ME, MI, ND, NE, NH, NJ, NY, OH, OK, PA, RI, TN, VT, WV, WY
154	1991	CO	AL, AR, CA, CT, DC, FL, GA, IA, ID, IL, IN, KS, KY, LA, MA, MD, ME, MI, MN, MO, MS, MT, NC, ND, NH, NJ, NV, NY, OH, OR, PA, RI, SC, TN, TX, VA, VT, WA, WI, WV
155	1991	CT	CA, CO, ND, NE, TN
156	1991	IA	IL, MN, MO, NE, WI
157	1991	IL	CA, CO, IA, ND, NE, TN
158	1991	KY	CA, CO, ND, NE
159	1991	LA	CA, CO, ND, NE
160	1991	MA	CA, CO, ND, NE, TN
161	1991	MI	CA, CO, ND, NE, TN
162	1991	MN	CO, IA, ND
163	1991	MO	IA
164	1991	ND	AZ, CA, CO, CT, DC, ID, IL, KY, LA, MA, ME, MI, MN, NE, NH, NJ, NM, NV, NY, OH, OK, OR, PA, RI, TN, TX, UT, VT, WA, WV, WY
165	1991	NE	AR, AZ, CA, CT, DC, IA, ID, IL, KY, LA, MA, ME, MI, ND, NH, NJ, NM, NV, NY, OH, OK, OR, PA, RI, TN, TX, UT, VT, WA, WV
166	1991	NJ	CA, CO, ND, NE, TN
167	1991	NY	CA, CO, ND, NE, TN
168	1991	OH	CA, CO, ND, NE, TN
169	1991	PA	CA, CO, ND, NE, TN
170	1991	RI	CA, CO, ND, NE, TN

continued on next page

Table A.1 – *Continued from previous page*

NO	YEAR	HOME	FOREIGN
171	1991	TN	AZ, CA, CO, CT, ID, IL, MA, ME, MI, ND, NE, NH, NJ, NM, NV, NY, OH, OK, OR, PA, RI, TX, UT, VT, WA, WY
172	1991	VT	CA, CO, ND, NE, TN
173	1991	WA	CO, ND, NE, TN
174	1991	WI	IA
175	1991	WV	CA, CO, ND, NE
176	1992	AR	KS
177	1992	CA	IN
178	1992	CT	IN
179	1992	IN	AZ, CA, CO, CT, DC, ID, LA, MA, ME, ND, NE, NH, NJ, NM, NV, NY, OK, OR, RI, TX, UT, VT, WA, WY
180	1992	KS	AR, CO, MO, NE, OK
181	1992	LA	IN
182	1992	MA	IN
183	1992	MI	MN
184	1992	MN	MI, OH
185	1992	MO	KS
186	1992	ND	IN
187	1992	NE	IN, KS
188	1992	NJ	IN
189	1992	NY	IN
190	1992	OH	MN
191	1992	RI	IN
192	1992	VT	IN
193	1992	WA	IN
194	1993	MN	MT
195	1993	MT	CO, ID, MN, ND, WY
196	1993	ND	MT
197	1994	CA	MN, NC, VA
198	1994	CT	MN, NC, VA
199	1994	IL	NC, VA
200	1994	IN	NC, VA
201	1994	KY	MN
202	1994	LA	MN
203	1994	MA	MN, NC, VA
204	1994	MI	NC, VA
205	1994	MN	AZ, CA, CT, DC, KY, LA, MA, ME, NC, NH, NJ, NM, NV, NY, OK, OR, PA, RI, TN, TX, UT, VA, VT, WV
206	1994	NC	AZ, CA, CO, CT, ID, IL, IN, MA, ME, MI, MN, ND, NE, NH, NJ, NM, NV, NY, OH, OK, OR, PA, RI, UT, VT, WA, WY
207	1994	ND	NC, VA
208	1994	NE	NC, VA
209	1994	NJ	MN, NC, VA
210	1994	NY	MN, NC, VA
211	1994	OH	NC, VA
212	1994	PA	MN, NC, VA
213	1994	RI	MN, NC, VA
214	1994	TN	MN
215	1994	VA	AZ, CA, CO, CT, ID, IL, IN, MA, ME, MI, MN, ND, NE, NH, NJ, NM, NV, NY, OH, OK, OR, PA, RI, TX, UT, VT, WA, WY
216	1994	VT	MN, NC, VA
217	1994	WA	NC, VA

continued on next page

Table A.1 – *Continued from previous page*

NO	YEAR	HOME	FOREIGN
218	1994	WV	MN
219	1995	AL	AZ, CA, CO, CT, IA, ID, IL, IN, KS, MA, ME, MI, MN, MO, MT, ND, NE, NH, NJ, NM, NV, NY, OH, OK, OR, PA, RI, UT, VT, WA, WI, WY
220	1995	AR	AZ, CA, CO, CT, GA, IA, ID, IL, IN, KY, MA, ME, MI, MN, MT, ND, NH, NJ, NM, NV, NY, OH, OR, PA, RI, UT, VT, WA, WI, WY
221	1995	CA	AL, AR, FL, GA, IA, KS, MD, MO, MS, MT, SC, WI
222	1995	CT	AL, AR, FL, GA, IA, KS, MD, MO, MS, MT, SC, WI
223	1995	FL	AZ, CA, CO, CT, IA, ID, IL, IN, KS, KY, MA, ME, MI, MN, MO, MT, ND, NE, NH, NJ, NM, NV, NY, OH, OK, OR, PA, RI, TX, UT, VT, WA, WI, WY
224	1995	GA	AR, AZ, CA, CO, CT, IA, ID, IL, IN, KS, MA, ME, MI, MN, MO, MT, ND, NE, NH, NJ, NM, NV, NY, OH, OK, OR, PA, RI, TX, UT, VT, WA, WI, WV, WY
225	1995	IA	AL, AR, AZ, CA, CO, CT, DC, FL, GA, ID, IN, KS, KY, LA, MA, MD, ME, MI, MS, MT, NC, ND, NH, NJ, NM, NV, NY, OH, OK, OR, PA, RI, SC, TN, TX, UT, VA, VT, WA, WV, WY
226	1995	IL	AL, AR, FL, GA, KS, MD, MS, MT, SC
227	1995	IN	AL, AR, FL, GA, IA, KS, MD, MO, MS, MT, SC
228	1995	KS	AL, AZ, CA, CT, DC, FL, GA, IA, ID, IL, IN, KY, LA, MA, MD, ME, MI, MN, MS, MT, NC, ND, NH, NJ, NM, NV, NY, OH, OR, PA, RI, SC, TN, TX, UT, VA, VT, WA, WI, WV, WY
229	1995	KY	AR, FL, IA, KS, MT
230	1995	LA	IA, KS, MO, MT, WI
231	1995	MA	AL, AR, FL, GA, IA, KS, MD, MO, MS, MT, SC, WI
232	1995	MD	AZ, CA, CO, CT, IA, ID, IL, IN, KS, MA, ME, MI, MN, MO, MS, MT, ND, NE, NH, NJ, NM, NV, NY, OH, OK, OR, RI, TX, UT, VT, WA, WI, WY
233	1995	MI	AL, AR, FL, GA, IA, KS, MD, MO, MS, MT, SC
234	1995	MN	AL, AR, FL, GA, KS, MD, MO, MS, SC
235	1995	MO	AL, AZ, CA, CO, CT, DC, FL, GA, ID, IN, LA, MA, MD, ME, MI, MN, MS, MT, NC, ND, NH, NJ, NM, NV, NY, OH, OR, PA, RI, SC, TX, UT, VA, VT, WA, WI, WV, WY
236	1995	MS	AZ, CA, CO, CT, DC, IA, ID, IL, IN, KS, MA, MD, ME, MI, MN, MO, MT, ND, NE, NH, NJ, NM, NV, NY, OH, OK, OR, PA, RI, UT, VT, WA, WI, WY
237	1995	MT	AL, AR, AZ, CA, CT, DC, FL, GA, IA, IL, IN, KS, KY, LA, MA, MD, ME, MI, MO, MS, NC, NE, NH, NJ, NM, NV, NY, OH, OK, OR, PA, RI, SC, TN, TX, UT, VA, VT, WA, WI, WV
238	1995	NC	IA, KS, MO, MT, WI
239	1995	ND	AL, AR, FL, GA, IA, KS, MD, MO, MS, SC, WI
240	1995	NE	AL, FL, GA, MD, MS, MT, SC, WI
241	1995	NJ	AL, AR, FL, GA, IA, KS, MD, MO, MS, MT, SC, WI
242	1995	NY	AL, AR, FL, GA, IA, KS, MD, MO, MS, MT, SC, WI
243	1995	OH	AL, AR, FL, GA, IA, KS, MD, MO, MS, MT, SC
244	1995	PA	AL, AR, FL, GA, IA, KS, MO, MS, MT, SC, WI
245	1995	RI	AL, AR, FL, GA, IA, KS, MD, MO, MS, MT, SC, WI
246	1995	SC	AZ, CA, CO, CT, IA, ID, IL, IN, KS, MA, ME, MI, MN, MO, MT, ND, NE, NH, NJ, NM, NV, NY, OH, OK, OR, PA, RI, TX, UT, VT, WA, WI, WY
247	1995	TN	IA, KS, MT, WI
248	1995	VA	IA, KS, MO, MT, WI
249	1995	VT	AL, AR, FL, GA, IA, KS, MD, MO, MS, MT, SC, WI
250	1995	WA	AL, AR, FL, GA, IA, KS, MD, MO, MS, MT, SC, WI
251	1995	WI	AL, AR, AZ, CA, CO, CT, DC, FL, GA, ID, KS, LA, MA, MD, ME, MO, MS, MT, NC, ND, NE, NH, NJ, NM, NV, NY, OK, OR, PA, RI, SC, TN, TX, UT, VA, VT, WA, WV, WY
252	1995	WV	GA, IA, KS, MO, MT, WI

B Data Appendix

B.1 Firm Establishment Data

B.1.1 Data source

Data of firm establishment location and employment come from Dun & Bradstreet (D&B). They were downloaded using the online platform, Mergent Data Explr, accessed through Princeton University Library. Individual data files by state-years were downloaded in CSV format. The platform is also able to produce annual data files in zipped CSV format. However, reading these zipped CSV files in such statistical software as R could result in incomplete loading of the dataset. The cause of this issue is unclear. I conjecture that the raw data files contain irregular characters that may be misinterpreted by the statistical software as the terminal character of the file. Therefore, a safe practice is to download raw data files in each year state by state, load the files as is into a software, and finally separate the lines by commas.

B.1.2 Ownership structure

D&B datasets contain detailed information about firm ownership structure. This information is essential for identifying boundaries of a firm and for measuring its operations across regions. Each observation in the dataset is an establishment, each assigned a unique identifier called DUNSNO. It is also associated with its headquarter ID by variable HQDUNSNO, its immediate parent ID by PARENTDUNSNO, and its ultimate parent ID by ULTDUNSNO. However, the original ownership information in the raw data may contain inconsistencies or errors, such as missing values, multiple parents, and infinite loops in ownership, etc. I follow the procedures below to construct a consistent ownership table for each year of the datasets. Since HQDUNSNO is the lowest level of a firm (or a group of establishments), it suffices to find the ultimate parents of these HQDUNSNOs.

1. **Preliminary ownership table.** To start, I construct a preliminary ownership table from the raw data file. First, assign an establishment's own id DUNSNO to its HQDUNSNO if the latter is missing, and then retain observations of HQDUNSNOs with non-missing values of either PARENTDUNSNO or ULTDUNSNO. This results in a table of triplets HQ-PARENT-ULT that contains firms with a higher level owner. HQDUNSNOs with PARENTDUNSNO and ULTDUNSNO both missing are assumed to be ultimately owned by HQDUNSNO itself, unless further modifications on the ownership structure occur in later stages.

2. **Immediate parent table.** For those HQs with non-missing PARENTDUNSNO in the preliminary ownership table, I extract a table of immediate parents (HQ-PARENT) after fixing multiple matches. Multiple matches are corrected according to data in adjacent years. I then apply this immediate parent table repeatedly to trace out higher levels of parents for each HQ until no higher parent can be found. Infinite loops may occur in this step when, for instance, two entities appear to be parents of each other. Typically, these incidents are manually fixed by removing the parent of one of the entities in the loop. To determine which entity's parent to remove, I include data from adjacent years to decide which one tends to be of higher level.

3. **Direct ultimate parents.** For observations with non-missing ULTPARENT in the preliminary ownership table, I reframe the table into duplets of direct ultimate parents so that each ULTPARENT is associated with all its subsidiaries in original columns of HQDUNSNO or PARENTDUNSNO. I then check if any entity is associated with multiple ULTDUNSNOs. Corrections are made according to the ownership structure suggested in the immediate parent table constructed in the previous step or data in adjacent years.

4. **Consistency check.** With the tables of immediate parents and direct ultimate parents at hand, I can assign ultimate owners to HQs in the preliminary ownership table in two ways, one tracing through immediate parents and the other through direct ultimate parents. Since not all immediate parents appear in the the column PARENTDUNSNO and some ULTDUNSNOs may still have PARENTDUNSNO, tracing the ownership through either approach would require using both the immediate parent table and the direct ultimate parent table. Specifically, to implement the first approach of tracing through immediate parents, I first link each HQ in the preliminary ownership table to a candidate ultimate parent using the immediate parent table. I then apply the direct ultimate parent table once to link this candidate ultimate owner to a second candidate ultimate owner. The immediate parent table is applied again to link the second candidate ultimate owner to a third. Similarly, to implement the second approach of tracing through direct ultimate parent, I first apply the direct ultimate parent table and then the immediate parent table. Finally, I check if these two approaches produce the same ultimate owner. If not, modifications will be made to immediate parents or direct ultimate parents.

B.1.3 Multistate firms

With the ownership table at hand, I now turn to construct a dataset of multistate firms.

1. Use the ownership table to identify the ultimate owner of each establishment.

2. If the ultimate owner is a non-financial non-government entity, then we are done. If the ultimate owner is a government entity (SIC 91–97), then it is excluded from our analysis with all its subsidiary establishments. If the ultimate owner is a financial company (SIC 60–67), then I find its highest level subsidiaries that are in the real sector and consider these subsidiaries as separate firms if there is more than one.

3. Exclude establishments that are financial entities controlled by these firms. Retain firms with operations in multiple states and at least 500 employees in total.

B.2 Merged DealScan-D&B Data

This section provides descriptive statistics on the merged DealScan-D&B dataset. The dataset is constructed by matching borrowers in DealScan to Dun & Bradstreet based on information on company name, industry and locations that are available in either databases. I retain the sample of loans in DealScan that are designated to US non-financial borrowers.

Figure B.1 shows the sample size of the merged dataset. Panel A plots the number of unique non-financial US borrowers in DealScan, panel B the number of loans, and Panel C the dollar amount of loans. The grey bars represent the entire DealScan sample of loans by US non-financial firms, while the black bars represent the fraction of loans whose borrowers can be matched to Dun & Bradstreet. As the figure shows, I was able to match almost all non-financial borrowers in DealScan to Dun & Bradstreet.

Table B.1 presents summary statistics for borrowing firms in the merged dataset in terms of their total employment, number of counties or states they operate in.

C Additional Results

Table C.1 presents coefficient estimates from Equation 1 for entry outcomes in one to ten years since deregulation.

Table C.2 includes bank-industry fixed effects to account for the effects of banks' industry specialisation on geographic expansion.

Table C.3 reports heterogeneous effects on entry outcomes in one to ten years since deregulation.

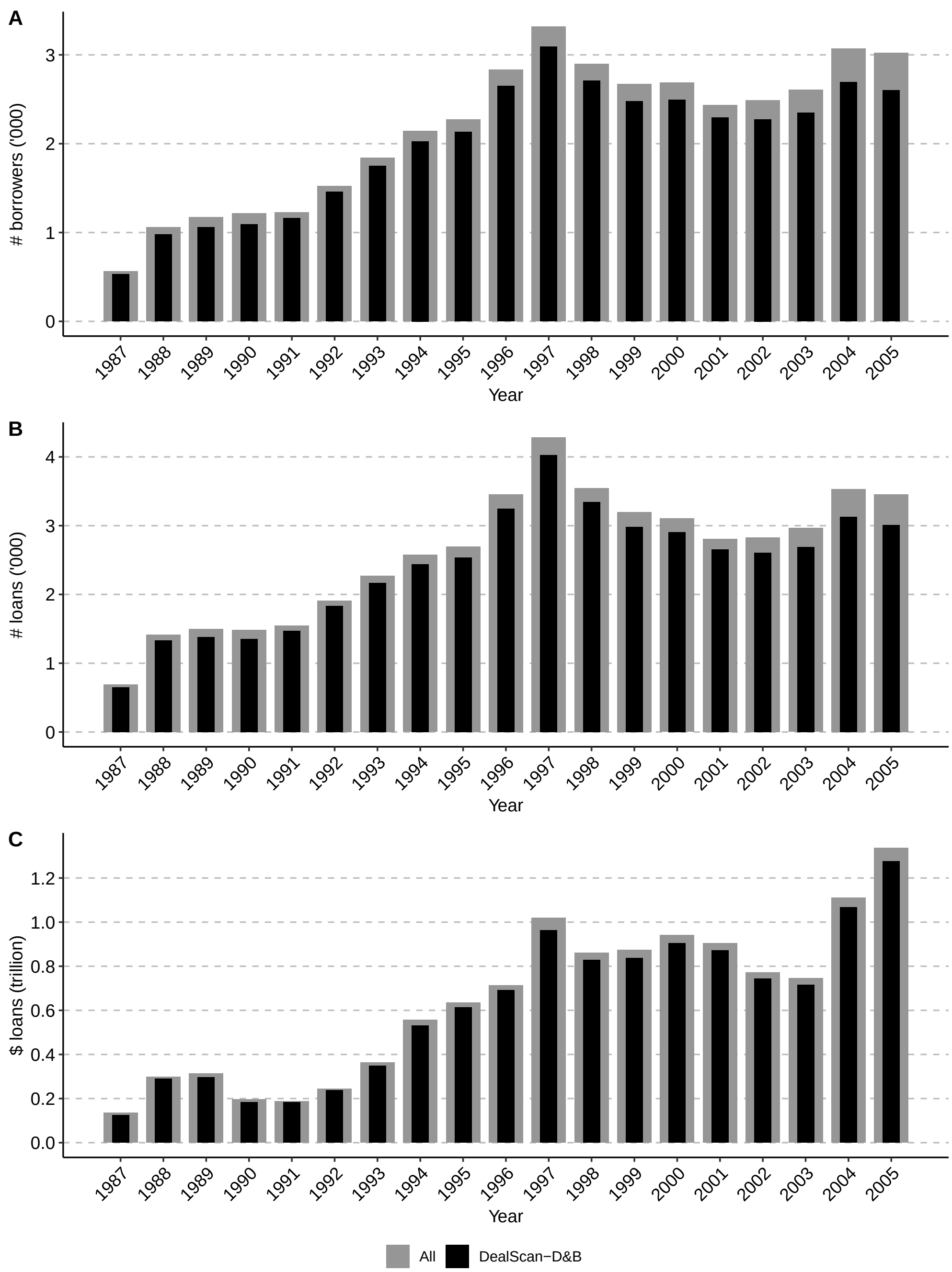

Figure B.1: Sample Size of the Merged DealScan-DunBradstreet Dataset

Notes: This figure plots the sample size of the merged DealScan-DunBradstreet dataset. Panel A plots the total number of unique US nonfinancial borrowers (in thousands) in DealScan, and the number of borrowers that can be matched to Dun & Bradstreet. Panel B plots the total number of loan packages (in thousands) in DealScan and those whose borrowers can be matched to Dun & Bradstreet. Panel C plots the total amount of loan packages (in trillions) in DealScan and those whose borrowers can be matched to Dun & Bradstreet.

Table B.1: Summary Statistics for Borrowers in Merged DealScan-D&B Dataset

Year	Variable	Observations	Mean	SD	Q25	Q50	Q75
1987	Employment	563	6347.3	17034.7	114	726	2821.4
	# Counties	563	32.2	68.1	2	7	20
	# States	563	10.1	12.5	1	4	10
1988	Employment	1058	5654.5	15370.9	117.8	716	2345.9
	# Counties	1058	29.3	59.1	2	7	19
	# States	1058	9.5	11.4	1	4	10
1989	Employment	1171	4586.5	12712.6	115	625	2142
	# Counties	1171	32.1	75.3	2	6	18
	# States	1171	9.4	11.7	1	4	10
1990	Employment	1214	5016.9	15137.9	108.5	623.5	2145.1
	# Counties	1214	29.1	62.5	2	6	17
	# States	1214	9	11.4	1	4	9
1991	Employment	1228	4937.1	15631.7	170.8	880	2553.3
	# Counties	1228	32.5	75.2	2	8	23
	# States	1228	9.6	11.4	2	4.5	11
1992	Employment	1525	5659.4	17006.2	203	1070	2931.2
	# Counties	1525	38.8	89.9	3	10	25
	# States	1525	10.6	12.1	2	5	12
1993	Employment	1842	6237.8	23562.3	198.2	929.5	2837.2
	# Counties	1842	36.6	79.5	2	9	26
	# States	1842	10.5	12.1	2	5	12
1994	Employment	2145	6881.9	23290.9	212	1000	2983.4
	# Counties	2145	43.1	98.2	3	10	26
	# States	2145	11.1	12.7	2	5	12
1995	Employment	2270	5893.2	21094.9	200	1008.5	2852.3
	# Counties	2270	41.6	98.3	3	9	26
	# States	2270	10.8	12.5	2	5	12
1996	Employment	2833	5201	19418	153	735	2187.4
	# Counties	2833	39.4	100.9	2	8	23
	# States	2833	10.2	12.2	2	5	11
1997	Employment	3320	5009.7	18448.5	152	703	2154.3
	# Counties	3320	37.7	96.9	2	8	20
	# States	3320	10	11.9	2	5	11
1998	Employment	2898	5069.1	20845.2	142.2	666.5	2018.7
	# Counties	2898	37.4	92.6	2	8	21
	# States	2898	10	11.8	2	5	11
1999	Employment	2673	6368.7	24829.7	168	804	2555.4
	# Counties	2673	45.2	110.1	2	9	26
	# States	2673	11.1	12.8	2	5	13
2000	Employment	2687	7257.9	24936.5	174	970	3195.8
	# Counties	2687	51	118.4	2	10	29
	# States	2687	11.8	13.4	2	6	14
2001	Employment	2436	8687.4	29332.9	200	1151.5	3806
	# Counties	2436	61	138.7	3	11	35
	# States	2436	12.8	14.2	2	6	15
2002	Employment	2487	7419.3	25414.7	211	1221	3588.2
	# Counties	2487	59.1	133.7	3	12	38
	# States	2487	12.9	14	2	7	16
2003	Employment	2609	6442.4	22316.9	175	1038	3193.4
	# Counties	2609	57.5	129.9	2	11	31
	# States	2609	12.4	13.9	2	6	15
2004	Employment	3069	5710.1	23895.5	125	916	2687.2
	# Counties	3069	54.2	132.1	2	10	30
	# States	3069	11.8	13.6	1	6	14
2005	Employment	3021	5937.4	26688	76	772	2522
	# Counties	3021	50.5	126.7	1	9	28
	# States	3021	11.2	13.3	1	5	13

Notes: This table reports summary statistics of the borrowers in the merged DealScan-D&B dataset. These include the total employment, the number of counties and states of operation.

Table C.1: Multistate Firms and Bank Entry: Results for all years

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Entry_{+1}	Entry_{+2}	Entry_{+3}	Entry_{+4}	Entry_{+5}	Entry_{+6}	Entry_{+7}	Entry_{+8}	Entry_{+9}	Entry_{+10}
Old Neighbours	0.592^{***}	0.926^{***}	1.180^{***}	1.44^{***}	1.57^{***}	1.84^{***}	2.01^{***}	2.01^{***}	2.26^{***}	2.72^{***}
	(0.157)	(0.266)	(0.296)	(0.351)	(0.397)	(0.430)	(0.460)	(0.465)	(0.527)	(0.668)
LogDist	-0.567^{***}	-0.748^{***}	-0.816^{***}	-1.02^{***}	-1.16^{***}	-1.38^{***}	-1.59^{***}	-1.69^{***}	-1.86^{***}	-2.02^{***}
	(0.159)	(0.168)	(0.171)	(0.195)	(0.216)	(0.269)	(0.288)	(0.290)	(0.314)	(0.329)
Bank-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Dep Var (%)	0.11	0.19	0.23	0.28	0.35	0.42	0.52	0.57	0.63	0.75
R^2 Adj	0.124	0.146	0.158	0.159	0.160	0.172	0.192	0.200	0.207	0.242
Observations	673,006	653,257	632,528	598,524	566,286	536,010	513,353	488,366	467,785	442,097

Notes: This table reports regression results of the linear probability model of bank entry in Equation 1. The dependent variable $Entry_{+h}$ is an indicator for whether bank controls a branch in the destination county, h years after deregulation. Variable *Old Neighbours* measures employment share in the destination county of a bank's original neighbouring firms, divided by 100. Variable *LogDist* measures the log geodesic distance between the bank's headquarters and the destination county (in miles), divided by 100. Standard errors are clustered at bank and county levels. Significance levels: ***1%, **5%, *10%.

Table C.2: Multistate Firms and Bank Entry: Controling for bank-industry specialisation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Entry_{+1}	Entry_{+2}	Entry_{+3}	Entry_{+4}	Entry_{+5}	Entry_{+6}	Entry_{+7}	Entry_{+8}	Entry_{+9}	Entry_{+10}
Old Neighbours	0.517^{***}	0.867^{***}	1.100^{***}	1.38^{***}	1.59^{***}	1.93^{***}	2.13^{***}	2.12^{***}	2.42^{***}	2.96^{***}
	(0.139)	(0.276)	(0.309)	(0.360)	(0.412)	(0.451)	(0.484)	(0.477)	(0.549)	(0.749)
LogDist	-0.559^{***}	-0.752^{***}	-0.827^{***}	-1.04^{***}	-1.19^{***}	-1.41^{***}	-1.60^{***}	-1.70^{***}	-1.85^{***}	-2.00^{***}
	(0.160)	(0.168)	(0.175)	(0.202)	(0.220)	(0.277)	(0.293)	(0.296)	(0.319)	(0.333)
Bank-SIC-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Dep Var (%)	0.11	0.18	0.22	0.28	0.34	0.42	0.52	0.57	0.62	0.74
R^2 Adj	0.193	0.200	0.201	0.198	0.202	0.213	0.239	0.250	0.259	0.293
Observations	679,197	659,250	638,348	604,067	571,530	540,987	518,118	492,919	472,122	446,186

Notes: This table reports regression results from Equation 1, controling for bank-2 digit SIC-year fixed effects. The dependent variable $Entry_{+h}$ is an indicator for whether bank controls a branch in the destination county, h years after deregulation. Variable *Old Neighbours* measures employment share in the destination county of a bank's original neighbouring firms, divided by 100. Variable *LogDist* measures the log geodesic distance between the bank's headquarters and the destination county (in miles), divided by 100. Standard errors are clustered at bank and county levels. Significance levels: ***1%, **5%, *10%.

Table C.3: Mutistate Firms and Bank Entry: Heterogeneous Effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Entry_{+1}	Entry_{+2}	Entry_{+3}	Entry_{+4}	Entry_{+5}	Entry_{+6}	Entry_{+7}	Entry_{+8}	Entry_{+9}	Entry_{+10}
Old N'bours	0.0726	0.00326	0.177	0.283	0.532	0.518	0.167	-0.343	-0.50	-1.06
	(0.185)	(0.298)	(0.353)	(0.430)	(0.495)	(0.590)	(0.654)	(0.628)	(0.733)	(0.883)
Old N'bours × C&I Loan	3.4000^{***}	6.05000^{**}	6.580^{**}	7.550^{**}	6.810^{**}	8.670^{**}	12.000^{***}	15.500^{***}	18.20^{***}	25.10^{***}
	(1.270)	(2.590)	(2.890)	(3.130)	(2.990)	(3.570)	(4.300)	(4.890)	(6.090)	(8.620)
LogDist	-0.5680^{***}	-0.75000^{***}	-0.817^{***}	-1.020^{***}	-1.160^{***}	-1.390^{***}	-1.590^{***}	-1.690^{***}	-1.87^{***}	-2.03^{***}
	(0.160)	(0.168)	(0.171)	(0.196)	(0.217)	(0.269)	(0.288)	(0.290)	(0.314)	(0.329)
Bank-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Dep Var (%)	0.11	0.19	0.23	0.28	0.35	0.42	0.52	0.57	0.63	0.75
R^2 Adj	0.124	0.146	0.158	0.159	0.16	0.172	0.192	0.20	0.207	0.242
Observations	672,990	653,241	632,512	598,508	566,270	535,994	513,337	488,366	467,785	442,097

Notes: This table reports regression results from Equation 1, allowing heterogenous effects across bank characteristics. The dependent variable $Entry_{+h}$ is an indicator for whether bank controls a branch in the destination county, h years after deregulation. Variable *Old Neighbours* is the employment share in the destination county by the bank's original neighbouring firms. Variable *C&I Loan* is the share of commercial and industrial loans out of bank's total assets. Standard errors are clustered at bank and county levels. Significance levels: ***1%, **5%, *10%.