Industry Specialization and Small Business Lending

Wenhua Di and Nathaniel Pattison[*][†]

Current Draft: March 2022

Abstract

This paper examines the rise and competitive impact of industry-specialized small business lenders. Using loan-level data with detailed industry codes from the Small Business Administration (SBA), we document a recent increase in lenders that originate loans nationally but to a limited number of industries. We then examine the impact of industry-specialized lending on credit availability and banking competition. Exploiting the staggered entry of a large, specialized lender, we find significant increases in total SBA-backed lending with no evidence of substitution from other lenders. We then explore potential mechanisms behind the increase in lending.

JEL G21, G23, L11
Keywords: Small Business Lending, Banking Competition, Specialization, Distance in Lending, Credit Access

[*]Di: Federal Reserve Bank of Dallas, 2200 N. Pearl St. Dallas, TX 75201, wenhua.di@dal.frb.org. Pattison (corresponding author): Southern Methodist University, ULEE 301E, Dallas, TX 75275, npattison@smu.edu.

[†]Disclaimer: The views in this paper are those of the authors and do not necessarily represent those of the Federal Reserve Banks or the Federal Reserve System. An earlier version of this paper was circulated as "Remote Competition and Small Business Lending: Evidence from SBA Lending." Acknowledgments: We thank SungJe Byun, Robert DeYoung, Scott Frame, Janet Garufis, Erik Mayer, Daniel Millimet, Pia Orrenius, Stephen Shore, Yichen Su, Joe Tracy, Gregory Udell, Mike Weiss, brown bag participants at the Federal Reserve Bank of Dallas and Southern Methodist University, and conference participants at the Federal Reserve System Applied Micro Conference, the Community Banking in the 21st Century research and policy conference at the Federal Reserve Bank of St. Louis, the Banking and Finance Workshop at the Federal Reserve Bank of Dallas, the Society for Government Economists at the 2018 ASSA Annual Meeting, the Stata Texas Empirical Microeconomics Conference, and the GSU Consumer Finance Seminar. We thank Benjamin Meier for his research assistance. Declarations of Interest: None.

1 Introduction

Two central challenges facing small business lenders are information opacity and business heterogeneity (Mills, 2019*b*). Information about a small business is difficult to acquire and communicate (Berger and Udell, 1995, Petersen and Rajan, 1994) and each small business is different, making information hard to generalize. Both challenges are closely tied to distance. Proximity aids in the collection and transfer of opaque information, leading to better risk assessment and fewer defaults among nearby borrowers (Petersen and Rajan, 2002, DeYoung, Glennon and Nigro, 2008, Agarwal and Hauswald, 2010). But lending only to nearby borrowers can heighten the challenge of business heterogeneity, since the smaller pool of potential borrowers limits the scope for a lender to specialize in certain types of business. In contrast, a lender covering a larger area could focus only on certain industries, perhaps developing industry-specific expertise or making industry-specific investments in underwriting or marketing, perhaps offsetting some disadvantages of distant lending.

Small business lenders face this trade-off between geographic specialization - lending locally to a variety of industries - and industry specialization - lending to specific industries across a variety of locations. Historically, lenders have nearly all chosen the former. The median distance between branches of small business lenders and their borrowers remains less than ten miles, and credit availability is tightly linked to the presence of nearby banks (Nguyen, 2019, Granja, Leuz and Rajan, 2018). In this paper, we document the recent rise of remote, industry-specialized lenders, i.e., institutions that lend nationally but specialize in a narrow set of industries. Based on loan-level data from the Small Business Administration (SBA) 7(a) program, the percent of SBA loans (in dollars) accounted for by these specialized lenders has increased from less than 2% in 2001 to more than 17% in 2017. These specialists develop expertise in advertising, underwriting, or monitoring in their specific industries, and we find evidence that specialists exhibit improved loan performance compared to other lenders in the same industry.

Do industry specialized lenders complement local lenders or do they substitute, competing with local lenders for the same borrowers? The answer to this question has important implications for lenders and small businesses. If remote, specialized lenders complement existing lenders, the growth in specialized lending may relax credit constraints common among small businesses and lead to growth within the targeted industries. Additionally, for the targeted industries, remote specialists may weaken the link dependence of small businesses on nearby branches (Nguyen, 2019), leading to an increase in small business lending within underbanked areas. On the other hand, if specialized lenders serve as substitutes, they may simply increase competition for existing borrowers. At best, substitution would result in little increase in the supply of credit, and in some cases, cream-skimming by new lenders could even lead to unraveling and a decline in credit availability (Detragiache, Tressel and Gupta, 2008, Gormley, 2014).

The focus of this paper is to examine the impact of industry-specialized lenders on the small business lending market, in particular the market for SBA-guaranteed loans. The primary challenge is that specialized lending has grown steadily and endogenously over time, making it difficult to separate the impact of specialized lending from other factors. A second challenge is that commonly used data on small business lending do not contain detailed industry information. Previous papers examining industry concentration among lenders had to group businesses into 10-25 broad categories such as agriculture, construction, or energy, which are too coarse to detect specialization in narrowly defined industries.

To address these challenges, we examine entry by a specific industry-specialized lender, Live Oak Bank, within the SBA 7(a) lending program. Live Oak Bank is a prominent industry-specialized lender, originating more than 80% of its loans to just six industries. The bank identifies industry-specific expertise as its primary advantage. Upon entering, Live Oak quickly accounts for around 50% of SBA 7(a) lending to these industries. These large, staggered shocks to the supply of credit allow us to identify the impact of a sudden increase in industry-specialized lending on the market for SBA-guaranteed small business loans. The fact that Live Oak operates within the SBA program also allows us to use the SBA's loan-level data containing the full NAICS industry code for each loan from more than 800 distinct industries. In addition to providing a unique opportunity to identify the impact of specialized lending, this setting is interesting in its own right. SBA lending is an important source of credit for constrained small businesses, providing around a quarter of loans to employer small businesses (Federal Reserve Banks, 2016-2019), and Live Oak is the SBA program's largest lender, originating around 6% of SBA-backed dollars.

Exploiting the entry of this industry-specialized lender into its six industries, we estimate the impact on total lending and the competitive effect on other lenders within the market for SBA-guaranteed loans. Our strategy compares the evolution of lending in these six treated industries to a group of control industries that Live Oak did not enter. We use the synthetic control method (Abadie and Gardeazabal, 2003, Abadie, Diamond and Hainmueller, 2010) to create a weighted average of control industries chosen to best match the treated industry's lending path during the pretreatment period.

We find that the industry-specialized lending significantly increased overall SBA-guaranteed lending to the treated industries. Across the six treated industries, annual loan originations rises by 30-110 percentage points relative to the synthetic control. Moreover, there is no substitution away from other SBA lenders. Other institutions' SBA lending to the targeted industries remains unchanged upon Live Oak's entry. These results are robust to a variety of extensions and robustness checks and indicate that Live Oak originates loans to new borrowers who would not have obtained an SBA loan otherwise. We lack the data to directly examine substitution from non-SBA commercial lending, but institutional features, existing evidence, and an empirical test all suggest that substitution from non-SBA bank lending is limited. One possibility is that specialized lending substitutes away from non-commercial alternatives, such as seller financing, which was historically common in several of the specialists' industries.

Next, we examine several mechanisms for the increase in lending that are related to industry specialization. Lenders themselves state that their ability to make these distant loans depends on their industry specialization and expertise. We find that specialized lenders focus on safer industries and industries where there is a weak relationship between distance and charge-offs. Consistent with expertise, we also find that specialized lenders experience better loan performance, especially for distant loans, than other lenders within the same industries. Live Oak Bank maintains similar charge-off rates to those of other lenders in the industries, despite significantly increasing total originations and lending at much greater distances.

Our main results focus on a single lender within the SBA program, so a natural question is whether these results generalize to other lenders and settings. Adapting the empirical strategy, we find similar average effects for other specialized lenders within the SBA program. For non-SBA lenders, detailed industry data on small business lending is not available. Other research and industry reports, however, indicate that the increase in industry-specialized lending is increasingly prominent outside of SBA lending as well. Karen Mills, former Administrator of the SBA, describes the emergence of industry-specialized lenders as a key innovation in small business lending more generally (Mills, 2019*a*) and trade publications have highlighted the recent growth of specialty lending.[1]. Additionally, Blickle, Parlatore and Saunders (2021) shows that industry-specialized commercial and industrial (C&I) lending exists even among large banks subject to stress testing. Similar to our setting, they find that there exists a subset of extreme specialists, that specialization is associated with improved loan performance, and that industry specialization is increasing over time.

The two main contributions of this paper are (i) to document the growth in industry-specialized small business lenders within the SBA program and (ii) to estimate the competitive impact of the largest specialized lender. Our paper adds to the literature examining sectoral specialization by banks. The existing literature generally focuses on the relationship between banks' sectoral specialization and their risk, often finding that sectoral specialization lowers risk, consistent with expertise. Winton (1999) and Stomper (2006) provide models of sectoral expertise and lending, and the related empirical literature mostly finds that specialization increases returns and reduces risks (Acharya, Hasan and Saunders, 2006, Hayden, Porath and Westernhagen, 2007, Boeve, Duellmann and Pfingsten, 2010, Jahn, Memmel and Pfingsten, 2016, Tabak, Fazio and Cajueiro, 2011, Beck, De Jonghe et al., 2013, Giometti and Pietrosanti, 2019). Others document the importance of lender specialization along other dimensions, including the export market (Paravisini, Rappoport and Schnabl, 2020), collateral type (Gopal, 2019), and loan types (e.g. consumer or commercial) in the purchase of failed banks (Granja, Matvos and Seru, 2017). Our paper is unique in that it examines industry specialization at a much finer level, with over 800 distinct industries, whereas the existing literature observes 20-40 broad sectors. We also corroborate some results related to risk for these highly specialized lenders, finding that lenders with a higher share of loans in an industry experience better loan performance in that industry.

Our paper also connects to the literature examining the competitive impact of different types of new entrants in banking, which finds mixed effects with the impact of new banking competition varying across contexts and types of lenders. One strand of this literature examines the entry of foreign (distant) banks and their impact on domestic (local) lenders after financial liberalization. Detragiache, Tressel and Gupta (2008) and Gormley (2014) develop theoretical models showing that competition from distant lenders can either increase or decrease aggregate lending. Empirical

[1] See American Banker (2013) and American Banker (2012) for examples of other niche lenders.

papers find that entry by foreign lenders sometimes reducing access to credit (Beck and Peria, 2010, Detragiache, Tressel and Gupta, 2008, Gormley, 2010) and sometimes increasing access to credit (Giannetti and Ongena, 2009, 2012, Bruno and Hauswald, 2013, Claessens and Van Horen, 2014). A related literature examines the impact of increased competition caused by interstate banking deregulation within the United States (see, e.g., Black and Strahan (2002) and Cetorelli and Strahan (2006)). Related to industry expertise, Karakaya, Michalski and Örs (2022) focuses on broader sector specialization within manufacturing uses interstate banking deregulation to identify the impact of entry by banks into new states. Lending grows in manufacturing sectors that the entering bank was more familiar with based on the industry composition of banks' home state. Others examine new competition from Fintech mortgage lenders (Buchak et al., 2018, Fuster et al., 2019) or peer-to-peer lenders (Tang, 2019, De Roure, Pelizzon and Thakor, 2019, Wolfe and Yoo, 2018, Jagtiani and Lemieux, 2017). Relative to these papers, we document the growth of a new type of lender (remote, industry-specialized small business lenders) and find that lenders specializing in certain sectors or industries can serve as complements to existing, local SBA small business lenders.

2 Setting and Data

Our analysis examines industry specialization within the market for Small Business Administration (SBA) 7(a) loans. The 7(a) program provides guarantees for small business loans. It is the SBA's largest funding program and is an important source of credit for small businesses. In 2017, SBA 7(a) originated more than 60,000 loans totaling $25.45 billion, which makes up 10% of SBA lending reported in the Community Reinvestment Act.[2] These SBA loans likely make up a larger share among employer small businesses, i.e. those with employees, and in certain industries where SBA lending is common. In the Small Business Credit Survey (Federal Reserve Banks, 2016-2019), 22-26% of employer small businesses seeking a loan or line of credit applied for an SBA loan. Of those that already held loans and did not apply in the last year, 17% held an SBA loan or line of credit.

To be eligible for a 7(a) loan, the borrower must run a for-profit small business that meets SBA

[2]These loan amounts are not directly comparable, as CRA data do not include loans for more than $1 million while SBA 7(a) statistics do and the CRA only collects information from banks with assets over $1 billion. These larger institutions represent 70% of all outstanding small business loans made by banks (Haynes and Williams, 2018). In the CRA, small business loans are defined as those with original amounts of $1 million or less and were reported on the institution's Call Report or Thrift Financial Report as either "Loans secured by nonfarm or nonresidential real estate" or "Commercial and industrial loans."

industry-specific size standards. The program targets credit-constrained businesses. Lenders must satisfy the "credit elsewhere" requirement by documenting why the borrower could not obtain a loan on reasonable terms without the SBA guarantee, and must also review the personal resources of any applicants owning more than 20% of the small business. The SBA-guaranteed loans can be used for working capital, expansions, to purchase a business or franchise, to buy commercial real estate, or to refinance debt.

Private lenders provide the capital for 7(a) loans. These lenders are mostly commercial banks, though there are also credit unions and other non-bank lenders. The private lenders make most decisions regarding the SBA loans subject to underwriting rules of the SBA such as a maximum interest rate and borrower requirements. The SBA provides the lender with a partial guarantee for the loan of up to 75-85%, depending on the loan size. In exchange, the lenders pay the SBA a fee that depends on the features of the loan and the amount guaranteed.

Although the loans are guaranteed, screening is still important. The SBA program serves a group of less creditworthy borrowers who could not obtain a loan on other terms, the guarantees are only partial, and the SBA monitors portfolio performance. The SBA can revoke Preferred Lender status for poor risk management or seek payment even for the guaranteed portion if a charge-off is attributable to technical deficiencies of the lender. Indeed, for the bank we examine in our empirical strategy, loan delinquencies, credit losses, and the possibility of repercussions from the SBA are consistently listed as the first risk factors in its annual report (Live Oak Bancshares, Annual Report, 2016, 2018). As evidence of screening, Federal Reserve Banks (2016-2019) show that approval rates for SBA loans are similar or slightly lower, on average, than the approval rates when small businesses apply for non-SBA personal or business loans. Additionally, DeYoung, Glennon and Nigro (2008), DeYoung et al. (2011), and Huang (2020) provide empirical evidence of the importance of credit-screening, default, and information asymmetries in lending through the SBA program.

Our main analysis uses data from the SBA Loan Data Report on all originated 7(a) loans between 2001 and 2017.[3] A key advantage of the data is that they contain the small businesses' industry and location, as well as each loan's amount, term, repayment status, and (starting in 2008) interest rate. We group businesses into industries by their 5-digit NAICS code, and businesses from

[3]We drop loans that were approved but canceled before origination.

more than 800 distinct industries obtain a 7(a) loan during our sample period. Using the lender identity and borrower location, we calculate the distance between each borrower and the closest branch of the institution making the loan.[4] To do so, we fuzzy match the SBA lender to bank branch networks in the Federal Deposit Insurance Corporation's (FDIC) Summary of Deposits. We match 92% of loans to branch networks, primarily missing loans from credit unions or other non-bank lenders, since only banks are included in the FDIC data. We then geocode the borrowers' addresses, matching 72%, and calculate the distance between the borrower and the closest branch of the lending institution. Although we only have the exact addresses geocoded for 72% of borrowers, all of our results using distance are robust to calculating distance using the borrower's county centroid, which is available for the full sample of bank loans. Internet Appendix B provides more details on the procedure for calculating distance.

3 Motivating Evidence

3.1 The Growth of Industry Specialization

We begin by documenting the existence and rise of remote, industry-specialized lenders within the SBA program. Figure 1 plots the relationship between each SBA lending institution's (log) median borrower-lender distance against its top-five industry share, defined as the share of the institution's loans extended to its five most common industries.[5] Two facts are evident. First, in all periods, there is a positive relationship between distant lending and industry concentration, reflecting a trade-off between geographic specialization and industry specialization. In the 2013-2017 period, institutions with a median borrower-lender distance less than 10 miles had an average top-five share of 23%, while lenders with a median borrower-lender distance of more than 100 miles had an average top-five share of 40%.[6] Second, the three periods in Figure 1 reveal an increasing number of

[4]Since lending decisions and monitoring may not be done at the local branch, one may want to measure distance between the borrower and the location where underwriting decisions are made, but these are not observed in the data. Granja, Leuz and Rajan (2018) also uses the location to the nearest branch, and finds that this measure is correlated with information and credit risk.

[5]Let S_{ijt} be the share of institution j's loans to industry i during period t. The top-five share for institution j during period t is the sum of its largest five S_{ijt} shares. Since we want to capture specialization, we drop the industry "limited-service restaurants" when calculating top-five share since that is the most common SBA industry and makes up 9.5% of all SBA loans. Among the other industries, none make up more than 2.2% of SBA loans.

[6]In Internet Appendix Tables A.1 and A.2, we show the positive relationship between distance and concentration is statistically significant and robust to additional controls and measures of distance. Additionally, to partially address the concern that this may be related to the SBA guarantee, we show that the relationship between distance and

institutions with a high degree of both distant lending and industry concentration. To highlight this growth, we classify a lender as a remote specialist if its median borrower-lender distance exceeds 100 miles and its top-five industry share exceeds 32% (the 90^{th} percentile during the 2001-2006 period) and mark these institutions as solid circles in Figure 1.[7]

Figure 2 shows the annual number of remote specialists (as defined above) and their share of total SBA lending between 2001 and 2017. The number of lenders classified as remote specialists increased from less than 10 to more than 40 over this period. Additionally, remote industry specialists make up a larger share of SBA lending, from less than 1.6% in 2001 to 17.4% in 2017. These graphs show a steady increase in industry-specialized lending that accelerates after 2012. A natural question is whether these trends in specialization are unique to SBA lending. Data containing loan counts by both the lender and industry are not available for most non-SBA small business lending, so we cannot examine this directly. Recently, however, Blickle, Parlatore and Saunders (2021) finds that industry-specialized C&I lending exists even among large banks subject to stress testing. Additionally, Karen Mills, former Administrator of the Small Business Administration, emphasizes that specialization by lenders in specific industries is a key innovation of emerging small business lenders within and outside of the SBA program (Mills, 2019*a*), and trade publications have also highlighted the general rise of niche or specialty lending.[8]

3.2 Potential Benefits of Industry Specialization

What advantages do industry-specialized lenders have over local local lenders for identifying profitable or low-risk borrowers? First, industry-specialized lenders can select industries with lower risks or less competitive markets. In Section 3, we show that Live Oak Bank, the subject of our case study, enters industries with low charge-off rates and industries where there is a weaker relationship between distant lending and loan performance. We also examine this for all specialized lenders in Internet Appendix C, which characterizes the specialists identified in Figure 1 and their chosen industries. We find that lenders specialize in a variety of industries, but they tend to have below-average charge-off rates. Second, industry specialization may facilitate expertise that offset the informational disadvantages of distant lending or help lenders appeal to new borrowers. For

lending is similar for loans with a low ($\leq 50\%$) or high ($> 50\%$) SBA guarantee (Internet Appendix Figure A.1).

[7] The qualitative patterns are not affected by using alternative thresholds.

[8] See American Banker (2013) and American Banker (2012) for examples of other niche lenders.

example, specialized lenders hire industry experts and develop industry-specific underwriting guidelines and performance metrics. In Internet Appendix C.2, we examine how the loan performance of specialized lender compares to other lenders. Consistent with expertise, we find that specialized lenders experience better loan performance than other lenders within the same industries. To provide a sense of the magnitude, these estimates imply that an industry share of 52% would offset the additional risk from lending to a borrower 100 miles away. The offsetting threshold increases with borrower-lender distance. This relationship between concentration and the probability of default remains similar when adding several geographic and loan controls.

These advantages of specialization are consistent with what specialized lenders themselves identify as their primary advantages: industry-specific expertise, practices, and investments. Live Oak Bank, which we examine in our empirical strategy, states, "We are one of the nation's top originators of small business loans primarily because our expertise in specific industries enables us to lend to business owners who haven't had access to capital in the past" (Live Oak Bank, n.d.). In particular, Live Oak develops expertise by hiring industry experts prior to lending. United Community Bank, another specialized SBA lender, reports that it mitigates the risk of "working with more borrowers it doesn't know well" by "originating SBA loans only within specific industries it has decided to cultivate after studying them carefully" (Schneider, 2016). Additionally, specialists use industry-specific underwriting criteria or collateral assessment to better evaluate credit risks. For example, specialists describe how general lenders do not understand the cash flow issues unique to specific industries,[9] or the off-balance-sheet assets (e.g., medical records, goodwill) unique to independent pharmacies and veterinary and dental practices.[10] As a result, local lenders may overestimate the risk of certain industries because they apply uniform, general underwriting criteria.

Finally, a specialist's industry-specific focus allows them to engage in industry-specific marketing (e.g. trade shows), build industry-specific networks (e.g. hiring industry insiders), and offer tailored advice in a way that is not feasible for general lenders. These unique potential advantages of industry-specialized lending may allow them to either identify new, profitable borrowers or to

[9] Concerning Live Oak's lending to Registered Investment Advisors (RIAs), "[O]ne of Live Oak's biggest advantages is that it understands the RIA industry and many banks don't ... A lot of lenders are uncomfortable with the RIA industry ... They don't understand this is a business without a lot of cash flow," states Jamie Carvallo, co-founder of Park Sutton Advisors LLC, quoted in Shidler (2013).

[10] First Financial Bank states, "Commercial banks are asset based lenders, and when it comes to a veterinary practice, the largest asset is usually an off-balance sheet asset – Patient files, Goodwill, etc. An SBA loan can be collateralized in different ways to make it possible to acquire the loan." (First Financial Bank, 2018*b*)

offer better rates and products to existing ones. The trade-off is that industry-specialized lenders must make more distant loans, making it difficult to collect soft information and to monitor businesses. Additionally, specialists are more at risk to industry-specific shocks, while geographically concentrated lenders are exposed to area-specific shocks. As stated by Live Oak Bank, "the risk associated with industry concentration is mitigated by the geographical diversity of the overall loan portfolio" (Live Oak Bancshares, Annual Report, 2016).

4 Empirical Strategy

Our goal is to examine the impact of remote, industry-specialized lenders on the availability of small business credit within the SBA program. Do specialized lenders complement or substitute for existing, primarily local lenders? What is their impact on the total amount of lending? As in the theoretical models of Detragiache, Tressel and Gupta (2008) and Gormley (2014), cream-skimming by new lenders could cause unraveling and a decline in credit availability. We investigate these questions by examining a case study of the largest remote, specialized SBA lender: Live Oak Bank. This case study of a prominent, remote specialized lender (and the largest SBA lender by volume) is of interest in its own right. Moreover, the size and entry strategy of Live Oak Bank provide a unique setting to estimate the competitive impact of entry by a remote lender, and provide evidence about the potential for specialized lending to complement existing lenders. In an extension, we adopt a similar strategy to assess whether the results of this case study generalize to other specialized SBA lenders.

4.1 Background: Live Oak Bank

Live Oak Bank was founded in 2007 as a niche lender, at focused exclusively on SBA lending, at first to veterinary practices but soon expanding to other industries. Our strategy will exploit Live Oak's staggered entry into these industries, which generates a sudden increase in industry-specialized lending, to assess its impact on lending markets. Live Oak operates almost exclusively in the market for government-guaranteed loans, predominately within the SBA 7(a) program and, to a lesser extent, with loans guaranteed by the U.S. Department of Agriculture.

As seen in Figure 1, Live Oak exhibits the two key features of remote, industry-specialized lenders. Live Oak gave 95% of its SBA loans to borrowers more than 100 miles from its single headquarters in North Carolina and 80% of its loans went to just six industries. Table 1 reports Live Oak's main industries (those with at least 50 loans), as well as Live Oak's post-entry share of SBA loans (number and dollar amount) in that industry as of 2017 and the month that Live Oak first originated a loan to the industry.

Our analysis examines entry into the six industries where Live Oak has given the most loans (as of 2017): veterinarians, dentists, investment advice establishments, pharmacies, broilers, and funeral homes. Live Oak's share of the total volume in dollars, at around 50%, is even greater. Live Oak's combination of size, industry concentration, and staggered entry generates sharp increases in total lending to these industries. When Live Oak enters, it provides a significant share of subsequent lending, ranging from 12% of SBA loans to offices of dentists to 58% of SBA loans to investment advice establishments. We exclude Live Oak's loans in its remaining industries because it either entered in mid-2015, so there is a short post-period, or because the loans made up only a small share of lending to that industry and so are unlikely to have had a measurable impact.

4.2 Sample Construction: Treatment and Control Industries

We use data from the SBA 7(a) Loan Data Report to construct annual counts of approved SBA 7(a) loans by industry (5-digit NAICS code) from 2001-2017.[11] We begin in 2001 because, prior to 2001, many 7(a) loans are missing the industry code. Of the initial 835 5-digit NAICS industries receiving SBA loans, we drop the industries where Live Oak has given a small number of loans (i.e. those not among the six primary Live Oak industries). Thus, the control industries face no competition from Live Oak. To ensure consistency in industry definitions, we also drop industries that had a change in their 5-digit NAICS code between 1997 and 2012, leaving 466 industries. Finally, we retain only the industries that have at least one SBA 7(a) loan approved for each year between 2001 and 2017. We also require the industries to average at least 20 loans per year during the period 2001-2006, so that the donor pool is similar in size to the industries that Live Oak enters. The final sample consists of a balanced panel from 2001-2017 of annual loan originations for 219 control industries and the six treated industries that Live Oak has entered. This forms the main sample for our analysis.

[11] We drop canceled loans and loans given to borrowers in the U.S. territories.

4.3 Synthetic Control Method

To estimate the effect of Live Oak's entry, we compare the path of total lending in the six entered (treated) industries to a comparison group of other control industries. For the comparison group, we use the synthetic control method (Abadie and Gardeazabal, 2003, Abadie, Diamond and Hainmueller, 2010) to construct a synthetic match for each treated industry. The synthetic match is a weighted combination of the control industries where the weights are chosen to best match the pretreatment lending trajectory of the treated industry.

Our setting is well suited for the synthetic control method. First, the synthetic control method requires large treatment shocks because the small number of treated units makes it difficult to distinguish small treatment effects from other idiosyncratic shocks (Abadie, 2021). Table 1 shows that, after entry, Live Oak originated 12-58% of SBA loans to these industries. Outside of Dentists, Live Oak's loans amount to a shock of one to three times the pretreatment standard deviation of lending within the treated industries. Second, the industries that Live Oak did not enter, provide a natural comparison group. These are loans to other small businesses that meet the SBA's requirements and, as discussed in the last section, we limit the donor pool of control industries to those that receive at least 20 loans annually during the pretreatment period.

Formally, consider a panel of I industries over T years with industry 1 as the single treated industry, which Live Oak enters in year $T_0 + 1$. Our outcome Y_{it} is the annual number of new SBA loans to industry i in year t, divided by the loans to industry i in 2006.[12] This normalization converts all outcomes to percentage changes relative to 2006, which allows us to compare growth in industries of different sizes. We choose 2006 as the base year because it is the year before Live Oak began lending.

Let Y_{it} be the observed SBA loan originations to industry i in year t and, using potential outcomes notation, let $Y_{1t}(1)$ and $Y_{1t}(0)$ be the potential loan originations to industry 1 during year t with and without treatment (Live Oak's entry). Our goal is to estimate the causal effect of entry on lending to industry 1, $\tau_{1t} = Y_{1t}(1) - Y_{1t}(0) = Y_{1t} - Y_{1t}(0)$ for periods $t > T_0$. We only observe $Y_{1t}(1)$ for the treated industry during the post-treatment period, so estimating the treatment effect requires an estimate of the counterfactual number of loans that would have been

[12]We report results using with the normalized loan counts as the outcome, but the results are similar when using unnormalized loan counts or loan volume in dollars.

given out if Live Oak had not entered, i.e., $Y_{1t}(0)$.

To estimate this counterfactual, we assume that the potential outcome under no treatment for all industries i follows the factor model

$$Y_{it}(0) = \delta_t + \lambda_t \mu_i + \varepsilon_{it} \tag{1}$$

where δ_t is a time fixed effect, λ_t is a vector of unobserved common factors, μ_i is a vector of unknown factor loadings, and ε_{it} is an unobserved, industry-level transitory shock with zero mean. As Abadie, Diamond and Hainmueller (2010) shows, if there is a set of weights $(w^*_{2t}, \ldots, w^*_{It})$, with $w^*_{it} \geq 0$ and $\sum_i w^*_{it} = 1$, such that a weighted combination of the outcomes of control industries equals the outcome of the treated industry for all pretreatment periods,

$$\sum_{i=2}^{I} w_i^* Y_{i1} = Y_{11}, \quad \sum_{i=2}^{I} w_i^* Y_{i2} = Y_{12}, \quad \ldots, \quad \sum_{i=2}^{I} w_i^* Y_{iT_0} = Y_{1T_0}, \tag{2}$$

then $\hat{\tau}_{1t} = Y_{1t} - \sum_{i=2}^{I} w_i^* Y_{it}$ for $t > T_0$, provides an asymptotically unbiased as the number of pretreatment periods grows. In practice, there is not a set of weights such that equations in (2) will hold exactly, so the estimation procedure chooses weights such that the equation holds approximately by solving the following optimization problem for each treated industry j, with treatment occurring in period $T_0^j + 1$:

$$\begin{aligned}
\{w_i^{j*}\}_{j \in \text{Treated}} = \underset{\{w_i^j\}_{i \in \text{Control}}}{arg\,min} & \sum_{t \leq T_0^j} [Y_{jt} - \sum_{i \in \text{Control}} w_i^j Y_{it}]^2 \\
s.t. \sum\nolimits_{i \in \text{Control}} w_i^j & = 1 \\
\text{and} \qquad w_i^j & \geq 0 \quad \forall i.
\end{aligned}$$

The weights $w_i^{j^*}$ minimize the pretreatment mean squared prediction error between the treated industry and the synthetic control.[13] With the optimal weights, the synthetic control for treated industry j is $\hat{Y}_{jt}(0) = \sum_{i \in \text{Control}} w_i^{j*} Y_{it}$. The estimated impact of Live Oak entering on the total loan volume in industry j is the overall treatment effect

$$\widehat{\tau}_{jt}^{\text{overall}} = Y_{jt} - \hat{Y}_{jt}(0).$$

This overall treatment effect reflects the combination of Live Oak's new lending and the com-

[13]In matching, we include all pretreatment outcomes Y_{jt} from the pre-treatment period as covariates and use the default procedure of "synth" in Stata, which uses a regression-based approach to obtain variable weights in the V-matrix of Abadie, Diamond and Hainmueller (2010). As discussed in detail in Kaul et al. (2015), this is equivalent to the minimization procedure above.

petitive effect of Live Oak's entry on other lenders

$$\underbrace{\widehat{\tau}_{jt}^{\text{overall}}}_{\text{Overall Effect}} = \underbrace{Y_{jt}^{\text{Live Oak}}}_{\text{Live Oak Lending}} + \underbrace{\widehat{\tau}_{jt}^{\text{comp}}}_{\text{Competitive Effect}} . \tag{3}$$

The direct effect of Live Oak's additional lending, $Y_{jt}^{\text{Live Oak}}$, is the number of loans that Live Oak originated to industry j in year t (normalized by total lending to industry j in 2006). These loans may have crowded out other SBA lenders, in which case the competitive effect $\widehat{\tau}_{jt}^{\text{comp}} < 0$. Alternatively, if Live Oak primarily complements existing lenders, $\widehat{\tau}_{jt}^{\text{comp}} \approx 0$. Following equation (3), we estimate $\widehat{\tau}_{jt}^{\text{comp}}$ as the difference between the overall treatment effect and Live Oak lending: $\widehat{\tau}_{jt}^{\text{overall}} - Y_{it}^{\text{Live Oak}}$.[14] Thus, we estimate both the overall effect on lending and the competitive effect on other lenders.

4.4 Identification

In this section, we discuss and evaluate the assumptions necessary to identify the treatment effects of Live Oak's entry. Our model representing the impact of entry on lending can be written as[15]

$$Y_{it} = \alpha_i + \delta_t + \sum_{m=0}^{M} \tau_{it} z_{i,t-m} + C_{it} + \epsilon_{it}$$

where α_i are industry fixed effects, δ_t are time fixed effects, and τ_{it} represents the industry-time-specific treatment effect of Live Oak's entry into industry i. The indicators $z_{i,t-m}$ equal one if industry i is treated as of period $t-m$. Our setting is a standard case of staggered adoption, in which the treatment z_{it} is binary and absorbing, i.e., $z_{it'} \geq z_{it}$ for all i and $t' \geq t$. The term ϵ_{it} is an industry-year shock that is uncorrelated with Live Oak's entry. Our goal is to estimate the treatment effects for the group of treated industries, i.e., τ_{it} for $i \in$ Treated.

The term C_{it} represents unobserved confounds that are correlated with entry decisions. For example, C_{it} would reflect scenarios where treated industries experience above-average growth or respond differently to the business cycle. When ignored, these confounding trends would lead to biased estimates of the treatment effect. With no restrictions on the confound C_{it}, any time-path of lending after entry could be explained by some pattern of confound shocks. Thus, to identify the

[14] This estimator is identical to estimating $\widehat{\tau}_{jt}^{\text{comp}}$ with a synthetic control on the outcome of normalized aggregate loan counts that exclude Live Oak's loans. The equivalence holds because the synthetic control is chosen using only on pretreatment observations and so is not affected by dropping Live Oak's lending, all of which occurs post-treatment.

[15] This section follows the notation and specification of Freyaldenhoven et al. (2021), which discusses many of these issues in a general setting.

treatment effects τ_{it} we must place restrictions on the confound C_{it}. The identifying assumption in our synthetic control strategy is that the confounds follow the structure

$$C_{it} = \lambda_t \mu_i$$

where λ_t is a vector of unobserved common factors and μ_i is a vector of unknown factor loadings. This structure allows, for example, each industry to have a different response (μ_i) to the aggregate economy (λ_t). It also accounts for the possibility that Live Oak enters industries based on their pre-existing trends, or that it enters industries less affected by the business cycle. Our assumption is more flexible than that of a difference-in-difference specification, which would impose that $C_{it} = 0$ or, by including linear trends, that $C_{it} = \mu_i t$.[16] However, our estimates would still be biased if the common factors do not fully control for confounds in the treated industries. For example, our assumption would be violated if Live Oak enters industries when they anticipate abnormal future growth that deviates from the factor model.

We investigate the possible violations of our identification assumption in several ways. First, we examine Live Oak's public documents and interviews about how they select the entered industries. Second, under the stronger identifying assumption of parallel trends across all industries, we diagnose potential bias by examining the pretrends of the treated and control industries. Third, our strategy exploits the exact timing of entry by Live Oak. We argue that the role of omitted variables are likely to be small relative to this large, discrete entry event. Finally, we include several additional robustness checks to address specific concerns.

Determinants of Entry

To investigate the validity of our identifying assumption, we examine the stated determinants of Live Oak's entry decisions in its annual reports, interviews, and articles. The bank's stated determinants of entry are historical repayment performance, the level of competition, and its ability to develop industry expertise through research and hiring experts.[17] The bank analyzes historical SBA data

[16] The synthetic control can account for unit-specific fixed effects when, in the pretreatment period, the weighted average of the outcomes for the synthetic control units exactly equal the average outcome for the treated unit (Arkhangelsky et al., 2019). In practice, this holds only approximately.

[17] "Our Emerging Markets group identifies new verticals by methodically analyzing payment records, level of competition, and most importantly, conducts a relentless search for a Domain Expert that not only understands the industry but also is a fit with our unique culture." (Live Oak Bancshares, Annual Report, 2018). Also, see Streeter (n.d.) and Bank To Bank (2016).

and payment records to select industries. Characteristics such as average industry risk are fixed within an industry and so are captured by industry fixed effects. Other components of risk may vary over time with historical trends or with macroeconomic shocks (e.g. cyclicality), and so are captured by the industry-specific factor loadings and time-varying factors. Thus, we control for the historical characteristics and trends that are the primary determinants of entry decisions. In its annual reports, interviews, and publicly available information, there is no indication that Live Oak chooses industries (verticals) based on temporary shocks or expectations of sudden growth in the industries. Rather than responding to short-term fluctuations, entry decisions are based on long-term trends or fixed characteristics and require industry-specific investments (e.g. hiring experts, developing expertise) that generate fixed costs of entry. This is consistent with our identification assumption, which assumes that, conditional on being in the treatment group, the *exact timing of entry* is not correlated with systematic deviations from the factor model.

Diagnosing Bias from Pretrends

We also diagnose the potential bias from such shocks by graphically examining pretrends among the treated industries under the stronger assumption from difference-in-differences that $C_{it} = 0$, i.e. there are no confounders. To do so, we estimate the standard event study model

$$Y_{it} = \alpha_i + \delta_t + \sum_{m=-10, m\neq -1}^{5} \beta_m \Delta z_{i,t-m} + \beta_{6+} z_{i,t-6} + \beta_{-11+}(1 - z_{i,t+10}) + \epsilon_{it}. \tag{4}$$

where Δ denotes the first difference operator. Note that $\Delta z_{i,t-m}$ is an indicator for whether industry i was treated exactly m periods before t, $z_{i,t-6}$ is an indicator for whether it was treated 6 or more periods before t, and $(1 - z_{i,t+10})$ is an indicator for whether it was treated more than 10 periods after t.[18] The parameters β_m can be interpreted as the cumulative treatment effects at different horizons for an event occurring at $m = 0$. To better detect short-term trends, we group the data into semi-annual bins so that time t reflects a six-month period when estimating the model.

Figure 3(a) reports the β_m coefficients from the event study, along with both pointwise confidence intervals and 95% simultaneous, sup-t confidence bands (Montiel Olea and Plagborg-Møller, 2019), which cover the entire parameter vector with 95% probability. The event study is estimated

[18] These binned endpoints are needed so that the omitted period consists only of $m = -1$ (see, e.g. Freyaldenhoven et al. (2021)).

on the main analysis sample, consisting of all treatment and control industries from 2001-2017. Figure 3(a) also plots the linear pretrend over the five years prior to entry.[19] The figure reveals stable coefficients leading up to entry in period 0, indicating that the treated and control industries were trending similarly up to the point of Live Oak's entry. Then, in the 1.5 years after entry, there is a 50 percentage point increase in lending to the treated industries.

A separate issue is that the dynamic treatment effects may be heterogeneous across industries, which complicates the interpretation of the estimates in specification (4). Panel (b) implements the doubly-robust event-time estimator of Callaway and Sant'Anna (2020), which provides estimates of a well-defined average treatment effect in a staggered adoption setting such as ours. Again, the estimates show smaller and stable coefficients prior to treatment then a sharp increase in the first 1.5 years after Live Oak's entry. Note that the synthetic control also addresses the potential for heterogeneous treatment effects by estimating the separate treatment effects $\widehat{\tau}_{jt}$ for each treated industry j and post-treatment period t, thereby avoiding issues present when estimating an average treatment effect in a two-way fixed effects model with staggered adoption (e.g. Callaway and Sant'Anna (2020), De Chaisemartin and d'Haultfoeuille (2020), Goodman-Bacon (2018)).

As seen in Figure 3, Live Oak's entry generates a sudden, roughly 50 percentage point increase in lending (relative to 2006 baseline) in the first 1.5 years after entry. Increases of a large magnitude are consistent with the summary statistics in Table 1, in which Live Oak makes roughly 30-50% of SBA loans to these industries after entry. Moreover, lending rises immediately after Live Oak's entry, then levels off in the subsequent periods. If confounding variables were to explain the pattern of coefficients, they must follow a similar pattern around the exact timing of Live Oak's entry. Given that Live Oak Bank does not report entering industries in response to immediate, short-term shocks, we think it is unlikely that confounders would follow this pattern around entry.

[19] The linear pretrend (β) in event time is estimated from the equation, as in Freyaldenhoven et al. (2021),

$$Y_{it} = \alpha_i + \delta_t + \beta \cdot r + \sum_{m=0, m\neq -1}^{5} \beta_m \Delta z_{i,t-m} + \beta_{6+} z_{i,t-6} + \beta_{-11+}(1 - z_{i,t+10}) + \epsilon_{it}.$$

where r equals event time m when $-10 \leq m \leq 0$ and $r = 0$ otherwise. The pretrend β is then extrapolated into the post-period.

Potential Bias from Spillovers

The identification assumption also requires no spillovers, meaning that Live Oak's entry into a treated industry does not affect the control industries. Given that general lenders often make loans to dozens or hundreds of industries, we expect that Live Oak's entry into a single industry is unlikely to have significant spillover effects on overall lending practices. If there are spillover effects, however, we think it is most likely that lenders may divert resources away from the industries Live Oak enters and into other non-treated industries. This would increase lending to the control industries, relative to the treated industry, potentially leading to a downward bias in our estimates of Live Oak's overall effect on lending and the competitive effect. Such a bias would reinforce the results we find in Section 5. There is, however, a concern that other remote lenders follow Live Oak into the treated industries in a way that is not captured by the trends allowed in the factor model. We address this concern directly by reporting results from the main sample and also from an alternative sample that excludes other remote loans (those with a distance more than 100 miles) from the industry loan counts.

4.5 Inference

To evaluate the statistical significance of the results, we use the permutation inference procedure of Abadie (2021) and Abadie and L'hour (2020). The procedure relies on the distribution of placebo treatment effects obtained by estimating a separate synthetic control for each of the control industries. Each industry is observed for periods $t = 1, \dots, T$, and let T_0^j be the last pretreatment period for treated industry j. For each treated industry j of the J treated industries, we estimate a synthetic control for each control industry i by assigning it the treatment timing $T_0^j + 1$. Let $J+1, \dots, J+I+1$ index the I control industries. $\widehat{Y}_{it}^j(0)$ is the predicted lending in period t produced by the placebo synthetic control for industry i when it is assigned treatment time T_0^j.

We summarize the treatment effects during the post-period with two test statistics. The first test statistic is the average treatment effect in the first three years of the post-treatment period:

$$\tau_i^j = \frac{1}{3} \sum_{t=T_0^j+1}^{T_0^j+3} (Y_{it} - \widehat{Y}_{it}^j(0)).$$

One concern with this test statistic is that some of the placebo synthetic controls may have a poor pretreatment fit, making the estimated placebo treatment effects less credible. For this reason, Abadie, Diamond and Hainmueller (2010) and Abadie (2021) suggest also using the ratio of the post-treatment fit to the pretreatment fit as another test statistic, where fit is measured by the root mean squared prediction error (RMSPE):

$$r_i^j = \frac{\left(\frac{1}{3}\sum_{t=T_0^j+1}^{T_0^j+3}\left(Y_{it}-\widehat{Y}_{it}^j(0)\right)^2\right)^{1/2}}{\left(\frac{1}{T_0^j}\sum_{t=1}^{T_0^j}\left(Y_{it}-\widehat{Y}_{it}^j(0)\right)^2\right)^{1/2}}.$$

The p-value based on the permutation distribution of each test statistic t (either τ or r) for treated industry j is

$$p_j(t) = \frac{1}{I+1}\left(1 + \sum_{i=J+1}^{J+I+1} \mathbb{1}\left(t_i^j \geq t_j^j\right)\right).$$

Finally, we also conduct joint inference using the method of Abadie and L'hour (2020), which extends the permutation methods to cases with multiple treated units. Let the true treated units be $D^{(0)} = \{1, \ldots, J\}$ and assign this group to iteration $b = 0$. We then form B random samples of J control industries $D^{(b)} = \{i_1^{(b)}, \ldots, i_J^{(b)}\}$ with control industry $i_j^{(b)}$ assigned the treatment timing of treated industry j. For each iteration $b = 1, \ldots, B$ and each $j \in \{1, \ldots, J\}$, we first compute the placebo treatment effect $\widehat{T}_j^{(b)} = \tau^j_{i_j^{(b)}}$ (and, in a separate procedure, $\widehat{T}_j^{(b)} = r^j_{i_j^{(b)}}$). Then, we calculate the ranks $R_1^{(0)}, \ldots, R_J^{(0)}, \ldots, R_1^{(B)}, \ldots, R_J^{(B)}$ associated with the absolute values of the $J \times (B+1)$ treatment effects $\widehat{T}_1^{(0)}, \ldots, \widehat{T}_J^{(0)}, \ldots, \widehat{T}_1^{(B)}, \ldots, \widehat{T}_J^{(B)}$. Using these rankings, we calculate the sum of ranks for each permutation $SR^{(b)} = \sum_{i=1}^{J} R_i^{(b)}$. The joint p-value is

$$p = \frac{1}{B+1}\sum_{b=0}^{B} \mathbb{1}\left(SR^{(b)} \geq SR^{(0)}\right).$$

5 Results

5.1 Main Results

Figure 4 plots the paths of each treated industry and its synthetic control. Internet Appendix Table A.3 shows the donor pool industries that make up the synthetic controls. In most cases, the synthetic control closely approximates the trajectory of lending during the pretreatment period prior to Live Oak's entry. However, the fit of the synthetic control is not equally good across all industries.

In particular, the MSPE for Broilers is 0.33, which is 16 times larger with the second largest MSPE. When there is no good pretreatment fit, synthetic controls are asymptotically biased and Abadie, Diamond and Hainmueller (2010) and Abadie (2021) recommend against using synthetic controls in such cases. For this reason, we report joint summary statistics that both include and exclude Broilers.

Turning to the post-period, the gaps between the treated industry and the synthetic control in Figure 4 indicate large increases in total lending upon Live Oak's entry. Each treated industry increases, often sharply, relative to the synthetic control. For most industries, lending to the synthetic control remains relatively flat while lending to the treated industry increases sharply. For Veterinarians, lending to the synthetic control declines sharply, as lending to many industries did at the start of the recession, while lending to the treated industry remains stable. Thus, Live Oak's lending caused Veterinarians to avoid the declines in lending present in other similar industries. In all cases, lending in the treated industry rises relative to the synthetic control as Live Oak's entry generated increases in total SBA lending to these industries.

Figure 5 plots these estimated overall treatment effects $\hat{\tau}_{jt}^{\text{overall}}$. A potential concern is that other remote lenders may have entered the same industries after Live Oak, which would generate increases in lending that we would mistakenly attribute to Live Oak's entry. To address this concern, Figure 5 also plots the treatment effects estimated from annual loan counts that exclude non-Live-Oak remote loans.[20] Except for "Broilers," all industries demonstrate a good pretreatment fit and a sharp growth in overall lending upon Live Oak's entry. As seen in the figure, the treatment effects are similar when non-Live-Oak remote loans are excluded, indicating that the growth in total lending is largely due to Live Oak's entry and not due to subsequent entry by other remote lenders.

The overall treatment effect estimates reflect the combination of Live Oak's new lending and the competitive effect of Live Oak's entry on other lenders. As discussed in Section 4, we can isolate the competitive effect by subtracting Live Oak's loans from the overall effect $\widehat{\tau}_{jt}^{\text{comp}} = \widehat{\tau}_{jt}^{\text{overall}} - Y_{it}^{\text{Live Oak}}$. Figure 6 plots the estimates of the competitive effect from synthetic controls using the full sample and the subsample excluding remote loans. The competitive effect is generally close to zero (or slightly positive), indicating that, upon Live Oak's entry, other SBA lenders continued lending

[20]Remote loans are those with a borrower-lender distance above 100 miles, with distance computed using county centroids. The small share of loans missing the county measure of borrower-lender distance (largely from credit unions and nonbanks) are also dropped.

similar amounts to the treated industries. There is no evidence of substitution away from other SBA lenders, suggesting that Live Oak's loans were given to borrowers who would not have otherwise received an SBA loan.

5.2 Statistical Significance

To assess the statistical significance of these treatment effects, we use the permutation-based inference outlined in Section 4.5. Figure 7 shows the distribution of these placebo treatment effects for each of the treated industries, along with the actual treatment effect estimate from the full sample in black.[21] In the figure, the treatment effects for the actually treated industries are not only positive, but also large relative to the distribution of placebo effects. Two test statistics for treated industry j – τ_j^j, the average treatment effect during the first three post-treatment years, and r_j^j, the ratio of the post- to pre-treatment RMSPEs – provide a formal comparison in Table 2. Across the six treated industries, the average treatment effect in Panel A column 1 is a 21-112 percentage point increase in annual lending, with most two-sided p-values significant at conventional levels (column 2).[22] When using r as the test statistic in Panel A columns 3 and 4, four of the treated industries have p-values below 0.1. The two that do not, Funeral Homes and Broilers, are those with the worst pretreatment fit (quantified Internet Appendix Table A.4) which explains their relatively low value for r_j. Across both test statistics, the joint inference p-values for overall significance are also highly significant and are similar when excluding Broilers, which has a poor pretreatment fit.

Despite these large increases in lending by Live Oak, Panel A columns 5-8 confirm that the competitive effects on other lenders are small, mostly positive, and statistically insignificant individually (except for r_j for Veterinarians) and jointly insignificant.[23] Panel B, which excludes loans from other remote lenders, largely corroborates the results of Panel A. The results in Panel B columns 3 and 4 are generally smaller and often have a p-value above 0.1. The joint test statistic,

[21] Following Abadie, Diamond and Hainmueller (2010), the plot excludes placebo industries with a poor pretreatment fit, i.e., a pretreatment MSPE more than 20 times that of the average MSPE among the treated industries.

[22] The exception is Dentists, with a p-value of 0.19. That the impact on Dentists is smaller is expected because Live Oak made up only 12% of the post-entry loans in that industry, while it made up at least 30% in the other treated industries.

[23] As seen in Figure 6, the significance for Veterinarians reflects that lending to Veterinarians became more volatile but, as seen in the τ statistic in column 5, experienced only a 1% fall in the average growth over the first three periods.

however, demonstrates that the effect sizes remain large relative to the permutation distribution, with a p-value less than 0.01. Overall, Table 2 shows that Live Oak's entry generated large and statistically significant increases in overall SBA lending with no indication of substitution away from existing lenders.

5.3 Substitution Within and Outside of SBA Lending

Additional Evidence on Substitution Within SBA Lending

The zero competitive effect indicates that Live Oak did not substitute for existing SBA lenders. To further investigate substitution within SBA lending, we check whether Live Oak's borrowers have previously obtained an SBA loan from another institution. At the time the Live Oak borrowers in our six industries obtain their first Live Oak loan, only 2.9% had a previous SBA loan from another institution in our 2001-2017 sample. For comparison, 13.8% of other SBA borrowers who originated a loan in 2014 had a previous SBA loan.[24] Of those with a previous loan, the size of their Live Oak loan exceeded the amount of their previous loan by an average of $813,000 (median $750,000). Thus, upon entry, Live Oak lends largely to new SBA borrowers and, in the few cases where a borrower has obtained a previous SBA loan, Live Oak originates large loans that may not have been approved by other SBA institutions.

Substitution From Non-SBA Lending

One possibility is that Live Oak caused substitution away from non-SBA lending. Institutional features, external evidence, and indirect evidence using a proxy for total lending, however, all suggest that Live Oak causing substitution between SBA and non-SBA lending is likely limited. First, the "credit elsewhere" test of the SBA 7(a) loan program requires SBA lenders to certify that the borrower could not obtain a loan on reasonable terms without an SBA guarantee. This credit elsewhere test does seem to be enforced, and lenders often refer borrowers to the SBA program after they fail to qualify for a conventional loan.[25] Moreover, as argued in Bachas, Kim and Yannelis

[24] We chose 2014 as the comparison year because it is the median year for Live Oak's loans.

[25] Temkin (2008) surveyed 23 banks that originate SBA loans about their application of the "credit elsewhere" requirement, and the surveys suggest that "the lenders are aware of the credit elsewhere requirement and adhere to the requirement." Lender representatives report that most SBA applicants are referred to the program if (i) the business shows insufficient net operating income to obtain a conventional loan, (ii) the collateral is limited, or (iii) the borrower does not have sufficient equity for the down payment.

(2021), lenders specialized in SBA lending, such as Live Oak, are most likely to comply with the credit elsewhere test as they would face the largest costs from violations, which could lead to exclusion from the SBA program. In addition to the requirement limiting substitution, other SBA loans are likely the closest substitutes with regards to loan features, collateral requirements, and loan durations.[26] Given that we find no substitution within the SBA program, it is likely that substitution from commercial lending outside of SBA lending is also limited. External evidence also suggests that SBA-guaranteed lending increases the supply of credit rather than substituting for non-SBA alternatives. Bachas, Kim and Yannelis (2021) examine heterogeneity in the elasticity of SBA lending with respect to the guarantee rate across areas, finding estimates consistent with limited substitution between SBA and non-SBA lending. Additionally, Brown and Earle (2017) finds that SBA lending leads to increases in employment, which would not occur if SBA loans simply crowded out non-SBA alternatives.

We also examine the impact of Live Oak's entry on a proxy for total industry lending from The Risk Management Association's (RMA) eStatement Studies. Financial institutions provide the RMA with financial statements collected from commercial borrowers or applicants, and the RMA collects information from hundreds of financial institutions including nine of the ten largest banks (Lisowsky, Minnis and Sutherland, 2017). The RMA's eStatement Studies publishes counts of the number of financial statements collected by industry, which provides and industry-specific proxy for total (SBA and non-SBA) lending activity. Small business borrowers provide financial statements (e.g. tax returns, income statements, balance sheets) as a part of the loan application and monitoring process, so these counts provide a proxy for lending in that industry.[27] Berger, Minnis and Sutherland (2017) shows a strong correlation between these financial statements and the size of bank's commercial and industrial lending portfolio. Live Oak is not a participant in the RMA survey during our sample period, so the RMA data provide a proxy for total industry lending excluding Live Oak, i.e., the competitive effect. Applying our same strategy to industry-year counts of financial statements in Internet Appendix D, we find no statistically significant declines in financial statements in the treated industries. The p-values indicate that more than 50% of the placebo industries experienced larger declines in lending. Overall, while we cannot directly examine

[26]Live Oak's 2017 Annual Report states that "[i]f we lose our status as a Preferred Lender, we may lose some or all of our customers to lenders who are SBA Preferred Lenders."

[27]See, for example, `https://www.nerdwallet.com/article/small-business/how-to-qualify-for-small-business-loans`.

non-SBA lending, the institutional features, external evidence, and the indirect test using financial statements all suggest that Live Oak's substitution from non-SBA lending is limited and unlikely to fully offset the observed growth in SBA lending within the treated industries.

Other Sources of Financing and Real Effects

It is possible that Live Oak's entry generated substitution away from non-commercial sources of financing, such as equity financing from friends and family or seller financing. In several of Live Oak's industries (pharmacies, dentists, veterinarians, funeral homes), commercial loans were generally unavailable for acquisitions. Much of the value of these businesses is in goodwill, which makes for poor collateral in the event of default, and some buyers had little wealth (e.g. new dentists and veterinarians with student loans). As a result, acquisitions were historically financed by the seller without a commercial loan from a bank. Live Oak and other specialized lenders have likely generated some substitution away from these seller-financed loans. Additionally, it is possible that Live Oak generated substitution away from non-bank lending, such as private equity or specialty non-bank lenders.

The increase in commercial financing through the SBA program could generate a measurable impact on employment or establishment counts in these industries. SBA lending provides nearly 20% of loans to employer small businesses (Federal Reserve Banks, 2016-2019), and perhaps a larger share in the treated industries, where SBA lending is more common.[28] Internet Appendix E and Table E.1 and uses the synthetic control approach to examine the impact of Live Oak's entry on employment and charge-off rates. There is some evidence of an increase in employment and establishment for investment advice agencies (p-values less than 0.1), which is the industry where Live Oak's impact was largest. Overall, the changes in employment and establishment counts are jointly insignificant. The lack of significant real effects is not surprising, as even significant increase in total SBA lending may not have a measurable impact on employment or establishments. Brown and Earle (2017) finds that SBA lending increased employment by only 3-3.5 jobs per million dollars in lending. Effect sizes of this magnitude would not be apparent in national employment counts. Additionally, many SBA loans are used for the purchase of existing practices or expansions

[28]First Financial Bank, and SBA lender, reports that SBA lending is the most common form of lending to independently owned veterinarian practices and pharmacies (First Financial Bank, 2018*a*,*b*).

of existing businesses, which would not change the establishment counts.

5.4 Extensions and Robustness

Sensitivity to Changes in Design

We summarize results from several diagnostic checks suggested by Abadie (2021) that examine the sensitivity of our results to the design of the synthetic control. We find that the results are robust to backdating the treatment timing by 1-3 years (Internet Appendix Figure A.2), dropping individual industries from the donor pool (Internet Appendix Figure A.3), and using alternative choices of predictors (average pre-treatment charge-off rates number of loans) for the synthetic control (Internet Appendix Figure A.4).[29] We also compare our synthetic control method with that of a simple difference-in-differences in Internet Appendix Figure A.5. Even with this simple comparison group, it is evident that total lending in each treated industry increases upon Live Oak's entry, although the parallel trends restriction fails to hold in the pretreatment period. This gives further support to the synthetic control strategy, which improves upon this simple average by selecting a weighted average of industries that better match the pretreatment lending path of each treated industry. To make the comparison between the two strategies precise, Internet Appendix Table A.4 shows that the MSPE from the simple average is 2.6 to more than 7,500 times larger than that of the synthetic control.

Other Specialized Lenders and External Validity

Given that this case study focuses on a single lender, a natural question is whether the results extend to other remote specialists. In this section, we extend the analysis to provide suggestive evidence about the other remote specialists identified in Figure 1. While Live Oak's staggered entry and lending volume make it uniquely suited for the synthetic control analysis, we can estimate the average impact of a broader set of remote, specialized lenders on total lending.

For industry j in year t, we estimate the following specification:

$$Loans_{jt} = \beta_0 + \beta_1 SpecLoans_{jt} + \delta_j + \tau_t + \epsilon_{jt}. \tag{5}$$

[29] One concern is that the gap in lending between the treated and control industries due to heterogeneity the cyclicality of the industries, especially during the Great Recession. These robustness checks also help address this concern about cyclicality, since they show that (i) the results are robust to selecting industries of similar risks (i.e., matching on charge-off rates) and (ii) the differences emerge only after Live Oak enters.

The outcome $Loans_{jt}$ is the total number of new SBA loans originated to industry j during year t, and the explanatory variable $SpecLoans_{jt}$ is the total number of new SBA loans originated to industry j during year t by remote, specialized lenders. The parameter of interest is β_1, which captures the impact of an increase in specialized lending on total lending. For example, if $\beta_1 \approx 0$, an additional lending by a specialized lender in an industry j does not alter the total number of loans to industry j, implying that specialized lending substitutes for other SBA lending. Alternatively, if $\beta_1 \approx 1$, it indicates that specialized lending complements other SBA lending and increases the total quantity of SBA loans. The primary concern with this exercise is that specialized lending may be correlated with ϵ_{jt}. For example, specialized lenders may enter industries that are growing quickly or trending differently. Although we allow for industry-specific linear trends in some specifications, residual correlation between $SpecLoans_{jt}$ and ϵ_{jt} would lead to biased estimates of β_1. Thus, unlike our synthetic control analysis of Live Oak Bank, we view this exercise as providing only suggestive evidence of the impact of other remote specialized lenders on total SBA lending.

Applying the definition of remote, specialized lenders from Section 3, we define a loan as a specialized loan if (i) it is from a lender with a median lending distance greater than 100 miles, a top-five industry share above 32%, and at least 50 total SBA loans and (ii) the loan is to an and industry in which the lender originates at least 10% of its SBA loans (all measured during the period 2013-2017).[30] Table 3 reports the estimates, with Panel A including industry and year fixed effects and Panel B adding industry-specific linear trends. As a benchmark, column 1 estimates the impact of Live Oak loans on total lending. Consistent with the lack of substitution found in the main analysis, the estimates in column 1 are close to one, indicating an additional loan from Live Oak increases total lending by roughly one loan. Although the estimate exceeds one in Panel A, it falls below one in Panel B, reflecting some sensitivity to the controls for industry-specific linear trends. Column 2 broadens the explanatory variable to include loans from all specialized lenders including Live Oak. Column 3 excludes all Live Oak loans from the outcome and explanatory variable. To avoid biased caused by other remote lenders entering the same industry, Column 4 drops Live Oak loans and all non-specialized remote loans (borrower-lender distance of more than 100 miles) from the counts forming the dependent variable. Estimates in all columns are close to one, indicating an increase in lending with little substitution, and generally statistically

[30] Internet Appendix Table A.5 lists the lenders and industries that are classified as specialized lenders.

significant. They remain similar when allowing for industry-specific linear trends in Panel B. This provides suggestive evidence that the estimates from Live Oak's entry may also be applicable to other remote specialized lenders.

6 Mechanisms and Loan Characteristics

The case study of Live Oak shows that entry by an industry-specialized lender can increase SBA lending, and we find no evidence of substitution away from other lenders. What explains the growth in lending and does industry specialization play a role? As discussed in Section 3, remote, specialized lenders themselves cite industry specialization as the key feature of their business model and a central factor allowing them to lend remotely and screen borrowers. Additionally, we find evidence consistent with industry expertise; specialized lenders experience better loan performance than other lenders within the same industries (Internet Appendix Table C.2).

In this section, we expand on this by investigating specific mechanisms that may explain the increase in lending caused by Live Oak Bank. We find that Live Oak Bank targeted a subset of industries with low charge-off rates and where the relationship between distance and charge-off rates is weak. This strategy of selecting safe industries that are well-suited for distant lending is only available to industry specialized lenders. We also find, consistent with industry expertise, Live Oak maintains similar charge-off rates to those of other lenders within these industries, despite significantly increasing total originations and lending distances. Live Oak also tends to originating larger, longer loans at lower interest rates than those offered by other lenders.

6.1 Industry Selection

One advantage available to industry specialists (compared to local lenders) is that they can target a subset of industries that are safer, better suited for distant lending, or less competitive. To investigate this, Table 4 compares the charge-off rates and interest rates in Live Oak's industries to those in other industries. The variable of interest is "LO industry," an indicator for whether the loan was originated in one of Live Oak's six industries. Importantly, the sample excludes all loans from Live Oak, so the estimates reflect are not confounded by Live Oak's lending. As seen in column 1, Live Oak enters safer industries. Live Oak's industries have three-year charge-off rates

0.69 percentage points lower than other industries, and this difference remains significant when loan-level controls for loan size and term are added in column 2. Columns 3 and 4 show that distant loans are also safer in these industries, with the interaction term LO industry $\times\, log(dist)$ nearly offsetting the positive relationship between distance and charge-off rates. Viewing the relationship between distance and charge-offs as a proxy for the importance of soft information in lending decisions, these results are consistent with Live Oak entering industries where soft information is less important.

Although these industries have lower charge-off rates, columns 5-8 show that the lower risk is not reflected in the interest rates charged by other lenders. Indeed, columns 7 and 8 show that interest rates rise more rapidly with distance in these industries, even though columns 3 and 4 show that charge-off rates rise more slowly. Thus, Live Oak entered industries that were lower risk, but where other lenders were not pricing this lower risk into interest rates.

6.2 Loan Performance, Characteristics, and Geography

Another advantage is that specialization may facilitate industry expertise. The evidence from industry-specialized lenders reported in Section 3 suggests this plays an important role. To investigate whether Live Oak's loans are consistent with industry expertise, we examine within-industry loan performance and characteristics. This analysis focuses on the years immediately after Live Oak's entry, and one may expect industry expertise to develop over time. Live Oak, however, credits its expertise to hiring an industry expert and developing an understanding of the industry before they enter, so the industry expertise would be available immediately upon entry.

Table 5 columns 1-4 investigate within-industry charge-off rates for Live Oak and other lenders in the sample of loans to the six treated industries.[31] The variable of interest is "Live Oak loan," an indicator for whether Live Oak originated the loan. Columns 1 (with industry fixed effects) and 2 (adding loan-level controls) show that Live Oak experiences similar charge-off rates to other lenders in these industries, despite significant increases and the number of borrowers and borrower-lender distances. Columns 3 and 4 add controls for the log of borrower-lender distance and its

[31] To focus on similar loans, we restrict the sample to loans for more than \$100,000 (in 2010 dollars) because 96% of Live Oak's loans are above \$100,000. Because the maximum SBA guarantee threshold changes at \$150,000, we also estimate these regressions on the sample of loans over \$150,000. We find similar estimates in Tables A.8 and A.9, except the relationship between distance and charge-offs is weaker (though still significant in Table A.9). We also repeat the regressions with county fixed effects in tables A.10 and A.11, and the results are very similar to those in the main table.

interaction with the indicator for Live Oak loans. For other lenders in these industries, there is the standard positive relationship between distance and charge-off rates. Live Oak, however, exhibits no significant relationship between distance and charge-off rates; the small positive coefficient on $log(dist)$ is completely offset by the interaction term Live Oak loan $\times \log(dist)$.[32] Online Appendix Figure A.6 reports results from semi-parametric versions of these regressions, and finds a similar relationship between distance and charge-off rates. These regressions show that Live Oak finds new, low-risk borrowers and maintains similar charge-off rates to those of other lenders, despite significantly increasing total originations.

We also compare interest rates in Table 5 columns 5-8. Live Oak's interest rates are 12.6 basis points lower than those of other lenders (column 5), or 6.9 basis points lower after controlling for loan size and term. Columns 7 and 8 reveal that these differences in interest rates are driven by distant loans. Other lenders increase rates by around 5 basis points for every 100 log point increase in distance, while Live Oak's interest rates do not vary with distance. In addition to these interest rate differences, Live Oak tends to originate larger, longer-term loans than those offered by other lenders. Since 2008, Live Oak's average loan size was $1.08 million (2010 dollars), compared with an average loan size of $459,000 for other lenders in the treated industries. Live Oak's average term was 209 months, compared with an average term of 149 months for other lenders. Together, these results suggest that the increase in total lending may be driven, in part, by Live Oak originating larger, longer loans at lower interest rates than those offered by other lenders.

In addition to providing different types of loans, Live Oak may expand lending if it originates distant loans in locations underserved by existing lenders. We find limited evidence for this channel. Live Oak's borrowers are not located farther from physical branches of SBA lenders than borrowers from local banks. The borrower-lender distance distributions of local and remote borrowers are very similar (Internet Appendix Figure A.7). Indeed, 99% of remote SBA borrowers are within 10 miles of a branch of a bank that grants SBA loans. Brown and Earle (2017) has found that having a high-volume, i.e., a Preferred Lenders Program (PLP) SBA lender within the county increases a business's access to the SBA program. We find that among borrowers in the six treated industries, Live Oak's borrowers were slightly more likely to have a branch of a PLP lender in their county.

[32] The results are similar if we calculate distances based on county centroids, which is available for all bank loans in the sample (Online Appendix Table A.6).

Thus, we do not find evidence that physical distance to an SBA lender explains the growth in lending after Live Oak's entry.

7 Conclusion

Remote, industry specialization offers a very different approach than the local, industry-diverse lending that has historically characterized small business finance. This paper documents recent growth in industry-specialized lenders, which grew from 2% of SBA originations to 17% in 2017. We then examine effects of entry by the largest of these remote, specialized lenders with the SBA program, Live Oak Bank. Upon Live Oak's entry into specific industries, total SBA lending increases sharply with no evidence of declines from other lenders. Examining specialization, we show that Live Oak (and other specialized lenders) target safer industries and, consistent with industry expertise, it experiences better loan performance within those industries, consistent with industry expertise. This setting demonstrates that the remote, industry-specific lending strategy has the potential to deepen commercial credit markets.

While our focus is within the SBA program, specialized lending is increasingly prominent outside of this setting. Industry experts and trade publications have highlighted the emergence of specialized or "vertical" small business lenders (Mills, 2019*a*, American Banker, 2013, 2012) and similar specialization exists among larger commercial lenders (Blickle, Parlatore and Saunders, 2021). Additional research is needed to understand the broader impact of industry-specialized lenders outside of the market for SBA-guaranteed loans. There are also implications for the broader economy. Growth in specialized lending may lead to changes in labor markets, entrepreneurship, and banking outcomes. If industry specialization increases the supply of loans to certain industries, it may alter the industrial composition of small businesses. Already, Live Oak Bank and other remote lenders have altered the industry composition of SBA 7(a) lending. For banking and risk management, specialized lenders are less exposed to regional economic downturns but more exposed to industry-specific risks, which affects credit risk and risk-sharing across the economy. To understand the trajectory of specialized lending and its potential scope, we need to know what makes certain industries or markets suitable for specialized lending. We leave these issues for future research.

References

Abadie, A, and J L'hour. 2020. "A penalized synthetic control estimator for disaggregated data."

Abadie, Alberto. 2021. "Using synthetic controls: Feasibility, data requirements, and methodological aspects." *Journal of Economic Literature*, 59(2): 391–425.

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2010. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association*, 105(490): 493–505. https://doi.org/10.1198%2Fjasa.2009.ap08746.

Abadie, Alberto, and Javier Gardeazabal. 2003. "The Economic Costs of Conflict: A Case Study of the Basque Country." *American Economic Review*, 93(1): 113–132. https://doi.org/10.1257/000282803321455188.

Acharya, Viral V, Iftekhar Hasan, and Anthony Saunders. 2006. "Should Banks Be Diversified? Evidence from Individual Bank Loan Portfolios." *The Journal of Business*, 79(3): 1355–1412. https://doi.org/10.1086/500679.

Agarwal, Sumit, and Robert Hauswald. 2010. "Distance and Private Information in Lending." *The Review of Financial Studies*, 23(7): 2757–2788. https://doi.org/10.1093/rfs/hhq001.

American Banker. 2012. "Specialty Lending Takes Off." *https://www.americanbanker.com/slideshow/specialty-lending-takes-off*, Accessed: 2019-09-01.

American Banker. 2013. "Niche Lending Gains Allure." *https://www.americanbanker.com/slideshow/niche-lending-gains-allure*, Accessed: 2019-09-01.

Arkhangelsky, Dmitry, Susan Athey, David A Hirshberg, Guido W Imbens, and Stefan Wager. 2019. "Synthetic difference in differences."

Bachas, Natalie, Olivia S Kim, and Constantine Yannelis. 2021. "Loan guarantees and credit supply." *Journal of Financial Economics*, 139(3): 872–894.

Bank To Bank. 2016. "Business Model: Chip Mahan, Live Oak Bank Interview."

Beck, Thorsten, and Maria Soledad Martinez Peria. 2010. "Foreign Bank Participation and Outreach: Evidence from Mexico." *Journal of Financial Intermediation*, 19(1): 52–73. https://doi.org/10.1016/j.jfi.2009.03.002.

Beck, Thorsten, Olivier De Jonghe, et al. 2013. "Lending Concentration, Bank Performance and Systemic Risk: Exploring Cross-country Variation." The World Bank.

Berger, Allen N, and Gregory F Udell. 1995. "Relationship Lending and Lines of Credit in Small Firm Finance." *Journal of Business*, 351–381. https://doi.org/10.1086/296668.

Berger, Philip G, Michael Minnis, and Andrew Sutherland. 2017. "Commercial Lending Concentration and Bank Expertise: Evidence from Borrower Financial Statements." *Journal of Accounting and Economics*, 64(2-3): 253–277. https://doi.org/10.1016/j.jacceco.2017.06.005.

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

Blickle, Kristian, Cecilia Parlatore, and Anthony Saunders. 2021. "Specialization in Banking." *FRB of New York Staff Report*, , (967).

Boeve, Rolf, Klaus Duellmann, and Andreas Pfingsten. 2010. "Do Specialization Benefits Outweigh Concentration Risks in Credit Portfolios of German Banks?"

Brown, J David, and John S Earle. 2017. "Finance and Growth at the Firm Level: Evidence from Sba Loans." *The Journal of Finance*, 72(3): 1039–1080. https://doi.org/10.1111/jofi.12492.

Bruno, Valentina, and Robert Hauswald. 2013. "The Real Effect of Foreign Banks." *Review of Finance*, 18(5): 1683–1716. https://doi.org/10.1093/rof/rft041.

Buchak, Greg, Gregor Matvos, Tomasz Piskorski, and Amit Seru. 2018. "Fintech, Regulatory Arbitrage, and the Rise of Shadow Banks." *Journal of Financial Economics*, 130(3): 453–483. https://doi.org/10.1016/j.jfineco.2018.03.011.

Callaway, Brantly, and Pedro HC Sant'Anna. 2020. "Difference-in-differences with multiple time periods." *Journal of Econometrics*.

Cetorelli, Nicola, and Philip E Strahan. 2006. "Finance as a barrier to entry: Bank competition and industry structure in local US markets." *The Journal of Finance*, 61(1): 437–461.

Claessens, Stijn, and Neeltje Van Horen. 2014. "Foreign Banks: Trends and Impact." *Journal of Money, Credit and Banking*, 46(s1): 295–326. https://doi.org/10.1111/jmcb.12092.

De Chaisemartin, Clément, and Xavier d'Haultfoeuille. 2020. "Two-way fixed effects estimators with heterogeneous treatment effects." *American Economic Review*, 110(9): 2964–96.

De Roure, Calebe, Loriana Pelizzon, and Anjan V Thakor. 2019. "P2p Lenders Versus Banks: Cream Skimming or Bottom Fishing?" https://doi.org/10.2139/ssrn.3174632.

Detragiache, Enrica, Thierry Tressel, and Poonam Gupta. 2008. "Foreign Banks in Poor Countries: Theory and Evidence." *The Journal of Finance*, 63(5): 2123–2160. https://doi.org/10.1111/j.1540-6261.2008.01392.x.

DeYoung, Robert, Dennis Glennon, and Peter Nigro. 2008. "Borrower-lender Distance, Credit Scoring, and Loan Performance: Evidence from Informational-opaque Small Business Borrowers." *Journal of Financial Intermediation*, 17(1): 113–143. https://doi.org/10.1016/j.jfi.2007.07.002.

DeYoung, Robert, W Scott Frame, Dennis Glennon, and Peter Nigro. 2011. "The Information Revolution and Small Business Lending: The Missing Evidence." *Journal of Financial Services Research*, 39(1-2): 19–33. https://doi.org/10.1007/s10693-010-0087-2.

Federal Reserve Banks. 2016-2019. "2016-2019 Small Business Credit Survey: Report on Employer Firms."

First Financial Bank. 2018*a*. "Small Business Administration (SBA) Lending 101: Pharmacy Edition."

First Financial Bank. 2018*b*. "Small Business Administration (SBA) Lending 101: Veterinary Edition."

Freyaldenhoven, Simon, Christian Hansen, Jorge Pérez Pérez, and Jesse M Shapiro. 2021. "Visualization, Identification, and Estimation in the Linear Panel Event-Study Design." National Bureau of Economic Research Working Paper 29170.

Fuster, Andreas, Matthew Plosser, Philipp Schnabl, and James Vickery. 2019. "The Role of Technology in Mortgage Lending." *The Review of Financial Studies*, 32(5): 1854–1899. https://doi.org/10.1093/rfs/hhz018.

Giannetti, Mariassunta, and Steven Ongena. 2009. "Financial Integration and Firm Performance: Evidence from Foreign Bank Entry in Emerging Markets." *Review of Finance*, 13(2): 181–223. https://doi.org/10.1093/rof/rfm019.

Giannetti, Mariassunta, and Steven Ongena. 2012. "Lending by Example: Direct and Indirect Effects of Foreign Banks in Emerging Markets." *Journal of International Economics*, 86(1): 167–180. https://doi.org/10.1016/j.jinteco.2011.08.005.

Giometti, Marco, and Stefano Pietrosanti. 2019. "Bank Specialization and the Design of Loan Contracts."

Goodman-Bacon, Andrew. 2018. "Difference-in-differences with variation in treatment timing."

Gopal, Manasa. 2019. "How Collateral Affects Small Business Lending: The Role of Lender Specialization."

Gopal, Manasa, and Philipp Schnabl. 2020. "The rise of finance companies and fintech lenders in small business lending." *Available at SSRN 3600068*.

Gormley, Todd A. 2010. "The Impact of Foreign Bank Entry in Emerging Markets: Evidence from India." *Journal of Financial Intermediation*, 19(1): 26–51. https://doi.org/10.1016/j.jfi.2009.01.003.

Gormley, Todd A. 2014. "Costly Information, Entry, and Credit Access." *Journal of Economic Theory*, 154: 633–667. https://doi.org/10.1016/j.jet.2014.06.003.

Granja, João, Christian Leuz, and Raghuram Rajan. 2018. "Going the Extra Mile: Distant Lending and Credit Cycles." National Bureau of Economic Research. https://doi.org/10.3386/w25196.

Granja, Joao, Gregor Matvos, and Amit Seru. 2017. "Selling failed banks." *The Journal of Finance*, 72(4): 1723–1784.

Hayden, Evelyn, Daniel Porath, and Natalja v Westernhagen. 2007. "Does Diversification Improve the Performance of German Banks? Evidence from Individual Bank Loan Portfolios." *Journal of Financial Services Research*, 32(3): 123–140. https://doi.org/10.1007/s10693-007-0017-0.

Haynes, George, and Victoria Williams. 2018. "Small Business Lending in the United States, 2016." U.S. Small Business Administration Office of Advocacy.

Huang, Ruidi. 2020. "The financial consequences of customer satisfaction: Evidence from yelp ratings and SBA loans." *Available at SSRN 3064343*.

Jagtiani, Julapa, and Catharine Lemieux. 2017. "Fintech Lending: Financial Inclusion, Risk Pricing, and Alternative Information." *Working Paper (Federal Reserve Bank of Philadelphia)*. doi:10.21799/frbp.wp.2017.17.

Jahn, Nadya, Christoph Memmel, and Andreas Pfingsten. 2016. "Banks' Concentration Versus Diversification in the Loan Portfolio: New Evidence from Germany." *Schmalenbach Business Review*, 17(1): 25–48. doi:10.1007/s41464-016-0006-7.

Karakaya, Neslihan, Tomasz K Michalski, and Evren Örs. 2022. "Banking integration and growth: Role of banks' previous industry exposure." *Journal of Financial Intermediation*, 49: 100944.

Kaul, Ashok, Stefan Klößner, Gregor Pfeifer, and Manuel Schieler. 2015. "Synthetic Control Methods: Never Use All Pre-intervention Outcomes Together with Covariates."

Lisowsky, Petro, Michael Minnis, and Andrew Sutherland. 2017. "Economic Growth and Financial Statement Verification." *Journal of Accounting Research*, 55(4): 745–794. https://doi.org/10.1111/j.1475-679x.2011.00411.x.

Live Oak Bancshares, Annual Report. 2016. "Live Oak Bancshares, 2015 Annual Report." *https://investor.liveoakbank.com/static-files/3530215f-7f54-4d44-ad52-ef2c22e85ecf*.

Live Oak Bancshares, Annual Report. 2018. "Live Oak Bancshares, 2017 Annual Report." *https://investor.liveoakbank.com/static-files/417a703e-1b87-4155-b723-0a5f24df2cc7*.

Live Oak Bank. n.d.. "Our Difference is Your Advantage." *https://www.liveoakbank.com/advantages/*, Accessed: 2019-09-01.

Mills, Karen. 2019*a*. "Small-Business Banking is About to Get a Whole Lot Better." *American Banker*.

Mills, Karen G. 2019*b*. *Fintech, Small Business & the American Dream: How Technology Is Transforming Lending and Shaping a New Era of Small Business Opportunity.* Springer.

Montiel Olea, José Luis, and Mikkel Plagborg-Møller. 2019. "Simultaneous confidence bands: Theory, implementation, and an application to SVARs." *Journal of Applied Econometrics*, 34(1): 1–17.

Nguyen, Hoai-Luu Q. 2019. "Are Credit Markets Still Local? Evidence from Bank Branch Closings." *American Economic Journal: Applied Economics*. https://doi.org/10.1257/app.20170543.

Paravisini, Daniel, Veronica Rappoport, and Philipp Schnabl. 2020. "Specialization in bank lending: Evidence from exporting firms." National Bureau of Economic Research.

Petersen, Mitchell A, and Raghuram G Rajan. 1994. "The Benefits of Lending Relationships: Evidence from Small Business Data." *The Journal of Finance*, 49(1): 3–37. https://doi.org/10.1111/j.1540-6261.1994.tb04418.x.

Petersen, Mitchell A, and Raghuram G Rajan. 2002. "Does Distance Still Matter? the Information Revolution in Small Business Lending." *The Journal of Finance*, 57(6): 2533–2570. https://doi.org/10.1111/1540-6261.00505.

Schneider, Howard. 2016. “Flourishing Forte.” *Independent Banker.*

Shidler, Lisa. 2013. “Fidelity’s M&a Program Reload Looks ’game-changing’ After It Partners with a Middleman to Get Uncle Sam to Guarantee Ria Deals.” *Riabiz.*

Stomper, Alex. 2006. “A Theory of Banks’ Industry Expertise, Market Power, and Credit Risk.” *Management Science*, 52(10): 1618–1633. https://doi.org/10.1287/mnsc.1060.0559.

Streeter, Bill. n.d.. “Don’t watch fintech. Wath the “Unbanker”.” *Banking Exchange.*

Tabak, Benjamin M, Dimas M Fazio, and Daniel O Cajueiro. 2011. “The Effects of Loan Portfolio Concentration on Brazilian Banks’ Return and Risk.” *Journal of Banking & Finance*, 35(11): 3065–3076. https://doi.org/10.1016/j.jbankfin.2011.04.006.

Tang, Huan. 2019. “Peer-to-Peer Lenders Versus Banks: Substitutes or Complements?” *The Review of Financial Studies*, 32(5): 1900–1938. https://doi.org/10.1093/rfs/hhy137.

Temkin, Kenneth. 2008. “An Analysis of the Factors Lenders Use to Ensure Their SBA Borrowers Meet the Credit Elsewhere Requirement.”

Winton, Andrew. 1999. “Don’t Put All Your Eggs in One Basket? Diversification and Specialization in Lending.” *SSRN Electronic Journal.* https://doi.org/10.2139/ssrn.173615.

Wolfe, Brian, and Woongsun Yoo. 2018. “Crowding Out Banks: Credit Substitution by Peer-to-Peer Lending.” *SSRN Electronic Journal.* doi:10.2139/ssrn.3000593.

Table 1: **Live Oak's Industries**

Industry	**Live Oak Loans**	**Share of Live Oak's Loans**	**Share of SBA Loans**	**Share of SBA Volume**	**Live Oak's Entry Month**
Veterinarians	1,455	0.25	0.33	0.49	06/2007
Offices of Dentists	1,038	0.18	0.12	0.27	03/2009
Investment Advice	814	0.14	0.58	0.75	02/2013
Pharmacies	799	0.14	0.30	0.56	11/2009
Broilers	520	0.09	0.37	0.60	04/2014
Funeral Homes	311	0.05	0.28	0.41	09/2011
Self-Storage	131	0.02	0.34	0.53	05/2015
Insurance Agencies	105	0.02	0.09	0.20	11/2015
Breweries	97	0.02	0.09	0.20	04/2015
Physicians	80	0.01	0.02	0.06	09/2012
Other	378	0.07	0.01	0.03	

This table shows the industries where Live Oak Bank has approved at least 50 loans, ordered by the number of loans. Industries with less than 50 loans are classified as "Other." "Share of Live Oak's Loans" is the share of Live Oak's 2007-2017 loans going to that industry. The columns "Share of SBA Loans" and "Share of SBA Volume" show Live Oak's post-entry share of SBA loans in each industry by number and by dollar amount, respectively. "Entry Month" is the month that Live Oak first approved a loan to that industry.

Table 2: **Average Treatment Effect and Inference**

	Overall Effect				Competitive Effect			
Industry	τ	p-val.	r	p-val.	τ	p-val.	r	p-val.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Sample:	All Loans							
Veterinarians	0.30	(0.06)	387.94	(0.01)	-0.01	(0.96)	173.81	(0.02)
Pharmacies	0.45	(0.05)	117.54	(0.03)	0.18	(0.26)	46.93	(0.11)
Dentists	0.21	(0.19)	69.89	(0.06)	0.13	(0.37)	50.09	(0.10)
Funeral Homes	0.61	(0.02)	4.81	(0.58)	0.18	(0.20)	1.46	(0.85)
Investment Advice	1.12	(0.02)	45.61	(0.05)	0.10	(0.35)	6.78	(0.39)
Broilers	1.04	(0.02)	1.90	(0.84)	0.29	(0.10)	0.56	(0.99)
Joint Inference		(<0.01)		(0.02)		(0.15)		(0.29)
Joint Inf. (excl. Broilers)		(<0.01)		(<0.01)		(0.30)		(0.10)
Panel B. Sample:	Excluding Other Remote Loans							
Veterinarians	0.26	(0.06)	4.68	(0.53)	-0.04	(0.68)	1.82	(0.84)
Pharmacies	0.59	(0.01)	176.64	(<0.01)	0.32	(0.07)	94.18	(0.01)
Dentists	0.38	(0.04)	15.11	(0.16)	0.31	(0.09)	12.49	(0.17)
Funeral Homes	0.41	(0.04)	7.73	(0.18)	-0.02	(0.81)	1.49	(0.83)
Investment Advice	0.92	(0.02)	20.51	(0.02)	-0.09	(0.41)	2.93	(0.57)
Broilers	0.61	(0.02)	1.76	(0.81)	-0.14	(0.28)	0.46	(0.99)
Joint Inference		(<0.01)		(0.02)		(0.22)		(0.71)
Joint Inf. (excl. Broilers)		(<0.01)		(0.01)		(0.29)		(0.47)

This table reports estimates of the overall effect on lending and the competitive effect on lending, as well as the corresponding p-values. Panel A shows estimates for the full sample, and Panel B shows estimates from the sample dropping non-Live-Oak remote loans. The test statistic τ_j^j is the average effect during the first three post-treatment years, and r_j^j is the ratio of the post- to pretreatment root MSPEs. The bottom two rows of each panel report p-values from the joint inference procedure using $B = 5,000$ random permutations. See Section 4.5 for details on the test statistics and inference procedures.

Table 3: **Impact of Other Remote Lenders on SBA Lending**

Outcome:	All SBA Loans	All SBA Loans	All SBA Loans (excl. Live Oak)	All SBA Loans (excl. Live Oak & other remote)
	(1)	(2)	(3)	(4)
Panel A: Industry and Year Fixed Effects				
Live Oak loans	1.264***			
	(0.142)			
Spec. loans		1.114***		
		(0.233)		
Spec. loans (excl. Live Oak)			1.236*	0.865
			(0.717)	(0.549)
Observations	4,199	4,199	4,199	4,199
Panel B: Industry and Year Fixed Effects, Industry-Specific Linear Trends				
Live Oak loans	0.784***			
	(0.208)			
Spec. loans		1.095***		
		(0.342)		
Spec. loans (excl. Live Oak)			1.383**	0.951**
			(0.546)	(0.455)
Observations	4,199	4,199	4,199	4,199

Sample consists of industry-year observations for 2001-2017, restricted to industries that average at least thirty loans per year during 2001-2008. The table reports estimates from equation (5). The outcome is the total number of SBA loans for each industry-year (excluding some loan types in columns 3 and 4) and the explanatory variable is the total number of loans from Live Oak (column 1) or all remote, specialized lenders (columns 2) or remote, specialized lenders excluding Live Oak (columns 3-4) in each industry-year. Panel A includes industry and year fixed effects, and Panel B adds controls for industry-specific linear trends. Standard errors are clustered at the industry level.

Table 4: **Live Oak's Industry Selection**

Sample: Excludes Live Oak loans								
Dependent variable:	Charge-off Indicator				Interest Rate			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LO industry	-0.00690***	-0.00387***			-0.00417	0.000626		
	(0.00116)	(0.00117)			(0.00933)	(0.00925)		
$log(dist)$			0.00312***	0.00290***			0.0421***	0.0384***
			(0.000172)	(0.000178)			(0.00137)	(0.00139)
LO industry$\times log(dist)$			-0.00280***	-0.00226***			0.00593*	0.0148***
			(0.000445)	(0.000444)			(0.00354)	(0.00347)
Observations	63,492	63,492	63,492	63,492	63,492	63,492	63,492	63,492
Mean of Dep. Var.	0.00879	0.00879	0.00879	0.00879	5.691	5.691	5.691	5.691
Year FE	X	X	X	X	X	X	X	X
Loan char.		X		X		X		X
Industry FE			X	X			X	X

The sample consists of loans for amounts over $100,000 (in 2010 dollars) that were originated between 2008-2017. Live Oak's loans are dropped and loans to Live Oak industries outside of the largest six are dropped. Interest rate data are available from 2008Q4 and observations missing the interest rate are dropped. The dependent variable is either an indicator for whether the loan was charged off within three years of origination or the loan's interest rate (in percentage points). LO industry is an indicator for whether the loan was originated to one of the six Live Oak industries. Loan characteristics include the share guaranteed and dummies for ventiles of the size of the loan and the term length. Industry fixed effects are indicators for the 5-digit NAICS code.

Table 5: **Live Oak's Charge-off Rates and Interest Rates**

Sample: Loans in the six treated industries								
Dependent variable:	Charge-off Indicator				Interest Rate			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Live Oak loan	-0.000444	0.000612	0.00253	0.00443	-0.126***	-0.0694***	-0.0201	0.0486
	(0.00110)	(0.00119)	(0.00649)	(0.00655)	(0.0176)	(0.0189)	(0.103)	(0.102)
$log(dist)$			0.000346*	0.000482**			0.0492***	0.0575***
			(0.000209)	(0.000219)			(0.00332)	(0.00343)
Live Oak loan$\times log(dist)$			-0.000673	-0.000866			-0.0481***	-0.0526***
			(0.000986)	(0.000989)			(0.0156)	(0.0154)
Observations	10,368	10,368	10,368	10,368	10,368	10,368	10,368	10,368
Mean of Dep. Var	0.00222	0.00222	0.00222	0.00222	5.655	5.655	5.655	5.655
Year FE	X	X	X	X	X	X	X	X
Loan char.		X		X		X		X
Industry FE	X	X	X	X	X	X	X	X

The sample consists of loans to the six treated industries for amounts over $100,000 (in 2010 dollars) that were originated between 2008-2017. Interest rate data are available from 2008Q4 and observations missing the interest rate are dropped. The dependent variable is either an indicator for whether the loan was charged off within three years of origination or the loan's interest rate (in percentage points). Live Oak loan is an indicator for whether Live Oak originated the loan. Loan characteristics include the share guaranteed and dummies for ventiles of the size of the loan and the term length. Industry fixed effects are indicators for the 5-digit NAICS code.

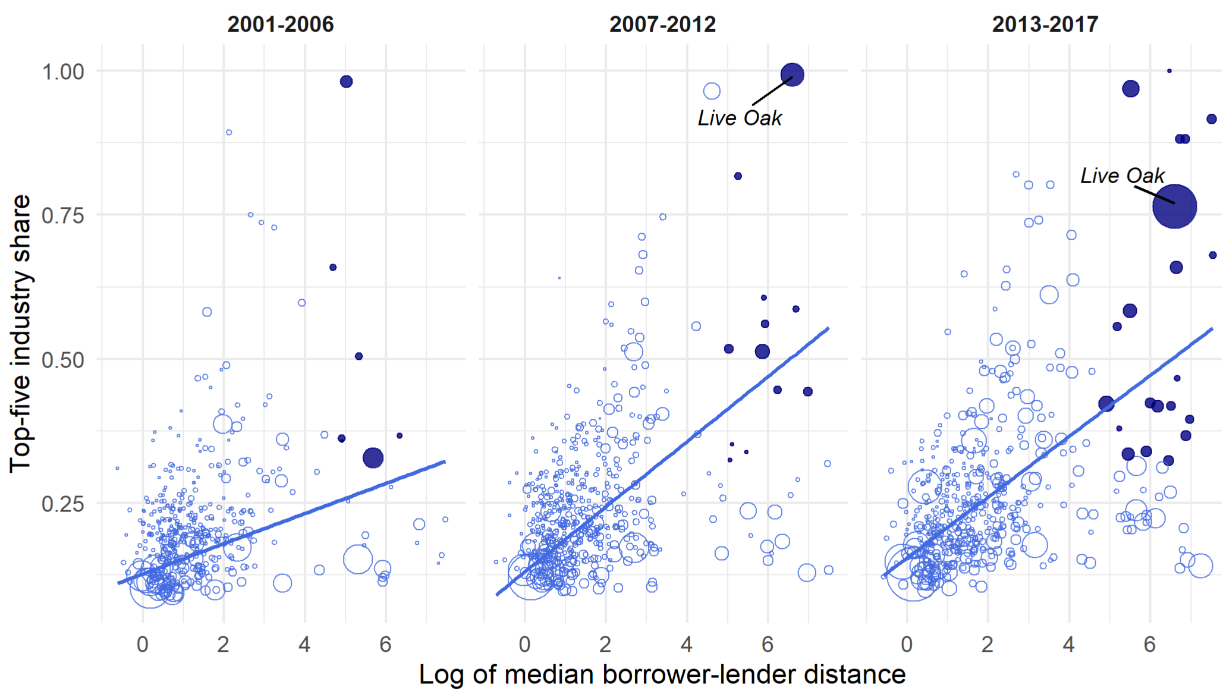


Figure 1: **SBA Lenders' Distance and Industry-Specialization**
These figures plot SBA institutions' (log) median borrower-lender distance against their top-five industry share for three periods. Each circle represents an institution and its size reflects the dollar amount of SBA loans it originated during the period. The sample is restricted to institutions originating at least 50 loans during the respective periods. The solid circles are remote, industry specialists (according to our classification in Section 3).

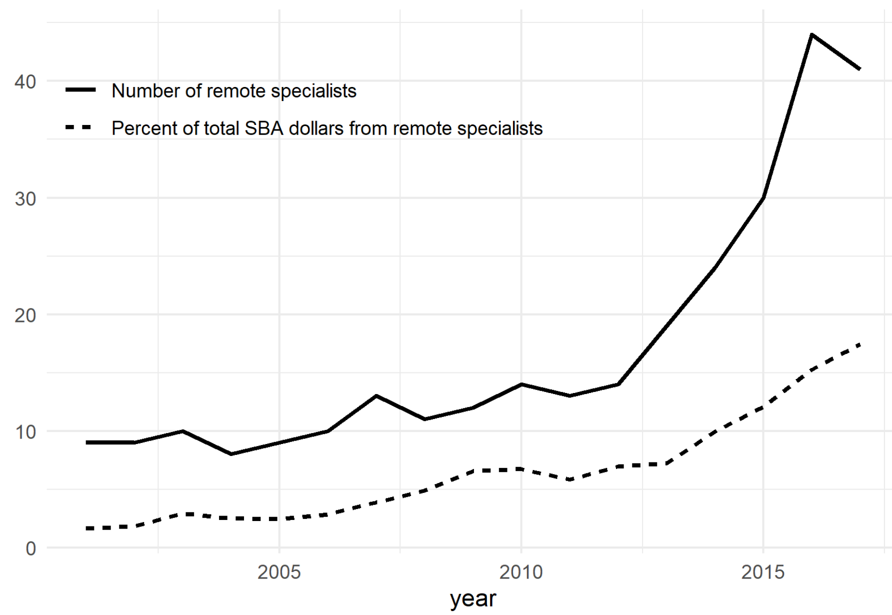
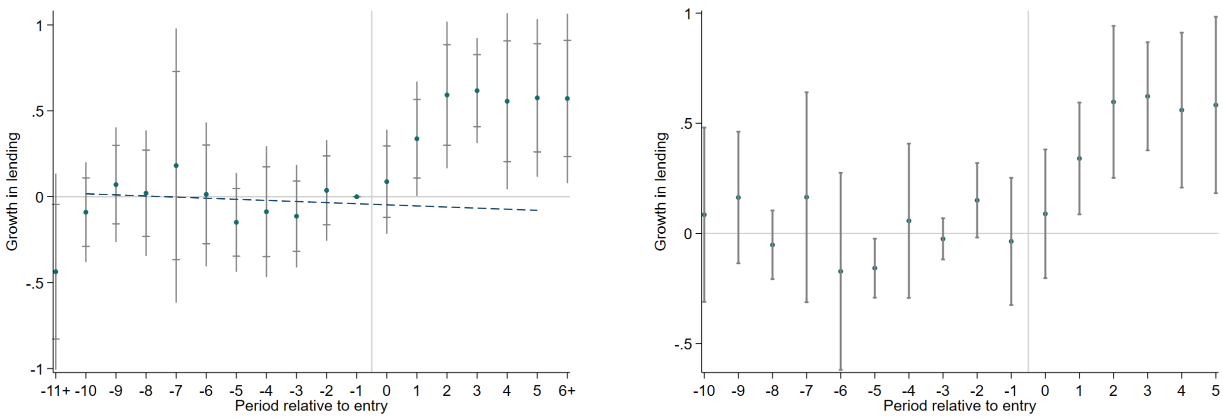


Figure 2: **Growth in Remote, Industry-Specialized Lenders**
The figure shows the number of SBA 7(a) remote, industry specialists (according to the classification in Section 3) and percent of SBA loan amounts originated by these specialists for each year from 2001-2017. We exclude institutions that originated fewer than 10 SBA loans in a year.



(a)

(b)

Figure 3: **Event study estimates**
The event study is estimated on the sample of all treatment and control industries in the main analysis sample from 2001-2017 (Section 4.2). Panel (a) reports estimates of β_m from specification (4), along with pointwise 95% confidence intervals (inner bars) and the 95% simultaneous, sup-t confidence bans of Montiel Olea and Plagborg-Møller (2019) (outer lines). The dashed blue line shows the linear pretrend over the five years prior to entry (see text for details). Panel (b) reports estimates of the dynamic average treatment on the treated using the estimator of Callaway and Sant'Anna (2020), along with their 95% simultaneous confidence bands.

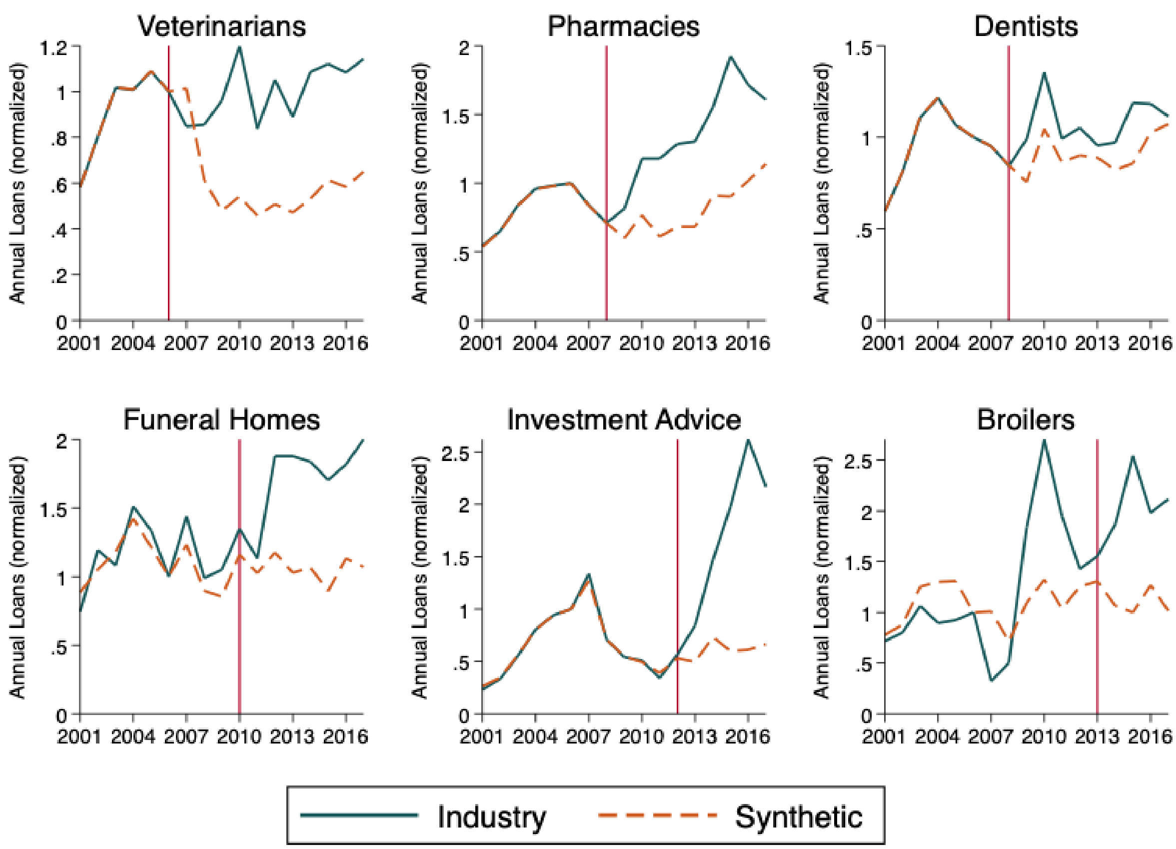


Figure 4: **Annual Lending: Treated vs. Synthetic Control**
This figure shows the growth in the annual number of SBA 7(a) loans in each industry (with 2006 loans normalized to one) for the treated industries and the synthetic control. The synthetic controls are formed by matching on all pretreatment years beginning in 2001, with no additional covariates. The vertical line shows the year before Live Oak entered.

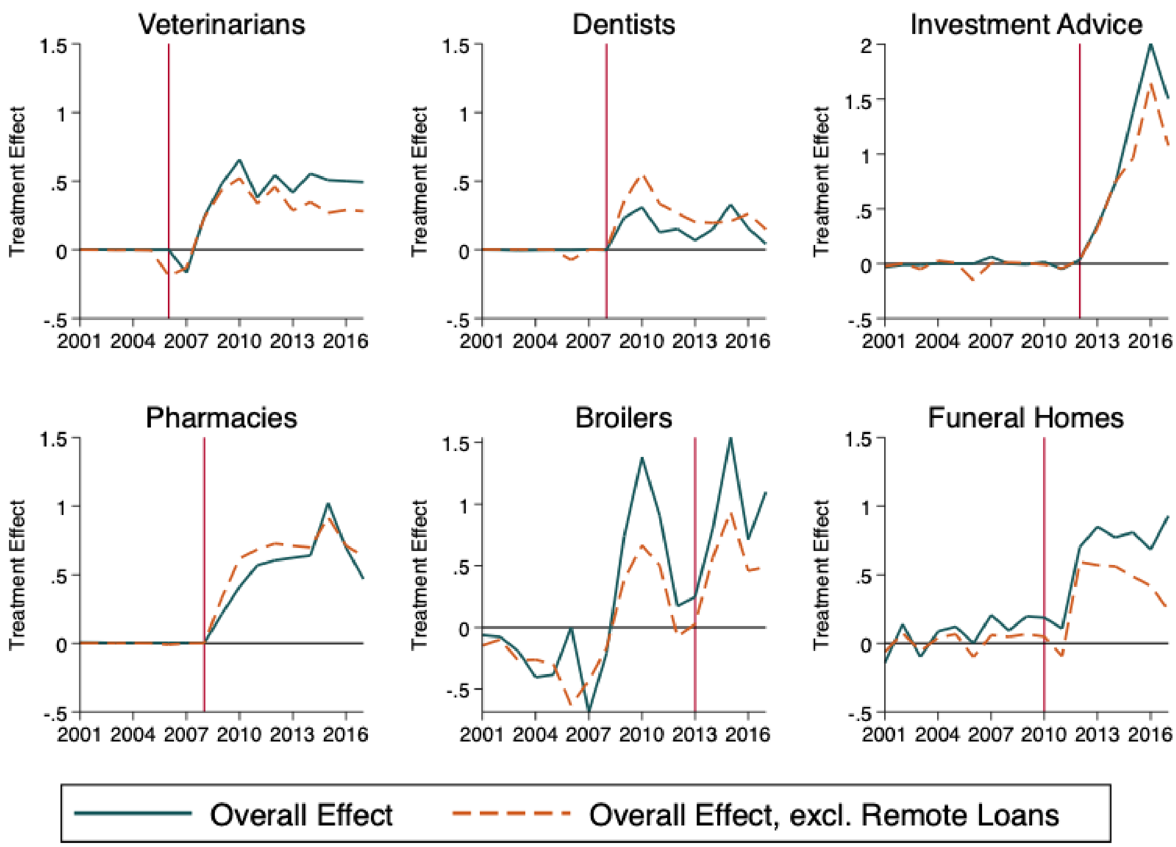


Figure 5: **Overall Effect of Entry**
Synthetic control estimates of the overall treatment effect on annual lending for each treated industry. The "Overall Effect" is for all loans, and "Overall Effect, excl. Remote Loans" excludes non-Live-Oak remote loans (loans with distance > 100 miles). The vertical line shows the year before Live Oak entered.

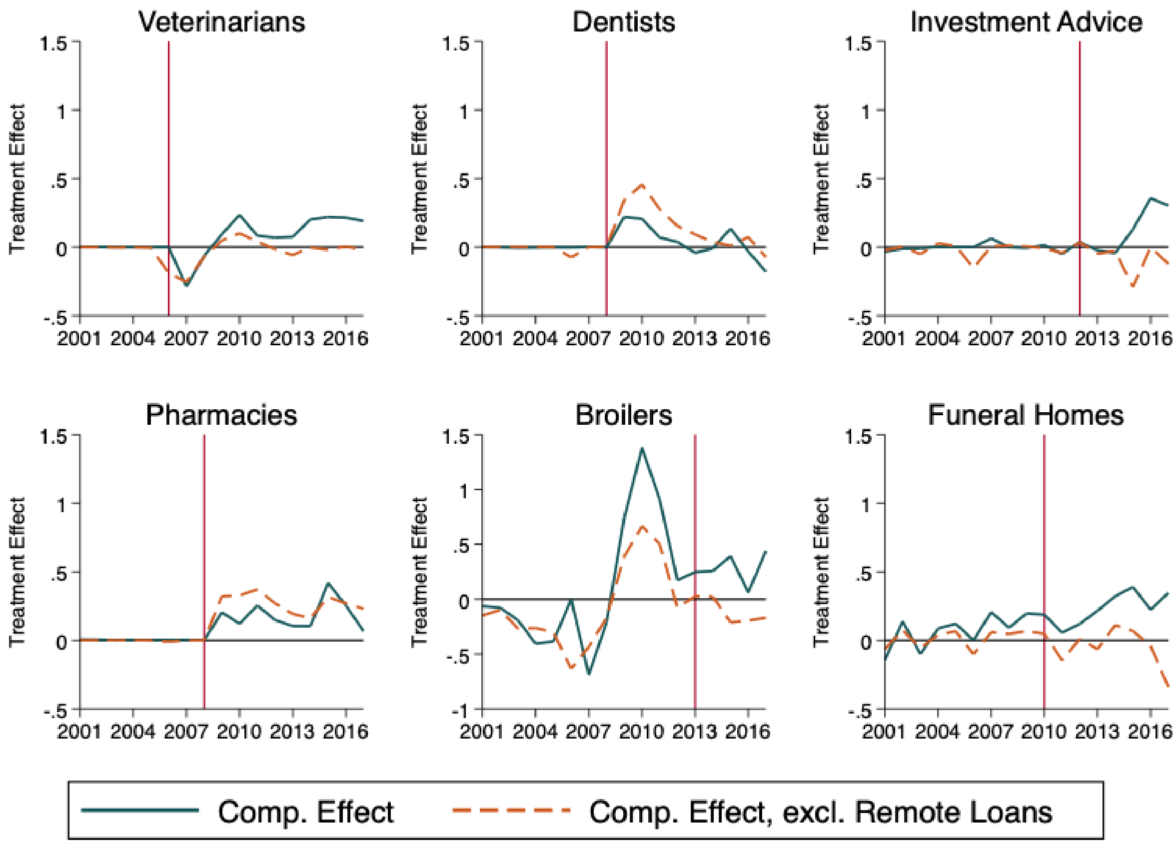


Figure 6: **Competitive Effect of Entry**

Synthetic control estimates of the competitive treatment effect on annual lending for each treated industry. The "Competitive Effect" is for all (non-Live-Oak) loans, and "Comp. Effect, excl. Remote Loans" excludes loans with borrower-lender distances > 100 miles. The vertical line shows the year before Live Oak entered.

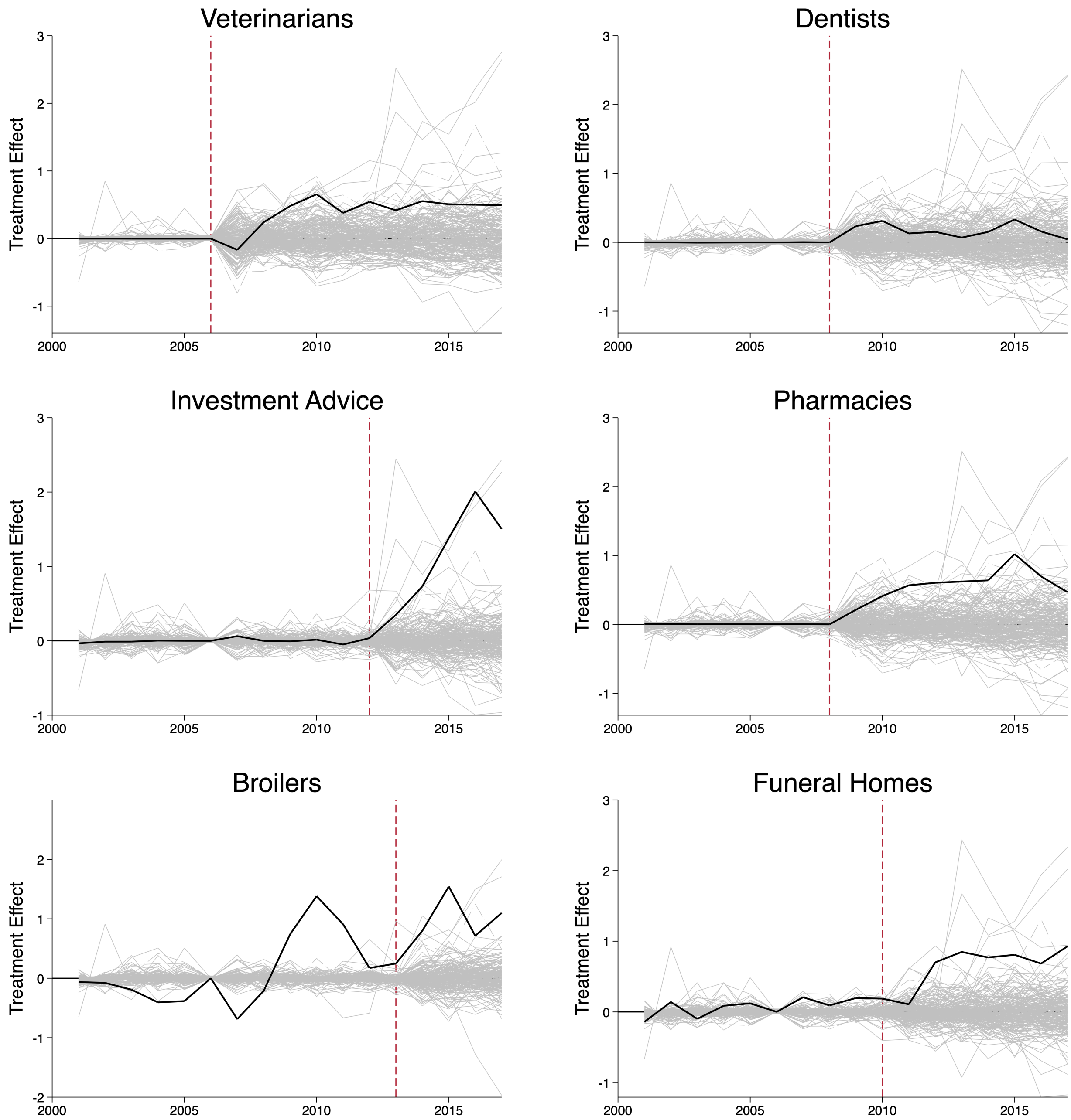


Figure 7: **Treated and Placebo Treatment Effects**
The bold line shows the gap for the industry that Live Oak entered, while the gray lines show the gap for the placebo industries. The figure omits industries with a pretreatment MSPE more than 20 times that of the average MSPE among the treated industries. The vertical line shows the year before Live Oak entered.

INTERNET APPENDIX

“Industry Specialization and Small Business Lending”

Wenhua Di   Nathaniel Pattison

Contents

List of Figures

List of Tables

A Appendix Tables and Figures

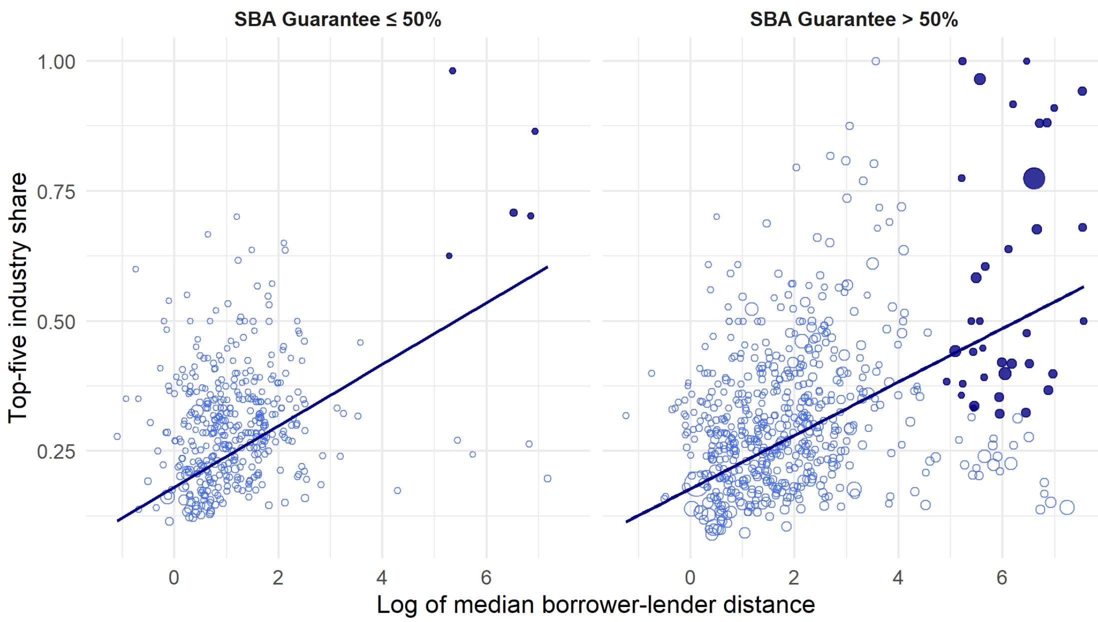


Figure A.1: **Distance and concentration by the SBA guarantee amount (2013-2017)** These figures plot institutions' (log) median borrower-lender distance against their top-five industry share for the period 2013-2017. The shares and distances are formed separately for loans with a low ($\leq$ 50%) or high ($>$ 50%) SBA guarantee. Each circle represents an institution and its size reflects the dollar amount of SBA loans it originated during the period. The sample is restricted to institutions originating at least 20 loans in the guarantee category during 2013-2017. The solid circles are remote, industry specialists (according to our classification in the text).

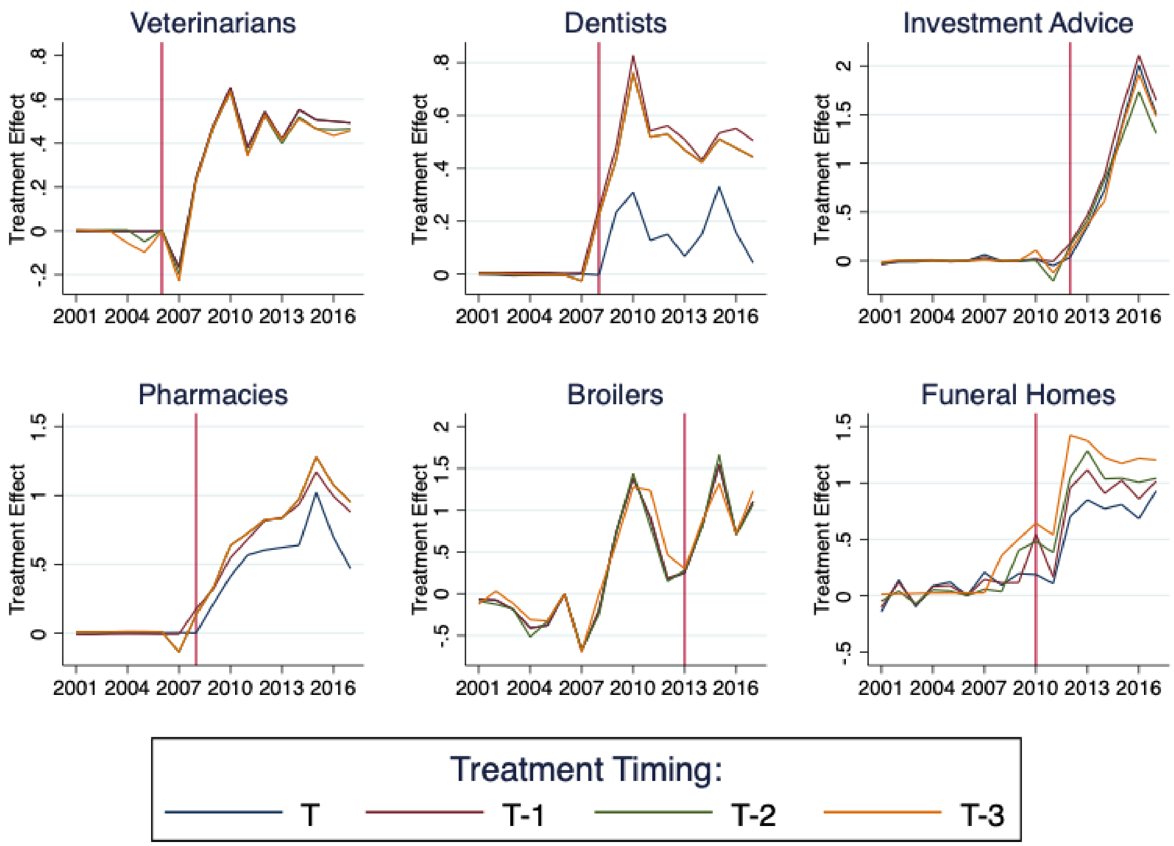


Figure A.2: **Sensitivity to Treatment Timing**

This figure reports treatment effects from four synthetic control estimates, varying the treatment timing $T-3$ through T, where T is the true treatment timing. The vertical line shows the year before Live Oak entered.

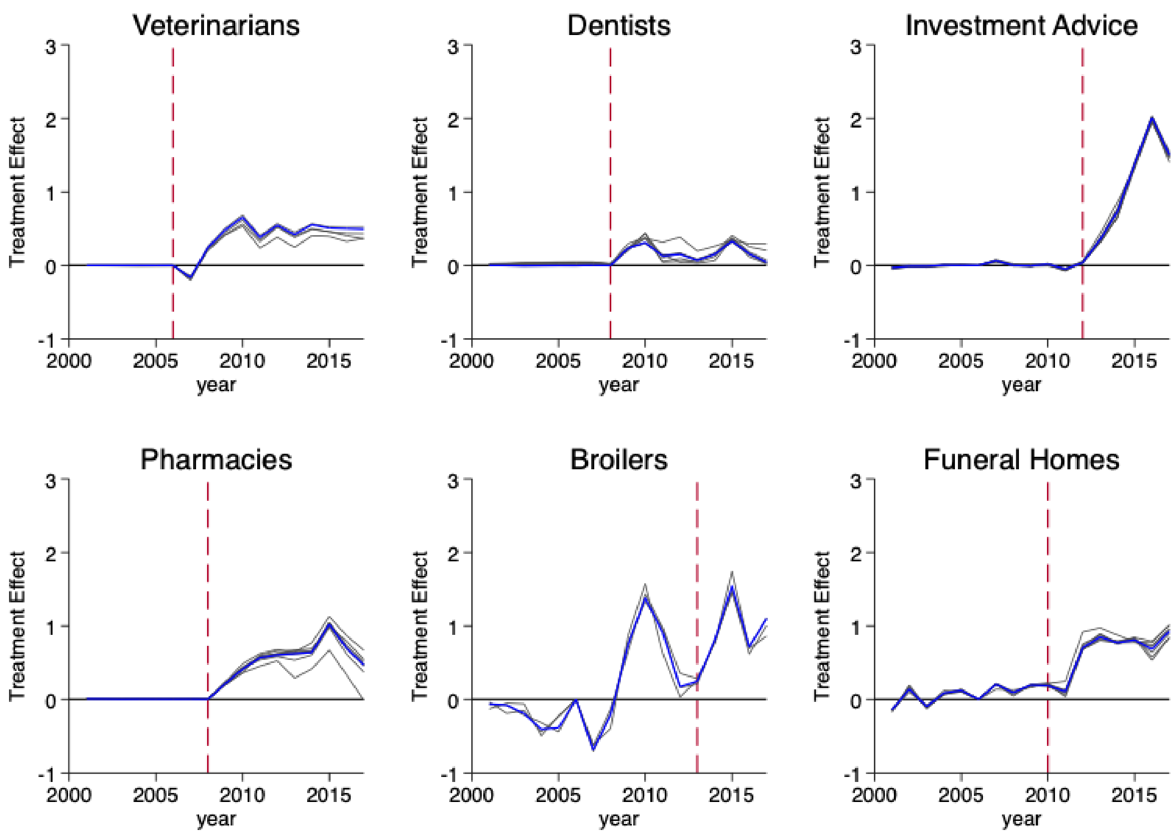


Figure A.3: **Sensitivity to the Donor Pool**

This figure reports leave-one-out treatment effect estimates. For each treated industry, we construct leave-one-out donor pools by iteratively dropping each control unit with a weight of at least 0.01 in the synthetic control of Figure 4 and re-estimating the treatment effects. The blue line shows the treatment effect when all industries are included. The vertical line shows the year before Live Oak entered.

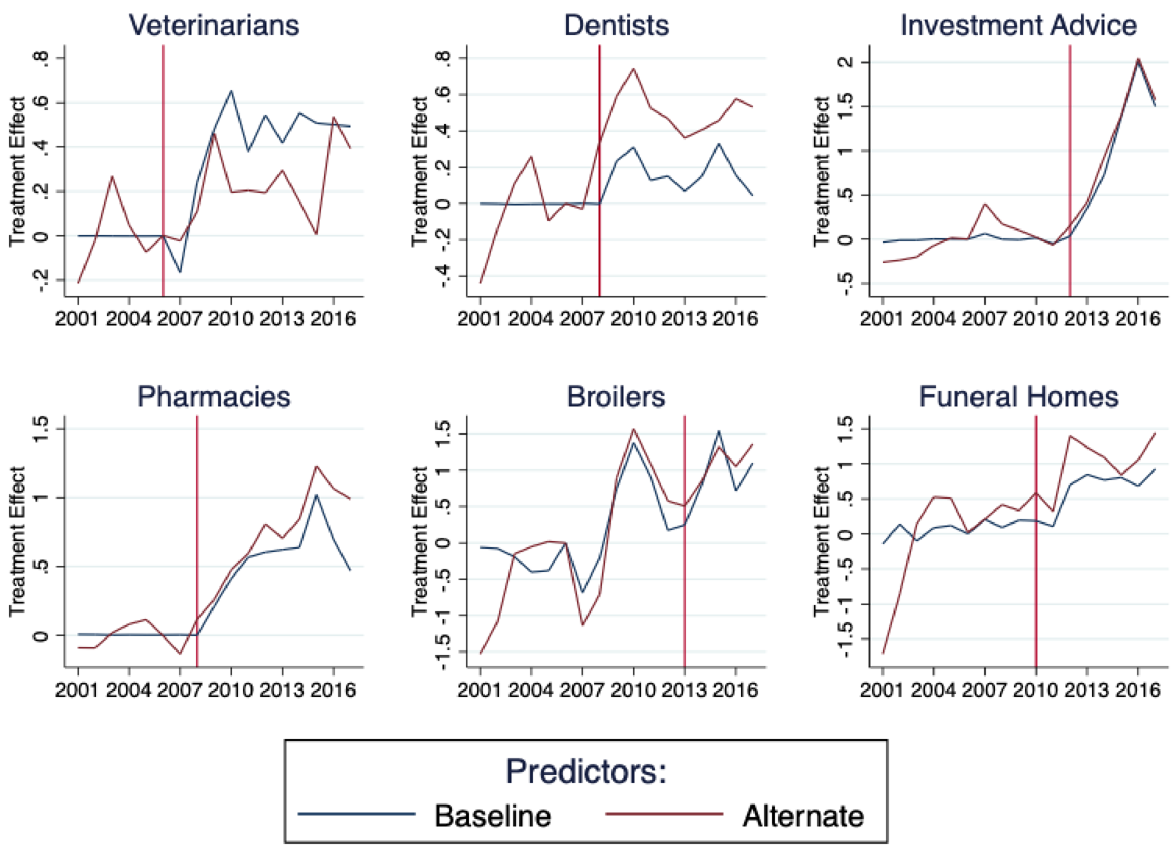


Figure A.4: **Sensitivity to Alternative Predictors**
The Baseline specification using all pre-treatment outcomes as controls. The Alternative matches units on the average pre-treatment controls for the number of observations, charge-off rate, and normalized lending, with each average taken over the pre-treatment period.

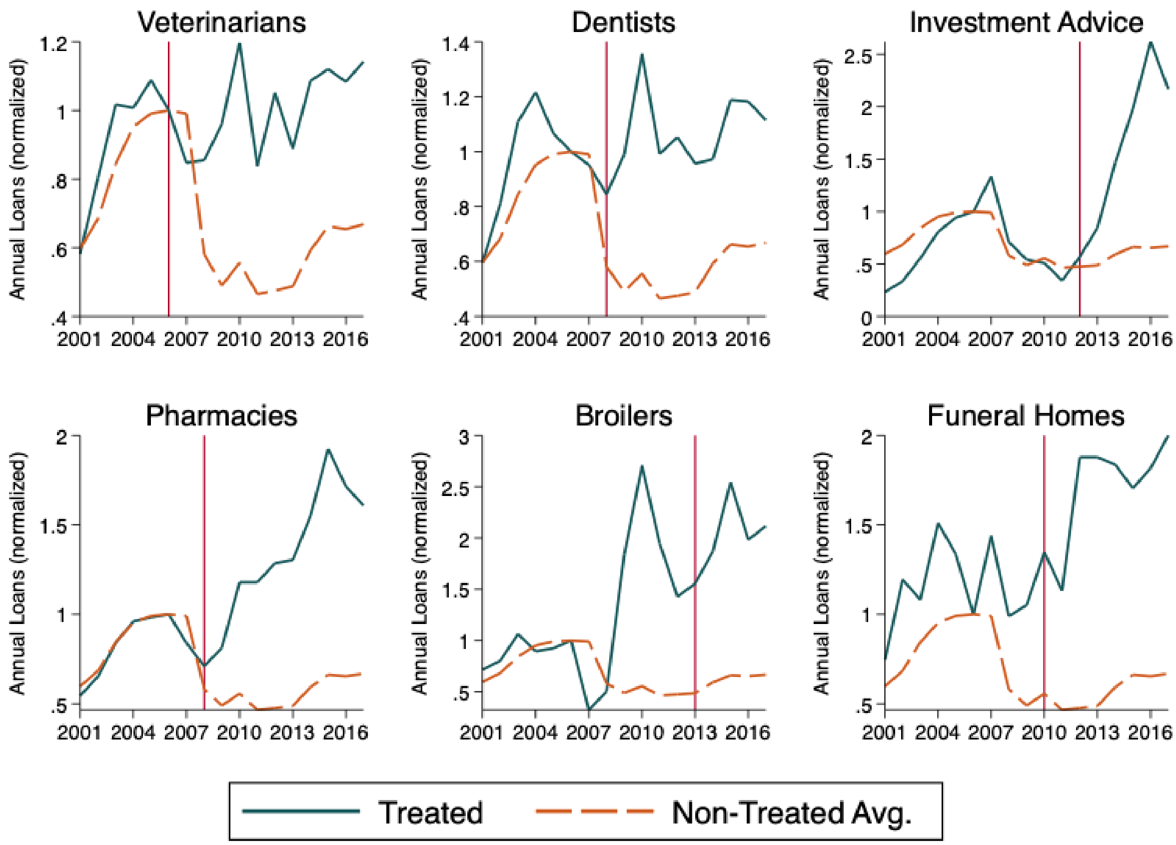
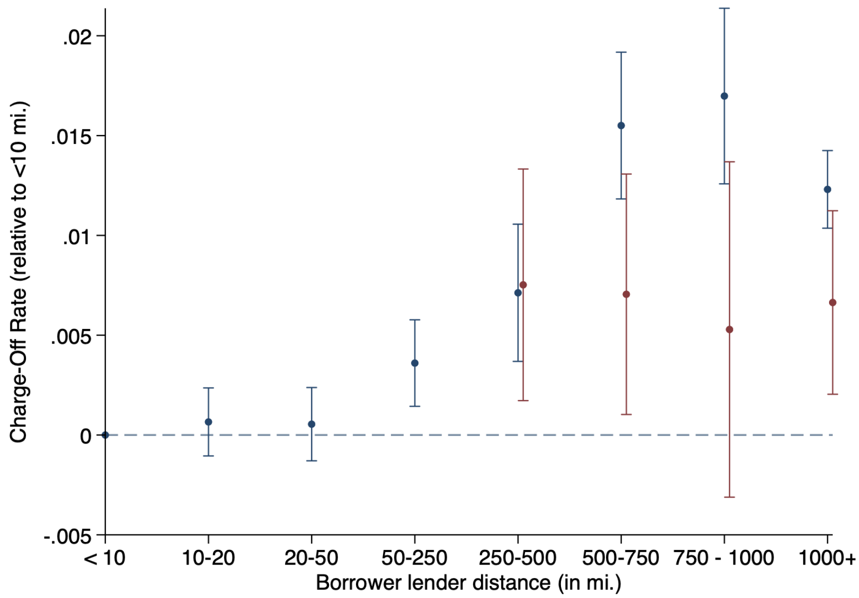
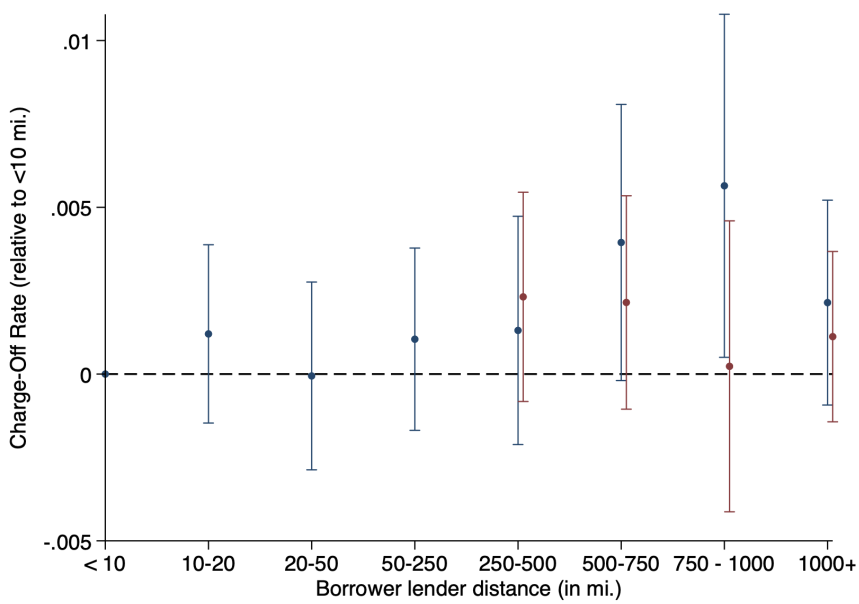


Figure A.5: **Annual Lending: Treated vs. Simple Comparison Group**
This figure shows the growth in the annual number of SBA 7(a) loans in each industry (with 2006 loans normalized to one) for the treated industries and the simple average of all control industries. In each panel, the non-treated group consists of the ten control industries whose average annual lending between 2001 and 2006 was closest to that of the treated industry. The vertical line shows the year before Live Oak entered.

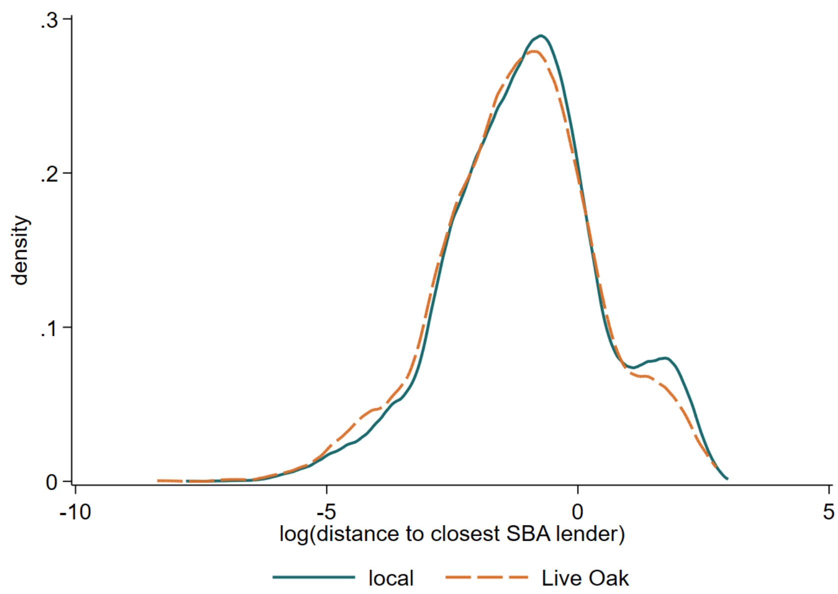


(a) All industries

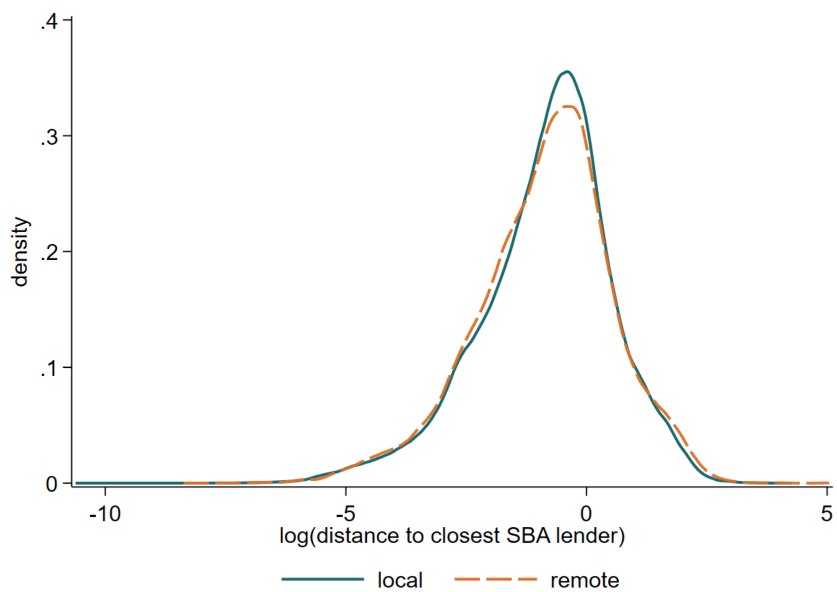


(b) Live Oak industries only

Figure A.6: **Distance and Charge-off Rates (Flexible Specification)**
We estimate the specification in Table 5 column 2 with the 3-year charge-off rate as the dependent variable, but replace the independent variable $log(dist)$ with a set of dummy variables for different distances (shown on the horizontal axis). We interact these dummies with an indicator for Live Oak and an indicator for non-Live-Oak lenders. These figures report the coefficients on these indicators, along with 95% confidence intervals, from a sample containing loans from all industries (figure a) and a sample containing only Live Oak's six industries. In both sets of regressions, Live Oak experiences better loan performance in distant loans, though the differences are only statistically significant in figure a. We do not report coefficients for Live Oak for loans of less than 250 miles, because Live Oak makes very few of these loans and the confidence interval spans the entire vertical axis.



(a) Comparison of local loans and Live Oak loans (in Live Oak industries)



(b) Comparison of local loans and remote loans

Figure A.7: **Distance to Closest SBA Branch**
This graph shows the similarity in the distribution of the distance between borrowers and the closest branch of any institution that grants SBA loans for local and remote between 2007 and 2017. The first figure compares local loans (from a lender within 100 miles) to Live Oak loans for borrowers in the six treated industries. The second figure compares local loans to remote loans (from a lender more than 100 miles away). Distance is calculated according to the procedure described in Section B, except it is the distance to the closest branch of any SBA lender.

Table A.1: **Institutions' Lending Distance and Portfolio Concentration**

	Dependent variable: Institution's Top Five Share					
	(1)	(2)	(3)	(4)	(5)	(6)
log(med. distance)	0.0244***	0.0304***	0.0140***	0.0131**		
	(0.00453)	(0.00398)	(0.00273)	(0.00544)		
Share 100+ mi.					0.222***	0.123***
					(0.0335)	(0.0251)
Observations	5,278	5,278	5,278	1,705	5,278	5,278
Mean Dep. Var.	0.430	0.430	0.430	0.318	0.430	0.430
Year FE	X	X	X	X	X	X
Inst. volume ventiles		X	X	X	X	X
Inst. FE			X	X		X
Balanced panel				X		

Observations are at the institution-year level from 2007-2017 and standard errors are clustered at the institution level. The dependent variable is the share of an institution's loan portfolio in its top five industries. Share 100+ mi. is the share of the institution's loans given to borrowers more than 100 miles from the closest branch. The sample is restricted to institution-year observations with at least 10 loans. Institution volume ventiles are ventile indicators for the number of SBA loans each year.

Table A.2: **Institutions' Lending Distance and Industry Concentration (HHI)**

	Dependent variable: Bank's Industry Concentration (HHI)					
	(1)	(2)	(3)	(4)	(5)	(6)
log(med. distance)	146.5***	162.0***	75.95***	42.41**		
	(23.96)	(23.09)	(14.74)	(17.80)		
Share 100+ mi.					1,264***	695.4***
					(192.8)	(142.3)
Observations	5,278	5,278	5,278	1,705	5,278	5,278
Mean Dep. Var.	985.6	985.6	985.6	686.5	985.6	985.6
Year FE	X	X	X	X	X	X
Inst. volume ventiles		X	X	X	X	X
Inst. FE			X	X		X
Balanced panel				X		

Observations are at the institution-year level from 2007-2017 and standard errors are clustered at the institution level. The sample is restricted to institution-year observations with at least 10 loans. The industry HHI for lender b in year t is defined as $HHI_{bt} = \sum_i S_{ibt}^2$, where S_{ibt} is the percent of lender b's loans given to industry i in year t. The HHI is increasing in industry concentration and takes a value from close to 0 (least concentrated) to 10,000 (all loans to a single industry). Institution volume ventiles are ventile indicators for the number of SBA loans each year.

Table A.3: **Industries Comprising Synthetic Controls (Donor Pool)**.

Industry	**Synthetic Makeup**	**Weight**
Broilers and Other Meat Type		
	Fluid Power Valve and Hose Fitting Manufacturing	0.14
	Logging	0.54
	Motion Picture Theaters (except Drive-Ins)	0.31
Pharmacies and Drug Stores		
	Continuing Care Retirement Communities	0.10
	Mobile Food Services	0.02
	Motion Picture Theaters (except Drive-Ins)	0.09
	Other Residential Care Facilities	0.04
	Precision Turned Product Manufacturing	0.04
	Recreational Vehicle Dealers	0.23
	Services for the Elderly and Persons with Disabilities	0.10
	Used Household and Office Goods Moving	0.02
Investment Advice		
	Audio and Video Equipment Manufacturing	0.09
	Child and Youth Services	0.04
	Direct Title Insurance Carriers	0.21
	Lessors of Other Real Estate Property	0.13
	Mortgage and Nonmortgage Loan Brokers	0.09
	Motor Vehicle Body Manufacturing	0.04
	Other Support Activities for Road Transportation	0.35
	Tour Operators	0.05
Veterinary Services		
	Fish and Seafood Markets	0.10
	Other Residential Care Facilities	0.04
	Photofinishing Laboratories (except One-Hour)	0.07
	Theater Companies and Dinner Theaters	0.05
Offices of Dentists		
	Agents and Managers for Artists, Athletes, Entertainers, and Other Pub	0.01
	All Other Miscellaneous Electrical Equipment and Component Manufacturi	0.16
	Fluid Power Valve and Hose Fitting Manufacturing	0.02
	Motion Picture Theaters (except Drive-Ins)	0.16
	Other Residential Care Facilities	0.26
	Other Support Activities for Air Transportation	0.09
	Packaging and Labeling Services	0.14
Funeral Homes and Funeral Services		
	All Other Miscellaneous Electrical Equipment and Component Manufacturi	0.16
	Child and Youth Services	0.09
	Logging	0.15
	Marinas	0.22
	Motion Picture Theaters (except Drive-Ins)	0.28
	Rendering and Meat Byproduct Processing	0.07
	Scenic and Sightseeing Transportation, Water	0.04

* Shows all industries with weight above 0.01. Industries with a weight of less than 0.01 are excluded.

Table A.4: **Comparison of Pre-Treatment Fit**

	(1) MSPE - Comparison Mean	(2) MSPE - Synthetic Control	(3) Ratio (1)/(2)
Veterinarians	.0095	1.2e-06	7,665.0
Offices of Dentists	.029	9.7e-06	2,982.5
Investment Advice	.044	.00079	55.3
Pharmacies	.0054	.000016	332.0
Broilers	.86	.33	2.6
Funeral Homes	.21	.02	10.6

This table compares the pre-treatment fit, measured by the mean squared prediction error (MSPE), for when the comparison group is either the simple average of all controls (column 1) or the synthetic control (column 2). Column 3 reports the ratio of the two MSPE calculations.

Table A.5: **Other Remote Lenders and Industries**

Bank	Industry	Year of First Loan	Bank's Share of Loans (2013-2017)	Bank's Loans in Industry (2001-2017)
Bank of George	Hotels (except Casino Hotels) and Motels	2015	0.83	90
Carver State Bank	Insurance Agencies and Brokerages	2016	0.93	70
Citizens Bank	Offices of Chiropractors	2015	0.15	34
Citizens Bank	Gasoline Stations with Convenience Stores	2001	0.11	28
Citizens Bank	Hotels (except Casino Hotels) and Motels	2001	0.10	77
Civis Bank	Gasoline Stations with Convenience Stores	2012	0.12	9
Civis Bank	Hotels (except Casino Hotels) and Motels	2010	0.12	14
Crestmark Bank	Insurance Agencies and Brokerages	2014	0.69	93
Crestmark Bank	Hotels (except Casino Hotels) and Motels	2014	0.13	18
FinWise Bank	Offices of Lawyers	2014	0.56	82
First Bank	Hotels (except Casino Hotels) and Motels	2016	0.10	20
First Chatham Bank	Child Day Care Services	2013	0.20	27
First Colorado National Bank	Hotels (except Casino Hotels) and Motels	2007	0.18	53
First Financial Bank	Broilers and Other Meat Type	2001	0.66	2592
First Financial Bank	Pharmacies and Drug Stores	2012	0.16	208
Meadows Bank	Retail Bakeries	2011	0.12	48
Meadows Bank	Child Day Care Services	2012	0.10	40
Mission Valley Bank	Hotels (except Casino Hotels) and Motels	2015	0.28	25
Mission Valley Bank	Funeral Homes and Funeral Services	2015	0.16	14
NOA Bank	Hotels (except Casino Hotels) and Motels	2009	0.38	163
NOA Bank	Gasoline Stations with Convenience Stores	2009	0.10	51
Spirit of Texas Bank, SSB	Beauty Salons	2009	0.27	462
Spirit of Texas Bank, SSB	Other Personal Care Services	2009	0.26	345
The MINT National Bank	Hotels (except Casino Hotels) and Motels	2014	0.67	74
The MINT National Bank	Gasoline Stations with Convenience Stores	2014	0.14	15
Titan Bank, National Association	Offices of Dentists	2013	0.30	27
United Community Bank	Offices of Dentists	2001	0.18	117
United Community Bank	Veterinary Services	2001	0.15	93
United Midwest Savings Bank	Offices of Dentists	2001	0.21	155

A lender is classified as a remote specialized lender if it has a median lending distance greater than 100 miles, a top-five industry share above 32%, and at least 50 total SBA loans (all measured during the period 2013-2017). We consider that lender to specialized in a specific industry if (in 2013-2017), at least 10% of that lenders loans are to the industry.

Table A.6: **Charge-off Rates and Interest Rates - County Distance Measure**

Sample: Loans in the six treated industries								
Dependent variable:	Charge-off Indicator				Interest Rate			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Live Oak loan	-0.000875	0.000515	-0.000830	0.00153	-0.123***	-0.0625***	0.129	0.207**
	(0.000825)	(0.000898)	(0.00502)	(0.00506)	(0.0141)	(0.0152)	(0.0846)	(0.0842)
$log(dist)$			0.000281	0.000470**			0.0659***	0.0763***
			(0.000195)	(0.000202)			(0.00328)	(0.00337)
Live Oak loan$\times log(dist)$			-0.000159	-0.000393			-0.0738***	-0.0797***
			(0.000770)	(0.000771)			(0.0130)	(0.0128)
Observations	15,569	15,569	15,569	15,569	15,569	15,569	15,569	15,569
Mean of Dep. Var	0.00193	0.00193	0.00193	0.00193	5.656	5.656	5.656	5.656
Year FE	X	X	X	X	X	X	X	X
Loan char.		X		X		X		X
Industry FE	X	X	X	X	X	X	X	X

This table repeats Table 5, using distance measured using the borrower's county-centroid. The sample consists of loans to the six treated industries for amounts over \$100,000 (in 2010 dollars) that were originated between 2008-2017. Interest rate data are available from 2008Q4 and observations missing the interest rate are dropped. The dependent variable is either an indicator for whether the loan was charged off within three years of origination or the loan's interest rate (in percentage points). Loan characteristics include the share guaranteed and dummies for ventiles of the size of the loan and the term length. Industry fixed effects are indicators for the 5-digit NAICS code.

Table A.7: **Live Oak Industry Selection - County Distance Measure**

Sample: Excludes Live Oak loans								
Dependent variable:	Charge-off Indicator				Interest Rate			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LO industry	-0.00685***	-0.00380***			-0.00837	-0.00247		
	(0.000929)	(0.000935)			(0.00762)	(0.00758)		
$log(dist)$			0.00317***	0.00313***			0.0517***	0.0478***
			(0.000167)	(0.000171)			(0.00135)	(0.00137)
LO industry×$log(dist)$			-0.00296***	-0.00241***			0.0121***	0.0217***
			(0.000433)	(0.000430)			(0.00350)	(0.00344)
Observations	90,969	90,969	90,969	90,969	90,969	90,969	90,969	90,969
Mean of Dep. Var.	0.00840	0.00840	0.00840	0.00840	5.693	5.693	5.693	5.693
Year FE	X	X	X	X	X	X	X	X
Loan char.		X		X		X		X
Industry FE			X	X			X	X

This table repeats Table 4, using distance measured using the borrower's county-centroid. The sample consists of loans for amounts over $100,000 (in 2010 dollars) that were originated between 2008-2017. Live Oak's loans are dropped. Interest rate data are available from 2008Q4 and observations missing the interest rate are dropped. The dependent variable is either an indicator for whether the loan was charged off within three years of origination or the loan's interest rate (in percentage points). Loan characteristics include the share guarantee and dummies for ventiles of the size of the loan and the term length. Industry fixed effects are indicators for the 5-digit NAICS code.

Table A.8: **Charge-off Rates and Interest Rates - Loans Above $150,000**

Sample: Loans in the six treated industries								
Dependent variable:	Charge-off Indicator				Interest Rate			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Live Oak loan	0.000418	0.00123	0.00215	0.00351	-0.126***	-0.0753***	0.0229	0.0900
	(0.000968)	(0.00104)	(0.00562)	(0.00566)	(0.0180)	(0.0192)	(0.103)	(0.103)
$log(dist)$			2.07e-05	0.000141			0.0501***	0.0574***
			(0.000188)	(0.000196)			(0.00345)	(0.00356)
Live Oak loan×$log(dist)$			-0.000274	-0.000428			-0.0546***	-0.0598***
			(0.000854)	(0.000856)			(0.0157)	(0.0155)
Observations	9,406	9,406	9,406	9,406	9,406	9,406	9,406	9,406
Mean of Dep. Var	0.00159	0.00159	0.00159	0.00159	5.638	5.638	5.638	5.638
Year FE	X	X	X	X	X	X	X	X
Loan char.		X		X		X		X
Industry FE	X	X	X	X	X	X	X	X

This table repeats Table 5, restricting the sample to loans above $150,000. The sample consists of loans to the six treated industries for amounts over $100,000 (in 2010 dollars) that were originated between 2008-2017. Interest rate data are available from 2008Q4 and observations missing the interest rate are dropped. The dependent variable is either an indicator for whether the loan was charged off within three years of origination or the loan's interest rate (in percentage points). Loan characteristics include the share guaranteed and dummies for ventiles of the size of the loan and the term length. Industry fixed effects are indicators for the 5-digit NAICS code.

Table A.9: **Live Oak Industry Selection - Loans Above $150,000**

Sample: Excludes Live Oak loans								
Dependent variable:	Charge-off Indicator				Interest Rate			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LO industry	-0.00431***	-0.00273***			0.0262***	0.0110		
	(0.000959)	(0.000971)			(0.00966)	(0.00965)		
$log(dist)$			0.000464***	0.000762***			0.0350***	0.0364***
			(0.000157)	(0.000158)			(0.00155)	(0.00155)
LO industry $\times log(dist)$			-0.000399	-0.000379			0.0137***	0.0162***
			(0.000374)	(0.000373)			(0.00371)	(0.00365)
Observations	51,436	51,436	51,436	51,436	51,436	51,436	51,436	51,436
Mean of Dep. Var.	0.00527	0.00527	0.00527	0.00527	5.654	5.654	5.654	5.654
Year FE	X	X	X	X	X	X	X	X
Loan char.		X		X		X		X
Industry FE			X	X			X	X

This table repeats Table 4, restricting the sample to loans above $150,000. The sample consists of loans for amounts over $100,000 (in 2010 dollars) that were originated between 2008-2017. Live Oak's loans are dropped. Interest rate data are available from 2008Q4 and observations missing the interest rate are dropped. The dependent variable is either an indicator for whether the loan was charged off within three years of origination or the loan's interest rate (in percentage points). Loan characteristics include the share guarantee and dummies for ventiles of the size of the loan and the term length. Industry fixed effects are indicators for the 5-digit NAICS code.

Table A.10: **Charge-off Rates and Interest Rates - County FE**

Sample: Loans in the six treated industries								
Dependent variable:	Charge-off Indicator				Interest Rate			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Live Oak loan	0.000524	0.00153	0.00328	0.00480	-0.124***	-0.0628***	-0.218	-0.0480
	(0.00125)	(0.00134)	(0.00943)	(0.00952)	(0.0195)	(0.0205)	(0.145)	(0.144)
$log(dist)$			0.000485**	0.000622**			0.0462***	0.0534***
			(0.000236)	(0.000248)			(0.00364)	(0.00376)
Live Oak loan×$log(dist)$			-0.000734	-0.000870			-0.0168	-0.0352
			(0.00142)	(0.00143)			(0.0219)	(0.0217)
Observations	10,353	10,353	10,353	10,353	10,353	10,353	10,353	10,353
Mean of Dep. Var	0.00222	0.00222	0.00222	0.00222	5.655	5.655	5.655	5.655
Year FE	X	X	X	X	X	X	X	X
County FE	X	X	X	X	X	X	X	X
Loan char.		X		X		X		X
Industry FE	X	X	X	X	X	X	X	X

This table repeats Table 5, including county fixed effects. The sample consists of loans to the six treated industries for amounts over \$100,000 (in 2010 dollars) that were originated between 2008-2017. Interest rate data are available from 2008Q4 and observations missing the interest rate are dropped. The dependent variable is either an indicator for whether the loan was charged off within three years of origination or the loan's interest rate (in percentage points). Loan characteristics include the share guaranteed and dummies for ventiles of the size of the loan and the term length. Industry fixed effects are indicators for the 5-digit NAICS code.

Table A.11: **Live Oak Industry Selection - County FE**

Sample: Excludes Live Oak loans								
Dependent variable:	Charge-off Indicator				Interest Rate			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LO industry	-0.00710***	-0.00402***			-0.0541***	-0.0385***		
	(0.00123)	(0.00123)			(0.00972)	(0.00958)		
$log(dist)$			0.00314***	0.00289***			0.0417***	0.0376***
			(0.000179)	(0.000184)			(0.00140)	(0.00142)
LO industry$\times log(dist)$			-0.00282***	-0.00221***			0.000708	0.0103***
			(0.000458)	(0.000457)			(0.00360)	(0.00353)
Observations	63,455	63,455	63,455	63,455	63,455	63,455	63,455	63,455
Mean of Dep. Var.	0.00879	0.00879	0.00879	0.00879	5.691	5.691	5.691	5.691
Year FE	X	X	X	X	X	X	X	X
County FE	X	X	X	X	X	X	X	X
Loan char.		X		X		X		X
Industry FE			X	X			X	X

This table repeats Table 4, including county fixed effects. The sample consists of loans for amounts over $100,000 (in 2010 dollars) that were originated between 2008-2017. Live Oak's loans are dropped. Interest rate data are available from 2008Q4 and observations missing the interest rate are dropped. The dependent variable is either an indicator for whether the loan was charged off within three years of origination or the loan's interest rate (in percentage points). Loan characteristics include the share guarantee and dummies for ventiles of the size of the loan and the term length. Industry fixed effects are indicators for the 5-digit NAICS code.

B Appendix: Matching Procedure

In this appendix, we describe the procedure used to construct a measure of borrower-lender distance.

B.1 Matching SBA Lenders to FDIC Summary of Deposits

The SBA 7(a) loan data contain the name and address of the institution that is currently assigned the loan. 5,815 institutions originated SBA loans between 2001 and 2017. For these institutions, we conduct a series of probabilistic matches using bank name, address, city, state, and zip code to link the SBA lending institutions to institutions in the 2017 FDIC Summary of Deposits. First, the matching procedure produces a match score between 0 and 1 based on the similarity of the text in the variables listed above, with more weight given to the bank name and address, since they are more likely to uniquely identify banks.[1] Of the 5,815 unique institutions, we find an exact match for 3,041. After checking for accuracy, we also count the roughly 800 additional institutions with a bigram match score greater than 0.98 as a match. For those with a score less than 0.98, we conduct a clerical review to determine whether the best match is accurate. After this first round of matching, we conduct a second round of matching and clerical review using different weights for the variables. We then manually match any unmatched institution that gave more than 100 SBA loans between 2001 and 2017 (provided that the institution is a bank and is not closed). Overall, we match 75% of the 5,815 institutions and these institutions provide 91.8% of SBA loans from 2001-2017. The majority of unmatched SBA institutions are credit unions or non-bank lenders, for which we do not have bank branch locations in the FDIC Summary of Deposit data, or they are closed banks whose assets were transferred.

B.2 SBA Lenders' Branch Locations

Having matched banks in the SBA data to banks in the FDIC Summary of Deposits, we construct historical branch networks. The FDIC Summary of Deposits contains annual counts and locations for bank branches from 1994-2017. For each matched SBA lender, we can therefore determine its branch locations at the time the loan was originated. The matches are imperfect, however, since the SBA 7(a) data contain the institution currently assigned the loan, rather than the institution that originated the loan. Bank closures, mergers, and acquisitions will generate differences between the banks currently assigned the loan and the bank that originated the loan. For example, BankBoston merged with Bank of America in 2004, and all of its branches were converted to Bank of America. Consequently, an SBA loan originated by BankBoston in 2001 may appear in the SBA data as currently held by Bank of America. To construct historical branch networks in light of these changes in bank structure, for each branch in each year from 2001-2017, we use the FDIC's Reports

[1] Specifically, we first standardize the bank names and addresses, then use reclink command in Stata. To assess similarity, reclink uses bigram comparison to score two strings based on the number of common 2-4 consecutive letter combinations. The first probabilistic match uses relative weights of 14 (out of 20) given to the name, 8 given to the address, 4 given to city, and 4 given to the zip code. The second match uses the same variables, but weights of 16,4,4, and 4. In both, we require state to match exactly.

of Structure Changes to determine the bank that holds that branch as of 2017. For example, we consider a branch to be a part of Bank of America's network if that branch is a Bank of America branch or would later become a Bank of America branch. That is, for a given year t, we consider a branch to be a part of an institution j's network in year t if that branch either (i) belongs to institution j in year t or (ii) would become a branch of institution j by 2017.

Another possible source of error is that banks may transfer loan assignments, even if there were no changes in bank structure. In order to gauge the error introduced by transfers of assignments, we compare loans of the top 100 lenders in FY2012 from the 2012 Coleman Report to the top 100 lenders in FY2012 based on who is currently assigned the loan. These top 100 lenders provided 59% of all SBA loans and 60% of SBA volume in FY2012. Of the top 100 lenders, we are able to match 70 in our 2017 data. The unmatched banks are due to name changes, closures, mergers, and acquisitions between 2012 and 2017. Of the matched banks, the number of loans attributed to them in our data is very similar to the loans attributed to them in the 2012 Coleman Report (see Figure B.1), suggesting that absent changes in bank structure, banks rarely transfer the assignment of SBA loans.

B.3 Borrower-Lender Distance

Starting with the 962,527 non-canceled SBA loans from 2001-2017 (and dropping the 179 that are missing industry info), we are able to match 885,166 to a lending institution in the FDIC Summary of Deposits. We then run these loans through the Census Geocoder, using the borrower's listed address, and are able to match 629,946 of the addresses to a latitude and longitude. Our results are also robust to using borrower-lender distance constructed using the centroid of the borrower's county, which is available for all borrowers. Then, based on the borrower's institution and year, we match each borrower to the historical branch network for that institution.[2] Finally, we calculate the (Haversine) distance between the borrower and (i) the closest branch of the institution that originated the loan and (ii) the closest branch of any SBA lender.[3]

[2]We drop the 1.5% of branches that are missing longitude and latitude data.

[3]The Haversine distance, which is the shortest distance over the earth's surface.

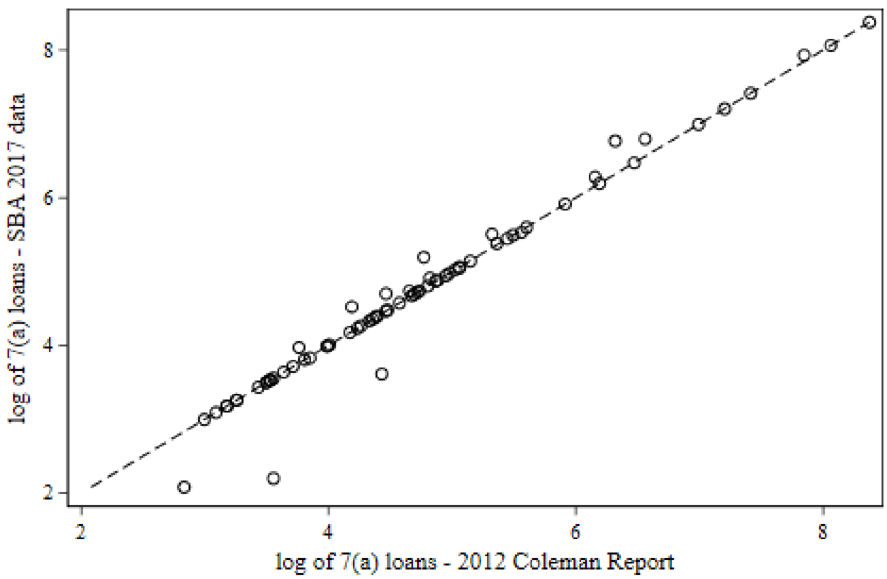


Figure B.1: **Difference between Counts at Origination in 2012 and Counts Recorded in 2017**

C Appendix: Characterizing Remote Specialists

C.1 Descriptive Characteristics

We briefly characterize the specialists and their chosen industries. Internet Appendix Table C.1 lists the 21 specialized lenders in the 2013-2017 period, along with their median borrower-lender distance and top-five share. We also list the institutions in which they specialize in Internet Appendix Table C.2. Among the 21 lenders classified as remote specialists in the 2013-2017 period, the average of the median borrower-lender distance is 677 miles and the average top-five share is 58%. Of the industries they specialize in, defined as those making up more than 10% of the lender's portfolio, hotels and gas stations are most commonly selected, and health professionals (chiropractors, dentists, pharmacists, and veterinarians) and financial or legal professionals (insurance agencies, investment advisers, and lawyers) are also common. There is also a variety of other industries, including funeral homes, bakeries, and daycare services.

These industries are likely selected, at least in part, because they have lower payment risk for lenders. The average three-year charge-off rate (from 2007-2012) for all industries receiving SBA loans was 7.5%, while the average charge-off rate for industries chosen by specialists (weighted by the number of specialists) is 2.8%. We also gathered industry characteristics from The Risk Management Association (RMA) and IBISWorld Industry Reports, which provide detailed information about market characteristics, industry conditions, and characterizes industries along ten dimensions. Compared to the fifteen most common industries in the SBA data, the specialists' industries tend to have higher capital intensity, greater regulation and greater industry assistance (defined as protection, direct or indirect government assistance, and support from associations and trade groups).

C.2 Industry Concentration and Loan Performance

If industry concentration facilitates expertise in lending to these industries, concentrated lenders may experience better loan performance within the industries where they focus. To investigate this idea, we examine whether loans from concentrated lenders perform better than loans from less concentrated lenders. As mentioned, concentrated lenders tend to focus on industries with lower charge-offs, which would lead to better loan performance even in the absence of expertise. So that our estimates will not be driven by this industry selection, our regressions will include industry fixed effects. Thus, our strategy compares within-industry across lenders.

Using the loan-level data, we estimate the following regression for a loan i from lender b to industry j originated in year t:

$$Chargeoff_{ibjt} = \alpha + \beta_0 log(dist_{ibjt}) + \beta_1 IndustryShare_{bjt} + X_{ibjt}\gamma + \delta_j + \tau_t + \epsilon_{ibjt} \tag{1}$$

where $Chargeoff_{ibjt}$ is an indicator for whether loan i from lender b originated to industry j during year t was charged off within three years of origination. The variable $log(dist_{ibjt})$ is the log of the distance between the borrower and the closest branch of the institution originating the loan. The main specification also includes loan-level controls for size and term length (X_{ibjt}) and industry (δ_j) and year (τ_t) fixed effects. Some specifications also include additional loan-level controls, state-by-year fixed effects, and institution-specific fixed effects.

Our measure of industry concentration, $IndustryShare_{bjt}$, is the share of total loans from lender b in year t that went to industry j. We focus on contemporaneous shares as our primary measure. If lenders build expertise (e.g. by hiring industry experts) then increase lending to the industry, current lending shares reflect expertise. However, if expertise are developed through past exposure to an industry, it may be more appropriate to use a lagged measure. In robustness checks, we find a similar effect using lagged shares. Moreover, contemporaneous and lagged shares are highly correlated - the coefficient of correlation is 0.92.[4] The coefficient of interest, β_1, captures the correlation between the probability that a loan in industry j is charged off within three years and the lender's $IndustryShare_{bjt}$. If β_1 is negative, it would reflect that lenders giving a larger share of their loans to an industry experience lower charge-off rates relative to other lenders. Since the specification includes industry fixed effects, β_1 reflects how the probability of charge-offs varies among loans given to the same industry. In some specifications, we add the interaction of the share of loans to an industry and borrower-lender distance, to examine whether industry concentration can mitigate the disadvantages of lending at a distance.

Table C.3 reports the results of specification (1). Consistent with the prior literature on distance and lending, the positive coefficient on the $log(dist)$ in Column 1 indicates that the probability of default increases with borrower-lender distance, controlling for loan characteristics (dummies for ventiles of loan size and term length). Column 2 adds the share of loans that a lender makes to the industry. The negative coefficient on the share in the industry indicates that having a greater share of loans to an industry is correlated with lower charge-off rates within that industry (relative to less concentrated lenders). To provide a sense of the magnitude, these estimates imply that an industry share of 52% would offset the additional risk of a 100-mile loan. The offsetting threshold is higher for more distant loans and lower for closer ones. This negative relationship between concentration and the probability of default remains similar when adding state-by-year fixed effects in Column 3. Column 4 includes the interaction of the "Share in industry" with the log of borrower-lender distance. The coefficient is negative and significant, suggesting that concentration in lending can mitigate the disadvantages of lending at a distance. Columns 5-8 repeat these specifications, but add institution fixed effects. The coefficients decrease in magnitude, but remain statistically significant. Thus, even within an institution, loan performance is better in the industries where the institution is more concentrated. However, adding institution fixed effects causes the interaction of industry share with $log(dist)$ to become statistically insignificant and slightly positive (column 8).

[4] An alternative measure concentration could be the *number* of loans a bank gave to the industry. This measure, however, would potentially conflate the effects of bank size and concentration. Instead, we adopt the common approach of using a measure that is comparable across banks of different sizes and then controlling directly for bank size in the regressions (Acharya, Hasan and Saunders, 2006, Hayden, Porath and Westernhagen, 2007, Berger, Minnis and Sutherland, 2017).

Table C.1: **List of Remote Lenders**

Institution	B-L distance	Top-5 Share	Industries	Share of lender's loans (%)	Share of SBA loans (%)	Ratio of column (5) to (6)
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Bank Of George	1,828	92	Hotels (except Casino Hotels) and Motels	83	1.7	49
			Lessors of Miniwarehouses and Self-Storage Units	2.8	0.23	12
Carver State Bank	640	100	Insurance Agencies and Brokerages	93	0.87	108
			Other Electronic and Precision Equipment Repair and Maintenance	4	0.25	16
Citizens Bank	399	42	Offices of Chiropractors	15	0.96	15
			Gasoline Stations with Convenience Stores	11	1.1	10
Civis Bank	187	38	Gasoline Stations with Convenience Stores	12	1.1	11
			Hotels (except Casino Hotels) and Motels	12	1.7	7.1
Crestmark Bank	830	88	Insurance Agencies and Brokerages	69	0.87	79
			Hotels (except Casino Hotels) and Motels	13	1.7	7.9
Evolve Bank & Trust	634	32	Veterinary Services	8.6	0.81	11
			Offices of Dentists	7.1	1.8	3.9
Finwise Bank	1,885	68	Offices of Lawyers	56	1.1	49
			Electronic Shopping	7.5	0.54	14
First Bank	363	34	Hotels (except Casino Hotels) and Motels	10	1.7	6.2
			Funeral Homes and Funeral Services	9.9	0.34	29
First Chatham Bank	671	42	Child Day Care Services	20	1.2	16
			Car Washes	6.7	0.77	8.7
First Colorado National Bank	1,062	40	Hotels (except Casino Hotels) and Motels	18	1.7	11
			Drycleaning and Laundry Services (except Coin-Operated)	6.5	0.48	14
First Financial Bank	249	97	Broilers and Other Meat Type	66	0.65	101
			Pharmacies and Drug Stores	16	0.68	24
Live Oak Banking Company	734	76	Investment Advice	19	0.53	36
			Offices of Dentists	18	1.8	9.9
Meadows Bank	233	34	Retail Bakeries	12	0.38	31
			Child Day Care Services	10	1.2	8.4
Mission Valley Bank	176	56	Hotels (except Casino Hotels) and Motels	28	1.7	17
			Funeral Homes and Funeral Services	16	0.34	47
Noa Bank	244	58	Hotels (except Casino Hotels) and Motels	38	1.7	23
			Gasoline Stations with Convenience Stores	10	1.1	9.1
Spirit Of Texas Bank, Ssb	769	66	Barber Shops	27	2	13
			Other Personal Care Services	26	0.73	36
T Bank, National Association	972	37	Car Washes	9.9	0.77	13
			Child Day Care Services	7.6	1.2	6.2
The Mint National Bank	947	88	Hotels (except Casino Hotels) and Motels	67	1.7	40
			Gasoline Stations with Convenience Stores	14	1.1	12
Titan Bank, National Association	781	47	Offices of Dentists	30	1.8	17
			Lessors of Nonresidential Buildings (except Miniwarehouses)	5.6	0.62	9
United Community Bank	136	42	Offices of Dentists	18	1.8	10
			Veterinary Services	15	0.81	18
United Midwest Savings Bank	480	42	Offices of Dentists	21	1.8	12
			Funeral Homes and Funeral Services	7.9	0.34	23

This table lists the 2013-2017 institutions in Figure 1 that are classified as remote specialists (according to our definition). Column 1 reports the institution's name. Columns 2 and 3 report the institution's median borrower-lender distance and its top-five share, calculated over 2013-2017. Column 4 lists the top two industries for each institution's and Column 5 lists the share of the institution's SBA loans going to that industry. For comparison, Column 5 lists the share of all SBA loans going to that industry. Finally, Column 7 shows the ratio of Column 5 to Column 6, which gives the share of the industry within each specialist institution relative to the industry's overall SBA share.

Table C.2: **List of Specialists' Industries**

Industry	Specialists (#)	Share of specialists' loans (%)	Share of SBA loans (%)	Ratio of column (3) to (4)	Charge-off rate (%)
(1)	(2)	(3)	(4)	(5)	(6)
Barber Shops	1	27	2	13	9.4
Broilers and Other Meat Type	2	39	0.65	60	0.73
Child Day Care Services	2	15	1.2	12	4.2
Funeral Homes and Funeral Services	1	16	0.34	47	1.2
Gasoline Stations with Convenience Stores	4	12	1.1	11	3.2
Hotels (except Casino Hotels) and Motels	9	31	1.7	19	0.97
Insurance Agencies and Brokerages	2	81	0.87	93	5.9
Investment Advice	1	19	0.53	36	9.2
Offices of Chiropractors	1	15	0.96	15	4.2
Offices of Dentists	4	22	1.8	12	0.85
Offices of Lawyers	1	56	1.1	49	3.5
Other Personal Care Services	1	26	0.73	36	9.3
Pharmacies and Drug Stores	2	15	0.68	21	1.7
Retail Bakeries	1	12	0.38	31	6.6
Veterinary Services	2	15	0.81	18	0.9
Overall SBA Average					7.5

This table reports the industries in which the institutions in Table C.1 specialize. The table includes any industry in which a specialist lender listed in Table C.1 originated at least 5% of its loans during the 2013-2017 period. Column 1 reports the industries and Column 2 reports the number of specialists giving at least 10% of its loans to the industry. Column 3 reports the share of the specialists' loans to that industry (or the average share when the number of specialists in that industry is greater than 1). For comparison, Column 4 reports the share of all 2013-2017 SBA loans that go to that industry, and Column 5 reports the ratio of Column 3 to Column 4. Finally, Column 6 reports the three-year charge-off rate for each industry during, calculated during the 2007-2012 period.

Table C.3: **Lender Portfolio Concentration and Loan Performance (within Industry)**

	Dependent variable: Indicator for Charge-off within 3 Years							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$log(dist)$	0.00476***	0.00500***	0.00437***	0.00521***	0.00208***	0.00210***	0.00252***	0.00208***
	(0.000362)	(0.000358)	(0.000332)	(0.000401)	(0.000388)	(0.000388)	(0.000369)	(0.000425)
Share in industry		-0.0441***	-0.0333***	-0.0391***		-0.0170***	-0.0174***	-0.0176***
		(0.00340)	(0.00284)	(0.00386)		(0.00428)	(0.00416)	(0.00507)
Share$\times log(dist)$				-0.00298**				0.000268
				(0.00144)				(0.00129)
Observations	255,874	255,874	255,874	255,874	255,874	255,874	255,874	255,874
Industry FE	X	X	X	X	X	X	X	X
Year FE	X	X	X	X	X	X	X	X
Loan char.	X	X	X	X	X	X	X	X
State-by-year FE			X				X	
Inst. FE					X	X	X	X

Using the loan-level data, we estimate the following regression for a loan i from lender b to industry j originated in year t:

$$Chargeoff_{ibjt} = \alpha + \beta_0 log(dist_{ibjt}) + \beta_1 IndustryShare_{bjt} + X_{ibjt}\gamma + \delta_j + \tau_t + \epsilon_{ibjt}$$

where $Chargeoff_{ibjt}$ is an indicator for whether loan i from lender b originated to industry j during year t was charged off within three years of origination. Observations are at the loan level from 2007-2014 and standard errors are clustered at the industry (5-digit NAICS) level. Loan characteristics include dummies for ventiles of the size of the loan and the term length. The state in the state-by-year fixed effects is determined by the location of the borrower's business.

D Appendix: An Indirect Test of the Impact on Total Lending

We empirically examine the impact of Live Oak's entry on a proxy for total industry lending from The Risk Management Association's (RMA) eStatement Studies.[5] Financial institutions provide the RMA with financial statements collected from commercial borrowers or applicants. Although participation is voluntary, hundreds of financial institutions including nine of the ten largest banks provide these statements to the RMA (Lisowsky, Minnis and Sutherland, 2017). The RMA's eStatement Studies publishes counts of the number of financial statements collected by industry (6-digit NAICS). Financial statements can be collected due to loan originations, applications, or monitoring, and so are an imperfect proxy for total loan originations. Still, these counts of financial statements provide, to our knowledge, the only industry-specific measure of total (SBA and non-SBA) lending activity and Berger, Minnis and Sutherland (2017) shows a strong correlation between these financial statements and the size of bank's commercial and industrial lending portfolio. Our RMA data includes a balanced panel of annual financial statement counts for 63 industries from 2001-2017 and the data contain five of the six treated industries (the industry Broilers is not available in the RMA data).[6] Live Oak is not a participant in the RMA survey during our sample period, so the RMA data provide a proxy for total industry lending excluding Live Oak, i.e., the competitive effect.

Using the RMA industry-specific statement reports, we form annual counts of financial statements by industry (normalized by financial counts in 2006) and estimate treatment effects using a synthetic control for each industry (Internet Appendix Figure D.1). If Live Oak caused substitution from non-SBA to SBA lending, we would expect financial statements from these other lenders to fall. Instead, for most treated industries, the actual number of financial statements closely tracks the number predicted by the synthetic control in the post-period. Table D.1 columns 1-4 report the average treatment effect estimates (τ_j^j), the RMSPE ratio (r_j^j), and p-values for the RMA outcomes. With the null hypothesis as a decline in financial statements, we report left-tailed p-values in column 2 and one-sided r_j measures in column 3.[7] As seen in columns 1 and 2, there are no statistically significant declines in financial statements in the treated industries and the p-values indicate that more than 50% of the placebo industries experienced larger declines in lending. Similarly, the one-sided r statistics are generally insignificant in columns 3 and 4, though Funeral Homes is significant at the 5% level (but its average effect τ in columns 1 and 2 is small and insignificant). The p-values based on the joint inference procedure for both test statistics are insignificant, indicating no significant overall change in financial statements within the treated industries.

[5] For more detailed information on the participants and coverage of RMA's eStatement Studies, see Berger, Minnis and Sutherland (2017) and Lisowsky, Minnis and Sutherland (2017).

[6] Because we manually code the data from RMA, we selected a subset of industries from the SBA sample: the industries with at least 20 SBA loans per year and are able to be uniquely mapped from the 5-digit NAICS our the SBA analysis to the 6-digit NAICS in the RMA data. The resulting sample is 63 industries with complete data for 2001-2017.

[7] The one-sided r_i^j measure replaces $\left(Y_{it} - \widehat{Y}_{it}^j(0)\right)$ in the numerator of r_i^j with just its negative part (Abadie, 2021). Positive values for the numerator are coded as zero.

Overall, while we cannot directly examine non-SBA lending, the institutional features, external evidence, and the indirect test using financial statements all suggest that Live Oak's substitution from non-SBA lending is limited and unlikely to fully offset the observed growth in SBA lending within the treated industries.

Table D.1: **Impact on Proxy for Total Lending**

Competitive Effect: RMA Financial Statements

Industry	τ (1)	p-val. (2)	r (3)	p-val. (4)
Veterinarians	-0.33	(0.11)	21.80	(0.25)
Pharmacies	0.04	(0.77)	5.58	(0.35)
Dentists	0.59	(0.95)	0.00	(1.00)
Funeral Homes	-0.03	(0.54)	40.63	(0.02)
Investment Advice	-0.05	(0.56)	2.26	(0.44)
Joint Inference				
Joint Inf. (excl. Broilers)		(0.65)		(0.23)

This table reports the test statistics τ_j^j and r_j^j, as well as the p-values from the respective permutation distributions, when estimating a synthetic control on the outcome of counts of RMA financial statements (normalized by statement counts in 2006). The bottom two rows of each panel report p-values from the joint inference procedure using $B = 5,000$ random permutations. See Section 4.5 for details on the test statistics and inference procedures.

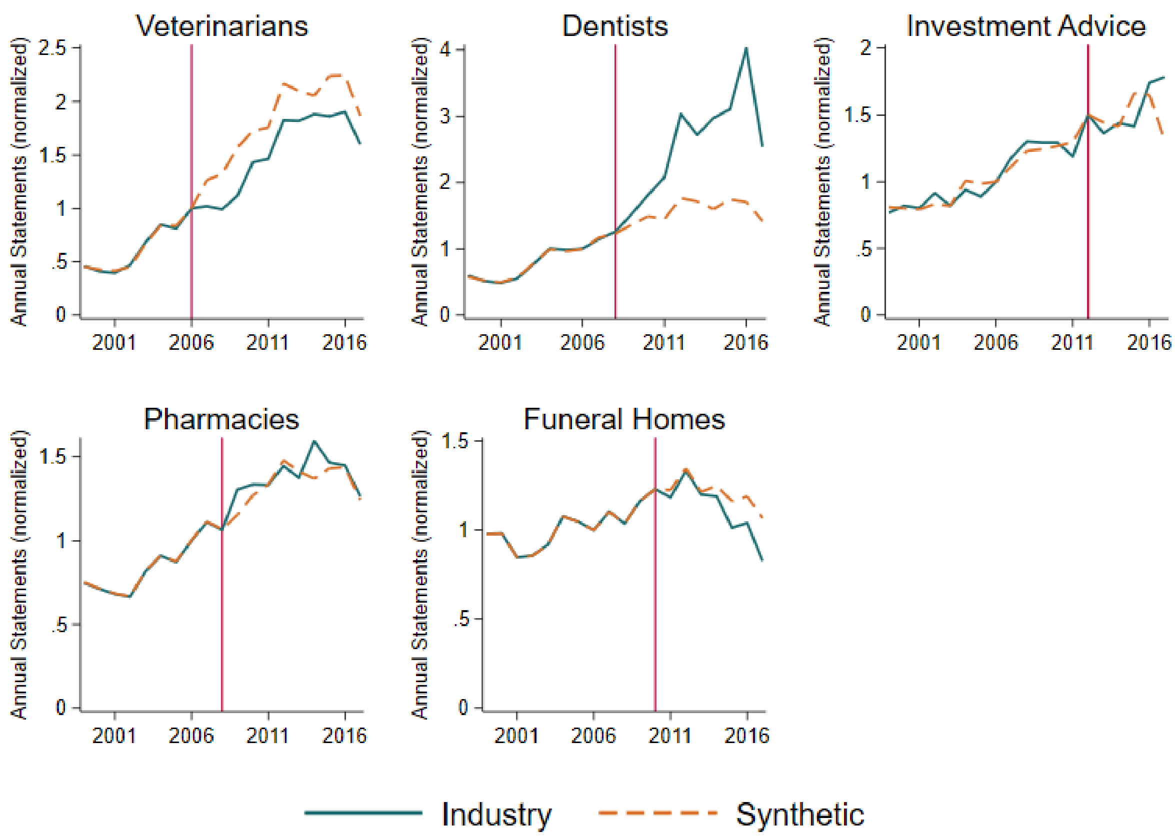


Figure D.1: **Synthetic Control using RMA Counts of Financial Statements**
This figure shows the change in counts of borrowers' financial statements collected by other lenders upon Live Oak's entry. The figure compares the number of statements collected in each industry (normalized by 2006 statement counts) that Live Oak enters to the normalized number of statements predicted by the synthetic control. The synthetic controls are formed by matching on all pre-treatment years beginning in 1999, with no additional covariates. The vertical line shows the year before Live Oak entered.
Source: The Risk Management Association's Annual eStatement Studies

E Appendix: Employment, Establishments, and Charge-offs

We examine the impact of Live Oak's entry on small business employment and charge-off rates. We use data on industry-level employment and establishment counts from the Quarterly Census of Employment and Wages (QCEW), which publishes data by NAICS code for workers in jobs covered by state unemployment insurance laws (95% of all jobs). We form national employment counts and establishment counts for businesses with 10-250 employees.[8] We drop industries where some employment counts for small businesses are not disclosed during the period 2001-2017, leaving a sample of 107 control industries and three treated industries (Veterinarians, Funeral Homes, and Broilers are dropped). We calculate 3-year charge-off rates by industry using information on charge-offs available in the SBA 7(a) data.[9] In the charge-off sample, we exclude Live Oak's loans in order to examine Live Oak's effects on the charge-off rates of other lenders and to investigate the possibility of cream-skimming by Live Oak would increase the charge-off rates of other lenders.

Table E.1 reports the average treatment effects, their p-values, and p-values for the r statistics for changes in employment (columns 1-3), establishments (columns 4-6), and charge-offs (columns 7-9).[10] Each outcome is normalized by the industry's 2006 values, so that the estimates of τ can be interpreted as percentage point changes (relative to the baseline of 2006). There is some evidence of an increase in employment and establishment for investment advice agencies (p-values less than 0.1), which is the industry where Live Oak's impact was largest. Overall, the changes in employment are jointly insignificant. Small effects on employment are consistent with the results of Brown and Earle (2017) which finds that SBA lending increased employment by only 3-3.5 jobs per million dollars in lending. Effect sizes of this magnitude would not be apparent in national employment counts. Finally, columns 7-9 show small and insignificant effects on charge-off rates of other lenders, consistent with the lack of cream-skimming.

[8] We choose these thresholds for small businesses because many counts for businesses with less than 10 or 250-499 employees are not disclosed.

[9] To calculate three-year charge-off rates for the full period 2001-2017, we expand the SBA data by merging charge-off data through 2020. Using business and lender names, addresses, and locations, we match 97% of loans to the more recent data.

[10] Internet Appendix Figure E.1 reports the synthetic controls.

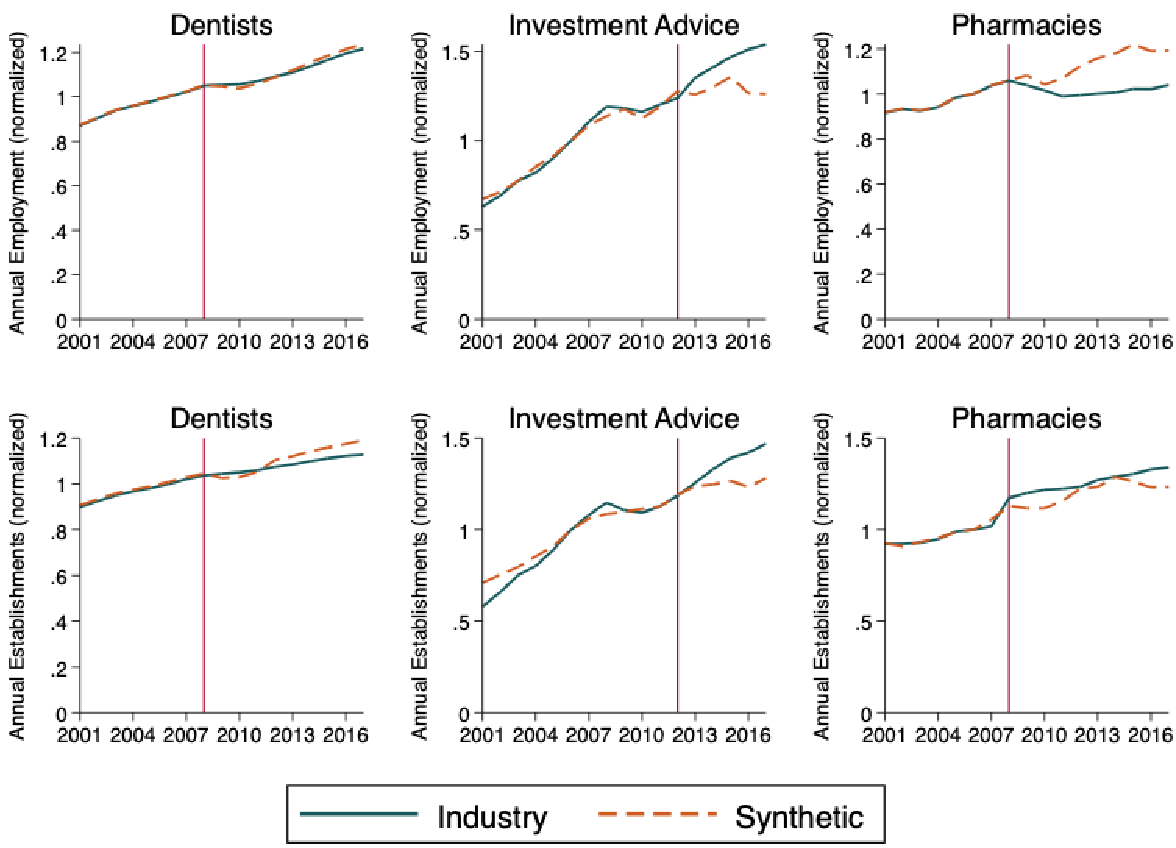


Figure E.1: **Impact on Employment and Establishment Counts**
Synthetic control estimates for employment (top row) and establishment counts (bottom row) of businesses with 10-250 employees (relative to 2006). Data are from the 2001-2017 Quarterly Census of Employment and Wages.

Table E.1: **Impact on Employment, Establishments, and Charge-offs**

	Industry Employment			Industry Establishments			3-Year Charge-off Rate ($\times$100)		
Industry	τ	p-val.	r p-val.	τ	p-val.	r p-val.	τ	p-val.	r p-val.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Veterinarians							-0.022	(0.49)	(0.16)
Pharmacies	-0.070	(0.38)	(0.20)	0.066	(0.28)	(0.86)	-0.005	(0.78)	(0.50)
Dentists	0.010	(0.81)	(0.72)	0.004	(0.96)	(0.94)	-0.012	(0.43)	(0.47)
Funeral Homes							-0.007	(0.49)	(0.47)
Investment Advice	0.141	(0.07)	(0.47)	0.103	(0.06)	(0.77)	-0.006	(0.55)	(0.67)
Broilers							-0.010	(0.41)	(0.67)
Joint Inference		(0.33)	(0.43)		(0.32)	(0.99)		(0.56)	(0.54)
Joint Inf. (excl. Broilers)		(0.31)	(0.42)		(0.32)	(0.99)		(0.60)	(0.45)

This table reports the test statistics τ_j^j p-values from the respective permutation distributions for τ_j^j and r_j^j for changes in employment counts, establishment counts, and three-year charge-off rates, all normalized so that 2006 values equals 1. Employment and establishment counts are from the QCEW for businesses with 10-250 employees. The bottom two rows of each panel report p-values from the joint inference procedure using $B = 5,000$ random permutations. See Section 4.5 for details on the test statistics and inference procedures.

References

Abadie, Alberto. 2021. "Using synthetic controls: Feasibility, data requirements, and methodological aspects." *Journal of Economic Literature*, 59(2): 391–425.

Acharya, Viral V, Iftekhar Hasan, and Anthony Saunders. 2006. "Should Banks Be Diversified? Evidence from Individual Bank Loan Portfolios." *The Journal of Business*, 79(3): 1355–1412. https://doi.org/10.1086/500679.

Berger, Philip G, Michael Minnis, and Andrew Sutherland. 2017. "Commercial Lending Concentration and Bank Expertise: Evidence from Borrower Financial Statements." *Journal of Accounting and Economics*, 64(2-3): 253–277. https://doi.org/10.1016/j.jacceco.2017.06.005.

Brown, J David, and John S Earle. 2017. "Finance and Growth at the Firm Level: Evidence from Sba Loans." *The Journal of Finance*, 72(3): 1039–1080. https://doi.org/10.1111/jofi.12492.

Hayden, Evelyn, Daniel Porath, and Natalja v Westernhagen. 2007. "Does Diversification Improve the Performance of German Banks? Evidence from Individual Bank Loan Portfolios." *Journal of Financial Services Research*, 32(3): 123–140. https://doi.org/10.1007/s10693-007-0017-0.

Lisowsky, Petro, Michael Minnis, and Andrew Sutherland. 2017. "Economic Growth and Financial Statement Verification." *Journal of Accounting Research*, 55(4): 745–794. https://doi.org/10.1111/j.1475-679x.2011.00411.x.