Peer Effects in Deposit Markets

By Kim Fe Cramer* and Naz Koont*

October 2021

We provide first empirical evidence that consumer peer effects matter for banks' deposit demand. Using a novel measure that depicts for each county how exposed peers are to a specific bank in a given year, we tightly identify the causal effect of peer exposure on deposit demand through a fixed effects identification strategy. We address key empirical challenges such as time-invariant homophily. We find that a one percent increase in a bank's peer exposure leads to a 0.05 percent increase in deposit market share. This effect has become stronger over time with the rise of the internet and social media, which facilitate cross-county communication. Peer exposure is especially relevant for smaller banks and customers that have access to the internet.

Keywords: Deposit Demand, Peer Effects, Banking
JEL Codes: G21, G41, G51

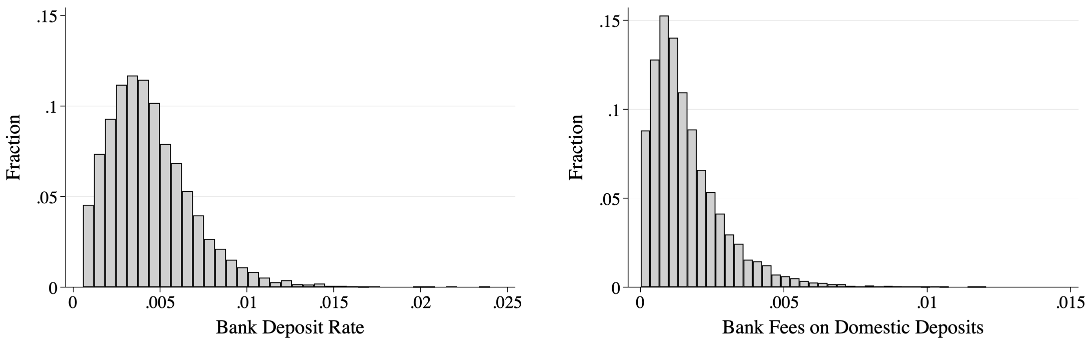
*Columbia Business School, Columbia University, 3022 Broadway, New York, NY 10027, USA (email: kcramer21@gsb.columbia.edu, www.kimfecramer.com; email: nkoont23@gsb.columbia.edu, www.nazkoont.com). We would like thank Xavier Giroud, Theresa Kuchler, Johannes Stroebel, and Jack Willis, as well as numerous seminar participants for their valuable comments. In particular, we would like to thank the Finance Faculty and Administrative Staff of Columbia Business School for their support. Researchers' own analyses calculated (or derived) based in part on data from The Nielsen Company (US), LLC and marketing databases provided through the Nielsen Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the Nielsen data are those of the researchers and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

I. Introduction

How do peer effects shape depositor behavior in the U.S. banking system? A customer's choice of where to open their bank account determines the rate of return they receive on their deposits and the service fees they are subject to. There is dispersion along these dimensions in the cross-section of banks (Figure 1), implying that customers have a meaningful decision to make. While bank choice is highly relevant, the complexity and opacity of bank account products implies that evaluating the trade-offs across different accounts is far from straightforward. This leads to lost returns and "hidden fees" for customers, with real economic consequences,[1] and makes deposit markets a natural setting for peer effects to play an important role. By relying on their social network, customers can potentially improve their decision, identifying a bank that best serves their needs. Indeed, survey evidence suggests that peer effects are highly important for bank choice. One in three U.S. citizens indicate that they speak to family and friends when researching where to open a checking account. They also assign these recommendations a high weight: personal recommendations are the third most important factor for choosing a provider, ranking at 24 percent, just behind fees and branch locality and far ahead of recommendations from professionals such as financial advisors (5 percent) (Forrester, 2018).

Understanding the role of peer effects in deposit markets is important for banks' business models, which rely heavily on their deposit franchise. Retail deposits represent over three-quarters of funding for U.S. commercial banks (Hanson et al., 2015). Further, banks maintain market power over their local deposit markets, allowing them to hedge their duration risk and obtain lower funding costs (Drechsler et al., 2018, 2021). Depending on their business model, banks may benefit from peer effects that increase customers' exposure to their services, or they may prefer more opacity about deposit market services in order to maintain their "sleepy" deposit base.

[1] Americans are spending a lot more at banks in ways that they cannot see https://www.ft.com/content/82232968-0433-357e-ace3-326fc7fc35f9. Banks quietly impose new fees to consumers https://www.nytimes.com/2011/11/14/business/banks-quietly-ramp-up-consumer-fees.html. Consumers face hidden overdraft charges from the nation's largest banks https://consumerfed.org/pdfs/CFAOverdraftStudyJune2005.pdf. All accessed on 11 July 2021.



(a) Dispersion in Bank Deposit Rates

(b) Dispersion in Bank Deposit Fees

Figure 1. Histograms for bank deposit rates and fees in 2018. Deposit rate is imputed by the rate expense per dollar of domestic deposits, and deposit fee is imputed by the fee income per dollar of deposits, using data from banks' Call Reports. Both variables are trimmed at the 1st and 95th percentile.

Despite the apparent prevalence of peer effects, we lack systematic empirical evidence that they matter for consumer choice in the complex and opaque deposit market. In order to fill this gap, we construct a novel measure that depicts for each county how exposed peers are to a specific bank in a given year. To identify peers, we utilize a county-to-county index of social connectedness based on the universe of Facebook friendship links (Bailey et al., 2018b). To measure how exposed peers in these other counties are to a given bank, we consider branch presence as reported in the Summary of Deposits (SOD) of the Federal Deposit Insurance Corporation (FDIC). Our peer exposure measure is bank-, county-, and year-specific. This granularity allows us to tightly identify the causal effect of peer exposure on deposit demand through a fixed effects identification strategy. In particular, we include county-year, bank-year, and bank-county fixed effects to control for many potential threats to identification. Importantly, we address the key empirical challenge of time-invariant homophily, that is, that peers tend to have similar preferences and thus may make similar choices even in the absence of influence from one another.[2] Additionally, we control for other drivers of deposit demand that vary on the county-bank-year level such as physical proximity to branches, and we conduct a variety of robustness tests.

[2]Homophily is a widely discussed empirical challenge in the peer effects literature (Manski, 1993).

We find that a one percent increase in peer exposure leads to a 0.05 percent increase in deposit market share for a bank. The effect is statistically significant at the one percent level and robust to various specifications, suggesting that peer effects are an important driver of deposit demand. We also find that the effect size has become stronger over time with the rise of the internet and social media, which facilitate cross-county communication. Conducting a heterogeneity analysis based on banks' Call Reports data, we find that peer exposure is especially relevant for smaller banks. These banks are potentially less likely to drive up demand through other determinants such as bank presence or advertising. Additionally, there is potentially more opacity for their products.

To study which consumers are more likely to rely on peer effects, we initially turn to the Survey of Consumer Finance (SCF). We find the clear prediction that younger people are more likely to be affected by peer effects: they are more likely to collect information from peers and to assign this information a higher weight when choosing an institution than older people. Predictions are ambiguous for richer or more educated people; they are more likely to collect information from peers but assign this information a lower weight. Motivated by the results of the SCF, we use our fixed effects identification strategy to analyze heterogeneity in importance of peer exposure for consumers. We measure county characteristics using the American Community Survey (ACS) of the Economic Census. Consistent with the narrative of peer effects, we initially observe that counties in which a higher share of the population has internet or computer access show a stronger effect. We also confirm that two groups for which peer effects are likely to be of high relevance indeed show stronger effects. Counties with higher shares of people speaking limited English or with a lower mean age show stronger peer effects. Finally, we learn that higher income and higher education are positive but insignificant predictors of the relevance of peer effects in the deposit market.

We conduct a battery of robustness tests to show that our findings are not driven by potential confounders. Those we control for include drivers of deposit demand such as a direct effect of bank presence, deposit rates and fees, and advertising campaigns. To obtain information on these dimensions, we utilize the SOD, banks' Call Reports, and

Nielsen data. We also test for robustness to identification threats arising from banks' specialization on certain industries or demographic segments. To control for industry trends, we utilize data from the Bureau of Labor Statistics. To control for demographic trends, we utilize the data from the ACS of the Economic Census. Finally, we show that our results are not driven by bank expansion or contraction along specializations.

We contribute to the literature on the role of peer effects in a wide range of financial and economic decisions, summarized by Kuchler and Stroebel (2020). With respect to financial decisions, researchers have provided evidence that peer effects impact investment behavior (Hong et al., 2004; Bursztyn et al., 2014; Ouimet and Tate, 2020; Kuchler et al., 2020a; Pool et al., 2015), housing market decisions (Bailey et al., 2018a, 2019), mortgage refinancing (Maturana and Nickerson, 2019), retirement plan decisions (Duflo and Saez, 2003; Beshears et al., 2015), and charitable giving (DellaVigna et al., 2012). Peer effects have also been shown to influence consumption decisions (Kuhn et al., 2011; Aral and Walker, 2012; Gilchrist and Sands, 2016; De Giorgi et al., 2016). This is the first paper to examine how peer effects shape decisions in deposit markets, which are of particular interest due to their reflection of banking relationships and outsize influence on bank funding.

That peer effects play an important role in bank choice is also in line with a large recent literature in banking that has emphasized depositors' rate insensitivity, suggesting that other factors[3] are more salient for bank choice (Dick, 2008; Egan et al., 2017; Drechsler et al., 2017; Xiao, 2020; Wang et al., 2020; Diamond et al., 2020). These studies use methods from empirical industrial organization (IO) (Berry, 1994; Berry et al., 1995) to examine drivers of deposit demand. Our empirical framework uses these insights to map deposit market shares to the discrete choices of depositors across differentiated banks and finds that bank peer exposure is a significant driver of demand.

[3]Other drivers of demand identified by the literature are presence of banks (Ho and Ishii, 2011) and advertisement (Honka et al., 2017).

The remainder of the paper is structured as follows. We describe our data sources in Section II. The construction of our peer exposure measure is shown in Section III.A followed by summary statistics in Section III.B. In Section IV.A, we specify our regression strategy, followed by a detailed discussion on identification in IV.B. We provide our main findings in Section V and evidence on robustness in Section VI. Results on heterogeneity are described over time in Section VII.A and for banks in Section VII.B. Section VII.C describes results from the SCF, motivating the heterogeneity on the county-level in Section VII.D. Finally, we conclude with Section VIII.

II. Data

To identify peers across the United States, we use the social connectedness index (SCI) provided by Bailey et al. (2018b). The authors construct this measure for county-pairs based on the universe of friendship links between Facebook users as of April 2016. The SCI is likely to reflect real-world social connections due to three reasons. First, Facebook is the worlds largest online social networking service. As of 2015, more than 68 percent of the U.S. adult population used Facebook. Second, the Facebook user population is relatively representative of the U.S. adult population; usage is nearly constant across income groups, education levels, as well as urban, rural, and suburban areas (Greenwood et al., 2016). Finally, Facebook is primarily used to connect real-world friends and acquaintances. Multiple other studies that use the social connectedness measure provide further evidence that it is a good proxy for real-world social connections (Bailey et al., 2018a, 2019, 2020a,b; Kuchler et al., 2020b). To construct the SCI, the authors map users to their county location using information such as the users' regular IP address. As a next step, they calculate the total number of friendship links between county i and county j ($\text{Friendships}_{i,j}$) and divide by the product of the respective total Facebook users (Users_i and Users_j) of the counties. Finally, the measure is scaled to have a fixed range by dividing through by the maximum value and multiplying by 1,000,000,000. The SCI reflects then the relative probability that a Facebook user in county i is friends with a Facebook user in county j.

$$\text{(1)} \qquad \mathrm{SCI}_{i,j} = \frac{\mathrm{Friendships}_{i,j}}{\mathrm{Users}_i \times \mathrm{Users}_j} \times \frac{1,000,000,000}{\mathrm{SCI}_{max}}$$

To measure how exposed peers are to a specific bank in a specific county in a given year, we obtain data on branch presence from the Federal Deposit Insurance Corporation (FDIC). The FDIC is an independent agency that was created by the Congress with the objective to insure deposits and examine financial institutions for safety, soundness, and consumer protection. The specific data set we use is the Summary of Deposits (SOD), an annual survey of branch office deposits. All institutions in the U.S. with branch offices are required to submit the survey, except institutions with only one main office. Besides providing information on branch locations, the SOD provides our main outcome of interest, the deposit market share of a specific bank in a specific county in a given year. Data is available on branches active between 1994 to 2020. We concentrate on the period between 1997 to 2018, for which data of the banks' Call Reports is available.

We complement the SCI Index and SOD data with multiple other data sets to conduct robustness tests and heterogeneity analysis. To conduct robustness tests related to bank rates and fees, we impute respective measures from bank Call Reports. The Consolidated Report of Condition and Income, referred to as the Call Reports, provides quarterly bank-level data for every U.S. insured commercial bank, including accounting information on deposits, incomes, expenses, and asset composition. Data from Call Reports is available for the period 1997 to 2018 and 92.76 percents of banks in our sample. To demonstrate that advertisement campaigns are not driving our results, we obtain data from NIELSEN Ad Intel, available for the years 2010 to 2018. Turning to heterogeneity, we utilize size measures from the Call Reports. To study for which consumers peer effects are particularly relevant, we turn towards the SCF. We utilize the 2019 version of the survey. It includes questions on how much individuals rely on peers for specific financial decisions, complemented by data on demographic and financial characteristics. Data is available for 5,777 households. We apply respective weights to make the sample representative of the U.S. population. Motivated by the results of the SCF, we use our fixed effects identification strategy to analyze heterogeneity in importance of peer effects for consumers. We measure county characteristics using the American Community Survey (ACS) of the

Economic Census. We obtain characteristics of the 5-year averages for the years 2010 to 2018. Characteristics are available for 99.53 percent of counties in our sample.

III. Peer Exposure

A. Measure

We are interested in measuring the peer exposure that a consumer has to a given bank in her local market at the time that she is choosing where to open her deposit account. We construct a novel measure that varies at the bank-county-year level. Initially, to identify peers of those living in county i, we utilize the $\mathrm{SCI}_{i,j}$ that captures the probability that an individual in county i is friends with any given person in county j. We multiply this probability with the population of county j to obtain the average number of friends that an individual in county i has in county j. Intuitively, the peer exposure that an individual faces towards a given bank scales with the number of friends that she has who have some information about this bank. Additionally, this conversion allows us to properly aggregate peer exposure across counties with different population of peers. In order to minimize the influence of population changes on our measure, in our base specification we use $\mathrm{Pop}_{j,2010}$, the county population reported in the 2010 Census. Notice that we focus only on out-of-county peers due to the granularity of the $\mathrm{SCI}_{i,j}$. While this data limitation ignores potentially important peers in the same county, our specification allows us to isolate a clean component of peer effects that is not likely to be confounded by other direct influences. If we would consider peers in the same county, any measure of how exposed these peers are is also likely to directly influence the demand of the individual of interest. As a next step, we want to measure how exposed peers in county j are to a given bank b. We thus multiply the number of friends that an individual in county i has in county j with the number of branches that bank b has in county j. Finally, we aggregate across all counties $j \neq i$. To summarize, we define $\mathrm{PeerExposure}_{b,i,t}$ for a given bank b in county i at year t as the weighted sum of out-of-county friends of an individual in county i, where the weights are the number of branches that bank b has in year t.

$$\text{PeerExposure}_{b,i,t} = \sum_{j \neq i} \text{SCI}_{i,j} \times \text{Pop}_{j,2010} \times \text{Branches}_{b,j,t} \tag{2}$$

Our measure $\text{PeerExposure}_{b,i,t}$ has several desirable properties of note. First, given the term $\text{SCI}_{i,j}$, the measure is increasing as the bank locates in counties that have higher social connectedness to county i, all else held constant, and therefore is indeed capturing the peer exposure of bank b in county i. Secondly, scaling the probability of friendship $\text{SCI}_{i,j}$ by $\text{Pop}_{j,2010}$ ensures that our measure is increasing in the population of counties in which the bank is present. This captures the notion that setting up a branch in a populous county will lead to higher peer exposure on average than setting up a branch in a sparsely-populated county, holding probability of friendship with a given individual in the county constant. Thirdly, due to the summation across counties $j \neq i$ and the branch weights $\text{Branches}_{b,j,t}$, our measure is increasing as bank b sets up more branches out-of-county, both in counties in which it is already present, and in new counties. This implies that all else equal, a bank with more branches will have higher peer exposure, and further that peer exposure never decreases as banks set up additional branches. The time variation in our measure is driven exactly by bank b opening and closing branches. Finally, all terms in our measure are independent from the behavior and branching decisions of other banks, avoiding confounding variation in the measure due to changes in the market structure of other counties.

For robustness, we consider multiple other specifications of peer exposure. First, we examine a simplified version that omits the population component. This essentially treats each county equally in terms of contribution towards peer effects, regardless of population size[4]. Second, instead of weighting by the number of branches, we weigh by the branch share of bank b in a given county, calculated as branches of bank b over the total number of branches of all banks in that county. Finally, we consider a version without the population component and with the branch share. We find that our findings are robust to all these different specifications of the measure.

[4]This alternative construction relatively over-weights the influence of rural counties.

B. Summary Statistics

Before examining variation in our peer exposure measure, we describe trends in the banking sector over our sample period. Table 1 provides summary statistics on the bank and county-level. We see that the number of banks in the U.S. has roughly halved from over 11,000 to less than 6,000 throughout the past two decades. The number of branches of a given bank has instead more than doubled from seven to 16. Similarly, over time the average bank has expanded from operating in two counties to operating in nearly five. Finally, the size of the average bank, measured by their asset size or deposit volume, has increased significantly over this time period. All of these trends capture the consolidation that has been occurring in the banking sector, leading to dynamic evolution of bank branch networks and the changing exposure of individuals to banks in their local market.

Next we turn towards summary statistics of the peer exposure measure. A given customer faces a choice *across banks* within a given market at a specific time. In the cross-section, banks with (1) more branches and that have branches (2) in more socially connected areas that are (3) more populous have higher peer exposure. In Figure 2, we observe that there is approximately a log-normal distribution of average log peer exposure across banks in a given year. Since we include year fixed effects in our main regression, we exploit how peer exposure for a given bank develops *over time*. This variation in peer

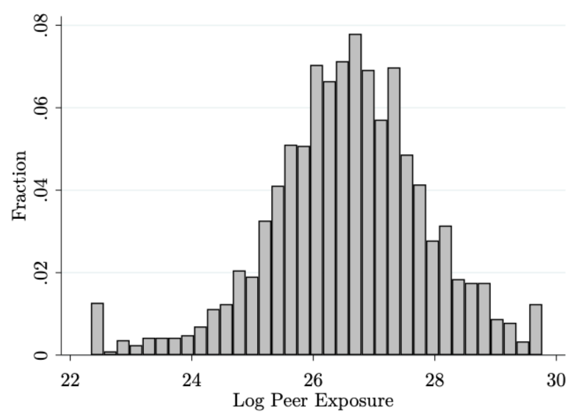


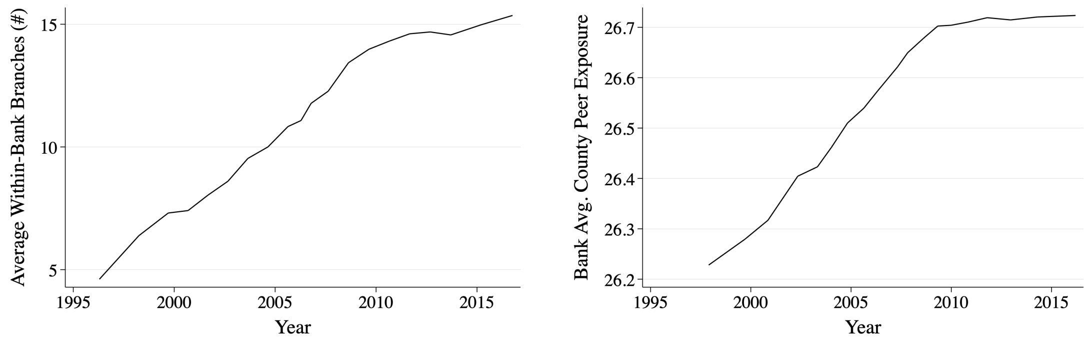
Figure 2. Dispersion of average log peer exposure across banks in 2018. The measure is winsorized at the 1st and 99th percentile.

exposure depends on (1) the evolution of the bank's branch network and (2) whether this evolution is occurring in more or less connected counties. In the general summary statistics, we have seen that there was a consolidation trend, with the average number of

Table 1: Summary Statistics

	1997	**2010**	**2018**
Banks			
Peer Exposure (e^{-12})	4.58	7.46	6.90
	(8.08)	(10.37)	(10.25)
Deposit Market Share	.1291	.1138	.1221
	(.1625)	(.1543)	(.1635)
Branches (#)	7.34	12.60	15.87
	(41.17)	(136.69)	(145.19)
Counties (#)	2.21	3.59	4.72
	(6.22)	(19.56)	(22.83)
Deposit Fee	.0046	.0034	.0016
	(.0025)	(.0023)	(.0013)
Deposit Rate	.0334	.0123	.0044
	(.0065)	(.0045)	(.0025)
Deposits (USD, millions)	302	1,019	2,212
	(2,438)	(16,857)	(32,529)
Assets (USD, millions)	533	1,875	3,169
	(6,211)	(36,474)	(50,083)
N	11,187	7,820	5,550
Counties			
Peer Exposure (e^{-12})	11.05	7.47	5.69
	(4.57)	(6.27)	(5.59)
Deposit Market Share	.2527	.2320	.2470
	(.2279)	(.2191)	(.2325)
Banks (#)	7.74	8.78	8.20
	(9.24)	(9.59)	(8.56)
Branches (#)	25.69	30.81	27.58
	(61.57)	(77.23)	(70.76)
N	3,210	3,210	3,210

This table presents summary statistics for banks and counties from 1997 to 2018. Standard deviation of variables is reported in parentheses. Peer exposure is constructed as described in section III.A. Deposit Fee is imputed by the fee income per dollar of deposits, and Deposit Rate is imputed by the rate expense per dollar of domestic deposits, using data from bank Call Reports.



(a) Within-Bank Branch Trend

(b) Within-Bank Peer Effects Trend

Figure 3. Average within-bank time trend of the number of branches and peer exposure. Peer exposure constructed as detailed in section III.A.

branches per bank increasing. Considering within-bank trends by exploiting bank fixed effects, we observe a steep increase in number of branches until a stagnation in the growth rate in 2010 (see Figure 3(a)). Figure 3(b) shows that the within-bank time trend of peer exposure reflects this trend in branch networks.

Finally, we can investigate the variation in peer exposure *across counties*. This variation depends on whether banks present in this county have (1) more branches in other counties and whether they these branches are located in areas with (2) higher social connectedness and (3) a larger population. The latter captures the increase the number of potential peers for a given probability of social connectedness. As seen in Figure 4, all three of these factors contribute to urban counties having a higher average level of peer exposure in comparison to rural areas: urban counties have larger banks and have a higher social connectedness to other urban areas with large populations.

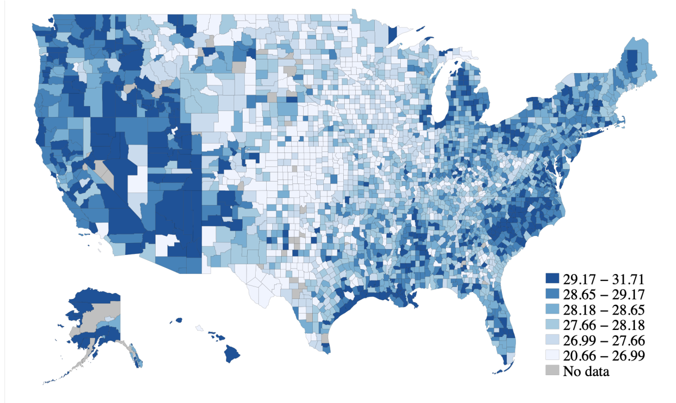


Figure 4. Average Peer Effects Across Counties in 2018. Constructed as detailed in section III.A. Darker colors correspond to regions with a higher average level of peer exposure.

IV. Empirical Analysis

A. Regression Specification

We are interested in how peer exposure to each bank in a given market affects a depositor's choice over these banks. In order to capture this notion using market-level deposit volume data, we follow the large recent deposit demand literature which uses tools from empirical industrial organization and particularly the methods in Berry et al. (1995) to map deposit market shares to the discrete choices of depositors across differentiated banks.[5] Following the literature, we define a market to be a county. In particular, we suppose that depositors in each county c at a time t have a fixed quantity of funds $F_{c,t}$ that they deposit at a bank b with branches in the county, or in an unobserved outside option such as a money market mutual fund. Preferences of depositors j in county c for bank b at time t follow a logit demand system, depicted in Equation 3.

$$U_{j,c,b,t} = \alpha_{PE}\text{LogPeerExp}_{c,b,t} + X_{c,b,t}\beta + \delta_{c,b,t} + \epsilon_{j,c,b,t} \tag{3}$$

Here α_{PE} measures customers' response to higher peer exposure to a given bank, $X_{c,b,t}$ denotes a vector of observable bank-county-time level bank characteristics, and $\delta_{c,b,t}$ captures unobservable bank-county-time characteristics. This formulation makes clear the

[5]Examples include Dick (2008); Xiao (2020); Egan et al. (2017); Diamond et al. (2020).

subtlety in measuring this elasticity. We expect that α_{PE} is significantly different than zero if peer effects matter due to the prevalence of information asymmetries in the deposit market. However, the direction of the response depends on whether the information provided by peers is positive or negative, and thus on the quality of a given bank. Under the assumptions of logit demand systems, the quantity of deposits invested in branches of bank b in market c at time t satisfies Equation 4.

$$(4) \qquad Q_{c,b,t} = F_{c,t} \frac{\exp(\alpha_{PE}\text{LogPeerExp}_{c,b,t} + X_{c,b,t}\beta + \delta_{c,b,t}}{1 + \sum_{c'} \exp(\alpha_{PE}\text{LogPeerExp}_{c',b,t} + X_{c',b,t}\beta + \delta_{c',b,t})}$$

Therefore, since the denominator in Equation 4 is constant across all banks in county c at time t, this demand system implies that α_{PE} and β can be estimated using the linear specification detailed in Equation 5. We regress the log of the deposit market share of a given bank b in county c at year t, denoted by $\text{LogDepMktShare}_{b,c,t}$, on the log of our peer exposure measure, denoted by $\text{LogPeerExp}_{b,c,t}$.

$$(5) \quad \text{LogDepMktShare}_{b,c,t} = \beta_0 + \beta_1 \text{LogPeerExp}_{b,c,t} + \eta_{c,t} + \eta_{b,t} + \eta_{b,c} + \gamma \text{X}_{b,c,t} + \epsilon_{b,c,t}$$

Due to the granularity of the outcome and our measure of peer exposure, we can include a rich set of fixed effects; county-year fixed effects ($\eta_{c,t}$), bank-year fixed effects ($\eta_{b,t}$), as well as bank-county fixed effects ($\eta_{b,c}$). As we will discuss, these fixed effects allow us to control for many threats to causal identification. However, they do not allow to control for variables that vary on the bank-county-year level. Consequently, we include other potential determinants of deposit demand with this variation discussed in the literature in a vector of controls ($\text{X}_{b,c,t}$). Following Diamond et al. (2020), we include lagged deposit market share to account for persistence in the stock of deposits. We cluster standard errors at the bank-county-year level. The main coefficient of interest is β_1, corresponding to α_{PE} in the utility specification. If peer effects play a role in deposit demand, we expect this coefficient to be different from zero and statistically significant. Before reporting the findings of our empirical strategy, we discuss what our specification allows us to control for and what are open potential confounders, corresponding to $\delta_{c,b,t}$ in our utility specification, which we address via additional robustness tests.

B. Identification Discussion

First, in our regression specification, *aggregate shocks* are not an identification issue. Assume for instance that there is a nationwide economic boom. As a result of the boom, households increase their deposits and banks open new branches, reflected in an increase in our peer exposure measure. Similarly, in an economic bust, households decrease their deposits and banks close branches, which translates into a decrease in our peer exposure measure. Crucially however, these aggregate shocks do not correlate with the outcome variable of interest. Even though overall deposits might increase or decrease, since by definition an aggregate shock affects all banks equally, deposit market shares do not move. Consequently, aggregate shocks do not pose a threat to identification in our specification.

Second, we control for *time-invariant county differences*. Assume for instance that there are two counties, New York County in New York and Kent County in Michigan. New York County is very urban. The average deposit market share of a single bank is low. Due to homophily, New York County is socially connected with many other counties that are very urban. In these counties, banks have a high number of branches, reflected in a high peer exposure through the branch component. Kent County on the other hand is rural. Here the average deposit market share of a single bank is high. Due to homophily, Kent County is socially connected with other counties that are rural. In these counties, banks have few branches, which translates into a low peer exposure through the branch component. If we would simply compare the deposit market share of a given bank in a given year in New York County (low market share, high peer exposure) to the deposit market share of a given bank in a given year in Kent County (high market share, low peer exposure), we would obtain a negative coefficient that picks up the underlying differences in how urban or rural counties are. To control for time-invariant county differences, we compare observations within a given county. We do not explicitly include county fixed effects, but county-year and bank-county fixed effects.

Third, we control for *time-varying county differences.* Assume that there are two counties, Kent County in Michigan and Allen County in Indiana. Now in 1998, the state of Michigan decides to deregulate its banking sector. As a result, new banks enter and the average deposit market share of a given bank decreases. Due to homophily, other states that Michigan is socially connected to also deregulate their banking sector. This allows banks to open more branches, increasing peer exposure of Kent County. On the other hand, there was no change in regulation for Allen County. Due to homophily, other states that Allen County is socially connected to also do not implement deregulation. Their number of branches does not significantly increase, not resulting in a surge in peer exposure. If we would simply compare deposit market shares of a given bank in Kent County in 1998 (low market share, high peer exposure) to deposits of a bank in Kent County in 1997 (high market share, low peer exposure), for instance by only including county-fixed effects, we would obtain a negative coefficient that picks up the deregulation. If we would simply compare deposit market shares of a given bank in Kent County in 1998 (low market share, high peer exposure) to deposit market share of a given bank in Allen County in 1998 (high market share, low peer exposure), for instance by only including year-fixed effects, we would also obtain a negative coefficient that picks up the deregulation. To control for time-varying county differences, we compare observations within a given county, within a given year. We include county-year fixed effects in our regression specification.

Fourth, we control for *time-invariant bank differences.* Assume that there are two banks, Big Bank and Small Bank. Big Bank has high deposit market shares and many branches. Small Bank has low deposit market shares and few branches. If we would simply compare deposit market shares of Big Bank in a given county in a given year (high market share, high peer exposure) with deposit market shares of Small Bank in a given county in a given year (low market share, low peer exposure), we would obtain a positive coefficient that captures the underlying bank differences. To control for time-invariant bank differences, we compare observations within a given bank. We do not explicitly include bank fixed effects, but control for bank-year fixed effects.

Fifth, we control for *time-varying bank differences.* Assume that there are two banks, Growth Bank and Stagnant Bank. Now in 1998, Growth Bank grows quickly, for instance because it hires new management. The bank increases its deposit market shares and sets up new branches. An increase in branches reflects in an increased peer exposure measure for Growth Bank. In contrast, Stagnant Bank experiences no growth activities. If we would simply compare deposit market shares of Growth Bank in a given county in 1998 (high market share, high peer exposure) to deposit market shares of Growth in a given county in 1997 (low market share, low peer exposure), for instance by only including bank-fixed effects, we would obtain a positive coefficient that picks up the growth activity. If we would simply compare deposit market shares of Growth Bank in a given county in 1998 (high market share, high peer exposure) to deposit market shares of Stagnant Bank in 1998 (low market share, low peer exposure), for instance by only including year-fixed effects, we would also obtain a positive coefficient that picks up the growth activity. To control for time-varying bank differences, we compare observations within a given bank, within a given year. In our regression specification, we include bank-year fixed effects.

Sixth, we control for *time-invariant bank-county* differences. Assume that there are two banks in New Haven in Connecticut, Education Bank and Other Bank. New Haven has many students, a customer segment Education Bank is specialized on. Consequently, Education Bank has a high deposit market share. New Haven is socially connected with other counties that have many students due to homophily. Also in these other counties, Education Bank is present, with many branches. Consequently, Education Bank in New Haven has a high peer exposure. Other Bank on the other hand is not specialized on students, has a low deposit market share in New Haven and low presence in socially connected counties. If we would simply compare deposit market shares of Education Bank in a given year in New Haven (high market share, high peer exposure) with deposit market shares of Other Bank in a given year in New Haven (low market share, low peer exposure), for instance by including only county fixed effects, we would obtain a positive coefficient that captures the bank-county match. Now assume that there are two counties, New Haven in Connecticut and Blanco County in Texas. Since Blanco County has few students, Education Bank has a low deposit market share and little presence in socially connected counties. If we would simply compare deposit market shares of Education Bank in a given year in New Haven (high market share, high peer exposure) with deposit market shares of Education Bank in Blanco County (low market share, low peer exposure), we would again obtain a positive coefficient that captures the bank-county match. To control for time-invariant bank-county differences, we compare observations within a given bank, within a given county. We include bank-county fixed effects in our regression.

Seventh, we control for a *direct effect of bank presence*. While peer effects potentially drive deposit demand, the literature has established that bank presence is a strong determinant (Ho and Ishii, 2011). We take multiple approaches to make sure that our effect is not driven by a direct effect of bank presence on deposit demand. Initially, we consider bank presence in the county of interest. High bank presence in the county of interest is likely to be valued by customers and therefore associated with a large deposit market share. Additionally, bank presence in the county of interest might positively correlate with bank presence in other counties and thereby with our peer exposure measure. To control for bank presence in the county of interest, we include two control variables: the total number of branches of a given bank in a given year and the branch density by population of a given bank in a given year. Next, we consider bank presence in other counties. Customers are likely to move between counties, for instance for their work commute. This means they are likely to factor in bank presence in surrounding counties into their decision which bank to choose. Additionally, bank presence in other counties directly correlates with peer exposure due to construction of our measure. We take two approaches to control for bank presence in other counties. First, we calculate a measure of physical proximity to branches[6]. This measure is constructed very similarly to our peer exposure measure, but the SCI component is replaced by a physical proximity component. Consistent with a strong relationship between social connectedness and physical proximity, the both measures have a correlation of 0.89. Importantly, social connectedness is shaped by other factors than physical proximity such as how urban an area is or whether there was migration in the past. Such variation helps us to distinguish the effect of peer exposure from simply physical proximity to branches. Additionally, we conduct robustness tests

[6] We follow the approach of Kuchler et al. (2020a).

to control for bank presence in other counties. We exclude any neighboring counties or counties in the commuting zone when constructing the peer exposure measure. Evidence from all these approaches indicates that a direct effect of bank presence on deposit market shares is not driving our results.

$$\text{PhysProximityBranches}_{b,i,t} = \sum_{j \neq i} \frac{1}{(1 + \text{Distance}_{i,j})} \times \text{Pop}_{j,2010} \times \text{Branches}_{j,b,t} \tag{6}$$

Our empirical strategy leaves only potential for confounders that vary on the bank-county-year level and are both correlated with deposit market share and peer exposure. In order to be correlated with peer exposure, the confounder must induce branch openings or closures. The other components of peer exposure do not vary on the bank-county-year level and are consequently controlled for. We provide an extensive battery of robustness tests to provide evidence that these kind of confounders are not driving our effect.

V. Findings

In our main regression specification we are interested in the role that peer effects play in deposit demand. Findings are reported in Table 2. Column 1 provides the result of the main specification in Equation 5. We find that a one percent increase in peer exposure increases the market share of the respective bank by 0.05 percent. This result indicates the *existence* of peer effects, and further that on average depositors are more likely to choose a bank for which they have higher peer exposure. In principle, peer effects could affect market share positively or negatively depending on the underlying mechanism. If peer effects serve to mitigate information frictions in understanding the characteristics of a given bank, then the effect that peer exposure has on a given bank should depend on the relative quality of that bank along rates and fees or other non-price dimensions. Peer effects may also be positive if they capture the additional utility that depositors get from using the same bank as their peers. A positive and significant coefficient is consistent with the alleviation of information frictions if peers tend to talk more about banks which they have a positive experience with as well as with a direct utility benefit of using the same bank as ones peers.

Table 2: Peer Exposure Main Regression

	LogDepMktShare				
	(1)	(2)	(3)	(4)	(5)
LogPeerExp	0.05***	0.05***	0.03***	0.03***	0.04***
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)
County x year FE	Yes	No	Yes	Yes	No
Bank x year FE	Yes	Yes	No	Yes	No
Bank x county FE	Yes	Yes	Yes	No	No
Lagged LogDepMktShare	Yes	Yes	Yes	Yes	Yes
Direct effect controls	Yes	Yes	Yes	Yes	Yes
R-squared	0.99	0.99	0.99	0.98	0.97
Observations	422,253	431,083	428,634	426,973	439,681
Banks	6,321	6,406	6,674	6,830	7,168
Counties	2,982	3,184	2,994	2,994	3,186
Years	21	21	21	21	21

This table presents the panel regression results on how peer exposure affects the deposit market share of a bank. Our sample consists of bank-county-year observations between 1997 and 2018. The dependent variable is LogDepMktShare is the logged deposit market share, which is the percentage of the amount in the deposit market held by a specific bank in a given county in a given year. The main independent variable is LogPeerExp, the logged peer exposure that is defined as $\sum_{j \neq i} \mathrm{SCI}_{i,j}$ x $\mathrm{Pop}_{j,2010}$ x $\mathrm{Branches}_{j,b,t}$. Lagged LogDepMktShare is the same variable as the outcome, but measured in the previous year. Direct effect controls include the total number of branches of a given bank in a given county, the density of branches by population of a given bank in a given county, as well as a measure of physical proximity to branches in other counties of a given bank in a given county. Standard errors are clustered by bank-county-year level and are reported in parentheses below each estimate. ***, **, and * indicate significance levels of 10%, 5%, and 1% respectively.

Our main result is highly robust to omitting fixed effects, suggesting that respective potential threats did not play a crucial role. In Column 2 to 4, we successively exclude fixed effects that account for time-varying county differences, time-varying bank differences, and time-invariant county-bank differences. Coefficients remain highly statistically significant and stable in size. For these specifications, we find that a one percent increase in peer exposure increases the market share of a bank by 0.03 to 0.05 percent. In Column 5, we exclude all fixed effects. Again, the coefficient remains highly statistically significant and stable. A one percent increase in peer exposure increases the market share of a bank by 0.04. That the finding is robust to exclusion of fixed effects will help us to address potential confounders, relaxing general fixed effects and including specific controls such as bank-year level deposit rates and fees.

VI. Robustness

A. A Direct Effect of Bank Presence

We conduct a battery of tests to demonstrate the robustness of our main result. Initially, we are interested in the potential threat that bank presence in nearby counties directly affects deposit market share of a bank. Customers might not only value a high bank presence in their own county but also in surrounding areas, for instance if they move between counties for work. Due to construction, bank presence in other counties also correlates with our peer exposure measure. In our main regression, we control for physical proximity to branches. In this robustness section, we take a second approach. We exclude counties in construction of our peer exposure measure that are located too close to counties of interest. Findings are depicted in Table 3. In Column 1, we exclude all adjacent counties when constructing the peer exposure measure. The coefficient remains statistically significant. A one percent increase in peer exposure, excluding adjacent counties, increases the deposit market share of the respective bank by 0.03 percent. As expected, the effect is slightly smaller than in the main specification as we consider a smaller sample of peers. In Column 2, we exclude all counties that are in the same commuting zone. Again, the coefficient remains statistically significant as expected. A one percent increase in peer exposure, excluding counties in the same commuting zones, increases the market share of the bank by 0.03 percent. Again, as expected, the effect is slightly smaller than in the main specification.

Table 3: Robustness: Direct Effect of Bank Presence

	LogDepMktShare	
	Exclude counties that are adjacent	Exclude counties in commuting zone
LogPeerExp	0.03***	0.03**
	(0.01)	(0.01)
County x year FE	Yes	Yes
Bank x year FE	Yes	Yes
Bank x county FE	Yes	Yes
R-squared	0.99	0.99
Observations	339,037	351,025
Banks	3,986	4,144
Counties	2,846	2,865
Years	21	21

This table shows that the main result is robust to a direct effect of bank presence. Our sample consists of bank-county-year observations between 1997 and 2018. The dependent variable is LogDepMktShare is the logged deposit market share, which is the percentage of the amount in the deposit market held by a specific bank in a given county in a given year. The main independent variable is LogPeerExp, the logged peer exposure that is defined as $\sum_{j \neq i} \text{SCI}_{i,j}$ x $\text{Pop}_{j,2010}$ x $\text{Branches}_{j,b,t}$. Lagged LogDepMktShare is the same variable as the outcome, but measured in the previous year. Direct effect controls include the total number of branches of a given bank in a given county, the density of branches by population of a given bank in a given county, as well as a measure of physical proximity to branches in other counties of a given bank in a given county. Physical proximity to branches excludes counties in that are adjacent or in the respective commuting zone respectively. Standard errors are clustered by bank-county-year level and are reported in parentheses below each estimate. ***, **, and * indicate significance levels of 10%, 5%, and 1% respectively.

B. *Rates and Fees*

Although the literature has found that bank depositors are relatively insensitive to prices, a key demand elasticity that is considered in frameworks of deposit demand is the elasticity with respect to rates and fees (Egan et al., 2017; Xiao, 2020). A bank with higher rates or lower fees might attract more depositors; the high market share might allow it to grow into (adjacent) counties by opening more branches. The bank-year level fixed effect in our main regression precludes us from including bank-level deposit rates imputed from the regulatory Call Reports, but it is of interest to confirm that given banks with identical rates and fees, depositors are more likely to choose the bank for which they have a higher peer exposure. We confirm this result below in Table 4. Column 1 shows initially that our result is robust to relaxing the bank-year fixed effect. Note that since county-year and bank-county FE are still included, we still control for time-invariant bank differences and time-varying aggregate shocks. In Column 2 and 3 we include bank-year level deposit rates and fees. In this specification, we additionally assume that there are

Table 4: Robustness: Inclusion of Rates and Fees

	LogDepMktShare		
	No rate or free control (1)	Rate control (2)	Fee control (3)
LogPeerExp	0.03***	0.04***	0.04***
	(0.01)	(0.01)	(0.01)
County x year FE	Yes	Yes	Yes
Bank x year FE	No	No	No
Bank x county FE	Yes	Yes	Yes
Lagged LogDepMktShare	Yes	Yes	Yes
Direct effect controls	Yes	Yes	Yes
R-squared	0.99	0.99	0.99
Observations	428,634	386,769	386,769
Banks	6,674	6,215	6,215
Counties	2,994	2,990	2,990
Years	21	21	21

This table shows that the main result is robust to inclusion of rates and fees. Our sample consists of bank-county-year observations between 1997 and 2018. The dependent variable is LogDepMktShare is the logged deposit market share, which is the percentage of the amount in the deposit market held by a specific bank in a given county in a given year. The main independent variable is LogPeerExp, the logged peer exposure that is defined as $\sum_{j \neq i} \text{SCI}_{i,j}$ x $\text{Pop}_{j,2010}$ x $\text{Branches}_{j,b,t}$. Lagged LogDepMktShare is the same variable as the outcome, but measured in the previous year. Direct effect controls include the total number of branches of a given bank in a given county, the density of branches by population of a given bank in a given county, as well as a measure of physical proximity to branches in other counties of a given bank in a given county. Standard errors are clustered by bank-county-year level and are reported in parentheses below each estimate. ***, **, and * indicate significance levels of 10%, 5%, and 1% respectively.

no other unobservable characteristics across banks in a given county that correlate with both peer exposure and deposit market shares. We find given banks with identical rates and fees, depositors are more likely to choose a bank with higher peer exposure.

C. Advertisement Campaigns

One potential threat to identification is that banks run advertisement campaigns. Assume that Flashy Bank runs a local advertisement campaign in the Designated Market Area (DMA) of Memphis including counties in Arkansas, Mississippi, Missouri, and Tennessee. DMAs are around 210 regions in the U.S. that are used to define advertisement markets. They are the smallest unit available at which companies can purchase advertisement. The advertisement campaign of Flashy Bank might increase the deposit market share in Carroll County. Additionally, if it is profitable enough, it might induce branch opening of Flashy Bank in other counties within the same DMA. This would reflect in an increase in the peer exposure measure of Carroll County. Note that the advertisement

Table 5: Robustness: Advertisement Campaigns

	LogDepMktShare			
		2010-2018		
	Excluding counties in same dma (1)	No ad control (2)	Dummy ad control (3)	Amount ad control (4)
LogPeerExp	0.02***	0.05**	0.05**	0.05**
	(0.01)	(0.02)	(0.02)	(0.02)
County x year FE	Yes	Yes	Yes	Yes
Bank x year FE	Yes	Yes	Yes	Yes
Bank x county FE	Yes	Yes	Yes	Yes
R-squared	0.99	0.99	0.99	0.99
Observations	310,944	200,705	200,705	200,705
Counties	2,741	2,949	2,949	2,949
Banks	3,179	4,192	4,192	4,192
Years	21	9	9	9

This table shows that the main result is robust to advertisement campaigns. Our sample consists of bank-county-year observations between 1997 and 2018. The dependent variable is LogDepMktShare is the logged deposit market share, which is the percentage of the amount in the deposit market held by a specific bank in a given county in a given year. The main independent variable is LogPeerExp, the logged peer exposure that is defined as $\sum_{j \neq i} \mathrm{SCI}_{i,j}$ x $\mathrm{Pop}_{j,2010}$ x $\mathrm{Branches}_{j,b,t}$. Lagged LogDepMktShare is the same variable as the outcome, but measured in the previous year. Direct effect controls include the total number of branches of a given bank in a given county, the density of branches by population of a given bank in a given county, as well as a measure of physical proximity to branches in other counties of a given bank in a given county. Standard errors are clustered by bank-county-year level and are reported in parentheses below each estimate. ***, **, and * indicate significance levels of 10%, 5%, and 1% respectively.

campaign must induce branch opening, which is a high barrier. We apply three strategies to show that our effect is robust to advertisement campaigns. The first strategy takes care of advertisement campaigns that are run in a single DMA. We exclude counties from the regression that are in the same DMA as the county of interest. Column 1 of Table 5 shows the results of this robustness test. We observe that a one percent increase in peer exposure, excluding peers in counties within the same DMA, increases the market share of the respective bank by 0.02 percent. As expected, the effect is slightly smaller than in the main specification as we consider a smaller sample of peers. Next, we address advertisement campaigns that are run in multiple DMAs. For this purpose we utilize the detailed NIELSEN Ad Intel data. It shows for all banks that ran advertisement campaigns in a given year in which DMAs they advertised and how much they spent. We match the SOD data and the Ad Intel data through fuzzy bank name matching. The Ad Intel data is only available from 2010 onwards, which is why in Column 2 we confirm that our effect holds in this period. In Column 3, we control for whether there was any advertisement spending for a given bank in a given county in a given year. In Column 4, we control for the respective amount. We find that our effect remains statistically significant and stable. This suggests that also advertisement campaigns run in multiple DMAs do not drive our effect. Note that any advertisement campaigns that are run nationwide are controlled for by bank-year fixed effects.

D. Bank Specialization on Industries

To understand the remaining identification threats, one has to consider bank specialization. Banks might specialize on certain industries or certain customer segments. Initially, any trend in these could give rise to an identification threat. To see this, first consider industry booms or busts. Assume that Soft Bank in the County of San Francisco is specialized on the software industry. Now there is an software industry boom in this county and Soft Bank's market share increases. At the same time, the County of San Francisco is socially connected with other counties that have a large software industry. As these areas also experience booms in the software industry, Soft Bank opens more branches there, resulting in increased peer exposure. Note that the industry shocks must induce branch opening or closure to pose a threat to identification. Optimally, we would want to control for the bank-industry match, comparing banks that equally rely on a certain industry. Since this data is unavailable, we control for industry trends, comparing counties that experience equal industry trends. We derive data from the Bureau of Labor Statistics (Quarterly Census of Employment and Wages) for the period of our sample, 1997 to 2018. We relax county-year fixed effects and instead control for industry-specific total number of establishments, total employment, employment location quotient relative to the U.S., total annual wages, and wage location quotient relative to the U.S. Location quotients compare the concentration of an industry within a county to the concentration of that industry nationwide. Industries are classified according to the North American Industry Classification System (NAICS)[7]. Results are depicted in Table 6. Column 1 demonstrates the main specification, just excluding county-year fixed effects. We see that the excluding county-year effects does not significantly alter the coefficient; a one percent increase in

[7] We control for the following industries: Professional and business services, education and health services, leisure and hospitality, goods-producing, natural resources and mining, construction, manufacturing, service-providing, trade, transportation, and utilities, financial activities, professional and business services, information.

Table 6: Robustness: Industry Boom and Busts

	LogDepMktShare					
		Controls for each industry				
	No industry control (1)	Number establ. (no.) (2)	Total employment (no.) (3)	Empl. quotient (ratio) (4)	Total wages (USD) (5)	Wage quotient (ratio) (6)
LogPeerExp	0.05***	0.05***	0.05***	0.05***	0.05***	0.05***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
County x year FE	No	No	No	No	No	No
Bank x year FE	Yes	Yes	Yes	Yes	Yes	Yes
Bank x county FE	Yes	Yes	Yes	Yes	Yes	Yes
Lagged LogDepMktShare	Yes	Yes	Yes	Yes	Yes	Yes
Direct effect controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry controls	No	Yes	Yes	Yes	Yes	Yes
R-squared	0.99	0.99	0.99	0.99	0.99	0.99
Observations	431,083	414,727	414,727	414,727	414,727	414,727
Banks	6,406	6,252	6,252	6,252	6,252	6,252
Counties	3,184	3,121	3,121	3,121	3,121	3,121
Years	21	21	21	21	21	21

This table shows that the main result is robust to industry booms and busts. Our sample consists of bank-county-year observations between 1997 and 2018. The dependent variable is LogDepMktShare is the logged deposit market share, which is the percentage of the amount in the deposit market held by a specific bank in a given county in a given year. The main independent variable is LogPeerExp, the logged peer exposure that is defined as $\sum_{j \neq i} \text{SCI}_{i,j}$ x $\text{Pop}_{j,2010}$ x $\text{Branches}_{j,b,t}$. Lagged LogDepMktShare is the same variable as the outcome, but measured in the previous year. Direct effect controls include the total number of branches of a given bank in a given county, the density of branches by population of a given bank in a given county, as well as a measure of physical proximity to branches in other counties of a given bank in a given county. Standard errors are clustered by bank-county-year level and are reported in parentheses below each estimate. ***, **, and * indicate significance levels of 10%, 5%, and 1% respectively.

peer exposure results in a 0.05 percent increase in deposit market share. For this specification, we can now include industry controls under the additional assumption that there are no other unobservable time-varying characteristics across counties that correlate with both peer exposure and deposit market shares. As expected, we observe that the coefficient remains highly statistically significant and very stable, so that holding fixed industry growth across counties, a given bank attracts higher deposit market share in counties with higher peer exposure. This suggests that general industry booms or busts seem to not be driving the effect.

E. Bank Specialization on Customer Segments

Banks might not only specialize on certain industries, but also on certain customers. Assume that Senior Bank targets retirees. In Orange County in California, the average

Table 7: Robustness: Demographic Trends

	LogDepMktShare			
		Demographic control		
	No demographic control	Medium income (USD)	High school or above (%)	Mean age (years)
	(1)	(2)	(3)	(4)
LogPeerExp	0.05**	0.05**	0.05**	0.05**
	(0.02)	(0.02)	(0.02)	(0.02)
County x year FE	No	No	No	No
Bank x year FE	Yes	Yes	Yes	Yes
Bank x county FE	Yes	Yes	Yes	Yes
Lagged LogDepMktShare	Yes	Yes	Yes	Yes
Direct effect controls	Yes	Yes	Yes	Yes
Demographic control	No	Yes	Yes	Yes
R-squared	0.99	0.99	0.99	0.99
Observations	203,729	203,630	203,703	203,703
Banks	4,240	4,239	4,240	4,240
Counties	3,178	3,178	3,178	3,178
Years	9	9	9	9

This table shows that the main result is robust to demographic trends. Our sample consists of bank-county-year observations between 2010 and 2018. This is the time frame for which demographic variables are available. The dependent variable is LogDepMktShare is the logged deposit market share, which is the percentage of the amount in the deposit market held by a specific bank in a given county in a given year. The main independent variable is LogPeerExp, the logged peer exposure that is defined as $\sum_{j \neq i} \mathrm{SCI}_{i,j}$ x $\mathrm{Pop}_{j,2010}$ x $\mathrm{Branches}_{j,b,t}$. Lagged LogDepMktShare is the same variable as the outcome, but measured in the previous year. Direct effect controls include the total number of branches of a given bank in a given county, the density of branches by population of a given bank in a given county, as well as a measure of physical proximity to branches in other counties of a given bank in a given county. Standard errors are clustered by bank-county-year level and are reported in parentheses below each estimate. ***, **, and * indicate significance levels of 10%, 5%, and 1% respectively.

population age increased over the past years, for instance because younger people are moving away. A very similar trend can be observed in counties that are socially connected to Orange County. As a result of this demographic trend, Senior Bank increases its deposit market share in Orange County and finds it profitable enough to open branches in socially connected counties, which increases peer exposure. Demographic trends could also arise from migration. For instance, if many people in the software industry move from New York County to the County of San Francisco, this might increase the deposit market share of Soft Bank in the County of San Francisco and potentially result in branch closure in New York County[8]. In our main specification, we control for general demographic trends with county-year fixed effects. However, county-year fixed effects do not capture the variation in bank-county-year level of how much business a bank makes in a

[8]Note that any general migration patterns unspecific to a customer segment targeted by banks do not pose a threat to identification, since they do not correlate with market shares, only overall deposit volumes.

certain county in a given year due to a demographic trend. As a robustness test, we relax our county-year fixed effects in Table 7 and control for various time-varying demographic measures. By doing so, we are able to explicitly control for demographic trends under the assumption that there are not other time-varying unobservable county characteristics that correlate with both our peer exposure measure and deposit market shares. Thus, we are testing whether holding fixed demographic composition across counties, a given bank attracts higher deposit market share in counties with higher peer exposure. We find that our coefficient of interest remains statistically significant and positive, supporting that bank specialization on customers does not drive our finding.

F. Expansion or Contraction Along Specializations

If banks expand or contract, this is captured by our bank-year fixed effects. However, banks might also expand or contract along specifications. Assume for example that Soft Bank grows and it does so specifically in counties with a high presence of software industry. In its growth process, it sets up branches both in the County of San Francisco and in socially connected counties with high software industry presence. This increases both the market share in the County of San Francisco and the measure of peer exposure. This industry-specific growth process varies on the bank-county-year level and is thus not captured by our fixed effects. Note that in this growth process, the correlation with deposit market share comes from a direct effect of newly opened branches of Soft Bank in the County of San Francisco. In contrast, for peer effects, only branches built in socially connected counties are of interest. We thus control for the number of branches opened (or closed) by Soft Bank in the County of San Francisco. In particular, we include a variable in our main regression specification that measures the change in branches for a given bank in the county of interest in Column 1 of Table 8. In our main specification, we already control for the *level* of bank presence, here we now control for the *trend*. We find that the coefficient remains highly statistically significant and stable. An increase in deposit market share could also arise due to a direct effect from banks set up in nearby counties. In Column 2, we thus control for the change in branches in neighboring counties. Again, the coefficient remains statistically significant and stable. Finally, note that the change

Table 8: Robustness: Expansion and Contraction Along Specializations

	LogDepMktShare		
	Control for change in branches		
	in county of interest (1)	in neighboring counties (2)	in county of interest due to merger (3)
LogPeerExp	0.05***	0.04***	0.05***
	(0.01)	(0.01)	(0.01)
County x year FE	Yes	Yes	Yes
Bank x year FE	Yes	Yes	Yes
Bank x county FE	Yes	Yes	Yes
Lagged LogDepMktShare	Yes	Yes	Yes
Direct effect controls	Yes	Yes	Yes
R-squared	0.99	0.99	0.99
Observations	422,253	422,253	422,253
Banks	6,321	6,321	6,321
Counties	2,982	2,982	2,982
Years	21	21	21

This table shows that the main result is robust to expansion and contraction along specializations. Our sample consists of bank-county-year observations between 1997 and 2018. The dependent variable is LogDepMktShare is the logged deposit market share, which is the percentage of the amount in the deposit market held by a specific bank in a given county in a given year. The main independent variable is LogPeerExp, the logged peer exposure that is defined as $\sum_{j \neq i} \mathrm{SCI}_{i,j}$ x $\mathrm{Pop}_{j,2010}$ x $\mathrm{Branches}_{j,b,t}$. Lagged LogDepMktShare is the same variable as the outcome, but measured in the previous year. Direct effect controls include the total number of branches of a given bank in a given county, the density of branches by population of a given bank in a given county, as well as a measure of physical proximity to branches in other counties of a given bank in a given county. Standard errors are clustered by bank-county-year level and are reported in parentheses below each estimate. ***, **, and * indicate significance levels of 10%, 5%, and 1% respectively.

in number of branches also captures bank mergers. Mergers are relatively frequent in our data. Of the 182,749 unique bank-year observations, around two percent indicate that the bank acquired branches that year. These activities can be viewed as an accelerated expansion (or contraction) process. We can also separately control for the change in branches due to mergers in the county of interest. The result is depicted in Column 3. Again, the coefficient remains statistically significant and stable. These robustness tests indicate that bank expansion or contraction along specializations does not drive the observed effect.

G. Robustness to Different Peer Exposure Specifications

Finally, we show that our results are robust to different specifications of our peer exposure measure. We develop three alternative specifications. Alternative A differs from our peer exposure measure by excluding the population component. In contrast to our original measure, this alternative is independent from population size of counties. The specification only captures the probability that the individual in county i is friends with any given person in county j. We test this specification to demonstrate that the effect is not simply driven by a correlation with this component of peer exposure. Alternative B does not examine the total number of branches but the share of bank b in county c at time t. This share is calculated by dividing the number of branches of bank b by the total number of branches of all banks in the county. In contrast to our original measure, this measure captures the importance of a bank in a county relative to other banks. Just as the total number of branches of the bank in the county, this relative importance is likely to be crucial for whether people talk about the bank to others. However, note that this measure is not independent from the action of other banks; peer exposure for instance decreases if another bank sets up more branches. Finally, Alternative C both excludes the population component and considers the branch share. For every regression specification we also correspondingly adapt the measure of physical proximity. Table 9 displays results for alternative peer exposure measures. We find that our findings are highly robust to these other specifications. Our original measure indicates that a one percent increase in peer exposure results in a 0.05 percent increase in deposit market share. Coefficient sizes for the three alternative measures vary between 0.06 and 0.07. All coefficients are highly statistically significant at the one percent level.

$$\text{Alternative A}_{b,i,t} = \sum_{j \neq i} \text{SCI}_{i,j} \times \text{Branches}_{j,b,t} \tag{7}$$

$$\text{Alternative B}_{b,i,t} = \sum_{j \neq i} \text{SCI}_{i,j} \times \text{Pop}_{j,2010} \times \frac{\text{Branches}_{j,b,t}}{\sum_{b=1}^{B} \text{Branches}_{j,b,t}} \tag{8}$$

$$\text{Alternative C}_{b,i,t} = \sum_{j \neq i} \text{SCI}_{i,j} \times \frac{\text{Branches}_{j,b,t}}{\sum_{b=1}^{B} \text{Branches}_{j,b,t}} \tag{9}$$

Table 9: Robustness: Different Peer Exposure Specifications

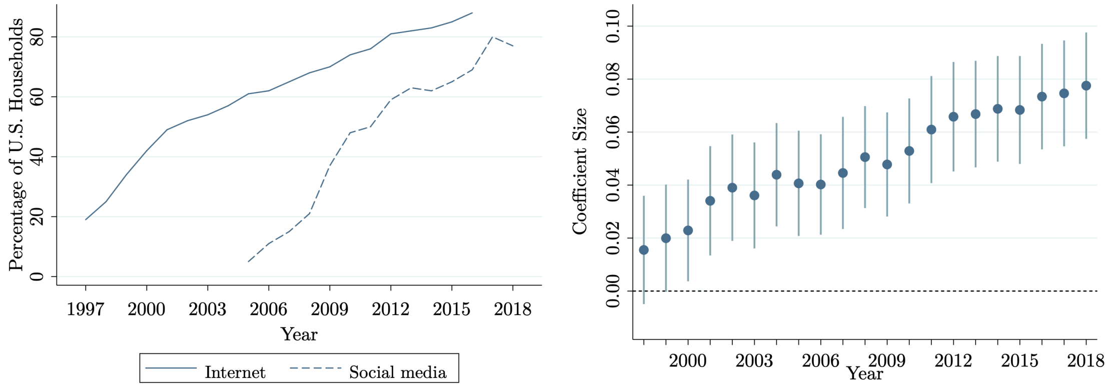
	LogDepMktShare		
	(1)	(2)	(3)
Alternative A	0.07***		
	(0.01)		
Alternative B		0.07***	
		(0.01)	
Alternative C			0.06***
			(0.01)
County x year FE	Yes	Yes	Yes
Bank x year FE	Yes	Yes	Yes
Bank x county FE	Yes	Yes	Yes
Lagged LogDepMktShare	Yes	Yes	Yes
Direct effect controls	Yes	Yes	Yes
R-squared	0.99	0.99	0.99
Observations	422,253	422,253	422,253
Banks	6,321	6,321	6,321
Counties	2,982	2,982	2,982
Years	21	21	21

This table shows that the main result is robust to different peer exposure specifications. Our sample consists of bank-county-year observations between 1997 and 2018. The dependent variable is LogDepMktShare is the logged deposit market share, which is the percentage of the amount in the deposit market held by a specific bank in a given county in a given year. The main independent variable is LogPeerExp, defined in Alternative A without the population component, in Alternative B with a branch share component, and in Alternative C without the population and with the branch share component. Lagged LogDepMktShare is the same variable as the outcome, but measured in the previous year. Direct effect controls include the total number of branches of a given bank in a given county, the density of branches by population of a given bank in a given county, as well as a measure of physical proximity to branches in other counties of a given bank in a given county. Standard errors are clustered by bank-county-year level and are reported in parentheses below each estimate. ***, **, and * indicate significance levels of 10%, 5%, and 1% respectively.

VII. Heterogeneity Analysis

A. Time

In the heterogeneity analysis, we target to gain an understanding for what time periods, for what kind of banks, and for what kind of customers peer effects are important. Initially, we turn towards a time-series analysis. Since 1997, the initial year under consideration, there has been a steep increase in the percentage of U.S. households who have access to the internet. In 2004, social media entered the stage and triggered a steep technology adoption among U.S. households (see Figure 5(a), from Comin and Hobijn (2004)). As households increasingly use the internet and social media, the potential for cross-county communication and thereby peer effects should strengthen. Indeed, we observe that the effect of peer exposure becomes more relevant over time. Figure 5(b) depicts the coef-



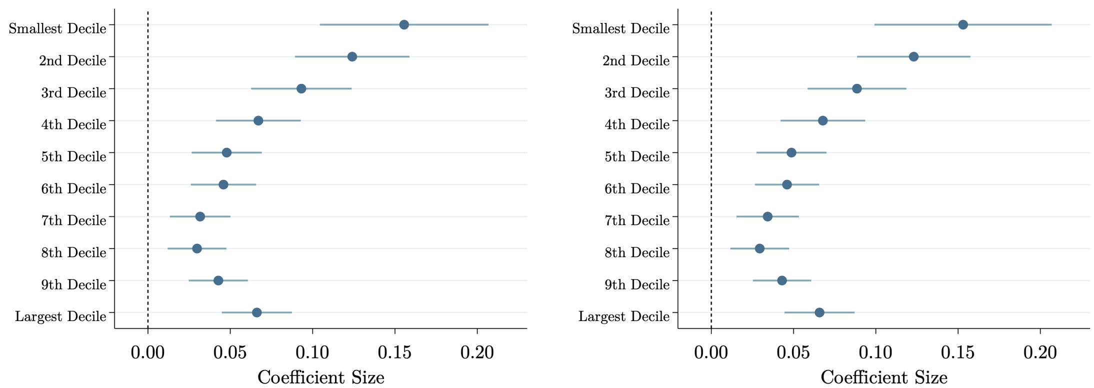
(a) Internet and Social Media (b) Peer Exposure

Figure 5. Time-Series Trends in Technologies and Peer Exposure Relevance.

ficients from our main regression; instead of our peer exposure measure, we include for each year a variable that is the value of the peer exposure measure in that given year and zero otherwise. Just as for internet access and social media usage, we observe a clear positive trend for the relevance of peer effects. This is in alignment with the fact that these technologies facilitate cross-county communication that we capture in our peer exposure measure.

B. Banks

Next, we focus on bank heterogeneity in terms of size. We utilize three bank-year size measures from the Call Reports: assets, liabilities, and number of employees. For each variable, we split the number of observations into deciles in each given year. We then run our main regression equation, including instead of our peer exposure measure ten variables that are the peer exposure variable multiplied with a dummy that is one if the observation falls into the respective decile and zero otherwise. Table A3 and Figure 6 describe results on heterogeneity along these dimensions. We see a clear pattern; peer effects appear to be more influential for smaller banks. Column 1 and 2 in Table A3 show that the difference between the coefficient of the smallest and the largest decile is statistically significant (at the one percent level for assets and at the five percent level for liabilities). Considering assets, a one percent increase in peer exposure increases the deposit market share



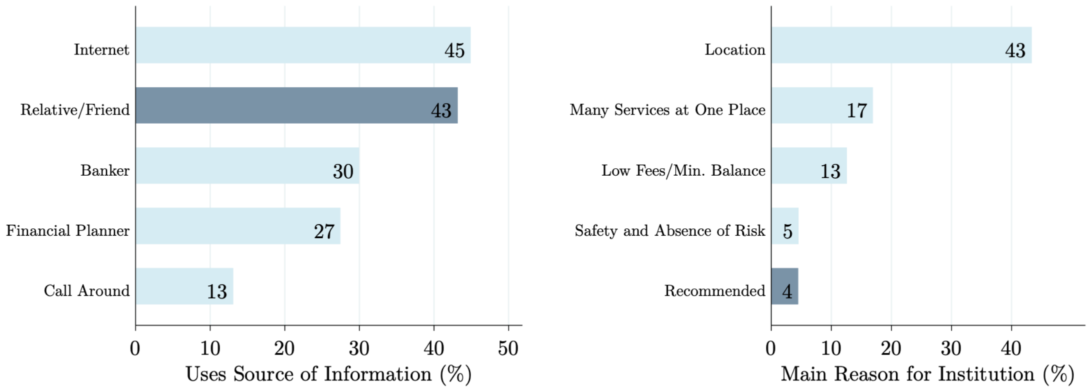
(a) Size Measure Assets (b) Size Measure Liabilities

Figure 6. Heterogeneity by Bank Size. Bank size measured utilizing Call Reports measures on assets and liabilities.

for banks in the smallest decile by 0.16 percent. In contrast, the same increase in peer exposure increases the deposit market share for banks in the largest decile by only 0.07. This goes hand in hand with the narrative that smaller banks have less means to invest in other drivers of deposit demand such as bank presence or advertisement campaigns. Additionally, opacity might be especially high for their products.

C. SCF

After analyzing for what kind of banks peer effects are of importance, we target to understand for which individuals peer effects are decisive. Initially, to inform our heterogeneity analysis on the county-level, we turn towards the SCF from 2019. This survey allows us to analyze which characteristics predict whether an individual relies on peers to make financial decisions. We distinguish between demographic characteristics such as income or education and financial characteristics such as whether the person has a checking account or mortgage. Table A4 provides summary statistics of these characteristics, adjusted by weights to make the sample representative of the U.S. population. We regress two outcomes from the survey on these demographics, using Equation 10. The first outcome is an indicator that is one if a person states that she uses relatives or friends as a source of information when making savings or investment decisions. The second outcome is an



(a) Savings or Investments Decisions

(b) Checking Account Institution

Figure 7. Survey of Consumer Finance (2019). Subfigure A depicts the share of individuals that use the respective source of information to make savings or investments decisions. Subfigure B depicts the share of individuals that cite the respective reason as the main factor for choosing their checking account institution. The sample of Subfigure B is conditional on having a checking account.

indicator that is one if a person responds that recommendations were the main reason for choosing her checking account provider. Initially, note that 43 percent of individuals replied that they use relatives and friends as a source of information for savings or investment decisions, just ranking after the internet (Figure 7(a)). Recommendations are less likely to be cited as the main reason for choosing a checking account institution; here factors as location rank higher. To examine which individuals are more likely to either use relatives or friends as a source of information or recommendations as the main reason to choose a checking account institution, we run the regression in Equation 10.

$$y_i = \beta_0 + \beta_1 \text{Characteristic}_i + \epsilon_i \tag{10}$$

Table A5 provides results. As expected, individuals that rank themselves high on search effort for savings and investment decisions are more likely to state that they use relatives or friends as a source of information. A clear pattern that emerges is that younger people both are more likely to collect information from their peers and to assign in a high weight in their decision. For the heterogeneity dimensions of income and education, the picture is more nuanced. While richer and more educated individuals are more likely to collect information from their peers, they are less likely to cite a recommendation as the main reason for choosing a checking account. A priori, it is consequently unclear whether this group should react stronger or weaker to peer exposure than individuals with lower income and education.

D. Counties

Motivated by the results of the SCF, we use our fixed effects identification strategy to analyze heterogeneity in importance of peer effects for consumers. We measure county characteristics using the ACS of the Economic Census. The ACS delivers data on various dimensions, including age, income, and education. For each demographic characteristics of interest, we split observations into tertiles in each given year. We then run our main regression equation, including instead of our peer exposure measure three variables that are the peer exposure variable multiplied with a dummy that is one if the observation falls into the respective tertile and zero otherwise. Table A6 and Figure 8 depict results

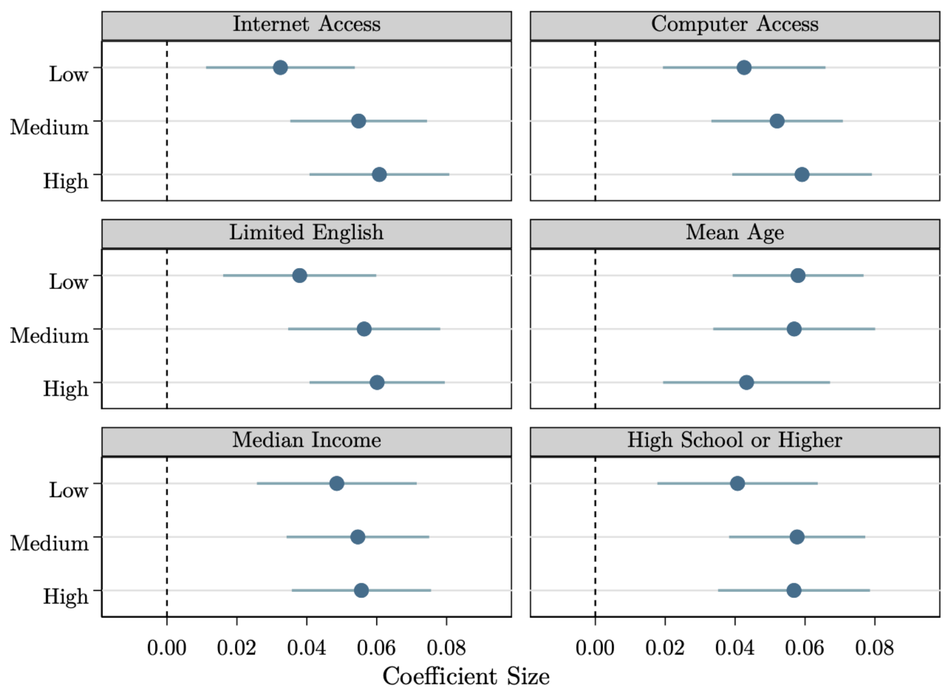


Figure 8. Heterogeneity by County Characteristics. Characteristics obtained from the ACS of the Economic Census. High school or higher represents the share of individuals in a given county that have at minimum a high school degree.

of the heterogeneity analysis. Initially, we observe that counties in which a high share of the population has internet access or computer access experience stronger reactions to peer exposure. This is in alignment with our narrative, since this technology is crucial for cross-county communication. The difference for internet access between the lowest and highest tercile is statistically significant at the five percent level. Next, we investigate splits according to the ability to speak English, age, income, and education. For people speaking limited English and of younger age, we expect lower effects; presumably they are both more likely to use peers as a source of information and assign this information a higher weight. For people with high income or education, the expected effect is ambiguous; they are more likely to use peers as a source of education, but less likely to assign this information a higher weight. For limited English speakers and young people, the county-level evidence confirms the evidence from the SCF. Counties in the tercile with the highest share of people with limited English knowledge have significantly stronger peer effects than counties in the tercile with the lowest share. The difference is significant at the 10 percent level. Counties in the tercile with the lowest mean age have lower weaker effects than counties in the tercile with the highest mean age, even though the difference is not statistically significant. Finally, we learn that higher income and higher education are positive but insignificant predictors of the relevance of peer effects in the deposit market.

VIII. Conclusion

In this paper we provide the first systematic empirical evidence that peer effects matter for consumer choice in deposit markets. To do so, we construct a novel time-varying and bank-county specific measure of peer exposure for potential depositors. The granular nature of our measure allows us to tightly identify the causal effect of peer effects on deposit demand through a fixed-effects identification strategy. In particular, we are able to include county-year, bank-year, and bank-county fixed effects to control for many potential threats to identification. Of note, we are able to address the key empirical challenge of time-invariant homophily, i.e. that peers tend to have similar preferences and thus may make similar choices even in the absence of influence from one another. Additionally, we control for other drivers of deposit demand that vary on the county-bank-year level and are consequently not captured by fixed effects such as local advertisement spending, and conduct a variety of robustness tests. We find that a one percent increase in bank peer exposure leads to a 0.05 percent increase deposit market shares. The effect is statistically significant at the one percent level and robust to various specifications, suggesting that peer effects are an important driver of deposit demand.

Appendix

Table A1: Robustness: Different Treatment of Outliers

	LogDepMktShare				
		Winsorizing		Triming	
	All	1st/99th	5th/95th	1st/99th	5th/95th
	(1)	(2)	(3)	(4)	(5)
LogPeerExp	0.08***	0.05***	0.04***	0.07***	0.07***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
County x year FE	Yes	Yes	Yes	Yes	Yes
Bank x year FE	Yes	Yes	Yes	Yes	Yes
Bank x county FE	Yes	Yes	Yes	Yes	Yes
R-squared	0.99	0.99	0.99	0.99	0.99
Observations	422,253	422,253	422,253	389,857	289,405
Banks	6,321	6,321	6,321	5,856	4,577
Counties	2,982	2,982	2,982	2,855	2,530
Years	21	21	21	21	21

This table shows that the main result is robust to different treatment of outliers. Our sample consists of bank-county-year observations between 1997 and 2018. The dependent variable is LogDepMktShare is the logged deposit market share, which is the percentage of the amount in the deposit market held by a specific bank in a given county in a given year. The main independent variable is LogPeerExp, the logged peer exposure that is defined as $\sum_{j \neq i} \mathrm{SCI}_{i,j}$ x $\mathrm{Pop}_{j,2010}$ x $\mathrm{Branches}_{j,b,t}$. Lagged LogDepMktShare is the same variable as the outcome, but measured in the previous year. Direct effect controls include the total number of branches of a given bank in a given county, the density of branches by population of a given bank in a given county, as well as a measure of physical proximity to branches in other counties of a given bank in a given county. Standard errors are clustered by bank-county-year level and are reported in parentheses below each estimate. ***, **, and * indicate significance levels of 10%, 5%, and 1% respectively.

Table A2: Heterogeneity by Year

	LogDepMktShare
LogPeerExp	
1998	0.02
	(0.01)
1999	0.02
	(0.01)
2000	0.02*
	(0.01)
2001	0.03***
	(0.01)
2002	0.04***
	(0.01)
2003	0.04***
	(0.01)
2004	0.04***
	(0.01)
2005	0.04***
	(0.01)
2006	0.04***
	(0.01)
2007	0.04***
	(0.01)
2008	0.05***
	(0.01)
2009	0.05***
	(0.01)
2010	0.05***
	(0.01)
2011	0.06***
	(0.01)
2012	0.07***
	(0.01)
2013	0.07***
	(0.01)
2014	0.07***
	(0.01)
2015	0.07***
	(0.01)
2016	0.07***
	(0.01)
2017	0.07***
	(0.01)
2018	0.08***
	(0.01)
P-value (1997 == 2018)	0.00
Lagged LogDepMktShare	Yes
Direct effect controls	Yes
County x year FE	Yes
Bank x year FE	Yes
Bank x county FE	Yes
R-squared	0.99
Observations	422,253
Banks	6,321
Counties	2,982
Years	21

This table shows heterogeneity by years. Variables are defined as in the Table 2. ***, **, and * indicate significance levels of 10%, 5%, and 1% respectively.

Table A3: Heterogeneity by Bank Size

	LogDepMktShare		
	(1)	(2)	(3)
LogPeerExp			
Smallest Decile	0.16***	0.15***	0.13***
	(0.03)	(0.03)	(0.04)
2nd Decile	0.12***	0.12***	0.10***
	(0.02)	(0.02)	(0.02)
3rd Decile	0.09***	0.09***	0.09***
	(0.02)	(0.02)	(0.02)
4th Decile	0.07***	0.07***	0.07***
	(0.01)	(0.01)	(0.02)
5th Decile	0.05***	0.05***	0.06***
	(0.01)	(0.01)	(0.01)
6th Decile	0.05***	0.05***	0.04***
	(0.01)	(0.01)	(0.01)
7th Decile	0.03***	0.03***	0.03***
	(0.01)	(0.01)	(0.01)
8th Decile	0.03***	0.03**	0.03***
	(0.01)	(0.01)	(0.01)
9th Decile	0.04***	0.04***	0.04***
	(0.01)	(0.01)	(0.01)
Largest Decile	0.07***	0.07***	0.07***
	(0.01)	(0.01)	(0.01)
P-value (Smallest == Largest)	0.01	0.01	0.12
Size measure	Assets	Liabilities	Employees
Lagged LogDepMktShare	Yes	Yes	Yes
Direct effect controls	Yes	Yes	Yes
County x year FE	Yes	Yes	Yes
Bank x year FE	Yes	Yes	Yes
Bank x county FE	Yes	Yes	Yes
R-squared	0.99	0.99	0.99
Observations	380,387	380,386	380,150
Banks	5,867	5,867	5,867
Counties	2,978	2,978	2,978
Years	21	21	21

This table shows heterogeneity by bank size. Our sample consists of bank-county-year observations between 1997 and 2018. The dependent variable is LogDepMktShare is the logged deposit market share, which is the percentage of the amount in the deposit market held by a specific bank in a given county in a given year. The main independent variable is LogPeerExp, the logged peer exposure that is defined as $\sum_{j \neq i} \text{SCI}_{i,j}$ x $\text{Pop}_{j,2010}$ x $\text{Branches}_{j,b,t}$. Lagged LogDepMktShare is the same variable as the outcome, but measured in the previous year. Direct effect controls include the total number of branches of a given bank in a given county, the density of branches by population of a given bank in a given county, as well as a measure of physical proximity to branches in other counties of a given bank in a given county. Bank size is measured utilizing Call Reports. Standard errors are clustered by bank-county-year level and are reported in parentheses below each estimate. ***, **, and * indicate significance levels of 10%, 5%, and 1% respectively.

Table A4: Survey of Consumer Finance (2019) Summary Statistics

	Mean	Sd
Demographic Characteristics		
Female	0.27	0.44
Age	51.73	17.50
White	0.68	0.47
Black	0.16	0.36
Hispanic/latino	0.11	0.31
Total income ($1,000)	101.00	235.04
Employed/self-employed	0.64	0.48
Unemployed	0.03	0.17
Retired	0.22	0.42
Highest degree none or high school	0.58	0.49
Highest degree college	0.27	0.44
Highest degree master/doctoral	0.15	0.36
Financial Characteristics		
Any checking account	0.93	0.25
Amount in checking account ($1,000)	7.73	30.72
Any mortgage	0.40	0.49
Any health insurance	0.92	0.26
Used online banking last year	0.78	0.42
Finance knowledge (index 1-10)	7.14	2.17
Risk taking (index 1-10)	4.30	2.70
Search effort savings/investments (index 1-10)	5.92	3.23
Search effort borrowing (index 1-10)	6.48	3.21
N	5,777	

This table shows summary statistics for the SCF from 2019.

Table A5: Survey of Consumer Finance (2019) Regressions (OLS)

	Info source relatives/friends savings/investment decisions	Main reason recommendation checking account institution
Demographic Characteristics		
Female	0.03**	0.03***
	(0.01)	(0.01)
Age	-0.01***	-0.00***
	(0.00)	(0.00)
White	0.02**	-0.01
	(0.01)	(0.01)
Black	-0.03**	0.00
	(0.02)	(0.01)
Hispanic/latino	-0.04**	0.01
	(0.02)	(0.01)
Total income ($1,000, log)	0.01***	-0.02***
	(0.00)	(0.00)
Employed/self-employed	0.09***	0.00
	(0.01)	(0.00)
Unemployed	0.04	0.01
	(0.03)	(0.01)
Retired	-0.13***	-0.01**
	(0.02)	(0.01)
Highest degree none or school	-0.07***	0.02***
	(0.01)	(0.01)
Highest degree college	0.05***	-0.01
	(0.01)	(0.00)
Highest degree master/doctoral	0.06***	-0.03***
	(0.02)	(0.01)
Financial Characteristics		
Any checking account	0.05*	-
	(0.02)	(-)
Amount in checking acc. ($1,000, log)	0.00*	-0.01***
	(0.00)	(0.00)
Any mortgage	0.04***	-0.03***
	(0.01)	(0.01)
Any health insurance	0.05***	-0.03***
	(0.02)	(0.01)
Used online banking last year	0.09***	-0.01
	(0.01)	(0.01)
Finance knowledge (index 1-10)	-0.01***	-0.01***
	(0.00)	(0.00)
Risk taking (index 1-10)	0.01***	-0.00*
	(0.00)	(0.00)
Search effort savings/investments (index 1-10)	0.02***	-0.00
	(0.00)	(0.00)
Search effort borrowing (index 1-10)	0.02***	-0.00
	(0.00)	(0.00)
N	5,777	5,433

This table shows results from Equation 10. The first column refers to the question whether the individual uses information from relatives or friends to make decisions about savings or investments. The second column refers to the question whether a recommendation was the main reason for choosing a checking account institution.

Table A6: Heterogeneity by County Characteristics

	LogDepMktShare					
	Internet access (1)	Computer access (2)	Limited English (3)	Mean age (4)	Median income (5)	High school or higher (6)
LogPeerExp						
Small	0.03**	0.04***	0.04***	0.06***	0.05***	0.04***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Medium	0.05***	0.05***	0.06***	0.06***	0.05***	0.06***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Large	0.06***	0.06***	0.06***	0.04***	0.06***	0.06***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
P-value (Small == Large)	0.01	0.18	0.05	0.24	0.54	0.23
Lagged LogDepMktShare	Yes	Yes	Yes	Yes	Yes	Yes
Direct effect controls	Yes	Yes	Yes	Yes	Yes	Yes
County x year FE	Yes	Yes	Yes	Yes	Yes	Yes
Bank x year FE	Yes	Yes	Yes	Yes	Yes	Yes
Bank x county FE	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.99	0.99	0.99	0.99	0.99	0.99
Observations	422,100	422,107	422,108	421,960	422,229	422,148
Banks	6,320	6,320	6,320	6,321	6,320	6,320
Counties	2,980	2,980	2,981	2,981	2,981	2,981
Years	21	21	21	21	21	21

This table shows heterogeneity by county characteristics. Our sample consists of bank-county-year observations between 1997 and 2018. The dependent variable is LogDepMktShare is the logged deposit market share, which is the percentage of the amount in the deposit market held by a specific bank in a given county in a given year. The main independent variable is LogPeerExp, the logged peer exposure that is defined as $\sum_{j \neq i} \text{SCI}_{i,j}$ x $\text{Pop}_{j,2010}$ x $\text{Branches}_{j,b,t}$. Lagged LogDepMktShare is the same variable as the outcome, but measured in the previous year. Direct effect controls include the total number of branches of a given bank in a given county, the density of branches by population of a given bank in a given county, as well as a measure of physical proximity to branches in other counties of a given bank in a given county. County characteristics are measured using the ACS of the Economic Census. High school or higher refers to the share of individuals who have at minimum a high school degree. Standard errors are clustered by bank-county-year level and are reported in parentheses below each estimate. ***, **, and * indicate significance levels of 10%, 5%, and 1% respectively.

References

Aral, S. and Walker, D. (2012). Identifying influential and susceptible members of social networks. *Science*, 337(6092):337–341.

Bailey, M., Cao, R., Kuchler, T., and Stroebel, J. (2018a). The economic effects of social networks: Evidence from the housing market. *Journal of Political Economy*, 126(6):2224–2276.

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

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

Bailey, M., Farrell, P., Kuchler, T., and Stroebel, J. (2020a). Social connectedness in urban areas. *Journal of Urban Economics*, page 103264.

Bailey, M., Gupta, A., Hillenbrand, S., Kuchler, T., Richmond, R. J., and Stroebel, J. (2020b). International trade and social connectedness. Technical report, National Bureau of Economic Research.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, pages 841–890.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, pages 242–262.

Beshears, J., Choi, J. J., Laibson, D., Madrian, B. C., and Milkman, K. L. (2015). The effect of providing peer information on retirement savings decisions. *The Journal of finance*, 70(3):1161–1201.

Bursztyn, L., Ederer, F., Ferman, B., and Yuchtman, N. (2014). Understanding mechanisms underlying peer effects: Evidence from a field experiment on financial decisions. *Econometrica*, 82(4):1273–1301.

Comin, D. and Hobijn, B. (2004). Cross-country technology adoption: making the theories face the facts. https://ourworldindata.org/rise-of-social-media (accessed 16 August 2021).

De Giorgi, G., Frederiksen, A., and Pistaferri, L. (2016). Consumption network effects. Working Paper 22357, National Bureau of Economic Research.

DellaVigna, S., List, J. A., and Malmendier, U. (2012). Testing for altruism and social pressure in charitable giving. *The quarterly journal of economics*, 127(1):1–56.

Diamond, W., Jiang, Z., and Ma, Y. (2020). A structural model of bank balance sheet synergies and the transmission of central bank policies. *Available at SSRN*.

Dick, A. A. (2008). Demand estimation and consumer welfare in the banking industry. *Journal of Banking & Finance*, 32(8):1661–1676.

Drechsler, I., Savov, A., and Schnabl, P. (2017). The deposits channel of monetary policy. *The Quarterly Journal of Economics*, 132(4):1819–1876.

Drechsler, I., Savov, A., and Schnabl, P. (2018). Banking on deposits: Maturity transformation without interest rate risk. Technical report, National Bureau of Economic Research.

Drechsler, I., Savov, A., and Schnabl, P. (2021). Banking on deposits: Maturity transformation without interest rate risk. *The Journal of Finance*, 76(3):1091–1143.

Duflo, E. and Saez, E. (2003). The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment. *The Quarterly journal of economics*, 118(3):815–842.

Egan, M., Hortaçsu, A., and Matvos, G. (2017). Deposit competition and financial fragility: Evidence from the us banking sector. *American Economic Review*, 107(1):169–216.

Forrester (2018). Forrester. How US Consumers Research and Buy Checking Accounts. https://www.forrester.com/report/How+US+Consumers+Research+And+Buy+Checking+Accounts+2018/RES131965 (accessed 10 February 2021).

Gilchrist, D. S. and Sands, E. G. (2016). Something to talk about: Social spillovers in movie consumption. *Journal of Political Economy*, 124(5):1339–1382.

Greenwood, S., Perrin, A., and Duggan, M. (2016). Social media update 2016. *Pew Research Center*, 11(2):1–18.

Hanson, S. G., Shleifer, A., Stein, J. C., and Vishny, R. W. (2015). Banks as patient fixed-income investors. *Journal of Financial Economics*, 117(3):449–469.

Ho, K. and Ishii, J. (2011). Location and competition in retail banking. *International Journal of Industrial Organization*, 29(5):537–546.

Hong, H., Kubik, J. D., and Stein, J. C. (2004). Social interaction and stock-market participation. *The journal of finance*, 59(1):137–163.

Honka, E., Hortaçsu, A., and Vitorino, M. A. (2017). Advertising, consumer awareness, and choice: Evidence from the us banking industry. *The RAND Journal of Economics*, 48(3):611–646.

Kuchler, T., Li, Y., Peng, L., Stroebel, J., and Zhou, D. (2020a). Social proximity to capital: Implications for investors and firms. Technical report, National Bureau of Economic Research.

Kuchler, T., Russel, D., and Stroebel, J. (2020b). The geographic spread of covid-19 correlates with structure of social networks as measured by facebook. Technical report, National Bureau of Economic Research.

Kuchler, T. and Stroebel, J. (2020). Social finance. Working Paper 27973, National Bureau of Economic Research.

Kuhn, P., Kooreman, P., Soetevent, A., and Kapteyn, A. (2011). The effects of lottery prizes on winners and their neighbors: Evidence from the dutch postcode lottery. *American Economic Review*, 101(5):2226–47.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The review of economic studies*, 60(3):531–542.

Maturana, G. and Nickerson, J. (2019). Teachers teaching teachers: The role of workplace peer effects in financial decisions. *The Review of Financial Studies*, 32(10):3920–3957.

Ouimet, P. and Tate, G. (2020). Learning from coworkers: Peer effects on individual investment decisions. *The Journal of Finance*, 75(1):133–172.

Pool, V. K., Stoffman, N., and Yonker, S. E. (2015). The people in your neighborhood: Social interactions and mutual fund portfolios. *The Journal of Finance*, 70(6):2679–2732.

Wang, Y., Whited, T. M., Wu, Y., and Xiao, K. (2020). Bank market power and monetary policy transmission: Evidence from a structural estimation. Technical report, National Bureau of Economic Research.

Xiao, K. (2020). Monetary transmission through shadow banks. *The Review of Financial Studies*, 33(6):2379–2420.