Cyberattacks on Small Banks and the Impact on Local Banking Markets

Fabian Gogolin
University of Leeds

Ivan Lim
Durham University

Francesco Vallascas*
Durham University

Abstract

Using a sized-matched difference-in-differences design, we document that successful cyberattacks decrease branch deposit growth rates at small US banks. This decrease is due to bank-specific reputational damages that erode the trust of bank customer and result in a "flight-to-reputation" within local markets leading to a reallocation of deposits to large banks. As a result of these damages, hacked banks attract riskier applicants in mortgage markets and are forced to lower credit standards. Our study implies that financial constraints to cybersecurity investments can reduce the competitiveness of small banks and lead to local banking markets increasingly dominated by large banks.

JEL Classification: G21, G28.
Keywords: Cybersecurity, Small Banks, Deposit Markets, Bank Lending.

*Corresponding Author. Francesco Vallascas (francesco.vallascas@durham.ac.uk), Durham University Business School, Durham University, Millhill Ln, Durham DH1 3LB, United Kingdom. We are grateful for the valuable comments of Maria Boutchkova, Angelica Gonzalez, Andreas Milidonis, Louis Nguyen and Ben Sila, the conference participants at the 2020 CSBS/FDIC/Fed Community Banking in the 21st Century Research and Policy Conference (Sept. 30-Oct. 1, 2020), and the seminar participants at the University of Edinburgh.

1 Introduction

The digital transformation of the economy exposes banks to new risks, in particular, cyber risks (Basel Committee on Banking Supervision, 2018; Duffie and Younger, 2019; Mester et al., 2019). FDIC Chairman Jelena McWilliams states that "*[c]ybersecurity is the biggest threat facing America's banks*"[1]. Echoing this view, a report by the Boston Consulting Group shows that financial firms are 300 times more likely to become targets of cyberattacks than other firms[2].

In the banking industry where contractual relationships are based on trust (Chen et al., 2019), cyberattacks can be extremely damaging. Indeed, successful cyberattacks generate reputational losses for target firms (Kamiya et al., 2020), and these losses can undermine customer trust and lead to a reduced customer base for the hacked banks (Chen et al., 2019; Kashyap and Wetherilt, 2019). Continuous investments in cybersecurity are becoming, therefore, necessary for banks to prevent cybercrimes and, in this way, safeguard customer trust. For instance, Deloitte (2019) shows that the average yearly cybersecurity investment by US banks has surpassed 10% of their IT budget, equivalent to $2,300 per employee.

However, the reoccurring investments needed to maintain a high level of cybersecurity require significant financial resources that might be available to large banks but not to small banks. In other words, small banks tend to be financially constrained and might find these onerous investments unsustainable over the long-term (Kashyap and Wetherilt, 2019; Paravisini, 2008). Consequently, small banks are likely to show an investment gap in cybersecurity and remain exposed to significant cyber risks to their business. According to a report by Nationwide, almost half of cybercrimes between 2012-2017 target US banks with assets below $1billion[3]. It is not surprising, therefore, that more than 70% of small

[1]See "Banks could get fined for cyber breaches, top regulator says", CNN (2019), available at https://edition.cnn.com/2019/08/01/investing/fdic-cyber-hack-fine/index.html

[2]https://markets.businessinsider.com/news/stocks/cyberattacks-impact-major-threats-to-financial-firms-not-prepared-2019-6-1028296130.

[3]See "5 Cybersecurity Myths Banks Should Stop Believing", Forbes (2019), available at https://www.forbes.com/sites/ronshevlin/2019/04/08/5-cybersecurity-myths-banks-should-stop-believing/#6c83bb1d630d

bankers have recently ranked cybersecurity as their top concern (Conference of State Bank Supervisors, 2019).

The consequences of inadequate investments in cybersecurity by small banks can be better understood in the context of theoretical models on the effects of investment and innovation gaps among rival firms (see, Bloom et al. (2013)). In these models, firms that underinvest in innovation become less competitive and lose market share to more innovative firms. In a similar vein, gaps in cybersecurity investments in small banks might induce customers to abandon these banks and shift to (large) rivals that are perceived as (digitally and technologically) safer. Following this argument, cybersecurity deficiencies can be costly for small banks and ultimately have an impact also on the structure of local banking markets. The economic implications of this structural change can be substantial given that small banks have a competitive advantage in lending to local businesses (Agarwal and Hauswald, 2010; Berger et al., 2005; Skrastins and Vig, 2019; Stein, 2002). More precisely, in local banking markets characterized by a reduced competitiveness of small banks, the access to credit of (small) local businesses may deteriorate, negatively impacting the development of local economies (Berger et al., 2017; Hakenes et al., 2015).

In this paper, we build on the above arguments and present the first attempt to document how breaches to cybersecurity create significant challenges to small banks by affecting their ability to retain and attract customers within a local market. Specifically, we offer a comprehensive assessment of depositor reactions to successful cyberattacks on small banks, how these attacks influence the allocation of deposits within local markets and impact on lending relationships in mortgage markets.

Our initial focus on deposit markets is motivated by two reasons. First, depositors are key bank stakeholders and their relationships with banks are based on trust (Chen et al., 2019). This trust can be broken when depositors' confidential personal and financial information is compromised. Therefore, depositors are directly affected by cyberattacks and should be the main bank customers to respond to cyberattacks. Second, deposit markets are a key source

of funding for small banks. Hence, if successful cyberattacks reduce the competitiveness of small banks in deposit markets, they might have negative consequences for the sustainability of the business model of these banks.

We base our analysis around exogenous cyberattacks involving small US banks covered in the Privacy Rights Clearinghouse (PRC) database over the period 2005-2017. We employ these exogenous events to capture cybersecurity breaches and implement difference-in-differences analyses to assess the response of bank customers. We start by documenting the effects of successful cyberattacks on the hacked small banks in the deposit market. Our analyses indicate that branches of these banks experience an economically significant 20 percentage point decline in the growth rate of their deposits compared to a control group of branches of similar sized banks operating in the same local market. Consequently, declines in deposit growth rates lead to a fall in the deposit market share of hacked small banks. Our results are robust to a number of alternative empirical settings, including the adoption of the estimation approach of Bertrand et al. (2004), as well as different sets of fixed effects and estimation windows.

We next show that depositors do not respond to other types of data breaches that occur endogenously as a result of day-to-day bank operations. As these breaches are generally narrower in nature and do not involve widespread losses of customer personal and financial information (Kamiya et al., 2020), this result indicates that our initial findings are motivated by depositors' concern over the soundness of a small bank's cybersecurity environment to external threats.

The decline in deposit growth in hacked banks is consistent with the view that cybersecurity soundness is important for the reputation of small banks, and with it, depositors' trust in these banks. To further assess the importance of depositor trust, we present two set of tests. First, we document that declines in deposit growth rates are less pronounced for banks with higher social capital. This result is consistent with explanations put forth by Lins et al. (2017) that the social capital of a firm reflects how trustworthy the

firm is perceived to be and should pay off when being trustworthy is most valuable, such as after a cyberattack (Akey et al., 2021). Second, we show that hacked banks increase deposit rates after a cyberattack. This is consistent with explanations that stakeholders demand better contractual terms to transact with firms affected by negative shocks to their trustworthiness and reputation (Kamiya et al., 2020; Karpoff, 2012). In addition, we do not find any evidence that depositor reactions are driven by bank fundamentals or can be explained by (potential) deteriorations in bank riskiness that might arise as a result of cyberattacks.

Recent studies document heterogeneous depositor responses to negative information regarding bank financial and social performance (see Chen et al. (2019) and Chen et al. (2020)). We find evidence of larger declines in deposit growth rates from depositors that are plausibly less knowledgeable about cyber risk. This complements findings by Chen et al. (2019) and Chen et al. (2020) who show that more technical disclosures such as bank earnings and regulatory ratings on community involvement illicit stronger response from more sophisticated depositors (presumably due to the ability to understand information). Our findings suggest instead that successful cyberattacks, which are more salient and directly impact depositors' personal welfare, are likely to incite larger responses from unsophisticated bank customers who might not be fully aware of the consequences and remediation processes following cyberattacks.

In addition to the effects highlighted above, cyberattacks on small banks can have widespread consequences for local deposit markets and lead to spillovers to other banks. In this respect, Kamiya et al. (2020) show negative value effects for non-financial firms operating in the same industry as those targeted by cyberattacks. Nevertheless, in contrast to Kamiya et al. (2020), our focus is on small and unlisted banking firms. As such, it is unlikely that depositors perceive successful cyberattacks on these banks as indicative of systemic industry-wide weaknesses in banks' exposure to cyber risks. Hence, the reputational damages suffered by hacked small banks need not result in loss of confidence to other institutions.

Instead, banks with a reputational advantage in local deposit markets would benefit from positive spillovers after the cyberattack via the reallocation of deposits towards their branches (Chen et al., 2019). Plausibly, this reallocation should favor large(r) institutions as they might be perceived as (digitally and technologically) safer by depositors. Positive spillovers towards these banks are then consistent with a "business stealing effect" in favor of rival firms with advantages in innovation (Bloom et al., 2013). Negative spillovers as in Kamiya et al. (2020) can then emerge for other small banks if perceived by depositors as equally vulnerable to cybercrimes.

We examine the spillover effects of cyberattacks by employing an alternative difference-in-differences setup. In this setup, we compare the evolution of deposit growth of the branches of untreated banks in counties where hacked small banks operate to the branches of the same untreated banks in adjacent counties where hacked small banks do not operate. We find positive spillovers only towards branches of large banks (but not small banks). Further, positive spillovers are more pronounced in large banks with an excellent reputation amongst customers and in more concentrated deposit markets wherein competition to provide banking services is expected to be lower. These results indicate a "flight-to-reputation" effect in local deposit markets after successful cyberattacks on small banks. A key implication of this finding is that the growing importance bank customers place on cybersecurity can result in large banks dominating local markets. Since large banks might be less inclined to supply small business lending, especially in times of crises, local businesses can face increasing financial frictions (Bord et al., 2018; Chen et al., 2017).

Finally, banks also engage in contractual relationships with households in mortgage markets. Although cyberattacks do not result in an immediate threat to potential borrowers, it might be argued that they signal bank reputational losses across all customer relationships (Akey et al., 2021). To assess the validity of this argument, we carry out two different tests. The first takes the borrower perspective and examines the effects of cyberattacks in terms of the number and composition of mortgage applications that affected banks receive.

The second focuses on the bank perspective and assesses the consequences on underwriting standards. We do not find that banks suffer from a decline in the number of mortgage applications after a cyberattack. However, treated banks are forced to originate riskier loans to maintain unchanged their mortgage approval rates.

We contribute to three streams of research. The first focuses on the effects of cyberattacks to corporations. This literature is primarily based on non-financial firms and documents that cyberattacks generate reputational damages that lead to reductions in shareholder value and risk appetite (Kamiya et al., 2020), decreased profitability (Akey et al., 2021) and higher audit fees (Li et al., 2020; Rosati et al., 2019). However, empirical investigations on the implications of cyberattacks on bank outcomes are almost non-existent. Eisenbach et al. (2020) simulate the externalities produced by cyberattacks through the wholesale payments network and show that damages to the five most active banks would affect more than a third of the network. Bouveret (2018) presents a cross-country overview of cyber risk in the financial industry and proposes a framework for its quantification. Aldasoro et al. (2020) document that cyber losses account for a significant portion of total operational value-at-risk. We present the first empirical investigation of the impact of cyberattacks on small banks. Further, we exploit the peculiarities of the banking industry to explore cyberattacks from a customer perspective as opposed to the shareholder perspective often taken in studies on non-financial firms. We show that cyberattacks result in reputational losses that reduce customer trust and lead to decreases in deposit growth rates. As the banking business is built on trust, cyberattacks might have long-term negative implications for small banks.

Second, we contribute to the literature on how depositors react to the disclosure of negative information by banks. A first group of studies focuses on the disclosure of financial information (Berger and Turk-Ariss, 2015; Chen et al., 2020; Iyer et al., 2016; Martinez Peria and Schmukler, 2001). The general consensus is that depositors react negatively to financial information highlighting negative bank performance, although there is heterogeneity in the response depending on the ability and incentives of depositors to monitor banks (Danisewicz

et al., 2018; Chen et al., 2020). More closely related to our analysis are studies on how depositors respond to negative non-financial information. Chen et al. (2019) document that banks are more likely to suffer from larger deposit outflows when they show poor social performance measured through CRA ratings and CRA ratings downgrades. Homanen (2018) finds a similar negative effect in banks that financed the 2016 Dakota Access Pipeline project which crossed major rivers and ancient burial grounds. We complement these studies by showing that cyberattacks also lead to negative responses by depositors. Further, we also document how such events influence the re-distribution of deposits in local deposit markets via spillover effects towards larger banks. In doing so, we are able to show that investments in cybersecurity are crucial in competing for funding and in affecting the structure of local deposit markets.

Finally, our study is also related to the literature on operational risks in banks. Earlier analyses show that most of the operational losses at US financial institutions are produced by failures in internal control systems (Chernobai et al., 2011). Along these lines, and more recently, Chernobai et al. (2020) document that operational risks are more pronounced in complex banks. Barakat et al. (2019) highlight the negative value effects arising from media announcements of operational risk events especially when the information on the event is opaque. Although frequently classified as part of operational risks, cyber risk shows key peculiarities related to the potential loss of confidentiality that could lead to damages to the integrity of data or systems (Eisenbach et al., 2020; Mester et al., 2019). These aspects are a potential concern for all stakeholders that engage in a contractual relationship with a bank and motivates our primary focus on deposit markets. Yet, contrary to existing studies on operational risks, we investigate events produced by external data breaches that are plausibly exogenous, allowing us to evaluate the causal implications of cyber risk on depositor behavior, deposit market structure and lending relationships.

2 Identification Strategy and Data

2.1 Treated and Control Banks

Our analysis is based on cyberattacks of small US commercial banks between 2005-2017[4]. We identify these attacks starting from a list of all data breach incidents involving financial institutions covered in the Privacy Rights Clearinghouse (PRC) database over the same period. This database includes breaches that are reported in a timely manner under State Security Breach Notification Laws (see Akey et al. (2021) and Kamiya et al. (2020)). Within the data breaches included in PRC, we first retain only breaches that affect financial firms. We next select events that satisfy the following three criteria: i) they are classified as a "HACK" by PRC; that is, they are caused by external parties (e.g., not by internal fraud, accidental disclosures and the loss of portable or stationary devices) and result in the loss of customer personal or financial information; ii) they target a small commercial bank (defined as a bank with total assets up to $10bln at the time of the data breach); iii) they affect banks for which we have detailed deposit data from the Summary of Deposits (SOD) provided by the FDIC. The first criterion ensures that the events are plausibly exogenous (Kamiya et al., 2020). Using this sampling procedure, we identify 16 cyberattacks on small US banks. We provide identifying information for our sample of cyberattacks in Table A1 of the Online Appendix.

The SOD data offers branch level information on deposits and allows for a tight geographic matching between branches of small banks targeted by a cyberattack and branches of unaffected banks. This sampling approach alleviates concerns that confounding geographical supply and demand factors might bias the analyses. One example of these geographical factors is the fraction of seniors across different geographical regions. Becker (2007) shows the volumes of deposits are higher in areas with more senior citizens. Hence, if seniors react differently to cyberattacks, and have different deposit trajectories, comparing branches of

[4]We do not include more recent cyberattacks in our sample because the implementation of our identification strategy requires three years of bank data after the attack has been reported.

treated and untreated banks from different geographic areas might yield biased results.

Our econometric setting is constructed around the state in which a cyberattack is reported according to PRC. Within this state, we identify all counties in which the affected banks operate branches. These branches represent our treated group. Next, within those counties, we form a control group of branches that are owned by similarly sized commercial banks. Constructing the control group from similarly sized banks is important, as previous research has documented that larger banks have advantages in deposit markets (Jacewitz and Pogach, 2018; Oliveira et al., 2015). To ensure a high degree of similarity in size between the treated and untreated small banks, our matching strategy is as follows. We divide treated banks with assets below the $10bln threshold into two size-based groups, i) treated banks with assets up to $1bln and ii) treated banks with assets between $1bln and up to $10bln. When we match branches of treated and untreated banks at the county level, the control group consists only of branches of banks falling into the same size group. In additional tests, we employ an alternative and tighter size matching between treated and untreated banks. We further discuss these additional tests in Section 3.1.

2.2 Econometric Method

Given the staggered nature of cyberattacks, we use a stacked difference-in-differences approach to estimate the causal impact of cyber risk on depositor behavior (Gormley and Matsa, 2011). We construct cohorts of treated branches for each event and stack the cohorts to estimate the average treatment effect. For each cohort, the control group consists only of bank branches that have not previously experienced a cyberattack. This choice allows us to more cleanly capture the treatment effect (Gormley and Matsa, 2011; Guo et al., 2019). We use an estimation window of 7 (-3;+3) years around each cyberattack for a total of 3,076 (12,384) observations belonging to branches of treated (untreated) banks. More formally, we estimate the following model:

$$\text{Ln(Deposits)}_{i,j,z,c,t} = \alpha + \beta_1 \text{ Treated} \times \text{Post} + \mathbf{BRANCH} + \mathbf{COUNTY} \times \mathbf{TIME} + \varepsilon_{i,j,z,c,t}, \quad (1)$$

Where Ln(Deposits) is the logarithmic transformation of deposits in thousands of US$ in branch i of bank j in county z, and belonging to a cohort c at time t. Treated is a dummy that equals one if a branch belongs to an hacked bank and zero otherwise; Post is a dummy equal to one in the post-shock window (up to 3 years after the shock). The difference-in-differences estimate of the coefficient of Treated × Post is the difference between how the dependent variable changes in the branches of treated banks (namely, banks affected by a cyberattack) and in the branches of control banks after the shock. Since we measure our dependent variable in log form, the estimated coefficient is approximately equivalent to the difference in the average growth rate of the US$ value of deposits in the groups of branches of treated and control banks from the pre- to the post-shock period. We estimate equation (1) with standard errors clustered at the commercial bank level to control for within bank correlation in the evolution of deposits. In section 3.2, we document that our results remain unchanged if we cluster the standard errors at the branch level.

The model includes branch (**BRANCH**) and county × year (**COUNTY** × **TIME**) fixed effects. The first set of fixed effects controls for branch-specific time-invariant omitted variables, such as the quality of its services, that could influence depositor behavior (e.g. Begley and Purnanandam (2021)). The inclusion of county × year fixed effects removes any time-varying county-level factors such as demographic and local business cycles (e.g., unemployment housing demand and shale gas discoveries) that could affect local deposit markets (Gilje et al., 2016; Mian and Sufi, 2014). With these two sets of fixed effects in place, we are comparing the changes in deposits in treated branches relative to the change in the control group of branches (belonging to similarly sized banks) in the same county in a given year.

[TABLE 1 HERE]

Initially, in equation (1), we do not include bank-specific control variables. In fact, any bank-specific control can also be affected by a cyberattack, making it difficult to interpret the coefficient of Treated × Post (Gormley and Matsa, 2011). Nevertheless, to mitigate concerns over omitted variables, we report two additional specifications with bank controls. First, we control only for bank size (measured by the logarithmic transformation of bank total assets in thousands of US$). In the final specification, we control for a range of other bank characteristics including the ratio between net income and total assets (ROA), tier 1 capital divided by risk weighted assets (Tier 1), the fraction of non-performing loans with respect to total loans as a proxy for credit risk (NPL), total loans divided by total assets (Loans) and the ratio between total assets and the number of employees (Productivity) that we employ as a proxy for bank productivity. The key summary statistics for all the variables employed in the main analysis are presented in Panel A of Table 1. Table A2 in the Online Appendix offers a more detailed description of the variables employed in the analysis and the related data sources.

2.3 Comparing the Treated and Control Group and Testing for Parallel Trends

Our empirical strategy requires that the untreated group represents an adequate counterfactual. This section presents several stylized facts to confirm that our setting satisfies this requirement.

2.3.1 Characteristics of Treated and Control Branches and Banks

We start by showing that the branches, and the related commercial banks, in the treated and control groups are sufficiently similar in their characteristics before the cyberattack. This comparison is important for two reasons. First, it allows us to alleviate concerns related to

the propensity of banks to be targets of cyberattacks conditional on their observable financial characteristics. For instance, Kamiya et al. (2020) show that firms that are more profitable are more likely to be targets of cyberattacks. Second, it also alleviates concerns that the two groups of banks differ along unobservable dimensions that might bias our results (Roberts and Whited, 2013).

Panel B of Table 1 reports the results of this comparison. Columns (2) and (3) present the average values of our dependent variable as well as bank controls for the treated and control group in the year before the cyberattack. Column (4) reports the normalized differences in branch and bank characteristics between the two groups computed as follows (Brown and Earle 2017; Nicoletti 2018):

$$\text{NDIFF} = \frac{\bar{x}_i - \bar{x}_j}{\sqrt{s_i^2 + s_j^2}}, \tag{2}$$

Where $\bar{x}_i$ (s_i^2) is the mean (variance) of a variable for the untreated group and $\bar{x}_j$ (s_j^2) is the mean (variance) of the same variable for the treated group. We note that the differences between the untreated and the treated group are below the threshold value of 0.25. Imbens and Wooldridge (2009) highlight that a value below this threshold is necessary to ensure that the two groups of observations are sufficiently homogeneous.

2.3.2 Parallel Trends Assumption

A key assumption when using difference-in-differences analyses is that, absent cyberattacks, treated and untreated branches would have shown a similar evolution in the (log transformation) of deposits (parallel trends assumption). However, this assumption cannot be directly validated because we are unable to observe the evolution of deposits in the treated group in the absence of a cyberattack. Instead, we rely on several conventional approaches in the literature to show that the parallel trends assumption is plausible. These approaches consider trend differentials between the treated and control group before the occurrence of

an exogenous event. If the two groups of branches follow similar trends in the evolution of deposits prior to the cyberattack, the parallel trends assumption is plausible.

We conduct three analyses to investigate pre-shock trend dynamics in the two groups. First, we follow Lemmon and Roberts (2010) and report the average one-year change in the dependent variable across the two groups of branches in each of the 3 years preceding the cyberattack. These average values are reported in the first two columns of Panel C in Table 1. In column (3), we test if these averages significantly differ between the two groups of branches using t-tests. For the parallel trends assumption to be plausible, the differences should not be statistically different from zero. The results in column (4) show this is the case.

Second, in Panel D of Table 1, we follow the approach of Bertrand and Mullainathan (2003) and Chen et al. (2018) to test for any pre-shock differentials in the evolution of the variable Ln(Deposits). We estimate regression specifications with Ln(Deposits) as the dependent variable and interact our Treated dummy with yearly dummies (D_{jt}) for each of the individual years around the cyberattack. All models include branch and county $\times$ year fixed effects and are estimated with and without bank controls. The parallel trends assumption is plausible if we observe no significant differences in the deposit dynamics of the two groups of branches in the years prior to the shock. Along these lines, we find that none of the coefficients of the interaction terms between the treated and year dummy variables before the cyberattack are statistically significant. Furthermore, the coefficients of the interaction terms are statistically significant in the years after the shock. This suggests that concerns of reverse causality are also unlikely.

[FIGURE 1 HERE]

Finally, Figure 1 plots the trend in Ln(Deposits) for the two groups of branches in the pre-cyberattack period. We estimate the trends from a linear specification that includes branch and county $\times$ year fixed effects as well as bank controls. The estimated values of

Ln(Deposits) in Figure 1 do not reveal any discernible differences in the trends of the two groups before the cyberattack. Overall, our tests suggest that the parallel trends assumption is likely to be valid in our setting.

3 Main Empirical Results

3.1 Cyberattacks and Deposits

This section presents our baseline results. Panel A of Table 2 shows a simple univariate difference-in-differences analysis to estimate the average treatment effect. We compute the average difference in Ln(Deposits) between the post and the pre-event period for groups of treated and untreated branches and then test whether these differences significantly differ between the two groups (using a t-test of equality of means). We find that, although both groups show a significant increase in Ln(Deposits) over the event window, the increase is significantly smaller for treated branches.

[TABLE 2 HERE]

In Panel B of Table 2, we extend the analysis to a multivariate setting based on equation (1). As mentioned earlier, the key coefficient is the interaction term Treated $\times$ Post that measures the change in the dependent variable Ln(Deposits) in the treated group from the pre-shock period to the post-shock period compared to the same change observed for the control group. In column (1), we report the estimates from a model that only includes branch and county $\times$ year fixed effects. In column (2), we control for bank size and lastly, in column (3), we add the remaining controls.

Throughout all specifications, and in line with the results from the univariate analysis, the coefficient of Treated $\times$ Post is negative and statistically significant at the 1% level. The coefficient ranges from -0.216 in column (3) to -0.250 in column (1). Ultimately, the results

consistently indicate that, compared to the control group, branches of banks affected by a cyberattack experience a decrease in the growth rate of their deposits. The magnitude of this decrease is also economically large: using the model in column (3), we find that treated branches report a deposit growth rate that is approximately 22 percentage points lower than the growth rate of the control group. Notably, none of the controls have a significant effect on the dependent variable.

A possible concern for our results is that the matching between treated and untreated banks does not fully remove the influence of unobserved bank heterogeneity due to size differentials. In Table A3 in the Online Appendix, we address this concern using a tighter size matching. Specifically, we divide the two size bins, banks up to \$1bln and banks from \$1bln to \$10bln, into quartiles. For instance, the first quartile of the first (second) size bin goes up to \$250mln (\$2.5bln). We then match banks in the treated group with untreated banks falling in the same quartile within each size category. As shown in Table A3 in the Online Appendix, our key findings remain largely unchanged. It should be noted that applying the tighter matching approach significantly reduces the number of observations that enter the regression analysis (we lose approximately 70% of observations). Therefore, we rely on the wider size bins in our main analysis[5].

In summary, we document a significant negative depositor reaction after cyberattacks that led to the loss of sensitive customer information. Crucially, this finding implies that depositors react to concerns over the soundness of small banks' cybersecurity systems and have expectations that their data be kept safe from malicious hackers. While other studies on non-financial firms have shown firm-level consequences of cyberattacks, taking the perspective of shareholders of large firms (see, for instance, Akey et al. (2021) and Kamiya et al. (2020)), we show how cyberattacks have implications on small firms through key stakeholders (namely, depositors).

[5]In this respect, it is important to note that the difference-in-differences model does not require similar deposit levels in the treated and untreated banks prior to the shock. It only requires similarity in trends as discussed in our analysis in the previous section.

3.2 Robustness Tests

3.2.1 Alternative Econometric Specifications

In the Online Appendix, we report additional specifications that document the robustness of our findings. First, we employ a different set of fixed effects (compared to our main specification in Table 2). In Panel A of Table A4, we follow Gormley and Matsa (2011) and replace branch fixed effects with branch × cohort fixed effects. Gormley and Matsa (2011) argue that allowing firm (branch) fixed effect to vary by cohort is a more conservative approach than using firm (branch) fixed effects. Despite the more conservative specification, we do not find any material change in our results. In Panel B of the same Table, we replace county × year fixed effects with state × year fixed effects to account for the possibility that demand factors in the deposit market are influenced by state-level variables. Again, our results remain largely unchanged.

We next address potential concerns related to standard errors. Bertrand et al. (2004) argue that biased standard errors might arise from the analysis of serially correlated outcomes. To mitigate this potential bias, we follow their approach and collapse the estimation period to one period before and one period after the shock by using the average values of Ln(Deposits) (as well as the other variables in the model) computed for the pre and post 3-year event window employed in our main test. This test, reported in Panel A of Table A5 confirms our main findings. In Panel B, we cluster the standard errors at the branch (and not at the bank) level to account for potential within-branch correlation in the evolution of deposits. Again, our findings remain unchanged.

Next, we test the robustness of our results to different estimation windows. To this end, we repeat the analyses using shorter estimation windows to reduce the possibility of noise biasing the treatment effects. It should also be noted that another advantage of using shorter estimation windows is that it partially alleviates issues of serially correlated outcomes as mentioned above. Panel A Table A6 uses a (-2;+2) years estimation window while Panel

B employs a (-1;+1) year window. Regardless of the estimation window we employ, our results remain intact.

3.2.2 Alternative Dependent Variables: Bank-County Level Analysis

In this section, we collapse our branch-county level observations to the bank-county level. While less granular, this approach allows us to reduce the possibility that any noise or outliers at the branch-level might be driving our results by aggregating the effects over a larger geographical region. Furthermore, it also allows us to understand the overall effect of cyberattacks on deposit growth rates in local markets.

We implement the analysis above using the sample of matched treated and untreated banks we have previously employed. More precisely, we use two dependent variables to assess the overall effect of the cyberattack at the county level. The first is the log transformation of the total amount of deposits of each bank in our sample in a given county. The second is the deposit market share of the same bank. Each model includes bank and county $\times$ year fixed effects, and is estimated with and without bank controls. We cluster standard errors at the bank level.

[TABLE 3 HERE]

Panel A of Table 3 show that, in line with our branch-level analysis, the deposits of hacked banks in a given county show a relative decline in the growth rate as compared to banks in the control group. Panel B of Table 3, where we focus on market share, confirms the negative effects of cyberattacks on hacked banks. The coefficients on Treated $\times$ Post indicate a decrease of approximately 1 percentage point in the county market share of treated banks compared to banks in the control group. This decline is economically meaningful given that the average county market share of a treated bank prior to cyberattacks is approximately 7.2%. In relative terms, the market share of treated banks decreases by approximately 14%.

Hence, the results in this section are consistent with our initial finding that the loss of sensitive customer information due to cyberattacks leads to significant slowdowns in the deposit growth rate of hacked small banks. As a result of this slowdown, these banks decrease their market share in local deposit markets.

3.3 Are Depositors Really Concerned about Cybersecurity?

3.3.1 Endogenous Data Breaches

Data breaches can also occur endogenously as a result of day-to-day bank operations. Examples of these types of events include the loss of portable and stationary devices by employees, paper documents, unintended disclosure of information and as well as fraud resulting from the malicious intent of insiders[6]. Importantly, these other breaches do not result in the same widespread loss of costumer information as hacks and are, therefore, considered narrower in nature (Kamiya et al., 2020). It follows that if our initial finding is indeed due to depositor concerns over the soundness of a small bank's cybersecurity environment, and the resulting risk of widespread losses of sensitive information, we should not observe a similar reaction for these other less severe data breaches.

To validate this conjecture, we identify 21 other breaches in the PRC database involving small banks. We next construct a control sample of untreated small banks using the matching method discussed in section 2.1 and re-estimate equation (1). We report the results of our analysis of other breaches in Table 4. In the first three columns (column (1) does not include bank controls, column (2) adds our size control while column (3) includes our full set of control variables), we do not find evidence of depositors reacting to other breaches

[6]In addition to breaches identified as HACK, PRC identifies seven other breach categories: CARD, defined as breaches that involve credit or debit card fraud such as skimming of devices at point-of-service terminals; INSD, a breach by an insider such as an employee, contractor or customer; PHYS, a breach involving loss or stolen paper documents; PORT, a breach that results from lost, discarded or stolen portable devices; STAT, a breach as a result of lost, inappropriately accessed, discarded or stolen computers or servers not designed for mobility; DISC, a breach as a result of an unintended disclose that does not involve hacking, intentional breaches or physical loss of information or hardware and finally; UNKN, breaches that PRC does not have sufficient information to appropriately classify.

as they do to cyberattacks. In columns (4) to (7), we arrive at the same conclusion when we re-estimate the model separately for different typologies of other data breaches from PRC: column (4) includes breaches by losses of portable devices and unintended disclosures; column (5) uses breaches caused by credit or debit card fraud; column (6) contains breaches by insiders and; column (7) breaches where the reason is unknown.

[TABLE 4 HERE]

The evidence in this section implies that cybersecurity concerns due to external threats on small banks drive our initial results. Further, we do not find any evidence of depositor reaction to other data breach events that are unrelated to a small bank's cybersecurity system. These findings highlight the importance of investments in cybersecurity for small banks.

3.3.2 Falsification Test

We implement a falsification test to further validate the causal interpretation of our results. We report the results in Table A7 in the Online Appendix. We assume that the cyberattacks occurred seven years prior to their actual date and re-estimate the difference-in-differences model 3 years before (after) the placebo date. By moving the event-window 7 years back, we avoid any overlap between the post-estimation window in the placebo test and the pre-estimation window in our original empirical setting. Since we create an artificial setting where we assume a false timing for the shocks to our treated banks, we should not observe any changes in deposit growth for the branches of treated banks.

To conduct the test, we interact a dummy (Treated Fake) equal to one for the banks that have suffered from a cyberattack in our original setting with a dummy (Post Fake) taking a value of one in the three years after the falsely-dated cyberattack. Consistent with our expectation, the analysis shows that the interaction term Treated Fake $\times$ Post Fake

does not enter significantly into any specification. This "non-result" further supports the interpretation that the effect we document in our main analysis (Table 2) likely captures the responses of depositors to cyberattacks and not intrinsic specificities of the small banks in the treated group.

4 Why do Depositors React to Cyberattacks?

Our analyses in the previous section show that cyberattacks lead to decreases in deposit growth rates. In this section, we conduct various tests to understand why depositors react to cyberattacks.

4.1 Loss of Trust

Why should cyberattacks lead to loss of trust in banks and lower deposit growth? Trust is defined as "*... the expectation that another person (or institution) will perform actions that are beneficial, or at least not detrimental to us regardless of our capacity to monitor those actions*" (Gambetta et al., 2000). Indeed, every financial transaction is, and has within itself, an element of trust (Arrow, 1972). Following this view, trust is pivotal to induce individuals to stipulate financial contracts (Guiso et al., 2006, 2008).

The contractual agreement between banks and depositors is built on the trust that banks will safeguard depositors' savings and their confidential information. When cyberattacks occur, the trust between banks and depositors is broken (Kamiya et al., 2020). Depositors might then react to the loss of trust by withdrawing their deposits (or by avoiding relationships with the affected banks). In line with this view, Sapienza and Zingales (2012) show, through survey evidence, that lower trust in banks increases the probability of deposit withdrawals. Similarly, Mourouzidou-Damtsa et al. (2019) find that higher levels of trust in a country are associated with higher levels of deposits due to better retention and loyalty of customers. Building on these arguments, we design two tests to show that a loss of trust is

a key driver behind the observed reduction in deposit growth at hacked small banks.

4.1.1 Small Bank Deposits and High Social Capital

Our first test relies on the idea that the social capital of a firm is closely related to how trustworthy a firm is perceived to be (Lins et al., 2017). If a firm's social capital helps build stakeholder trust, it should pay off in times when being trustworthy is more valuable (Putnam, 1993, 2000). This is the case after a cyberattack has occurred (Akey et al., 2021). In line with this argument, Lins et al. (2017) finds that firms with high social capital perform better in crises and are able to raise more debt. Accordingly, we expect banks with higher social capital to experience a less severe decline in the growth rate of their deposits after a cyberattack.

Following Hasan et al. (2017, 2020), we proxy for a firms' social capital using the social capital of the county in which the bank is headquartered. Specifically, we use the interpolated values of a county-level social capital index constructed by Rupasingha et al. (2006) and maintained by NRCRD at Pennsylvania State University[7]. Treated banks are sorted into high (low) social capital groups if they are above (below) the median social capital index as measured the year before cyberattacks (Treated High (Low) Social Capital Index). We employ the following specification (see Irani and Oesch (2016)):

$$\begin{aligned} \text{Ln(Deposits)}_{i,j,z,t} = \alpha &+ \beta_1(\text{Treated High Social Capital} \times \text{Post}) \\ &+ \beta_2(\text{Treated Low Social Capital} \times \text{Post}) \\ &+ \mathbf{BRANCH} + \mathbf{COUNTY} \times \mathbf{TIME} + \varepsilon_{i,j,z,c,t}, \end{aligned} \tag{3}$$

where β_1 (β_2) measures the differential impact of a cyberattack for the group of branches of banks which are headquartered in counties with high (low) social capital. In line with our baseline model, we estimate equation (3) with and without bank controls.

[7]The social capital index is created using principal component analysis of 4 different factors (voter turnout, census response rate, density of social and non-profit organizations). The index is available for years 1997, 2005, 2009 and 2014. It can be obtained from: https://aese.psu.edu/nercrd/community/social-capital-resources.

The results are displayed in Panel A of Table 5. The coefficient of the interaction terms Treated High Social Capital × Post and Treated Low Social Capital × Post are both negative and significant, but affected banks with a higher social capital experience a smaller decline in deposit growth rates following a cyberattack. The difference between the two groups of hacked banks is statistically (according to a t-test of equality) and economically significant; the relative decline in growth rates is approximately 27 percentage points smaller in hacked banks with high social capital (column 3). This result is aligned with our expectation that depositors' initial trust levels have differential effects in mitigating reputational damages from data breaches.

[TABLE 5 HERE]

In Panel B of Table 5, we repeat the analysis using a bank's Community Reinvestment Act (CRA) rating as an alternative measure of bank social capital. Chen et al. (2019) employ CRA ratings as a proxy for a bank's social performance and highlight the importance that depositors assign to this measure. We re-estimate equation (3) by separating the branches of treated banks with an "outstanding" rating (the highest rating achievable in the four-tiered system) from the remaining treated banks. Consistent with our earlier findings, we do not find a decline in deposits for small banks with "outstanding" CRA ratings; declines in deposit growth rates are only observed for banks with lower ratings (social performance).

The results in this section on the moderating role of a bank's social capital imply that the loss of trust experienced by depositors is a key driver of the slowdown in deposit growth of hacked banks. This result is consistent with the evidence on the value assigned by stakeholders to corporate trustworthiness (measured via social capital) around negative events (Lins et al., 2017).

4.1.2 Deposit Funding Costs

Our second test investigates if depositors alter their transactional terms with affected banks after cyberattacks. As Kamiya et al. (2020) and Karpoff (2012) document, stakeholders might demand better contractual terms to transact with firms negatively affected by shocks to their trustworthiness and reputation. Consequently, depositors might require higher interest rates for their deposits to maintain (or establish) contractual relationships with affected banks.

We use branch-level deposit rates from RateWatch to test our assertion[8]. We identify rates offered on: 1) all deposit products (All rates); 2) certificate of deposits (CD rates), 3) money market deposit accounts (MM rates) and; 4) savings accounts (SAVS rates). Initially, we focus on all interest rate products to understand if there is any overall change in total funding costs. Next, we focus on CD, MM and SAVS rates as these are the three most representative categories of time and savings deposits used by bank customers (Drechsler et al., 2017, 2018).

To conduct our analysis, we estimate the following difference-in-differences model:

$$\begin{aligned} \text{Ln(Rates)}_{p,i,j,z,c,t} = \alpha + \beta_1 \text{Treated} \times \text{Post} + \mathbf{BRANCH} \times \mathbf{PRODUCT} \\ + \mathbf{COUNTY} \times \mathbf{TIME} + \varepsilon_{p,i,j,z,c,t}, \end{aligned} \tag{4}$$

Where p is the product belonging to branch i of bank j in county z, and belonging to a cohort c at time t. Rates is the logarithmic transformation of the rates offered on deposit products as described earlier. As before, our key explanatory variable is Post $\times$ Treated and measures the change in deposit rates from the pre to the post shock period (as defined in equation (1)) in the group of treated branches compared to the control group. However, differently to our baseline model in (1), we define Post as equal to zero (one) for the 36-months before (after) the month in which a cyberattack occurred[9]. Our choice is motivated by the

[8]RateWatch collects weekly branch level data since 2001 on rates offered for various products (e.g., Certificate of Deposits, Money Market Deposits, Savings Accounts, Interest Checking Accounts) of different nominal amounts and maturities. The dataset covers over 50% of bank branches in the U.S.

[9]Deposit rates are available at weekly intervals. However, we define Post using monthly data because it seems unplausible that depositors receive immediate notification of the hack and react accordingly.

higher frequency at which we can observe deposit rates.

It is important to note that equation (4) includes branch × product fixed effects where product is defined as the unique deposit product that is offered by a branch. For instance, branches offer numerous CD products with varying maturities and principal amounts. Similarly, different principal amounts are offered on MM and Savings accounts. The inclusion of branch × product fixed effects ensures that we are comparing the rates offered on similar products (with similar nominal values and maturities) at treated branches before and after cyberattacks relative to our group of control branches.

Table 6 shows the results. Columns (1) to (3) in Panel A report the findings for the rates offered on all products (Ln(All rates)). As indicated by the positive and statistically significant coefficient on Treated × Post in the first three columns of Panel A, we observe an increase in overall deposit rates after cyberattacks. This finding supports our assertion that cyberattacks leads to a loss of trust in depositors and indicates that depositors require higher rates of return in order to continue or initiate contractual relationships with hacked banks.

[TABLE 6 HERE]

Looking at the results for certificates of deposits (Ln(CD rates), money market deposits (Ln(MM rates)) and savings accounts (Ln(SAVS rates)), we only observe increases in rates offered on certificates of deposits. We interpret these findings as consistent with depositors losing trust in banks after cyberattacks. In fact, as CDs require depositors to commit to a pre-specified maturity, depositors might be more reluctant to place their cash with the bank. This is less problematic for MM and SAVS accounts because they only require depositors to commit to a principal amount.

It is worth noting that an increase in the rates paid on (some) deposits does not imply that the increase is sufficient to avoid declines in deposit growth. In fact, if depositors have heterogeneous utility preferences related to a loss in trust (e.g., Guiso et al. (2004)), the

decrease in deposits that can be offset is constrained by what affected banks can afford in terms of funding costs. In other words, it might be too costly for banks to increase rates to attract (or retain) all depositors that would, in the absence of a cyberattack, have selected the affected bank. Yet, subsequent to a major reputational event such as a cyberattack, there may not even exist a rate at which some depositors would continue to contract with the affected bank. These arguments, therefore, motivate a decline in deposit growth rates at affected banks even in the presence of an increase in the rates for some typologies of deposits.

Taken together, the results discussed in this section provide additional support for the loss in trust by depositors as a key reason behind declines in deposit growth rates at hacked small banks.

4.2 Depositor Sophistication and Response to Cyberattacks

Next, we investigate whether heterogeneity in depositor sophistication matters for the reaction to cyberattacks. Ex-ante, it is unclear if sophisticated depositors who are more informed about cybersecurity and its associated risks should react more strongly to cyberattacks. On the one hand, Chen et al. (2020) document that negative bank performance is primarily understood and penalized by more sophisticated depositors. Similarly, Chen et al. (2019) show that "sophisticated" depositors react more negatively to the disclosure of negative information on bank social performance. Following this line of thought, sophisticated (informed) depositors should show a stronger response to cyberattacks because they are better able to understand and judge the negative consequences of cyberattacks.

On the other hand, the evidence of Chen et al. (2019) and Chen et al. (2020) may not be valid for our setting. For example, the events we consider do not directly raise concerns over bank (social) performance (which might only be understood by sophisticated depositors) but instead more directly affect bank depositors through the exposure of their personal and financial information. Furthermore, Duffie and Younger (2019) and Eisenbach et al. (2020) suggest that the consequence of cyber risk for depositors can be framed within

theories of bank runs. As a result, what we observe might not necessarily be driven by the ability of depositors to adequately interpret information on issues related to cybersecurity. Additionally, cyberattacks might cause more anxiety in unsophisticated depositors as they are less likely to understand the exact ramifications of cyberattacks (Solove and Citron, 2017). Therefore, the reaction of depositors to cyberattacks might be driven by uninformed (unsophisticated) depositors that overreact to the shock.

To understand the role of depositor sophistication in relation to our results, we differentiate depositors on the basis of their degree of "digital literacy" that we measure using several socioeconomic characteristics of the local deposit market.

The first measure is based on estimates of the percentage of broadband subscriptions in a county provided by Tolbert and Mossberger (2020). The second is from Form 477 on internet access connections per thousands of households at the county level provided by the Federal Communication Commission[10]. As in equation (3), we define counties with values of digital literacy above (below) the median to measure the differential impact of cyberattacks on depositors with varying levels of digital sophistication.

[TABLE 7 HERE]

The results reported in Table 7 Panel A show that the relative decline in deposits in the treated group is stronger in counties where depositors show (plausibly) low levels of digital literacy. In particular, we find that only the coefficient of Treated Low Digital Literacy $\times$ Post is negative and significant across all specifications regardless of the proxy we employ. In Panel B, we provide further support for the conclusion above by repeating the analysis with more indirect proxies of digital literacy. The first is the median household income in a county taken from the US Census bureau (with higher values denoting more digital literacy). The second is the per capita income from dividends, interests and rents with larger values indicating more depositor financial sophistication, and implicitly, higher digital

[10]The data is available at https://www.fcc.gov/general/broadband-deployment-data-fcc-form-477.

literacy from Bureau of Economic Analysis. Using these alternative measures, we continue to find significantly larger declines in deposit growth rates in counties with (plausibly) lower levels of digital literacy.

Our results indicate that the negative consequences of cyberattacks on deposit growth rates are driven by unsophisticated depositors who might not be able to fully understand the ramifications and consequences of these attacks and, consequently, the actions needed to avoid or limit financial losses. Our results imply that these depositors seem more likely to lose trust in affected small banks.

4.3 Are Depositors Reacting to Bank Risk?

It might be argued that the relative decline in deposit growth rates in the treated group is not due to a loss of trust in banks (primarily by unsophisticated depositors) but is the consequence of a reaction to bank fundamentals wherein depositors differentiate between riskier and safer banks. For instance, if cyberattacks lead to significant IT and remediation costs as well as regulatory penalties, bank fragility might increase. As a result, depositors would penalize riskier banks due to concerns over future bank soundness (e.g., Martinez Peria and Schmukler (2001)). However, this explanation is unlikely to explain the observed effect. For instance, Kamiya et al. (2020) observe that the direct out-of-pocket costs (e.g., investigation and remediation costs, legal and regulatory penalties) resulting from cyberattacks only account for approximately 1% of the loss in market value. This implies that the remaining value losses are due to damages to a firms' reputation and a loss of trust. Further, the fact that the deposit declines are stronger for unsophisticated depositors makes it unlikely that what we observe reflects depositors' ability to differentiate between weak and strong treated banks.

To corroborate the arguments above, we present two tests to show that concerns about bank risk are unlikely to drive the relative decline in deposit growth rates of the branches of treated banks. The results are reported in Table 8.

[TABLE 8 HERE]

First, in the first three columns of Panel A, we split treated banks by their riskiness (as measured by the log of Z-score) the year before the cyberattacks[11]. We denote treated banks to be riskier (less risky) if their Z-score is below (above) the median in the year before the cyberattack. As observed, the coefficients on the interaction of the post dummy with the two treated groups are similar and not statistically different. The last three columns of Panel A show similar results when riskier banks are defined as banks jointly having NPL and Tier 1 ratios above (below) the sample median. The results of these two tests suggest that depositor reaction is not due to concerns over bank risk.

Second, in Panel B of Table 8, we sequentially interact our vector of bank controls (Size, ROA, NPL, Tier 1, Loan, Productivity) with Post to investigate if any change in bank characteristics post-shock are significant in dampening the economic significance of Treated $\times$ Post. If bank performance or risk were crucial in depositors' reaction, we should observe significant decreases in the economic significance of the coefficient of Treated $\times$ Post after we control for these variables. As observed, the only bank characteristic that leads to a small detectable change (reduction) in the economic magnitude of Treated $\times$ Post is Size (column (1)). The inclusion of other bank controls (such as tier 1 capital, which captures bank stability) does not seem to induce any significant change in Treated $\times$ Post; the coefficient on the interaction of interest remains fairly stable throughout the different specifications. This indicates that the expected evolution of bank risk is unlikely to be the reason for declines in deposit growth rates in the group of treated banks.

[11] Z score is calculated as ROA plus the equity ratio divided by the standard deviation of ROA (that we compute using a 3-year window prior to the cyber shock).

5 Spillovers Effects in Local Deposit Markets

Our baseline analysis does not consider the possibility of spillover effects within local deposit markets[12]. In empirical settings involving companies operating in the same industry (and often in the same geographic markets), the assumption of no spillovers might not necessarily be valid. For instance, Kamiya et al. (2020) show negative spillovers at the industry level after successful cyberattacks on non-financial firms.

Nevertheless, differently from Kamiya et al. (2020), we focus on small unlisted banks and not on large listed firms. Therefore, it seems unlikely that depositors would perceive successful cyberattacks on these small banks as indicative of a negative industry wide shock that impacts their confidence in all other institutions. Instead, the negative reputational consequences of these attacks are likely to remain bank specific. Under an "equilibrium framework" for deposit markets, at least part of the withdrawn or not-deposited funding at affected small banks would then be reallocated to other banks operating in the same local market. In this respect, the reallocation should primarily favor larger banks that should benefit from reputational advantages since they might be seen as (digitally and technologically) safer by customers (e.g., Chen et al. (2017)).

The positive spillovers would also be consistent with the view that investment and innovation gaps between competing firms generate a business stealing effect in favor of those firms with an innovation advantage (Bloom et al., 2013). A key implication of this finding would then be the increasing presence of large banks in local deposit markets. Since large banks might be less inclined to supply small business lending, especially in times of crises, local businesses could then face increasing financial frictions (Bloom et al., 2013; Chen et al., 2017).

Furthermore, if any negative spillovers would emerge as in Kamiya et al. (2020), they

[12]The assumption of a lack of spillover effects is rooted in any conventional difference-in-differences framework that excludes interferences across units by formally requiring that the Stable Unit Treatment Value Assumption (SUTVA) holds. This assumption postulates that there are no indirect effects arising from treatment related to externalities that can influence the control group after treatment (Boehmer et al., 2020).

are plausibly more likely to impact on other (unaffected) small banks as their cybersecurity environment might be perceived by depositors as equally vulnerable. This perception might then motivate panic-based bank runs but only in these small institutions (e.g., Calomiris and Mason (2003) and Diamond and Dybvig (1983)).

In the next two sections, we elaborate on the arguments above and test for two different typologies of spillovers in local deposit markets: a) towards large banks (that is, banks with total assets over $10bln) and; b) towards small banks. To test for the presence of these (potential) spillover effects, we compare the evolution of deposits in the branches of untreated banks in the counties where the affected banks operate with the branches of the same untreated banks operating in adjacent counties (where hacked banks do not operate). Our identification strategy is graphically presented in Figure 2, where we illustrate examples of the treated (in red) and untreated (in blue) counties for our spillover analysis.

More precisely, for each affected county where untreated large (small) banks operate (our original control group), we select the adjacent counties where branches of the same untreated large (small) banks operate and estimate a similar model to equation (1). For instance, Figure 2a graphically illustrates the cyberattack on Salem Five Savings Bank in Massachusetts in 2016. This treated bank had branches in the counties of Middlessex and Norfolk but not, for instance, in the counties of Worchester and Bristol. In the spillover test, Middlessex and Norfolk are still treated counties whereas Worchester and Bristol are categorized as untreated counties.

[FIGURE 2 HERE]

More generally, the branches of unaffected banks that operate in the treated counties (in red) are considered (indirectly) "treated" whereas the branches of these banks in adjacent counties (in blue) are defined as untreated. By focusing on adjacent counties, we ensure that the two groups of branches are likely to be affected by similar observable and unobservable economic and social conditions (Huang, 2008). Furthermore, this setting alleviates concerns

of omitted bank characteristics driving our results since treated and untreated branches belong to the same bank.

5.1 Spillover Effects towards Large Banks

This section examines if successful cyberattacks on small banks produce spillovers towards larger banks. Before conducting the test, we start by showing in Panel A of Table 9 that there is no evidence of trend differentials in deposit growth prior to the shock between our treated (branches belonging to large banks in affected counties) and control group (branches belonging to the same large banks but residing in adjacent unaffected counties). This suggests that similar to our main analysis, the parallel trends assumption is also likely to hold for this test.

The regression results of our analysis, reported in Panel B of Table 9, consistently indicate an increase in the deposit growth rates at branches of large banks located in the counties affected by cyberattacks compared to the branches of the same banks in unaffected adjacent counties. The differential increase in deposit growth rate in the two groups of branches is approximately equal to 15 percentage points.

[TABLE 9 HERE]

Our findings, therefore, are consistent with the importance of trust established in our main analysis and highlight a "flight-to-reputation" effect in local deposit markets associated with cyberattacks on small banks. To offer further support for this interpretation, we extend the analysis in Panel C of Table 9 by separating our sample of "treated" branches of large banks using the customer reputation score assigned to these banks from the annual survey conducted by American Banker. We define banks ranked in the top 5 of the survey as having an exceptional reputation with their customers[13]. We find support for our interpretation of

[13]American Banker surveys bank customers on their perception of bank reputation. As this survey mainly

a "flight-to-reputation" effect, as we observe that positive spillovers are significantly larger for large banks with exceptional reputation according to their customers.

In summary, large banks benefit from reputational advantages in terms of cybersecurity and are then able to attract more deposits after small banks are hacked. This finding can be understood in the context of theoretical models in which investment and innovation gaps among firms generate a business stealing effect in favor of more innovative firms (Bloom et al., 2013).

5.2 Spillover Effects towards (untreated) Small Banks

Next, we examine spillovers to (untreated) small banks. To conduct this test, we define small banks in our initial control group as indirectly treated. The untreated group consists of the branches of these banks that reside in unaffected adjacent counties. As before, we show in Panel D of Table 9 that the parallel trends assumption is likely to be plausible; there are no differences in the dynamics of deposit growth in the treated and untreated group prior to the event.

The results of our analysis are presented in Panel E of Table 9. Across all specifications, we find no evidence that the growth rate of deposits at (indirectly) treated branches is significantly different from the growth rate observed for untreated branches of the same small banks.

Two conclusions emerge from these tests. First, it is unlikely that our initial estimates of the average treatment effects are biased; small banks in our initial control group do not show any significant changes in deposit growth rates after cyberattacks. Second, and more generally, there is no evidence that a cyberattack on a small bank has any spillover effects to similar untreated small banks.

covers large banks, we are unable to employ this test for our sample of small banks. Because we only have survey data from 2010 to 2017, and our sample begins in 2005, we use the reputation scores from 2010 for years 2005 to 2009. In Table A8 in the Online Appendix, we present the results of our additional results based on the top 10 or 15 banks in terms of reputation. Our results and conclusions remain unchanged.

5.3 Market Structure and Spillover Effects

The degree of deposit market concentration could also affect the spillover effects. In particular, it might be argued that in local markets where there is less competition, large banks are more likely to gain from their reputational advantage. Alternatively, it could be suggested that untreated small banks are more likely to capture some of the diverted deposit flows from affected banks when they face less competitive pressure.

[TABLE 10 HERE]

To test if there are heterogenous spillover effects to both untreated large and small banks, we estimate an equation similar to (3) where we categorize the treated counties by high (above the median) and low (below the median) degrees of deposit market concentration. We measure market concentration using the total market share of the top 3 (CR3) and top 5 (CR5) banks in a county (e.g., Bushman and Wittenberg-Moerman (2012) and Ross (2010)). We report the results of these tests in Table 10 . Panel A shows that the spillovers in favor of large banks are significantly more pronounced when local deposit markets are more concentrated; namely, when bank customers might have fewer alternatives in terms of deposit reallocation. In Table A9 in the Online Appendix we show that this conclusion also holds if we measure deposit market concentration using the Herfindahl-Hirschman Index (HHI). In contrast, we do not observe any differential spillover effects in the group of small banks regardless of the structure of local deposit markets.

Ultimately, this section supports the view that reputational advantage of larger banks enables them to capture the deposits of small hacked banks. Further, this effect is more pronounced in more concentrated markets wherein competition to provide banking services is expected to be weaker. Additionally, regardless of deposit market structure, there is no evidence of spillovers to untreated small banks.

6 Do Cyberattacks Matter for Bank Borrowers?

Besides deposit markets, banks engage in contractual relationships with households in mortgage markets. While cyberattacks do not pose immediate threats to potential borrowers, they might still damage a bank's reputation in mortgage markets and its competitive position (Akey et al., 2021).

We examine the consequences of cyberattacks in relation to mortgage lending in two steps. First, we take the perspective of mortgage applicants and test whether potential borrowers shy away from banks that have suffered cyberattacks and whether the characteristics of these borrowers change. If cyberattacks lead to reputational damages in mortgage markets, we should observe that less risky applicants that have more market alternatives opt for other lenders. It follows that cyberattacks should result in a decrease in the quality of applicants at affected small banks. Second, we analyze a bank's response to borrower behavior in terms of underwriting standards. To maintain their market position, affected banks might be forced to approve riskier loans, resulting in a consequent deterioration of their lending standards.

We base our analysis on loan data from the Home Mortgage Disclosure Act (HMDA) database collected by the Federal Financial Institutions Examination Council (FFIEC)[14]. Each loan application in the HMDA dataset contains information on borrower demographics, loan characteristics, the decision undertaken, the geographical location of the property the year in which the loan application decision is made, and the lender's identifier. However, the HMDA data does not enable us to track the loans submitted to individual branches. As such, our analysis is conducted at the bank-county-year level.

We drop from our sample loan applications where the lender does not have a branch in

[14]HMDA is a loan-level dataset that covers all mortgage applications that have been reviewed by qualified financial institutions, both private and public. HMDA requires an institution to disclose any mortgage lending if it has at least. one branch in any metropolitan statistical area and meets the minimum size threshold. For instance, in 2010, this reporting threshold is $39 million in book assets. The annual reporting criteria can be accessed at: https://www.ffiec.gov/hmda/reporterhistory.htm. Due to the low reporting requirements, the HMDA dataset covers the majority of lenders and accounts for nearly 90% of the U.S. mortgage market (Cortés et al., 2016). We winsorize the variables Applicant Income and Loan Amount at the 5% tails to minimize reporting errors.

the county of the mortgage's location because they are likely to be loans that were submitted to independent mortgage brokers (Cortes, 2015). Given that our initial tests focus on the response of potential borrowers of a bank that are located in geographic proximity to where a cyberattack occurred, retaining these observations would introduce noise into the analysis. We then aggregate HMDA loan-level variables to the bank-county-year and estimate the following difference-in-differences model:

$$\begin{aligned} \text{Lending}_{i,z,t} = \alpha + \beta_1 \text{Treated} \times \text{Post} + \mathbf{BANK} + \mathbf{COUNTY} \times \mathbf{TIME} \\ + \mathbf{CONTROLS}\varepsilon_{i,z,t}, \end{aligned} \tag{5}$$

Where Lending is one of the following variables 1) Num. Loans (the log transformation of the total number of loans submitted in a bank-county-year); 2) Submitted LTI (the average loan amount requested divided by the average income of the applicant in a bank county-year); 3) Approval Rate (number of approved loans/total loans submitted at the bank-county-year level); 4) Approved LTI (the bank-county-year average of loan amount requested in approved loans/applicant income). The first two variables, therefore, take the borrowers' perspective while the remaining variables take the bank's perspective. We use Loan-to-Income ratios as a proxy for the riskiness of a borrower as higher ratios indicate a lower capacity of borrowers to repay these loans, leading to higher borrower defaults (Campbell and Cocco, 2015; Dell'Ariccia et al., 2012).

Our key explanatory variable is Treated × Post and measures the change in one of the lending variables from the pre to the post shock period, defined as in equation (1), in the group of treated banks as compared to the control group. In all specifications, we include a vector of controls consisting of borrower/loan control variables such as Ln(Applicant Income), Avg Female, Avg Native American, Avg Asian, Avg African-American, Avg Hawaiian Native, Avg Conventional, Avg FHA and Avg VA. We provide detailed definition of these variables in Table A2 in the Online Appendix.

[TABLE 11 HERE]

Table 11 shows the results of our analysis. In columns (1) and (2), we do not find evidence of an overall decline in the number of mortgage applications in the sample of the affected banks compared to the control group. However, in columns (3) and (4) we observe a relative increase in the Loan-to-Income ratio of submitted loans for banks that have experienced a cyberattack. The results presented in columns (5) to (8), taking the lender's perspective, suggest that the approval rate of affected banks does not change. However, there is some evidence of an increase in the Loan-to-Income ratio of loans that have been approved.

The results indicate, therefore, that small banks are more likely to attract riskier borrowers subsequent to a cyberattack and are forced to relax their lending standards to maintain their approval rate. These results imply that the negative reputational effects for banks experiencing a cyberattack extend beyond deposit markets and also adversely affect banks in the mortgage market.

7 Conclusion

Cybersecurity is a rising concern for regulators and bankers. Unlike large banks which have a wide range of human and financial resources to strengthen their IT infrastructure against cyberattacks, small banks are likely to be more susceptible. Indeed, CEOs of small community banks have indicated that cyber risks are a major threat to their business (Conference of State Bank Supervisors, 2019). In this paper, we document the validity of this view by identifying the negative business consequences for small banks after cyberattacks and the observed follow-on spillover effects on the distribution of deposits across banks in local markets.

We show that the branches of small banks affected by cyberattacks experience a significant slowdown in the growth rate of their deposits compared to branches of unaffected similarly sized banks. Consequently, this decline leads to a significant decrease in the deposit market share of these banks. The negative effects of cyberattacks in local deposit markets seem to

be driven by the loss of trust by depositors as a result of these attacks. Consistent with this interpretation, we show that the slowdown in deposit growth for affected banks is more pronounced when banks have lower social capital prior to the shock. Affected banks are also forced to increase rates on time deposit products. The loss of trust in small banks is especially evident in the group of depositors with plausibly less knowledge about cyber risks.

Underscoring the key role of trust in deposit markets, we show that cyberattacks generate positive deposit spillovers to branches of unaffected large banks, in particular, those that have an excellent reputation, operating in geographically proximate locales. By contrast, we do not find any evidence of deposit spillovers to smaller unaffected banks. Essentially, cyberattacks led to a "flight-to-reputation" effect as larger, more reputable banks are likely to be seen by depositors as more secure against cyberattacks.

In a final set of tests, we find that the negative effects produced by a cyberattack on target banks also extend to the mortgage market. Hacked small banks attract relatively riskier applicants subsequent to the cyberattack compared to similar but untreated banks. The results also indicate that they are forced to relax their lending standards to maintain approval rates.

Overall, our findings document that cyberattacks lead to significant bank-specific reputational damages undermining the trust of bank customers. This results in a reduced competitive position of small banks due to deteriorated contractual relationships with depositors and borrowers. Therefore, financial constraints that impede cybersecurity investments has the potential to significantly undermine the pivotal role that small banks play in local economies.

Ultimately, our study highlights the need for sectorial cybersecurity initiatives that can complement and support small bank-specific investments in cybersecurity strategies. Yet, equally important appear initiatives to increase depositor awareness of cybersecurity and the implementation of cost recovery options to reduce the negative reputational effects arising from cyberattacks on small banks.

References

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

Akey, P., S. Lewellen, I. Liskovich, and C. Schiller (2021). Hacking corporate reputations. *Working Paper, Rotman School of Management*.

Aldasoro, I., L. Gambacorta, P. Giudici, and T. Leach (2020). Operational and cyber risks in the financial sector. *Working Paper, BIS*.

Arrow, K. J. (1972). Gifts and exchanges. *Philosophy & Public Affairs*, 343–362.

Barakat, A., S. Ashby, P. Fenn, and C. Bryce (2019). Operational risk and reputation in financial institutions: Does media tone make a difference? *Journal of Banking & Finance 98*, 1–24.

Basel Committee on Banking Supervision (2018). Cyber-resilience: Range of practices.

Becker, B. (2007). Geographical segmentation of US capital markets. *Journal of Financial Economics 85*(1), 151–178.

Begley, T. A. and A. Purnanandam (2021). Color and credit: Race, regulation, and the quality of financial services. *Journal of Financial Economics, forthcoming*.

Berger, A. N., C. H. Bouwman, and D. Kim (2017). Small bank comparative advantages in alleviating financial constraints and providing liquidity insurance over time. *The Review of Financial Studies 30*(10), 3416–3454.

Berger, A. N., N. H. Miller, M. A. Petersen, R. G. Rajan, and J. C. Stein (2005). Does function follow organizational form? Evidence from the lending practices of large and small banks. *Journal of Financial Economics 76*(2), 237–269.

Berger, A. N. and R. Turk-Ariss (2015). Do depositors discipline banks and did government actions during the recent crisis reduce this discipline? An international perspective. *Journal of Financial Services Research 48*(2), 103–126.

Bertrand, M., E. Duflo, and S. Mullainathan (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics 119*(1), 249–275.

Bertrand, M. and S. Mullainathan (2003). Enjoying the quiet life? Corporate governance and managerial preferences. *Journal of Political Economy 111*(5), 1043–1075.

Bloom, N., M. Schankerman, and J. Van Reenen (2013). Identifying technology spillovers and product market rivalry. *Econometrica 81*(4), 1347–1393.

Boehmer, E., C. M. Jones, and X. Zhang (2020). Potential pilot problems: Treatment spillovers in financial regulatory experiments. *Journal of Financial Economics 135*(1), 68–87.

Bord, V. M., V. Ivashina, and R. D. Taliaferro (2018). Large banks and small firm lending. *Working Paper, National Bureau of Economic Research*.

Bouveret, A. (2018). *Cyber risk for the financial sector: A framework for quantitative assessment*. International Monetary Fund.

Bushman, R. M. and R. Wittenberg-Moerman (2012). The role of bank reputation in "certifying" future performance implications of borrowers' accounting numbers. *Journal of Accounting Research 50*(4), 883–930.

Calomiris, C. W. and J. R. Mason (2003). Consequences of bank distress during the great depression. *American Economic Review 93*(3), 937–947.

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

Chen, B. S., S. G. Hanson, and J. C. Stein (2017). The decline of big-bank lending to small business: Dynamic impacts on local credit and labor markets. *Working Paper, National Bureau of Economic Research*.

Chen, P.-Y., Y. Hong, and Y. Liu (2018). The value of multidimensional rating systems: Evidence from a natural experiment and randomized experiments. *Management Science 64*(10), 4629–4647.

Chen, Q., I. Goldstein, Z. Huang, and R. Vashishtha (2020). Bank transparency and deposit flows. *Working Paper*.

Chen, Y.-C., M. Hung, and L. L. Wang (2019). Depositors' responses to public nonfinancial disclosure. *Working Paper, Hong Kong University of Science and Technology*.

Chernobai, A., P. Jorion, and F. Yu (2011). The determinants of operational risk in us financial institutions. *Journal of Financial and Quantitative Analysis 46*(6), 1683–1725.

Chernobai, A., A. Ozdagli, and J. Wang (2020). Business complexity and risk management: Evidence from operational risk events in us bank holding companies. *Journal of Monetary Economics, forthcoming*.

Conference of State Bank Supervisors (2019). Community banking in the 21st century.

Cortés, K., R. Duchin, and D. Sosyura (2016). Clouded judgment: The role of sentiment in credit origination. *Journal of Financial Economics 121*(2), 392–413.

Danisewicz, P., D. McGowan, E. Onali, and K. Schaeck (2018). Debt priority structure, market discipline, and bank conduct. *The Review of Financial Studies 31*(11), 4493–4555.

Dell'Ariccia, G., D. Igan, and L. U. Laeven (2012). Credit booms and lending standards: Evidence from the subprime mortgage market. *Journal of Money, Credit and Banking 44*(2-3), 367–384.

Deloitte (2019). Pursuing cybersecurity maturity at financial institutions.

Diamond, D. W. and P. H. Dybvig (1983). Bank runs, deposit insurance, and liquidity. *Journal of Political Economy 91*(3), 401–419.

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

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

Duffie, D. and J. Younger (2019). *Cyber runs*. Brookings.

Eisenbach, T. M., A. Kovner, and M. J. Lee (2020). Cyber risk and the us financial system: A pre-mortem analysis. *Federal Reserve Bank of New York, Staff Report, 909*.

Gambetta, D. et al. (2000). *Can we trust trust*. Oxford University Press.

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

Gormley, T. A. and D. A. Matsa (2011). Growing out of trouble? Corporate responses to liability risk. *The Review of Financial Studies 24*(8), 2781–2821.

Guiso, L., P. Sapienza, and L. Zingales (2004). The role of social capital in financial development. *American Economic Review 94*(3), 526–556.

Guiso, L., P. Sapienza, and L. Zingales (2006). Does culture affect economic outcomes? *Journal of Economic Perspectives 20*(2), 23–48.

Guiso, L., P. Sapienza, and L. Zingales (2008). Trusting the stock market. *The Journal of Finance 63*(6), 2557–2600.

Guo, B., D. Pérez-Castrillo, and A. Toldrà-Simats (2019). Firms' innovation strategy under the shadow of analyst coverage. *Journal of Financial Economics 131*(2), 456–483.

Hakenes, H., I. Hasan, P. Molyneux, and R. Xie (2015). Small banks and local economic development. *Review of Finance 19*(2), 653–683.

Hasan, I., C. K. Hoi, Q. Wu, and H. Zhang (2017). Social capital and debt contracting: Evidence from bank loans and public bonds. *Journal of Financial and Quantitative Analysis 52*(3), 1017–1047.

Hasan, I., C.-K. S. Hoi, Q. Wu, and H. Zhang (2020). Is social capital associated with corporate innovation? evidence from publicly listed firms in the us. *Journal of Corporate Finance 62*, 101623.

Homanen, M. (2018). Depositors disciplining banks: The impact of scandals. *Working Paper, Chicago Booth Research Paper* (28).

Huang, R. R. (2008). Evaluating the real effect of bank branching deregulation: Comparing contiguous counties across us state borders. *Journal of Financial Economics 87*(3), 678–705.

Imbens, G. W. and J. M. Wooldridge (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature 47*(1), 5–86.

Irani, R. M. and D. Oesch (2016). Analyst coverage and real earnings management: Quasi-experimental evidence. *Journal of Financial and Quantitative Analysis 51*(2), 589–627.

Iyer, R., M. Puri, and N. Ryan (2016). A tale of two runs: Depositor responses to bank solvency risk. *The Journal of Finance 71*(6), 2687–2726.

Jacewitz, S. and J. Pogach (2018). Deposit rate advantages at the largest banks. *Journal of Financial Services Research 53*(1), 1–35.

Kamiya, S., J.-K. Kang, J. Kim, A. Milidonis, and R. M. Stulz (2020). Risk management, firm reputation, and the impact of successful cyberattacks on target firms. *Journal of Financial Economics, forthcoming*.

Karpoff, J. M. (2012). *Does reputation work to discipline corporate misconduct*. Oxford University Press Oxford.

Kashyap, A. K. and A. Wetherilt (2019). Some principles for regulating cyber risk. In *AEA Papers and Proceedings*, Volume 109, pp. 482–87.

Lemmon, M. and M. R. Roberts (2010). The response of corporate financing and investment to changes in the supply of credit. *Journal of Financial and Quantitative Analysis 45*(3), 555–587.

Li, H., W. G. No, and J. E. Boritz (2020). Are external auditors concerned about cyber incidents? evidence from audit fees. *Auditing: A Journal of Practice & Theory 39*(1), 151–171.

Lins, K. V., H. Servaes, and A. Tamayo (2017). Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis. *The Journal of Finance 72*(4), 1785–1824.

Martinez Peria, M. S. and S. L. Schmukler (2001). Do depositors punish banks for bad behavior? Market discipline, deposit insurance, and banking crises. *The Journal of Finance 56*(3), 1029–1051.

Mester, L. J. et al. (2019). Cybersecurity and financial stability.

Mian, A. and A. Sufi (2014). What explains the 2007–2009 drop in employment? *Econometrica 82*(6), 2197–2223.

Mourouzidou-Damtsa, S., A. Milidonis, and K. Stathopoulos (2019). National culture and bank risk-taking. *Journal of Financial Stability 40*, 132–143.

Oliveira, R. d. F., R. F. Schiozer, and L. A. d. C. Barros (2015). Depositors' perception of "too-big-to-fail". *Review of Finance 19*(1), 191–227.

Paravisini, D. (2008). Local bank financial constraints and firm access to external finance. *The Journal of Finance 63*(5), 2161–2193.

Putnam, R. (1993). The prosperous community: Social capital and public life. *The American Prospect 13*.

Putnam, R. D. (2000). *Bowling alone: America's declining social capital*. Springer.

Roberts, M. R. and T. M. Whited (2013). Endogeneity in empirical corporate finance. In *Handbook of the Economics of Finance*, Volume 2, pp. 493–572. Elsevier.

Rosati, P., F. Gogolin, and T. Lynn (2019). Audit firm assessments of cyber-security risk: Evidence from audit fees and sec comment letters. *The International Journal of Accounting 54*(03).

Ross, D. G. (2010). The "dominant bank effect:" How high lender reputation affects the information content and terms of bank loans. *The Review of Financial Studies 23*(7), 2730–2756.

Rupasingha, A., S. J. Goetz, and D. Freshwater (2006). The production of social capital in US counties. *The Journal of Socio-Economics 35*(1), 83–101.

Sapienza, P. and L. Zingales (2012). A trust crisis. *International Review of Finance 12*(2), 123–131.

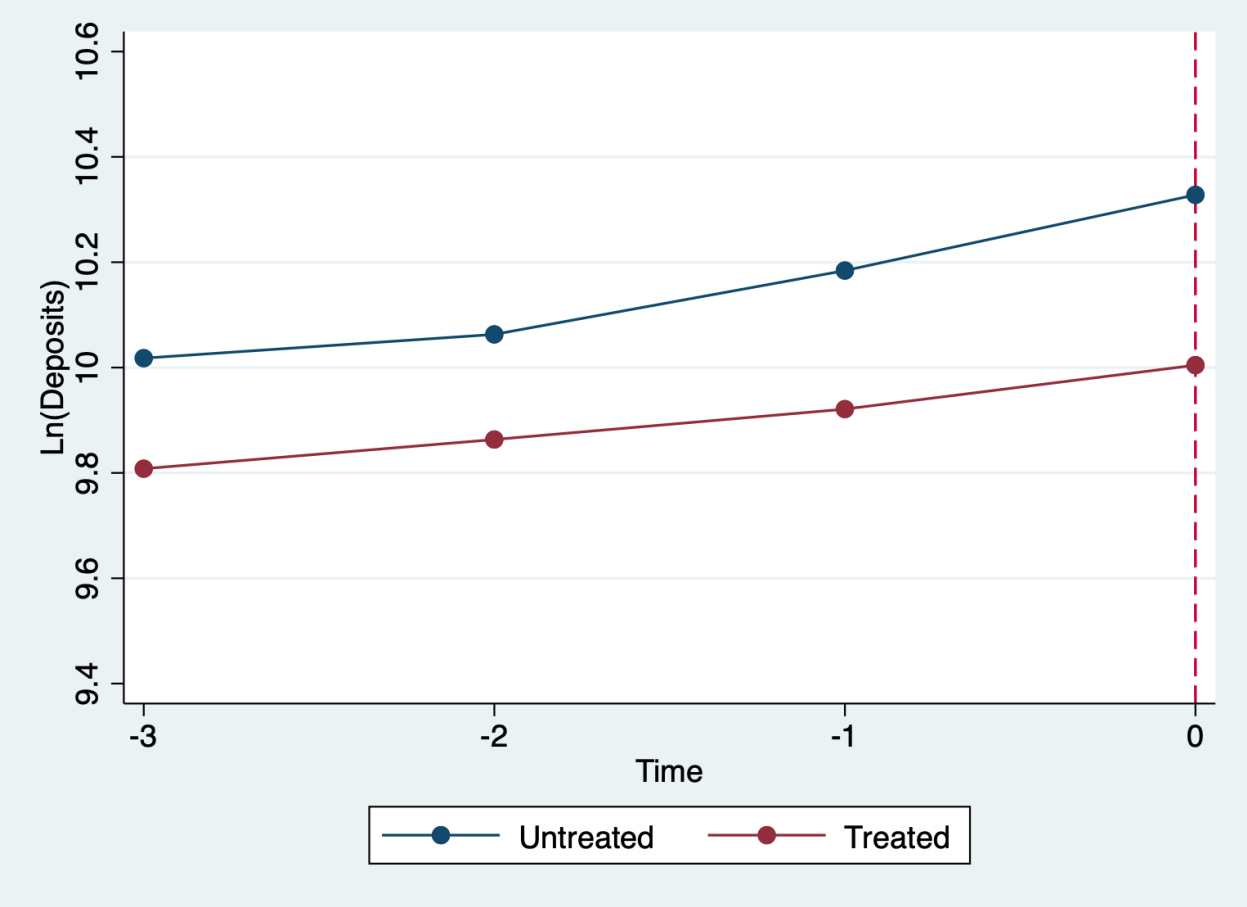
Skrastins, J. and V. Vig (2019). How organizational hierarchy affects information production. *The Review of Financial Studies 32*(2), 564–604.

Solove, D. J. and D. K. Citron (2017). Risk and anxiety: A theory of data-breach harms. *Tex. L. Rev. 96*, 737.

Stein, J. C. (2002). Information production and capital allocation: Decentralized versus hierarchical firms. *The Journal of Finance 57*(5), 1891–1921.

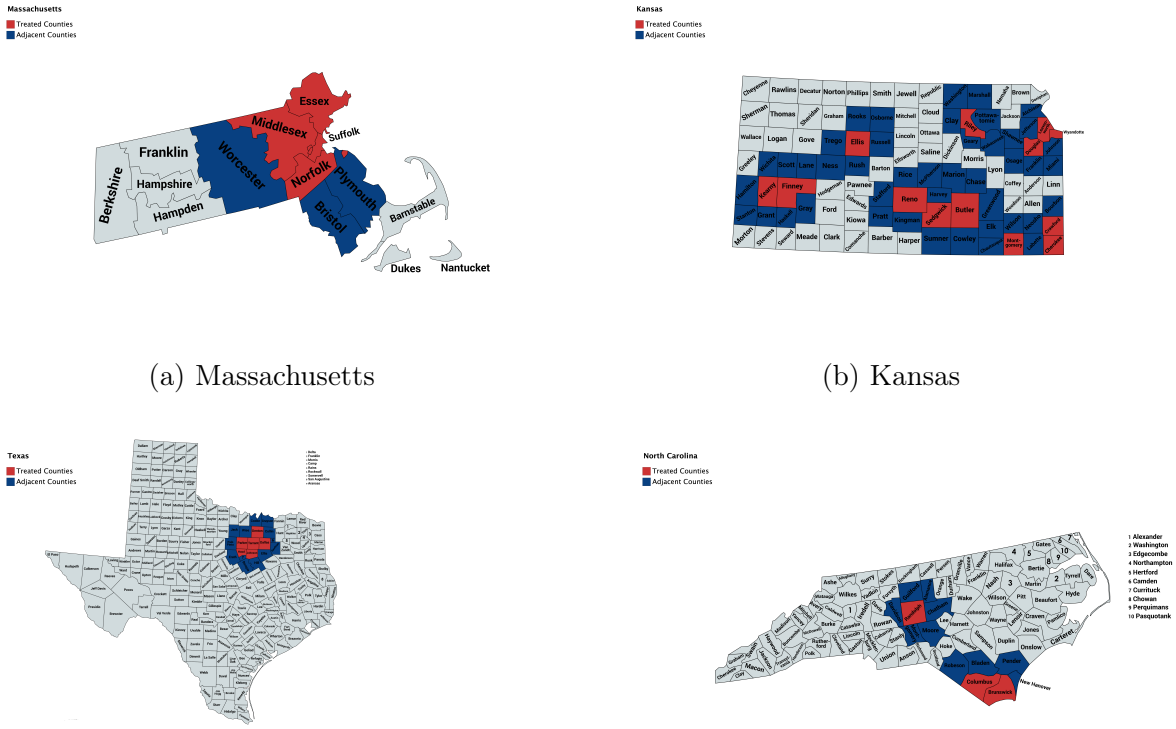
Tolbert, C. and K. Mossberger (2020). U.S. Current Population Survey & American Community Survey Geographic Estimates of Internet Use, 1997-2018.

Figure 1
Evolution of deposits in the pre-shock period



This figure plots the trend in Ln(Deposits) for branches of treated and untreated banks in the 3-year period before the cyberattack. We estimate and plot Ln(Deposits) using a linear model that accounts for branch and county fixed effects and bank controls (Size, ROA, Tier 1, NPL, Loans and Productivity). Size is measured as the logarithmic transformation of bank total assets in thousands of US$. ROA is the ratio between net income and total assets, Tier 1 is total tier 1 capital divided by risk weighted assets, NPL is the fraction of non-performing loans with respect to total loans, Loans is constructed as total loans divided by total assets and Productivity is defined as the ratio between total assets and the number of employees.

Figure 2
Graphic Representation of the Spillover Analysis



(a) Massachusetts

(b) Kansas

(c) Texas

(d) North Carolina

This figure graphically illustrates examples of the treated (in red) and untreated (in blue) counties for our spillover analysis. Treated counties in red are where untreated large (small) banks operate (our original control group). Untreated blue counties are adjacent counties where branches of the same untreated large (small) banks operate. Part (a) graphically illustrates the spillover analysis for the cyberattack of Salem Five Savings Bank in Massachusetts in 2016. This treated bank had branches in the counties of Middlessex and Norfolk but not, for instance, in the counties of Worchester and Bristol. In the spillover test, Middlessex and Norfolk are still treated counties whereas Worchester and Bristol are categorized as untreated counties. In a similar way, Part (b) illustrates the cyberattack of Commerce Bank in Kansas in 2007, Part (c) shows the cyberattack of OmniAmerican Bank in Texas in 2008 and Part (d) displays the cyberattack on Security Savings Bank in North Carolina in 2006.

Table 1
Descriptive Statistics and Parallel Trends

The table below reports descriptive statistics and tests of the parallel trend assumption for our sample of cyberattacks on small banks. Cyberattacks are identified using the breach classification "HACK" by Privacy Rights Clearinghouse (PRC). Panel A provides descriptive statistics of the main variables used in the analyses. Ln(Deposits) is the logarithmic transformation of deposits in thousands of US$. Size is measured as the logarithmic transformation of bank total assets in thousands of US$. ROA is the ratio between net income and total assets, Tier 1 is total tier 1 capital divided by risk weighted assets, NPL is the fraction of non-performing loans with respect to total loans, Loans is constructed as total loans divided by total assets and Productivity is defined as the ratio between total assets and the number of employees. Panel B reports a comparison of the characteristics of treated and control branches and treated and untreated banks in the year prior to a cyberattack. Columns (2) and (3) present the average values of our dependent variable and bank controls while column (4) reports the normalized differences in branch and bank characteristics between the two groups. Panel C reports the average one-year change in the dependent variable across the two groups of branches in each of the 3 years preceding the cyberattack. The average values are reported in column (1) and (2). The differences in average values are reported in column (3) while column (4) reports T-tests on differences in the average values. Panel D reports pre-shock differentials to show the evolution of Ln(Deposits). The regression specifications is estimated with Ln(Deposits) as the dependent variable. Treated is a dummy that equals one if a branch belongs to a bank that has suffered from an exogenous cyberattack in the sample period and zero otherwise. The variable Treated is interacted with yearly dummies for each of the individual years around the cyberattack and with and without bank controls (Size, ROA, Tier 1, NPL, Loans and Productivity). All models include branch and county × year fixed effects. Standard errors are clustered at the bank-level and are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels.

Panel A	Descriptive Statistics					
	Obs. (1)	Mean (2)	Median (3)	SD (4)	25th (5)	75th (6)
Ln(Deposits)	15,460	10.080	10.601	2.316	9.709	11.247
Hack	15,460	0.199	0.000	0.399	0.000	0.000
Post	15,460	0.453	0.000	0.498	0.000	1.000
Size	15,334	14.821	14.905	0.969	14.212	15.556
ROA	14,730	0.010	0.009	0.008	0.007	0.013
NPL	14,730	0.012	0.007	0.018	0.003	0.013
Tier 1	14,730	0.134	0.117	0.052	0.103	0.149
Loan	15,082	0.651	0.670	0.144	0.561	0.760
Productivity	15,080	5.348	4.783	2.930	3.197	6.741

Panel B	Pre-Shock Characteristics				
	N (1)	Treated (A) (2)	Untreated (B) (3)	Normalized Diff. (A-B) (4)	T-test (A-B) (5)
Ln(Deposits)	2,328	10.095	10.038	-0.024	0.6436
Size	243	13.986	13.727	-0.129	0.4627
ROA	243	0.002	0.002	0.003	0.7818
NPL	243	0.014	0.016	0.109	0.6119
Tier 1	243	0.139	0.156	0.195	0.3867
Loan	242	0.661	0.674	0.069	0.7274
Productivity	231	4.823	5.641	0.248	0.2655

Panel C	Parallel Trends			
	Treated (A) (1)	Untreated (B) (2)	Diff. (A-B) (3)	T-value (4)
Δ Ln(Deposits)$_{t-3}$	0.085	0.092	-0.007	0.826
Δ Ln(Deposits)$_{t-2}$	0.080	0.121	-0.041	0.190
Δ Ln(Deposits)$_{t-1}$	0.143	0.143	0.000	0.999

Table 1 (cont.)
Descriptive Statistics and Parallel Trends

Panel D	Pre-Shock Trend Differentials		
	Ln(Deposits)		
	(1)	(2)	(3)
Treated × Dummy (t-3)	0.117	0.109	0.010
	(0.085)	(0.084)	(0.063)
Treated × Dummy (t-2)	0.115	0.107	0.100
	(0.079)	(0.076)	(0.072)
Treated × Dummy (t-1)	0.072	0.069	0.055
	(0.075)	(0.072)	(0.065)
Treated × Dummy (t+1)	-0.153***	-0.152***	-0.149***
	(0.049)	(0.049)	(0.052)
Treated × Dummy (t+2)	-0.151***	-0.148***	-0.141**
	(0.055)	(0.055)	(0.057)
Treated × Dummy (t+3)	-0.238***	-0.232***	-0.233***
	(0.058)	(0.058)	(0.057)
Size Control	No	Yes	Yes
Other Bank Controls	No	No	Yes
Branch FE	Yes	Yes	Yes
County x Year FE	Yes	Yes	Yes
Observations	15460	15334	14382
Adjusted R^2	0.935	0.936	0.936

Table 2
Do Depositors Respond to Cyberattacks?

The table below reports difference-in-differences regression results of cyberattacks on small banks. Cyberattacks are identified using the breach classification "HACK" by Privacy Rights Clearinghouse (PRC). Ln(Deposits) is the logarithmic transformation of the branch-level deposits in US dollar. Panel A shows the results of a univariate difference-in-differences analysis to estimate the average treatment effect. The T-test of equality of means compares the average difference in Ln(Deposits) between the post and the pre-event period for groups of treated and untreated branches and then test whether these differences significantly differ between the two groups. Panel B reports the results of a multivariate analysis (based on equation (1)). Treated is a dummy that equals one if a branch belongs to a hacked bank and zero otherwise; Post is a dummy equal to one in the post-shock window (up to 3 years after the shock). The difference-in-differences estimate of the coefficient of Treated x Post is the difference between how the dependent variable changes in the branches of treated banks (namely, banks affected by a cyberattack) and in the branches of control banks after the shock. Size is measured as the logarithmic transformation of bank total assets in thousands of US$. ROA is the ratio between net income and total assets, Tier 1 is total tier 1 capital divided by risk weighted assets, NPL is the fraction of non-performing loans with respect to total loans, Loans is constructed as total loans divided by total assets and Productivity is defined as the ratio between total assets and the number of employees. All models include branch and county × year fixed effects. Standard errors are clustered at the bank-level and are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels.

Panel A	Ln(Deposits)		
	Treated	Untreated	Diff-in-diff
	(1)	(2)	(3)
Average Diff. Pre-Post	0.163**	0.371***	-0.209***
T-value	(3.734)	(18.140)	(4.490)
Panel B	**Ln(Deposits)**		
	(1)	(2)	(3)
Treated × Post	-0.250***	-0.241***	-0.216***
	(0.086)	(0.084)	(0.077)
Size		0.062	0.080
		(0.066)	(0.085)
ROA			3.547
			(3.547)
NPL			1.218
			(1.200)
Tier 1			-0.026
			(0.597)
Loan			-0.132
			(0.229)
Productivity			0.001
			(0.017)
Branch FE	Yes	Yes	Yes
County x Year FE	Yes	Yes	Yes
Observations	15460	15334	14382
Adjusted R^2	0.935	0.936	0.936

Table 3
Do Depositors Respond to Cyberattacks? Bank-County Evidence

The table below reports two sets of difference-in-differences regression results of cyberattacks on small banks. Cyberattacks are identified using the breach classification "HACK" by Privacy Rights Clearinghouse (PRC). In Panel A, Ln(Deposits) is the logarithmic transformation of the total amount of deposits of each bank in our sample in a given county. In Panel B, market share is the deposit market share of each bank in our sample in a given county. Treated is a dummy that equals one if a bank belongs to the treated group and zero otherwise; Post is a dummy equal to one in the post shock window (up to 3 years after the shock). The difference-in-differences estimate of the coefficient of Treated x Post is the difference between how the dependent variable changes in treated banks (namely, banks affected by a cyberattack) and the control banks after the shock. Bank controls include: Size, ROA, NPL, Tier 1, Loan and Productivity. Size is measured as the logarithmic transformation of bank total assets in thousands of US$. ROA is the ratio between net income and total assets, Tier 1 is total tier 1 capital divided by risk weighted assets, NPL is the fraction of non-performing loans with respect to total loans, Loans is constructed as total loans divided by total assets and Productivity is defined as the ratio between total assets and the number of employees. Variable definitions, details on the construction of variables and sources are provided in Table A2 in the Online Appendix. All models include bank and county × year fixed effects. Standard errors are clustered at the bank-level and are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels.

Panel A	Bank-county Ln(Deposits)		
	(1)	(2)	(3)
Treated × Post	-0.279***	-0.253***	-0.258***
	(0.084)	(0.086)	(0.076)
Size Control	No	Yes	Yes
Other Bank Controls	No	No	Yes
Bank FE	Yes	Yes	Yes
County x Year FE	Yes	Yes	Yes
Observations	2710	2679	2502
Adjusted R^2	0.741	0.742	0.744
Panel B	Bank-county Market Share		
	(1)	(2)	(3)
Treated × Post	-0.014***	-0.011***	-0.011***
	(0.002)	(0.002)	(0.002)
Size Control	No	Yes	Yes
Other Bank Controls	No	No	Yes
Bank FE	Yes	Yes	Yes
County x Year FE	Yes	Yes	Yes
Observations	2710	2679	2502
Adjusted R^2	0.937	0.947	0.952

Table 4
Do Depositors Respond to Other Data Breaches?

The table below reports difference-in-differences regression results for a sample of other breaches on small banks. Other breaches are classified into seven categories by Privacy Rights Clearinghouse (PRC): CARD, defined as breaches that involve credit or debit card fraud such as skimming of devices at point-of-service terminals; INSD, a breach by an insider such as an employee, contractor or customer; PHYS, a breach involving loss or stolen paper documents; PORT, a breach that results from lost, discarded or stolen portable devices; STAT, a breach as a result of lost, inappropriately accessed, discarded or stolen computers or servers not designed for mobility; DISC, a breach as a result of an unintended disclose that does not involve hacking, intentional breaches or physical loss of information or hardware and finally; UNKN, breaches that PRC does not have sufficient information to appropriately classify. Ln(Deposits) is the logarithmic transformation of the branch-level deposits in US dollar. Other Breaches is a dummy that equals one if a branch belongs to a breached bank and zero otherwise; Post is a dummy equal to one in the post-shock window (up to 3 years after the shock). Results for the aggregated other breaches are reported in columns (1) to (3). In columns (4) to (7), other breaches are separated into breach categories as classified by PRC. The difference-in differences estimate of the coefficient of Other Breaches x Post is the difference between how the dependent variable changes in the branches of treated banks (namely, banks affected by an endogenous breach) and in the branches of control banks after the breach. Bank controls include: Size, ROA, NPL, Tier 1, Loan and Productivity. Size is measured as the logarithmic transformation of bank total assets in thousands of US$. ROA is the ratio between net income and total assets, Tier 1 is total tier 1 capital divided by risk weighted assets, NPL is the fraction of non-performing loans with respect to total loans, Loans is constructed as total loans divided by total assets and Productivity is defined as the ratio between total assets and the number of employees. Variable definitions, details on the construction of variables and sources are provided in Table A2 in the Online Appendix. All models include branch and county × year fixed effects. Standard errors are clustered at the bank-level and are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels.

	Other Breaches (Aggregate) Ln(Deposits)			Other Breaches Ln(Deposits)			
				PORT & DISC	CARD	INSD	UNKN
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Other Breaches × Post	0.018	0.054	0.033	0.103	-0.018	-0.135	-0.079
	(0.051)	(0.047)	(0.046)	(0.100)	(0.068)	(0.116)	(0.091)
Size Control	No	Yes	Yes	Yes	Yes	Yes	Yes
Other Bank Controls	No	No	Yes	Yes	Yes	Yes	Yes
Branch FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County x Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	11760	11477	10984	2947	4099	3132	620
Adjusted R^2	0.909	0.914	0.914	0.856	0.907	0.942	0.989

Table 5
Can Cyberattacks Reduce Depositor Trust?

The table below reports difference-in-differences regression for heterogeneity in depositor responses to cyber attacks on small banks. Cyberattacks are identified using the breach classification "HACK" by Privacy Rights Clearinghouse (PRC). Ln(Deposits) is the logarithmic transformation of the branch-level deposits in US dollar. Heterogenous depositor responses to cyberattacks are measured and conditional on two different measures of social capital. Panel A reports results for measures constructed for a banks' social capital. The social capital score is based on the county in which the bank is headquartered. The score is computed using interpolated values of a county-level social capital index constructed by Rupasingha et al. (2006) and maintained by NRCRD at Pennsylvania State University. Treated banks are sorted into high (low) social capital groups if they are above (below) the median social capital index as measured the year before cyberattacks (Treated High (Low) Social Capital). Panel B reports results for an alternative measure constructed for a banks' social capital; Community Reinvestment Act (CRA) rating. Banks are divided into those with "outstanding" CRA ratings and those with a rating below outstanding. Treated banks are sorted into high (low) social capital groups if they have (do not have) an "outstanding" CRA rating as measured the year before cyberattacks (Treated High (Low) Social Capital). Treated is a dummy that equals one if a branch belongs to a hacked bank and zero otherwise; Post is a dummy equal to one in the post-shock window (up to 3 years after the shock). The difference-in-differences estimate of the coefficient of Treated High (Low) Social Capital × Post is the difference between how the dependent variable changes in the branches of treated banks with high (low) social capital (namely, banks affected by a cyberattack) and in the branches of control banks after the shock. Bank controls include: Size, ROA, NPL, Tier 1, Loan and Productivity. Size is measured as the logarithmic transformation of bank total assets in thousands of US$. ROA is the ratio between net income and total assets, Tier 1 is total tier 1 capital divided by risk weighted assets, NPL is the fraction of non-performing loans with respect to total loans, Loans is constructed as total loans divided by total assets and Productivity is defined as the ratio between total assets and the number of employees.Variable definitions, details on the construction of variables and sources are provided in Table A2 in the Online Appendix. All models include branch and county × year fixed effects. Standard errors are clustered at the bank-level and are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels.

Panel A	Social Capital (HQ) Ln(Deposits)		
	(1)	(2)	(3)
Treated High Social Capital × Post	-0.093**	-0.090**	-0.083**
	(0.038)	(0.039)	(0.042)
Treated Low Social Capital × Post	-0.396***	-0.388***	-0.353***
	(0.100)	(0.099)	(0.093)
Size Control	No	Yes	Yes
Other Bank Controls	No	No	Yes
Branch FE	Yes	Yes	Yes
County x Year FE	Yes	Yes	Yes
High-Low	0.303***	0.298***	0.271***
Observations	15460	15334	14382
R^2	0.950	0.950	0.951
Panel B	CRA Rating Ln(Deposits)		
	(1)	(2)	(3)
Treated High Social Capital × Post	-0.048	-0.043	-0.036
	(0.041)	(0.043)	(0.047)
Treated Low Social Capital × Post	-0.499***	-0.492***	-0.457***
	(0.092)	(0.091)	(0.088)
Size Control	No	Yes	Yes
Other Bank Controls	No	No	Yes
Branch FE	Yes	Yes	Yes
County x Year FE	Yes	Yes	Yes
High-Low	0.451***	0.449***	0.422***
Observations	15460	15334	14382
Adjusted R^2	0.936	0.936	0.937

Table 6
Can Cyberattacks Reduce Depositor Trust? Evidence From Deposit Rates

The table below reports difference-in-differences regression results of cyberattacks on small banks. Branch-level deposit rates are from RateWatch. Panel A reports results for the rates offered on all deposit products Ln(All rates) and rates offered on certificate of deposits Ln(CD rates). Panel B reports results for the rates offered on money market deposit accounts Ln(MM rates) and rates offered on savings accounts Ln(SAVS rates). Treated is a dummy that equals one if a branch belongs to a hacked bank and zero otherwise; Post is a dummy equal to one in the post-shock window (up to 36 months after the shock). The shorter time period is motivated by the higher frequency interval at which the deposit rate data is available from RateWatch. The difference-in-differences estimate of the coefficient of Treated x Post is the difference between how the dependent variable changes in the branches of treated banks (namely, banks affected by a cyberattack) and in the branches of control banks after the shock. Bank controls include: Size, ROA, NPL, Tier 1, Loan and Productivity. Size is measured as the logarithmic transformation of bank total assets in thousands of US$. ROA is the ratio between net income and total assets, Tier 1 is total tier 1 capital divided by risk weighted assets, NPL is the fraction of non-performing loans with respect to total loans, Loans is constructed as total loans divided by total assets and Productivity is defined as the ratio between total assets and the number of employees. Variable definitions, details on the construction of variables and sources are provided in Table A2 in the Online Appendix. All models include branch × product and county × year fixed effects. Standard errors are clustered at the bank-level and are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels.

Panel A	All rates Ln(All rates)			Certificate of Deposits Ln(CD rates)		
	(1)	(2)	(3)	(4)	(5)	(6)
Treated × Post	0.029***	0.030**	0.029**	0.057***	0.055**	0.050**
	(0.011)	(0.015)	(0.014)	(0.018)	(0.022)	(0.020)
Size Control	No	Yes	Yes	No	Yes	Yes
Other Bank Controls	No	No	Yes	No	No	Yes
Branch x Product FE	Yes	Yes	Yes	Yes	Yes	Yes
County x Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1864462	1851165	1850989	1135870	1130117	1130104
R^2	0.955	0.955	0.955	0.959	0.959	0.960
Panel B	Money Market Deposits Ln(MM rates)			Savings Accounts Ln(SAVS rates)		
	(1)	(2)	(3)	(4)	(5)	(6)
Treated × Post	-0.020	0.001	-0.000	-0.001	-0.002	0.002
	(0.026)	(0.027)	(0.027)	(0.018)	(0.022)	(0.019)
Size Control	No	Yes	Yes	No	Yes	Yes
Other Bank Controls	No	No	Yes	No	No	Yes
Branch x Product FE	Yes	Yes	Yes	Yes	Yes	Yes
County x Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	223421	222376	222376	119209	118719	118719
R^2	0.919	0.918	0.923	0.931	0.928	0.929

Table 7
Which Depositors Are More Sensitive to Cyberattacks?

The table below reports difference-in-differences regression for heterogeneity in depositor responses to cyber attacks on small banks. Cyberattacks are identified using the breach classification "HACK" by Privacy Rights Clearinghouse (PRC). Ln(Deposits) is the logarithmic transformation of the branch-level deposits in US dollar. Heterogenous depositor responses are measured and conditional on two direct measures of digital literacy and two indirect measures of digital literacy. Panel A reports results for measures constructed based on the percentage of broadband subscriptions in a county provided by Tolbert and Mossberger (2020) and Form 477 on internet access connections per thousands of households at the county level, provided by the Federal Communication Commission. Panel B reports results for measures constructed based on the median household income in a county taken from the US Census bureau (with higher values denoting more digital literacy). The second is the per capita income from dividends, interests and rents with larger values indicating more depositor financial sophistication, and implicitly, higher digital literacy from the Bureau of Economic Analysis. Banks are sorted into high (low) digital literacy groups if they are above (below) the median digital literacy measured the year before a cyberattack (Treated High (Low) Digital Literacy). Treated is a dummy that equals one if a branch belongs to a hacked bank and zero otherwise; Post is a dummy equal to one in the post-shock window (up to 3 years after the shock). The difference-in-differences estimate of the coefficient of Treated (High) Low Digital Literacy x Post is the difference between how the dependent variable changes in the branches of treated banks in counties with high (low) digital literacy (namely, banks affected by a cyberattack) and in the branches of control banks after the shock. Bank controls include: Size, ROA, NPL, Tier 1, Loan and Productivity. Size is measured as the logarithmic transformation of bank total assets in thousands of US$. ROA is the ratio between net income and total assets, Tier 1 is total tier 1 capital divided by risk weighted assets, NPL is the fraction of non-performing loans with respect to total loans, Loans is constructed as total loans divided by total assets and Productivity is defined as the ratio between total assets and the number of employees. Variable definitions, details on the construction of variables and sources are provided in Table A2 in the Online Appendix. All models include branch and county × year fixed effects. Standard errors are clustered at the bank-level and are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels.

Panel A	Digital Literacy					
	Broadband Ln(Deposits)			Form 477 Ln(Deposits)		
	(1)	(2)	(3)	(4)	(5)	(6)
Treated High Digital Literacy × Post	-0.063	-0.057	-0.050	-0.063	-0.058	-0.050
	(0.039)	(0.041)	(0.043)	(0.039)	(0.041)	(0.044)
Treated Low Digital Literacy × Post	-0.514***	-0.507***	-0.455***	-0.521***	-0.514***	-0.481***
	(0.116)	(0.116)	(0.106)	(0.098)	(0.098)	(0.095)
Size Control	No	Yes	Yes	No	Yes	Yes
Other Bank Controls	No	No	Yes	No	No	Yes
Branch FE	Yes	Yes	Yes	Yes	Yes	Yes
County x Year FE	Yes	Yes	Yes	Yes	Yes	Yes
High-Low	0.452***	0.449***	0.405**	0.459***	0.456***	0.431***
Observations	15460	15334	14382	15460	15334	14382
Adjusted R^2	0.935	0.936	0.937	0.936	0.936	0.937
Panel B	Indirect Digital Literacy					
	Median household income Ln(Deposits)			Income form dividends, interests and rents Ln(Deposits)		
	(1)	(2)	(3)	(4)	(5)	(6)
Treated High (Indirect) Digital Literacy × Post	-0.106***	-0.0583*	-0.0497	-0.157***	-0.117***	-0.113***
	(0.030)	(0.030)	(0.032)	(0.034)	(0.037)	(0.036)
Treated Low (Indirect) Digital Literacy × Post	-0.288***	-0.514***	-0.481***	-0.231***	-0.501***	-0.450***
	(0.104)	(0.151)	(0.155)	(0.038)	(0.052)	(0.054)
Size Control	No	Yes	Yes	No	Yes	Yes
Other Bank Controls	No	No	Yes	No	No	Yes
Branch FE	Yes	Yes	Yes	Yes	Yes	Yes
County x Year FE	Yes	Yes	Yes	Yes	Yes	Yes
High-Low	0.181*	0.456***	0.432***	0.074	0.384***	0.337***
Observations	15460	15334	14382	15460	15334	14382
Adjusted R^2	0.934	0.936	0.936	0.934	0.936	0.936

Table 8
Can Bank Risk Explain the Depositor Response?

The table below reports difference-in-differences regression for heterogeneity in depositor responses to cyber attacks on small banks. Cyberattacks are identified using the breach classification "HACK" by Privacy Rights Clearinghouse (PRC). Ln(Deposits) is the logarithmic transformation of the branch-level deposits in US dollar. Heterogenous depositor responses are measured and conditional on two measures of bank risk. In columns (1) to (3) of Panel A, treated banks are divided by bank riskiness (measured by the log of Z-score) in the year before a cyberattack. Treated banks are denoted as riskier (less risky) if their Z-score is below (above) the median value in the year before the cyberattack. In column (4) to (6) riskier banks are defined as those that jointly have NPL and Tier 1 ratios above (below) the sample median. Treated banks are sorted into high (low) risk groups if they are above (below) the median risk measures (Treated Hack High (Low) Risk). Treated is a dummy that equals one if a branch belongs to a hacked bank and zero otherwise; Post is a dummy equal to one in the post-shock window (up to 3 years after the shock). The difference-in-differences estimate of the coefficient of Treated Hack High (Low) Risk x Post is the difference between how the dependent variable changes in the branches of treated banks with high (low) risk (namely, banks affected by a cyberattack) and in the branches of control banks after the shock. Panel B reports results of the baseline regression where bank controls are sequentially (columns (1) to (6)) and jointly (column (7)) interacted with Post. Size is measured as the logarithmic transformation of bank total assets in thousands of US$. ROA is the ratio between net income and total assets, Tier 1 is total tier 1 capital divided by risk weighted assets, NPL is the fraction of non-performing loans with respect to total loans, Loans is constructed as total loans divided by total assets and Productivity is defined as the ratio between total assets and the number of employees. Variable definitions, details on the construction of variables and sources are provided in Table A2 in the Online Appendix. All models include branch and county × year fixed effects. Standard errors are clustered at the bank-level and are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels.

Panel A	Ln(Z Score)			NPL & Tier 1		
	Ln(Deposits)			Ln(Deposits)		
	(1)	(2)	(3)	(4)	(5)	(6)
Treated Hack High Risk × Post	-0.236**	-0.221**	-0.191**	-0.377**	-0.372**	-0.365**
	(0.097)	(0.094)	(0.081)	(0.174)	(0.175)	(0.174)
Treated Hack Low Risk × Post	-0.278*	-0.264*	-0.256*	-0.381***	-0.370***	-0.321***
	(0.141)	(0.141)	(0.143)	(0.116)	(0.114)	(0.102)
Size Control	Yes	Yes	Yes	Yes	Yes	Yes
Other Bank Controls	Yes	Yes	Yes	Yes	Yes	Yes
Branch FE	Yes	Yes	Yes	Yes	Yes	Yes
County x Year FE	Yes	Yes	Yes	Yes	Yes	Yes
High-Low	0.042	0.044	0.066	0.038	0.034	-0.015
Observations	15400	15274	14334	14328	14202	13272
Adjusted R^2	0.935	0.936	0.936	0.935	0.935	0.936

Panel B	Fundamentals Ln(Deposits)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treated × Post	-0.185**	-0.210***	-0.216***	-0.206***	-0.217***	-0.233***	-0.212***
	(0.074)	(0.074)	(0.077)	(0.077)	(0.074)	(0.076)	(0.068)
Size × Post	-0.056						-0.009
	(0.048)						(0.045)
ROA × Post		6.869					6.093
		(4.165)					(3.766)
NPL × Post			1.895				1.401
			(2.545)				(2.668)
Tier 1 × Post				0.604			1.269
				(0.810)			(0.972)
Loan × Post					0.395		0.522
					(0.308)		(0.327)
Productivity × Post						-0.013	-0.018
						(0.013)	(0.011)
Branch FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County x Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	14382	14382	14382	14382	14382	14382	14382
Adjusted R^2	0.936	0.936	0.936	0.936	0.936	0.936	0.937

Table 9
Cyberattacks and the Reallocation of Bank Deposits

The table below reports difference-in-differences regression results for spillover effects following cyberattacks on small banks. Cyberattacks are identified using the breach classification "HACK" by Privacy Rights Clearinghouse (PRC). Ln(Deposits) is the logarithmic transformation of the branch-level deposits in US dollar. The table presents tests for two different typologies of spillovers in local markets: a) towards large banks and; b) towards small banks. To test for the presence of spillover effects, we compare the evolution of deposits in the branches of untreated banks in the counties where the affected banks operate to the branches of the same untreated banks operating in adjacent counties (where no cyberattacks have occurred). In Panels A to C, Treated is a dummy that equals one if a branch belongs to a large hacked bank operating in counties where small banks have been hacked; Treated is a dummy that equals zero (the control group) if it belongs to branches of the same large bank that operate in adjacent counties (where no cyberattacks have occurred). In Panels D and E, Treated is a dummy that equals one if a branch belongs to a small bank that has not been hacked operating in counties where small banks have been hacked; Treated is a dummy that equals zero (the control group) if it belongs to branches of the same unhacked small banks that operate in adjacent counties (where no cyberattacks have occurred). Post is a dummy equal to one in the post-shock window (up to 3 years after the shock). Panels A and D provide an analysis of potential trend differentials in deposit growth prior to the shock between the treated and the control group for the spillover model. It reports the average one-year change in the dependent variable across the respective two groups of branches in each of the 3 years preceding the cyberattack. The average values are reported in column (1) and (2). The differences in average values are reported in column (3) while column (4) reports T-tests of statistical significance on differences in the average values. Panel B and E formally examine spillovers to large (small) banks. The difference-in-differences estimate of the coefficient of Treated x Post is the difference between how the dependent variable changes in the branches of treated banks (large unaffected banks and small unaffected banks) and in the branches of control banks (branches belonging to the large unaffected banks and small banks operating in unaffected adjacent counties) after the shock. Panel C reports heterogenous depositor results for measures constructed for the reputation or large banks. The Top 5 Reputation score is based on information provided by bank customer on the reputation of banks conducted by American Banker. Treated banks are sorted into Top 5 (Non-Top 5) Reputation groups if they are ranked in the Top 5 (not in the Top 5) of the survey (Treated Hack Hack Top 5 (Non-Top 5) Reputation). The difference-in-differences estimate of the coefficient of Treated Hack Top 5 (Non-Top 5) Reputation x Post is the difference between how the dependent variable changes in the branches of treated banks with Top 5 (Non-Top 5) reputation (namely, banks affected by a cyberattack) and in the branches of control banks after the shock. Bank controls include: Size, ROA, NPL, Tier 1, Loan and Productivity. Variable definitions, details on the construction of variables and sources are provided in Table A2 in the Online Appendix. All models include branch and county × year fixed effects. Standard errors are clustered at the bank-level and are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels.

Panel A	Parallel Trends (>10Bln)			
	Treated (A)	Untreated (B)	Diff. (A-B)	T-value
	(1)	(2)	(3)	(4)
Δ Ln(Deposits)$_{t-3}$	0.075	0.072	0.003	0.733
Δ Ln(Deposits)$_{t-2}$	0.078	0.078	-0.001	0.956
Δ Ln(Deposits)$_{t-1}$	0.093	0.102	-0.009	0.317
Panel B	Large Bank Spillover (> 10Bln) Ln(Deposits)			
	(1)	(2)	(3)	
Treated × Post	0.150**	0.150**	0.152**	
	(0.057)	(0.058)	(0.067)	
Size Control	No	Yes	Yes	
Other Bank Controls	No	No	Yes	
Branch FE	Yes	Yes	Yes	
County x Year FE	Yes	Yes	Yes	
Observations	37603	37587	34696	
Adjusted R^2	0.897	0.898	0.904	
Panel C	Large Bank Reputation Spillover (> 10Bln) Ln(Deposits)			
	(1)	(2)	(3)	
Treated Hack Top 5 Reputation × Post	0.374***	0.398***	0.417***	
	(0.108)	(0.112)	(0.117)	
Treated Hack Non-Top 5 Reputation × Post	0.147**	0.147**	0.149**	
	(0.056)	(0.057)	(0.066)	
Size Control	No	Yes	Yes	
Other Bank Controls	No	No	Yes	
Branch FE	Yes	Yes	Yes	
County x Year FE	Yes	Yes	Yes	
High-Low	0.227***	0.251***	0.268***	
Observations	37603	37587	34696	
Adjusted R^2	0.897	0.898	0.904	

Table 9 (cont.)
Cyberattacks and the Reallocation of Bank Deposits

Panel D	Parallel Trends (<10Bln)			
	Treated (A)	Untreated (B)	Diff. (A-B)	T-value
	(1)	(2)	(3)	(4)
Δ Ln(Deposits)$_{t-3}$	0.097	0.095	0.002	0.845
Δ Ln(Deposits)$_{t-2}$	0.092	0.094	-0.002	0.852
Δ Ln(Deposits)$_{t-1}$	0.072	0.087	-0.014	0.104

Panel E	Small Bank Spillover (<10Bln) Ln(Deposits)		
	(1)	(2)	(3)
Treated × Post	0.028	0.024	0.023
	(0.053)	(0.054)	(0.055)
Size Control	No	Yes	Yes
Other Bank Controls	No	No	Yes
Branch FE	Yes	Yes	Yes
County x Year FE	Yes	Yes	Yes
Observations	32165	31756	29539
Adjusted R^2	0.921	0.925	0.931

Table 10
Local Market Concentration, Cyberattacks and The Reallocation of Bank Deposits

The table below reports difference-in-differences regression results for heterogeneity in spillover effects following cyberattacks on small banks. Cyberattacks are identified using the breach classification "HACK" by Privacy Rights Clearinghouse (PRC). Ln(Deposits) is the logarithmic transformation of the branch-level deposits in US dollar. The table presents tests for two different typologies of spillovers in local markets: a) towards large banks and; b) towards small banks. To test for the presence of heterogeneity in spillover effects, we compare the evolution of deposits in the branches of untreated banks in the counties where the affected banks operate to the branches of the same untreated banks operating in adjacent counties (where no cyberattacks have occurred) conditional on market structure. In Panel A, Treated is a dummy that equals one if a branch belongs to a large hacked bank operating in counties where small banks have been hacked; Treated is a dummy that equals zero (the control group) if it belongs to branches of the same large bank that operate in adjacent counties (where no cyberattacks have occurred). In Panels B, Treated is a dummy that equals one if a branch belongs to a small bank that has not been hacked operating in counties where small banks have been hacked; Treated is a dummy that equals zero (the control group) if it belongs to branches of the same unhacked small banks that operate in adjacent counties (where no cyberattacks have occurred). Post is a dummy equal to one in the post-shock window (up to 3 years after the shock). Heterogenous depositor responses are measured and conditional on two measures of market concentration; the deposit share of the Top 3 (CR3) and Top 5 (CR5) banks in a county. Banks are sorted into high (low) market concentration groups if they are above (below) the median market concentration measured the year before a cyberattack (Treated Hack High (Low) Market Concentration). The difference-in-differences estimate of the coefficient of Treated Hack High (Low) Market Concentration x Post is the difference between how the dependent variable changes in the branches of treated banks in high (low) levels of market concentration (namely, banks affected by a cyberattack) and in the branches of control banks after the shock. Bank controls include: Size, ROA, NPL, Tier 1, Loan and Productivity. Size is measured as the logarithmic transformation of bank total assets in thousands of US$. ROA is the ratio between net income and total assets, Tier 1 is total tier 1 capital divided by risk weighted assets, NPL is the fraction of non-performing loans with respect to total loans, Loans is constructed as total loans divided by total assets and Productivity is defined as the ratio between total assets and the number of employees. Variable definitions, details on the construction of variables and sources are provided in Table A2 in the Online Appendix. All models include branch and county × year fixed effects. Standard errors are clustered at the bank-level and are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels.

Panel A	Large Banks (Over 10Bln)					
	C3 Ln(Deposits)			C5 Ln(Deposits)		
	(1)	(2)	(3)	(4)	(5)	(6)
Treated Hack High Market Concentration × Post	0.298***	0.283***	0.258***	0.285***	0.276***	0.251***
	(0.082)	(0.082)	(0.089)	(0.087)	(0.086)	(0.092)
Treated Hack Low Market Concentration × Post	0.140**	0.142**	0.146**	0.139**	0.141**	0.147**
	(0.056)	(0.058)	(0.067)	(0.057)	(0.059)	(0.068)
Size Control	No	Yes	Yes	No	Yes	Yes
Other Bank Controls	No	No	Yes	No	No	Yes
Branch FE	Yes	Yes	Yes	Yes	Yes	Yes
County x Year FE	Yes	Yes	Yes	Yes	Yes	Yes
High-Low	0.158**	0.141**	0.111*	0.146*	0.135*	0.103
Observations	37603	37587	34696	37603	37587	34696
Adjusted R^2	0.898	0.898	0.904	0.898	0.898	0.904
Panel B	Small Banks (Under 10Bln)					
	C3 Ln(Deposits)			C5 Ln(Deposits)		
	(1)	(2)	(3)	(4)	(5)	(6)
Treated Hack High Market Concentration × Post	0.049	0.054	0.046	0.048	0.045	0.038
	(0.058)	(0.060)	(0.060)	(0.057)	(0.058)	(0.059)
Treated Hack Low Market Concentration × Post	0.007	0.002	0.003	-0.008	-0.010	-0.002
	(0.051)	(0.052)	(0.052)	(0.053)	(0.053)	(0.054)
Size Control	No	Yes	Yes	No	Yes	Yes
Other Bank Controls	No	No	Yes	No	No	Yes
Branch FE	Yes	Yes	Yes	Yes	Yes	Yes
County x Year FE	Yes	Yes	Yes	Yes	Yes	Yes
High-Low	0.042	0.052	0.043	0.040	0.055	0.056
Observations	32165	31756	29539	32165	31756	29539
Adjusted R^2	0.921	0.925	0.931	0.921	0.925	0.931

Table 11
Cyberattacks and Mortgage Lending

The table below reports difference-in-differences regression results cyberattacks on small banks. Cyberattacks are identified using the breach classification "HACK" by Privacy Rights Clearinghouse (PRC). Loan data are from the Home Mortgage Disclosure Act (HMDA) database collected by the Federal Financial Institutions Examination Council (FFIEC). The HMDA loan-level variables are aggregated to the bank-county-year level. The dependent variable is one of the following: 1) Num. Loans (the log transformation of the total number of loans submitted in a bank-county-year); 2) Submitted LTI (the average loan amount requested divided by the average income of the applicant in a bank county-year); 3) Approval Rate (number of approved loans/total loans submitted at the bank-county-year level); 4) Approved LTI (the bank-county-year average of loan amount requested in approved loans/applicant income). Treated is a dummy that equals one if a branch belongs to an hacked bank and zero otherwise; Post is a dummy equal to one in the post-shock window (up to 3 years after the shock). The difference-in-differences estimate of the coefficient of Treated x Post is the difference between how the dependent variable changes for treated banks (namely, banks affected by a cyberattack) and in control banks after the shock. Bank controls include: Size, ROA, NPL, Tier 1, Loan and Productivity. Loan controls include: Ln(Applicant Income), Avg Female, Avg Native American, Avg Asian, Avg African-American, Avg Hawaiian Native, Avg Conventional, Avg FHA and Avg VA. Size is measured as the logarithmic transformation of bank total assets in thousands of US$. ROA is the ratio between net income and total assets, Tier 1 is total tier 1 capital divided by risk weighted assets, NPL is the fraction of non-performing loans with respect to total loans, Loans is constructed as total loans divided by total assets and Productivity is defined as the ratio between total assets and the number of employees. Variable definitions, details on the construction of variables and sources are provided in Table A2 in the Online Appendix. All models include bank and county × year fixed effects. Standard errors are clustered at the bank-level and are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels.

	Mortgage Lending							
	Ln(Num. Loans)		Submitted LTI		Approval Rate		Approved LTI	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treated × Post	0.103	0.106	0.166**	0.167**	-0.019	-0.020	0.111**	0.111*
	(0.148)	(0.148)	(0.073)	(0.073)	(0.021)	(0.023)	(0.055)	(0.057)
Ln(Num. Loans)			-0.636***	-0.638***	-0.033	-0.035	-0.747***	-0.747***
			(0.105)	(0.105)	(0.023)	(0.023)	(0.056)	(0.057)
Ln(Total Loan Applied)			0.591***	0.593***	0.036	0.037	0.703***	0.703***
			(0.104)	(0.104)	(0.023)	(0.023)	(0.051)	(0.052)
Approval Rate							0.132	0.130
							(0.094)	(0.093)
Size Control	No	Yes	No	Yes	No	Yes	No	Yes
Other Bank Controls	No	Yes	No	Yes	No	Yes	No	Yes
Loan Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bank FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County x Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2033	2033	2033	2033	2033	2033	1992	1992
R^2	0.817	0.818	0.882	0.883	0.744	0.745	0.872	0.873

Online Appendix

(Not For Publication)

Table of Contents

Table A1
Sample Description

The table below provides a description of the 16 cyberattacks on small banks used in the analyses. Cyberattacks are identified using the breach classification "HACK" by Privacy Rights Clearinghouse (PRC). Column (2) provides the date that the cyberattack was reported. Column (3) displays the RSSDID of the bank. Column (4) shows assets size (in millions USD) the year before the hack. Column (5) provides the state in which the cyberattack occurred. For each affected State, Column (6) reports the number of counties in which affected banks operate branches. The information on bank size is from the Summary of Deposits (SOD).

ID	Report Date	RSSDID	Assets (t-1)	Affected State	Affected Counties
(1)	(2)	(3)	(4)	(5)	(6)
1	May 19, 2006	682563	9595562	Texas	17
2	May 25, 2006	853372	313698	North Carolina	3
3	November 20, 2006	181758	52180	Louisiana	2
4	May 21, 2007	174572	3683951	New Jersey	10
5	October 10, 2007	500050	1293771	Kansas	4
6	January 24, 2008	975984	1021318	Texas	3
7	June 10, 2008	991340	3509342	Indiana	8 (10)
8	August 28, 2008	816603	2395586	Rhode Island	3 (4)
9	September 10, 2008	621076	321851	Ohio	1
10	January 12, 2010	799612	1569436	New York	1
11	November 16, 2010	616193	124537	New Hampshire	1 (2)
12	January 31, 2013	997847	278904	Wisconsin	1
13	July 17, 2014	790534	2471993	Florida	1
14	January 4, 2016	618807	3517028	Massachusetts	4 (5)
15	January 12, 2016	119779	745395	Massachusetts	1
16	January 12, 2016	128904	8803622	Massachusetts	7 (11)

Table A2
Variable descriptions

This table provides descriptions for the variables used in this study. The table reports variable names (Column 1), its definition (Column 2) and source (Column 3).

Variable Name	Definition	Source
Branch-county-year		
Ln(Deposits)	Logarithmic transformation of the nominal amount of deposits in thousands of US$	SOD
Ln(All rates)	Logarithmic transformation of the rates (in %) offered on deposit products	RateWatch
Ln(CD rates)	Logarithmic transformation of the rates (in %) offered on certificate of deposit products	RateWatch
Ln(MM rates)	Logarithmic transformation of the rates (in %) offered on money market products	RateWatch
Ln(SAVS rates)	Logarithmic transformation of the rates (in %) offered on savings products	RateWatch
Bank-county-year		
Market Share	Ratio of bank deposits in a county to total deposits in the county	SOD
Ln(Num. Loans)	Logarithmic transformation of the total nominal number of mortgage loans submitted to a bank-county-year	HMDA
Submitted LTI	Ratio of the average loan amount to applicant income of all loans submitted to a bank-county-year	HMDA
Approval Rate	Ratio of the loans that were approved to all loans that were submitted to a bank-county-year	HMDA
Approved LTI	Ratio of the average loan amount to applicant income of all loans approved in a bank-county-year	HMDA
Ln(Applicant Income)	Logarithmic transformation of the average applicant income of loans submitted to a bank-county-year	HMDA
Ln(Total Loan Applied)	Logarithmic transformation of the average nominal loan amount submitted to a bank-county-year	HMDA
Avg Female	Ratio of the number of Female loan applicants to all loans submitted to a bank-county-year	HMDA
Avg Native American	Ratio of the number of Native American loan applicants to all loans submitted to a bank-county-year	HMDA
Avg Asian	Ratio of the number of Asian loan applicants to all loans submitted to a bank-county-year	HMDA
Avg African-American	Ratio of the number of African-American loan applicants to all loans submitted to a bank-county-year	HMDA
Avg Hawaiian Native	Ratio of the number of Hawaiian Native loan applicants to all loans submitted to a bank-county-year	HMDA
Avg Conventional	Ratio of the number of conventional loans to all loans submitted to a bank-county-year	HMDA
Avg FHA	Ratio of the number of FHA insured loans to all loans submitted to a bank-county-year	HMDA
Avg VA	Ratio of the number of VA insured loans to all loans submitted to a bank-county-year	HMDA
Bank-year		
Size	Logarithmic transformation of the nominal amount of total assets measured in thousands of US$	SNL Financial
ROA	Ratio of net income to total assets	SNL Financial
NPL	Ratio of non-performing loans to total loans	SNL Financial
Tier 1	Ratio of tier 1 capital to total risk weighted assets	SNL Financial
Loan	Ratio of loans to total assets	SNL Financial
Productivity	Ratio of total assets to total number of employees	SNL Financial
Ln(Z Score)	Logarithmic transformation (ROA + Equity / σROA)	SNL Financial
Equity	Ratio of Equity to total assets	SNL Financial

Table A2 cont.
Variable descriptions

Variable Name	Definition	Source
Bank-year cont.		
σROA	Standard deviation of ROA from a rolling 12-quarter window	SNL Financial
High (Low) Social Capital	Dummy that equals 1 (0) if the social capital index (sk) is above (below) the sample median. The sk index is created using principal component analysis of 4 different factors (voter turnout, census response rate, density of social and non-profit organizations). The index is available for years 1997, 2005, 2009 and 2014. Annual values are interpolated	NRCRD
High (Low) CRA Rating	Dummy that equals 1 (0) if the CRA rating is "oustanding" (not "oustanding")	FFIEC
High (Low) Digital Literacy (Broadband)	Dummy that equals 1 (0) if the % broadband subscriptions in the county is above (below) the sample median.	Tolbert and Mossberger (2020)
High (Low) Digital Literacy (Form 477)	Dummy that equals 1 (0) if the % of households in the county with internet access is above (below) the sample median.	FCC Form 477
High (Low) Financial Literacy (Median household income)	Dummy that equals 1 (0) if median household income in the county is above (below) the sample median.	US Census Bureau
High (Low) Financial Literacy (Income from div, interests and rents)	Dummy that equals 1 (0) if the per capita income derived from dividends, interests and rents is above (below) the sample median.	Bureau of Economic Analysis
High (Low) Market Concentration (C3)	Ratio of deposits of the top 3 banks in a county to total deposits in the county	SOD
High (Low) Market Concentration (C5)	Ratio of deposits of the top 5 banks in a county to total deposits in the county	SOD

Table A3
Tighter Size Matching

The table below reports difference-in-differences regression results of cyberattacks on small banks using a tighter size match. Cyberattacks are identified using the breach classification "HACK" by Privacy Rights Clearinghouse (PRC). Ln(Deposits) is the logarithmic transformation of the branch-level deposits in US dollar. Banks are divided into quartiles within two size bins; the first size bin are banks up to $1bln and the second size bin are banks from $1bln to $10bln. For instance, the first quartile of the first (second) size bin goes up to $250mln ($2.5bln). We then match banks in the treated group with untreated banks falling in the same quartile within each size bin. Treated is a dummy that equals one if a branch belongs to a hacked bank and zero otherwise; Post is a dummy equal to one in the post-shock window (up to 3 years after the shock). The difference-in-differences estimate of the coefficient of Treated x Post is the difference between how the dependent variable changes in the branches of treated banks (namely, banks affected by a cyberattack) and in the branches of control banks after the shock. Bank controls include: Size, ROA, NPL, Tier 1, Loan and Productivity. Size is measured as the logarithmic transformation of bank total assets in thousands of US$. ROA is the ratio between net income and total assets, Tier 1 is total tier 1 capital divided by risk weighted assets, NPL is the fraction of non-performing loans with respect to total loans, Loans is constructed as total loans divided by total assets and Productivity is defined as the ratio between total assets and the number of employees. Variable definitions, details on the construction of variables and sources are provided in Table A2 in the Online Appendix. All models include branch and county × year fixed effects. Standard errors are clustered at the bank-level and are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels.

	Tighter Size Matching Ln(Deposits)		
	(1)	(2)	(3)
Treated × Post	-0.327***	-0.302***	-0.274***
	(0.097)	(0.090)	(0.087)
Size Control	No	Yes	Yes
Other Bank Controls	No	No	Yes
Branch FE	Yes	Yes	Yes
County x Year FE	Yes	Yes	Yes
Observations	4152	4149	3989
R^2	0.965	0.965	0.965

Table A4
Alternative Fixed Effects

The table below reports difference-in-differences regression results of cyberattacks on small banks using different fixed effects. Cyberattacks are identified using the breach classification "HACK" by Privacy Rights Clearinghouse (PRC). Ln(Deposits) is the logarithmic transformation of the branch-level deposits in US dollar. In Panel A, branch fixed effects (originally employed in our main results in Table 2 Panel B) are replaced with branch × cohort fixed effects. In Panel B county × year fixed effects (originally employed in our main results in Table 2 Panel B) are replaced with state x year fixed effects. Treated is a dummy that equals one if a branch belongs to a hacked bank and zero otherwise; Post is a dummy equal to one in the post-shock window (up to 3 years after the shock). The difference-in-differences estimate of the coefficient of Treated × Post is the difference between how the dependent variable changes in the branches of treated banks (namely, banks affected by a cyberattack) and in the branches of control banks after the shock. Bank controls include: Size, ROA, NPL, Tier 1, Loan and Productivity. Size is measured as the logarithmic transformation of bank total assets in thousands of US$. ROA is the ratio between net income and total assets, Tier 1 is total tier 1 capital divided by risk weighted assets, NPL is the fraction of non-performing loans with respect to total loans, Loans is constructed as total loans divided by total assets and Productivity is defined as the ratio between total assets and the number of employees. Variable definitions, details on the construction of variables and sources are provided in Table A2 in the Online Appendix. Standard errors are clustered at the bank-level and are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels.

Panel A	Branch × Cohort Fixed Effects Ln(Deposits)		
	(1)	(2)	(3)
Treated × Post	-0.248***	-0.237***	-0.211***
	(0.086)	(0.083)	(0.076)
Size Control	No	Yes	Yes
Other Bank Controls	No	No	Yes
Branch x Cohort FE	Yes	Yes	Yes
County x Year FE	Yes	Yes	Yes
Observations	15460	15334	14382
R^2	0.950	0.950	0.951
Panel B	State × Year Fixed Effects Ln(Deposits)		
	(1)	(2)	(3)
Treated × Post	-0.223**	-0.213**	-0.188**
	(0.090)	(0.087)	(0.077)
Size Control	No	Yes	Yes
Other Bank Controls	No	No	Yes
Branch FE	Yes	Yes	Yes
State x Year FE	Yes	Yes	Yes
Observations	15460	15334	14382
Adjusted R^2	0.935	0.935	0.936

Table A5
Alternative Standard Errors

The table below reports difference-in-differences regression results of cyberattacks on small banks. Cyberattacks are identified using the breach classification "HACK" by Privacy Rights Clearinghouse (PRC). Ln(Deposits) is the logarithmic transformation of the branch-level deposits in US dollar. Panel A reports results following Bertrand et al. (2004) using observations that are collapsed to one period before and one period after the shock by using the average values of Ln(Deposits) (as well as the other variables in the model) computed for the pre and post 3-year event window employed in our main test. Panel B reports results with standard errors clustered at the branch level (and not at the bank level originally employed in our main results in Table 2 Panel B). Treated is a dummy that equals one if a branch belongs to a hacked bank and zero otherwise; Post is a dummy equal to one in the post-shock window (up to 3 years after the shock). The difference-in-differences estimate of the coefficient of Treated × Post is the difference between how the dependent variable changes in the branches of treated banks (namely, banks affected by a cyberattack) and in the branches of control banks after the shock. Bank controls include: Size, ROA, NPL, Tier 1, Loan and Productivity. Variable definitions, details on the construction of variables and sources are provided in Table A2 in the Online Appendix. All models include branch and county × year fixed effects. Standard errors are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels.

Panel A	Bertrand et al. (2004) Model Ln(Deposits)		
	(1)	(2)	(3)
Treated × Post	-0.206**	-0.209**	-0.218**
	(0.086)	(0.086)	(0.086)
Size Control	No	Yes	Yes
Other Bank Controls	No	No	Yes
Branch FE	Yes	Yes	Yes
County x Year FE	Yes	Yes	Yes
Observations	4438	4413	4161
Adjusted R^2	0.929	0.930	0.927
Panel B	**Branch-Clustered Standard Errors Ln(Deposits)**		
	(1)	(2)	(3)
Treated × Post	-0.250***	-0.241***	-0.216***
	(0.086)	(0.084)	(0.077)
Size Control	No	Yes	Yes
Other Bank Controls	No	No	Yes
Branch FE	Yes	Yes	Yes
County x Year FE	Yes	Yes	Yes
Observations	15460	15334	14382
Adjusted R^2	0.935	0.936	0.936

Table A6
Alternative Estimation Window

The table below reports difference-in-differences regression results of cyberattacks on small banks using different estimation windows. Cyberattacks are identified using the breach classification "HACK" by Privacy Rights Clearinghouse (PRC). Ln(Deposits) is the logarithmic transformation of the branch-level deposits in US dollar. Panel A uses an alternative (-2;+2) years estimation window while Panel B employs a (-1;+1) year window. Treated is a dummy that equals one if a branch belongs to aa hacked bank and zero otherwise; Post is a dummy equal to one in the post-shock window (up to 3 years after the shock). The difference-in-differences estimate of the coefficient of Treated × Post is the difference between how the dependent variable changes in the branches of treated banks (namely, banks affected by a cyberattack) and in the branches of control banks after the shock. Bank controls include: Size, ROA, NPL, Tier 1, Loan and Productivity. Size is measured as the logarithmic transformation of bank total assets in thousands of US$. ROA is the ratio between net income and total assets, Tier 1 is total tier 1 capital divided by risk weighted assets, NPL is the fraction of non-performing loans with respect to total loans, Loans is constructed as total loans divided by total assets and Productivity is defined as the ratio between total assets and the number of employees. Variable definitions, details on the construction of variables and sources are provided in Table A2 in the Online Appendix. All models include branch and county × year fixed effects. Standard errors are clustered at the bank-level and are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels.

Panel A	Alternative Event Window (-2;+2) Years		
	Ln(Deposits)		
	(1)	(2)	(3)
Treated × Post	-0.235***	-0.228***	-0.221***
	(0.085)	(0.083)	(0.082)
Size Control	No	Yes	Yes
Other Bank Controls	No	No	Yes
Branch FE	Yes	Yes	Yes
County x Year FE	Yes	Yes	Yes
Observations	11041	10951	10679
Adjusted R^2	0.943	0.943	0.941
Panel B	Alternative Event Window (-1;+1) Year		
	Ln(Deposits)		
	(1)	(2)	(3)
Treated × Post	-0.213***	-0.205***	-0.184**
	(0.077)	(0.075)	(0.077)
Size Control	No	Yes	Yes
Other Bank Controls	No	No	Yes
Branch FE	Yes	Yes	Yes
County x Year FE	Yes	Yes	Yes
Observations	6594	6557	6303
Adjusted R^2	0.951	0.951	0.948

Table A7
Falsification Test

The table below reports difference-in-differences regression results for our falsification test of cyberattacks on small banks. Cyberattacks are identified using the breach classification "HACK" by Privacy Rights Clearinghouse (PRC). Ln(Deposits) is the logarithmic transformation of the branch-level deposits in US dollar. In this table it is assumed that the cyberattacks occurred seven years prior to their actual date. The regression equation is re-estimated the difference-in-differences model 3 years before (after) the placebo date. By moving the event-window 7 years back, there is no overlap between the post-estimation window in the placebo test and the pre-estimation window in the original empirical setting. The variable of interest is the interaction between Treated Fake × Post Fake. Treated Fake is a dummy equal to one for the banks that have suffered from a cyberattack in our original setting with a dummy; Post Fake is a dummy equal to one in the three years after the falsely-dated cyberattack. Bank controls include: Size, ROA, NPL, Tier 1, Loan and Productivity. Size is measured as the logarithmic transformation of bank total assets in thousands of US$. ROA is the ratio between net income and total assets, Tier 1 is total tier 1 capital divided by risk weighted assets, NPL is the fraction of non-performing loans with respect to total loans, Loans is constructed as total loans divided by total assets and Productivity is defined as the ratio between total assets and the number of employees. Variable definitions, details on the construction of variables and sources are provided in Table A2 in the Online Appendix. All models include branch and county × year fixed effects. Standard errors are clustered at the bank-level and are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels.

	Falsification Test Ln(Deposits)		
	(1)	(2)	(3)
Treated Fake × Post Fake	-0.445	-0.254	-0.047
	(0.327)	(0.199)	(0.045)
Size Control	No	Yes	Yes
Other Bank Controls	No	No	Yes
Branch FE	Yes	Yes	Yes
County x Year FE	Yes	Yes	Yes
Observations	13903	11064	7887
R^2	0.924	0.939	0.966

Table A8
Local Market Concentration, Cyberattacks and the Reallocation of Deposits

The table below reports difference-in-differences regression results for spillover effects following cyberattacks on small banks. Cyberattacks are identified using the breach classification "HACK" by Privacy Rights Clearinghouse (PRC). Ln(Deposits) is the logarithmic transformation of the branch-level deposits in US dollar. The table presents tests for spillovers in local markets towards large banks. To test for the presence of spillover effects, we compare the evolution of deposits in the branches of untreated banks in the counties where the affected banks operate to the branches of the same untreated banks operating in adjacent counties (where no cyberattacks have occurred). Treated is a dummy that equals one if a branch belongs to a large hacked bank operating in counties where small banks have been hacked; Treated is a dummy that equals zero (the control group) if it belongs to branches of the same large bank that operate in adjacent counties (where no cyberattacks have occurred). Post is a dummy equal to one in the post-shock window (up to 3 years after the shock). The difference-in-differences estimate of the coefficient of Treated × Post is the difference between how the dependent variable changes in the branches of treated banks (large unaffected banks and small unaffected banks) and in the branches of control banks (branches belonging to the large unaffected banks operating in unaffected adjacent counties) after the shock. The table reports heterogenous depositor results for measures constructed for the reputation or large banks. The Top 10 and Top 15 Reputation score is based on information provided by bank customer on the reputation of banks conducted by American Banker. Treated banks are sorted into Top 10 (Non-Top 10) and Top 15 (Non-Top 15) Reputation groups if they are ranked in the Top 10 (not in the Top 10) or Top 15 (Non-Top 15) of the survey (Treated Hack Hack Top 10 / 15 (Non-Top 10 / 15) Reputation). The difference-in-differences estimate of the coefficient of Treated Hack Top 10 / 15 (Non-Top 10 / 15) Reputation × Post is the difference between how the dependent variable changes in the branches of treated banks with Top 10 / 15 (Non-Top 10 / 15) reputation (namely, banks affected by a cyberattack) and in the branches of control banks after the shock. Bank controls include: Size, ROA, NPL, Tier 1, Loan and Productivity. Size is measured as the logarithmic transformation of bank total assets in thousands of US$. ROA is the ratio between net income and total assets, Tier 1 is total tier 1 capital divided by risk weighted assets, NPL is the fraction of non-performing loans with respect to total loans, Loans is constructed as total loans divided by total assets and Productivity is defined as the ratio between total assets and the number of employees. Variable definitions, details on the construction of variables and sources are provided in Table A2 in the Online Appendix. All models include branch and county × year fixed effects. Standard errors are clustered at the bank-level and are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels.

Panel A	Ln(Deposits)		
	(1)	(2)	(3)
Treated Hack Top 10 Reputation × Post	0.277***	0.292***	0.313***
	(0.084)	(0.088)	(0.096)
Treated Hack Non-Top 10 Reputation × Post	0.146**	0.145**	0.146**
	(0.056)	(0.057)	(0.066)
Size Control	No	Yes	Yes
Other Bank Controls	No	No	Yes
Branch FE	Yes	Yes	Yes
County x Year FE	Yes	Yes	Yes
High-Low	0.131**	0.1463**	0.167**
Observations	32165	31756	29539
Adjusted R^2	0.921	0.925	0.931
Panel B	Ln(Deposits)		
	(1)	(2)	(3)
Treated Hack Top 15 Reputation × Post	0.276***	0.286***	0.303***
	(0.083)	(0.086)	(0.092)
Treated Hack Non-Top 15 Reputation × Post	0.143**	0.143**	0.143**
	(0.056)	(0.057)	(0.065)
Size Control	No	Yes	Yes
Other Bank Controls	No	No	Yes
Branch FE	Yes	Yes	Yes
County x Year FE	Yes	Yes	Yes
High-Low	0.132**	0.143**	0.160***
Observations	37603	37587	34696
Adjusted R^2	0.897	0.898	0.904

Table A9
Spillover Analysis: Alternative Measure of Market Concentration

The table below reports difference-in-differences regression results for heterogeneity in spillover effects following cyberattacks on small banks. Cyberattacks are identified using the breach classification "HACK" by Privacy Rights Clearinghouse (PRC). Ln(Deposits) is the logarithmic transformation of the branch-level deposits in US dollar. The table presents tests for two different typologies of spillovers in local markets: a) towards large banks and; b) towards small banks. To test for the presence of heterogeneity in spillover effects, we compare the evolution of deposits in the branches of untreated banks in the counties where the affected banks operate to the branches of the same untreated banks operating in adjacent counties (where no cyberattacks have occurred) conditional on market structure. In Panel A, Treated is a dummy that equals one if a branch belongs to a large hacked bank operating in counties where small banks have been hacked; Treated is a dummy that equals zero (the control group) if it belongs to branches of the same large bank that operate in adjacent counties (where no cyberattacks have occurred). In Panel B, Treated is a dummy that equals one if a branch belongs to a small bank that has not been hacked operating in counties where small banks have been hacked; Treated is a dummy that equals zero (the control group) if it belongs to branches of the same unhacked small banks that operate in adjacent counties (where no cyberattacks have occurred). Post is a dummy equal to one in the post-shock window (up to 3 years after the shock). Heterogenous depositor responses are measured and conditional on the Herfindahl-Hirschman Index (HHI) of deposit market concentration. Banks are sorted into high (low) deposit market concentration groups if they are above (below) the median market concentration measured the year before a cyberattack (Treated Hack High (Low) Market Concentration). The difference-in-differences estimate of the coefficient of Treated Hack High (Low) Market Concentration × Post is the difference between how the dependent variable changes in the branches of treated banks in high (low) levels of market concentration (namely, banks affected by a cyberattack) and in the branches of control banks after the shock. Bank controls include: Size, ROA, NPL, Tier 1, Loan and Productivity. Size is measured as the logarithmic transformation of bank total assets in thousands of US$. ROA is the ratio between net income and total assets, Tier 1 is total tier 1 capital divided by risk weighted assets, NPL is the fraction of non-performing loans with respect to total loans, Loans is constructed as total loans divided by total assets and Productivity is defined as the ratio between total assets and the number of employees. Variable definitions, details on the construction of variables and sources are provided in Table A2 in the Online Appendix. All models include branch and county × year fixed effects. Standard errors are clustered at the bank-level and are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels.

Panel A	Large Banks (>10Bln) Ln(Deposits)		
	(1)	(2)	(3)
Treated Hack High Market Concentration × Post	0.279***	0.280***	0.286**
	(0.088)	(0.090)	(0.110)
Treated Hack Low Market Concentration × Post	0.045	0.045	0.051
	(0.051)	(0.050)	(0.049)
Size Control	No	Yes	Yes
Other Bank Controls	No	No	Yes
Branch FE	Yes	Yes	Yes
County x Year FE	Yes	Yes	Yes
High-Low	0.234**	0.235*	0.235*
Observations	37603	37587	34696
Adjusted R^2	0.897	0.898	0.904
Panel B	Small Banks (<10Bln) Ln(Deposits)		
	(1)	(2)	(3)
Treated Hack High Market Concentration × Post	0.099	0.095	0.110
	(0.083)	(0.087)	(0.087)
Treated Hack Low Market Concentration × Post	-0.010	-0.011	-0.018
	(0.060)	(0.060)	(0.059)
Size Control	No	Yes	Yes
Other Bank Controls	No	No	Yes
Branch FE	Yes	Yes	Yes
County x Year FE	Yes	Yes	Yes
High-Low	0.109	0.106	0.128
Observations	32165	31756	29539
Adjusted R^2	0.921	0.925	0.931