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

Successful cyberattacks decrease branch deposit growth rates at small US banks. This decrease is due to the loss of trust of depositors in the ability of the hacked small banks to protect their private information and not to concerns over bank stability. The loss of trust results in a "flight-to-reputation" within local markets leading to increased deposit flows to large banks and negative spillovers for some unhacked small banks. As a result, local deposit markets become more dominated by large banks and the access to credit to small borrowers decreases. Ultimately, weak cybersecurity systems in small banks influence their ability to attract and retain customers, lead to a change in market structure and reduce access to credit of small local businesses.

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 Nagpurnanand Prabhala, 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), the 2021 Financial Stability Conference and the seminar participants at the University of Edinburgh.

1 Introduction

Small banks play a pivotal and unique role in the economy (Behr et al., 2013; Berger et al., 2017; Degryse and Van Cayseele, 2000). They facilitate the access to finance for small firms that have otherwise limited funding opportunities and, in this way, they offer a crucial contribution to the development of the local economy (Berger et al., 2017; Hakenes et al., 2015). However, the digital transformation of the economy brings new challenges to the business model of small banks by exposing the banking industry to new risks, in particular cyber risks (Basel Committee on Banking Supervision, 2018; Duffie and Younger, 2019; Mester et al., 2019).

Specifically, banks need continuous investments in cybersecurity to mitigate their exposure to cyber risks and avoid successful cyberattacks that can undermine customer trust (Chen et al., 2019; Kamiya et al., 2020; Kashyap and Wetherilt, 2019)[1]. The reoccurring investments necessary to maintain a high level of cybersecurity require significant financial resources that might not be sufficiently available to small banks (Kashyap and Wetherilt, 2019; Paravisini, 2008)[2]. Consequently, these banks might suffer from an investment gap in cybersecurity and remain more exposed to significant cyber risks to their business than larger banks. Not surprisingly, therefore, more than 70% of small bankers have ranked cybersecurity as their top concern (Conference of State Bank Supervisors, 2019). Indeed, a report by Nationwide indicates that almost half of cybercrimes between 2012-2017 target US banks with assets below $1billion.

The potential business consequences of cybersecurity deficiencies in 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; Klette and Kortum, 2004). 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

[1]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

[2]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.

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 on the structure of local banking markets. The economic implications of this structural change can be substantial given the competitive advantage of small banks in lending to local businesses (Agarwal and Hauswald, 2010; Berger et al., 2005; Skrastins and Vig, 2019; Stein, 2002).

In this paper, we build on the above arguments and present the first attempt to document how cyberattacks create significant business challenges for small banks by affecting their ability to retain and attract customers within a local market. More precisely, we offer a comprehensive analysis of how depositors respond to successful cyberattacks on small banks by assessing the direct effects of these attacks on deposit growth at hacked banks, and their indirect effects (spillovers) on non-hacked large and small banks. Jointly, these effects indicate the reallocation of deposits and the impact on local banking markets. Additionally, we assess whether the direct and indirect effects of cyberattacks result in a loss of competitiveness of small banks with negative consequences for small business lending.

Our main 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 by cyberattacks compromising depositors' confidential financial and personal information. Therefore, depositors are directly affected by cyberattacks. Second, deposit markets are a key source of funding for small banks and facilitate their lending business. 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 (plausibly) exogenous cyberattacks involving small US banks covered in the Privacy Rights Clearinghouse (PRC) database over the period 2005-2017. We employ these events in size-matched stacked difference-in-differences analyses that control for differences between hacked and non-hacked banks and allow a cleaner identification of the treatment effect in the presence of staggered events (Baker et al., 2022; Gormley and Matsa, 2011).

We start by documenting that the branches of hacked small 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 deposit market. 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 offer and contrast two potential explanations for the depositor response to a successful cyberattack. Both explanations are grounded on the stipulation that cyberattacks can undermine the trust of depositors in banks (Kamiya et al., 2020; Mourouzidou-Damtsa et al., 2019; Sapienza and Zingales, 2012). As a result, depositors might review existing contractual relationships with the hacked bank and decide to withdraw their deposits (or might avoid to establish new relationships with the hacked banks). The first explanation focuses on the loss of trust of depositors in the ability of the hacked small banks to protect their private information (*information loss explanation*). The second explanation is centered around the loss of trust in the stability of the small bank sustained by the attack, and consequently in the bank's ability to protect savings (*stability concern explanation*). We present several analyses that support the information loss explanation of the slowdown in deposit growth for hacked banks.

First, we document larger declines in deposit growth in local markets where depositors are plausibly more exposed to the risk of identity theft or when they are less knowledgeable about cyber risk. Our findings suggest, therefore, that successful cyberattacks, which more saliently and directly impact depositors' personal welfare, are likely to incite stronger responses from customers more exposed to the risk of a malicious use of the stolen personal information and from unsophisticated bank customers who might not be fully aware of the consequences and remediation processes following a cyberattack.

Next, we show that depositor reactions are not driven by bank default risk or can be explained by (potential) deteriorations in bank fundamentals that might arise as a result of cyberattacks. Finally, we show the effects of a cyberattack on the contractual

relationships between households and banks in mortgage markets. These relationships represent an ideal setting to disentangle the two possible explanations of our finding as they expose households to the risk of private information loss from a cyberattack but not to a financial loss. We find that hacked banks attract riskier borrowers after a cyberattack and are forced to originate riskier mortgages to maintain unchanged mortgage approval rates. These effects are, therefore, consistent with the information loss explanation of our results thereby hacked banks respond to the greater reluctance of customers in establishing relationships in the mortgage market by relaxing their origination criteria.

Having documented the negative consequences of cyberattacks on bank deposits and mortgage origination of small banks and their motivation, we progress by examining if cyberattacks generate spillover effects on non-hacked banks with potential consequences on the structure of local deposit markets. We expect opposite sign of spillovers for large and small unhacked banks. Specifically, our prior is that depositors are unlikely to perceive successful cyberattacks on small and unlisted banks as indicative of industry-wide weaknesses in cybersecurity that affect their trust on larger banks. Instead, large banks might be perceived as digitally and technologically safer by depositors because of their financial capacity to continuously invest in cybersecurity. These banks can then benefit from positive spillovers via the reallocation of deposits towards their branches The positive spillover would be consistent with a "business stealing effect" favoring more innovative of rival firms and potentially increasing the exit risk for less innovative firms (Bloom et al., 2013; Klette and Kortum, 2004).

Negative spillovers, with a consequent decline in branch deposit growth, may instead affect other unhacked small banks operating in local deposit markets where branches of the hacked small banks are present. The negative spillovers might result from depositors' perception that other small banking firms have similar vulnerabilities in their cybersecurity systems as the hacked small banks. As a result, these unhacked small banks still expose depositors to a high risk of private information loss from cyberattacks.

We employ an alternative difference-in-differences setup to examine the spillover

effects of cyberattacks on large and small banks. 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. In line with our prediction, we find positive spillovers towards branches of large banks. The spillovers are more pronounced for large banks with an excellent reputation amongst customers but do not depend on the systemic importance of these banks. The results indicate, therefore, a "flight-to-reputation" effect in local deposit markets after successful cyberattacks on small banks that differ from the "fight-to-safety" effect due to the systemic importance of large banks previously highlighted by banking studies (see, for instance, Farhi and Tirole, 2012).

For unhacked small banks, we do not observe any general negative spillover effect. However, we find negative spillovers when these banks have more geographic overlaps with the branch network of the hacked banks. This finding is consistent with the importance of depositor's local network for the spread of negative spillovers in local deposit markets (Iyer et al., 2016) and with the presence of a local dimension in negative spillovers leading to contagion risk in the US banking industry (Addoum et al., 2020; Pino and Sharma, 2019).

We finally document that the highlighted direct and indirect effects of cyberattacks result in significant market-wide consequences. We show that the growing importance bank customers place on cybersecurity leads to in an increased market share of large banks in local deposit markets and a reduced competitiveness of small banks. This change in the structure of local deposit markets is accompanied by a reduced access to credit by very small borrowers, consistently with the view that large banks tend to be less inclined to supply small business lending (see, for instance, Bord et al., 2018; Chen et al., 2017).

We contribute to three streams of research. The first focuses on the effects of cyberattacks on 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). Empirical investigations on the implications of cyberattacks on bank outcomes are more limited. Eisenbach et al. (2020) simulate the externalities produced by cyberattacks through the wholesale payments network while Bouveret (2018) proposes a framework for the quantification of cyber risk in the financial industry. Aldasoro et al. (2020) document that cyber losses account for a significant portion of total operational value-at-risk. More closely related to our work is a contemporaneous study by Engels et al. (2021) focusing on the implications of different forms of data breaches for bank deposits and their cost. Differently from this study, we build a cleaner identification strategy to focus on the impact of plausibly exogenous cyberattacks on small banks, and on the consequences for the competitiveness of these banks in local deposit markets, through a joint examination of the direct and indirect effects of the attacks. Additionally, our analysis offers evidence beyond the deposit markets by documenting the implications of the attacks for mortgage lending by the hacked banks and for the aggregate small business lending in local markets.

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 et al., 2005; 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 documenting deposit outflows for banks disclosing negative non-financial information linked to poor social performance (Chen et al., 2019; Homanen, 2018). We complement these studies by showing that cyberattacks not only lead to negative responses by depositors, but result in the re-distribution of deposits in local deposit markets via positive spillover effects towards larger banks and negative spillovers to some unhacked small banks. We document that the effects have implications for the structure of local deposit markets and lead to a decline in the growth of loans to very small borrowers.

Finally, our study is 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). 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 motivate our primary focus on deposit markets.

2 Identification Strategy and Data

2.1 Treated and Control Banks

Our analysis is based on cyberattacks targeting small US commercial banks between 2005-2017[3]. We identify these attacks starting from a list of all data breach incidents 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; 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 target a small commercial bank (defined following the Federal Reserve's classification as a bank with total assets up to $10bln at the time of the data breach); ii) they are classified as a "HACK" by PRC; that is, they are caused by external parties and result in the loss of customer personal or financial information; iii) they affect banks with detailed branch deposit data in the Summary of Deposits (SOD) provided by

[3]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.

the FDIC and accounting data from call reports. Using this sampling procedure, we identify 16 cyberattacks on small US banks. We provide detailed information of the sampled cyberattacks in Table A1 of the Online Appendix.

The second criterion ensures that the events are plausibly exogenous (Kamiya et al., 2020), whereas the third criterion allows us to have detailed geographic data on branch deposits that we employ as one of the key inputs of our matching strategy between hacked small banks and non-hacked banks (control group). In particular, the matching is based on i) geographic location of branches of the two groups of banks and ii) size similarity between these banks.

The geographic matching alleviates concerns that confounding geographical supply and demand factors might bias the analyses[4]. Matching by bank size is instead important because, as documented by previous research, large banks have advantages in deposit markets (Jacewitz and Pogach, 2018; Oliveira et al., 2015). For instance, large banks have lower funding costs due to the too-big-to-fail subsidy (Jacewitz and Pogach, 2018).

To implement our geographic matching between the two groups of banks, we first identify the state in which a cyberattack is reported according to PRC. Within the identified state, we select all counties in which the affected banks operate branches according to the SOD data. These branches represent our treated group. Next, for each county in which the hacked bank operates, we form a control group of branches owned by commercial banks of similar size.

To ensure size similarity between the treated and untreated small banks, our initial 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 untreated banks falling into the same size group as the hacked bank. In section 2.3,

[4]One example of geographical drivers affecting deposit markets is the fraction of seniors across different geographical regions. Becker (2007) shows that 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 treated and untreated banks from different geographic areas might yield biased results.

we show that this simple matching approach is sufficient to ensure similarity between treated and untreated banks across several bank characteristics including profitability and risk. Furthermore, in additional tests discussed in section 3, we show that our analysis remains valid for a much tighter size matching between treated and untreated banks.

2.2 Econometric Method

We use a stacked difference-in-differences approach to estimate the causal impact of cyberattacks on depositor behavior (Gordon et al., 2011; Baker et al., 2022). We construct cohorts of treated branches for each event and for each cohort we include as a control group only bank branches that have not previously experienced a cyberattack. We then stack the cohorts to estimate the average treatment effect. This approach allows us to more cleanly capture the treatment effect (Gormley and Matsa, 2011; Guo et al., 2019) and avoid identification problems in staggered difference-in-differences settings as highlighted by Baker et al. (2022).

Initially, we restrict the estimation window to 7 (-3;+3) years around each cyberattack, although we also report results for smaller estimation windows to remove potential confounding factors. The largest sample we employ includes a total of 3,076 (12,384) observations belonging to branches of treated (untreated) banks. Our model takes the following functional form:

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

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 i in a given cohort c belongs to an hacked bank j and zero otherwise; Post is a dummy equal to one in a cohort c in the post-shock window (up to 3 years after the shock). The coefficient β_1 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 untreated banks (those not affected by a cyberattack) after the shock. Since the dependent variable is the logarithmic transformation of branch deposits, 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. In equation (1) we cluster standard errors at the bank level to control for within bank correlation in the evolution of deposits. 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 while county × year fixed effects remove any time-varying county-level factors such local business cycles (e.g., unemployment housing demand and shale gas discoveries) that could affect local deposit market (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, equation (1) does not include bank-specific control variables because a cyberattack can affect these controls, 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 computed from annual call reports and measured with 1-year lag to reduce endogeneity concerns. In the first specification, we add bank size (the logarithmic transformation of bank total assets in thousands of US$). In the second specification we control also for profitability, via the ratio between net income and total assets (ROA), capital adequacy, using tier 1 capital divided by risk weighted assets ($Tier\ 1$), credit risk, defined as non-performing loans scaled by total loans (NPL), asset composition, that is, total loans divided by total assets ($Loans$), and bank productivity, via the ratio between total assets and the

number of employees (*Productivity*). Panel A of Table1 provides summary statistics for the variables we employ. Table A2 in the Online Appendix offers a description of these variables 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 some banks to be targets of cyberattacks conditional on their observable 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.

The findings discussed above might not completely rule out the possibility that the two groups of banks have different ex-ante probabilities of being hacked. Panel C offers evidence against this argument. This Panel presents the results of a logit model wherein we estimate whether the probability of being hacked can be predicted by bank characteristics measured with a 1-year lag. We find that none of the selected bank characteristics are a significant predictor of the likelihood to be hacked. Overall, this section suggests that our simple size matching is sufficient to achieve a high degree of homogeneity between the two groups of banks in terms of potential exposure to cyberattacks.

2.3.2 Parallel Trends Assumption

A key assumption of our difference-in-differences analysis is that, absent the shock (cyberattack), treated and untreated branches would have shown a similar evolution in the (log transformation) of deposits (parallel trends assumption). 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. However, the literature offers different options to examine whether the parallel trends assumption is plausible. In particular, if the two groups of branches follow similar trends in the evolution of deposits prior to the cyberattack, the parallel trends assumption is deemed to be reasonable.

We conduct two 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 D 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.

[FIGURE 1 HERE]

Second, 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 × 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 seems plausible in our setting.

3 Direct Effects of Cyberattacks

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 show the results for equation (1). As mentioned previously, the coefficient of interest is the interaction term Treated × Post. The coefficient 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 × 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 in Panel A, the coefficient of Treated × 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 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 branches in control group. Notably, none of the controls have a significant effect on the dependent variable.

One concern is that unobserved heterogeneity due to size differentials between the two groups of banks is not entirely removed by our matching strategy. In the Online Appendix, we address this concern using a tighter size matching[5]. We divide the two size bins we have employed in our matching (up to $1bln and 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, results based on this alternative matching confirm our initial findings[6].

In summary, while 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; Kamiya et al., 2020), we show the direct effects of cyberattacks on small firms through key stakeholders (namely, depositors).

3.2 Additional Tests

In Table 3, we report additional specifications that document the robustness of our results. First, in column (1) we address concerns related to standard errors. Bertrand

[5]However, 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.

[6]Notably, 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.

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 using the average values of our dependent variable *Ln(Deposits)* (as well as the other variables employed in our main test) computed for the pre and post 3-year event window. The results confirm our main findings.

[TABLE 3 HERE]

Next in column (2) we use a 5 (-2;+2) year estimation window, while in the next column we employ a 3 (-1;+1) year window. The use of shorter estimation windows reduces the possibility of noise biasing the treatment effects and also partially alleviates issues emerging around serially correlated outcomes, as discussed above. In both settings, our results remain intact. In column (4) we use a different set of fixed effects (as compared to our main specification in Table 2). Specifically, we follow Gormley and Matsa (2011) and replace branch fixed effects with branch $\times$ 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. Following this approach, we do not find any material changes to our results.

In column (5) 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 and it also allows us to understand the overall effect of cyberattacks on deposit growth rates in local markets. To implement this analysis, we use the logarithmic transformation of the total amount of deposits of each treated and control bank in our sample in a given county as the dependent variable and re-estimate a modified specification of equation (1). Specifically, we include bank and county $\times$ year fixed effects and cluster standard errors at the bank level. Consistent with our main analysis, we observe evidence of a relative decline in deposit growth for treated banks.

Finally, in the last column of Table 3, we implement a falsification test to validate

the causal interpretation of our findings. We assume that the cyberattacks occurred seven years prior to their actual date and re-estimate the difference-in-differences model based on 3 years before (after) the new identified event data. 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 initial approach. In this artificial setting, we should not observe any changes in deposit growth for the treated branches. 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 (placebo date) cyberattack. Consistent with our expectation, the interaction term Treated_Fake × Post_Fake is not significant. This "non-result" further supports the interpretation that the negative effect on deposit growth rates, documented in our main analysis (Table 2), is likely the result of depositor responses to cyberattacks.

Hence, the results in this section are consistent with our initial finding that successful cyberattacks lead to a significant slowdown in the deposit growth of hacked small banks.

3.3 What Drives the Depositor Response? Private Information Loss Versus Small Bank Stability

We show decreases in deposit growth for hacked small banks. In this section, we conduct various tests to understand what drives this result. In particular, we contrast two possible explanations that have different policy implications. Both explanations are based on the fact that contractual agreements between banks and depositors are built on trust; namely, banks are expected to safeguard depositors' savings and their confidential information. When a successful cyberattack occurs, the trust of depositors in the hacked bank is damaged (Kamiya et al., 2020). Depositors might then react 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. The two explanations discussed below, however, differ in the rationale for the loss of trust by depositors.

The first explanation places emphasis on the loss of trust in the ability of the hacked small banks to keep private data safe. Under this perspective, depositors are not concerned about the stability of the small bank but react to the cyberattacks because they are exposed to the loss of private information (*information loss explanation*). This explanation highlights, therefore, the importance to ensure that small banks have ex-ante adequate cybersecurity systems in place to protect private data of their customers.

The second explanation focuses on the possibility that depositors lose trust in the stability of the small bank and, consequently, in its ability to protect their savings. Accordingly, the successful cyberattack is perceived as an indication of fragility of the hacked small bank and might result in the risk of losing the availability of funding due to future bank distress (*stability concern explanation*). Along these lines, theoretical work suggests that deposits flow from distressed banks towards healthier institutions (Egan et al., 2017). This explanation would highlight the need to reassure investors ex-post of the soundness of the small hacked banks.

The following sections present a series of tests aiming at understanding if private information loss or concerns over bank stability drive the depositor response to cyberattacks.

3.3.1 Exposure to Identity Theft, Depositor Digital Sophistication and Response to Cyberattacks

We initially exploit depositor heterogeneity in terms of ex-ante exposure to identity theft and digital sophistication. The events we consider directly affect bank depositors and signal the risk of a threat through third parties. If concerns over data loss, and not bank stability, are driving our results, cyberattacks should lead to a stronger response in markets where depositors are more exposed to the risk of identity theft through the

use of their private information. Additionally, the attacks might cause more anxiety in digitally unsophisticated depositors as they are less likely to understand the exact ramifications of cyberattacks (Solove and Citron, 2017). Therefore, we should observe a stronger reaction by digitally unsophisticated depositors that overreact to the shock. In contrast, if we are capturing broader concerns in terms of bank fragility, there is no obvious reason to expect that our results differ by the exposure to identity theft risk or level of depositor digital sophistication.

[TABLE 4 HERE]

To account for the exposure of depositors to the risk of identity theft, we rely on data from the Federal Trade Commission (FTC). Through their periodic Consumer Sentinel Network Report, the FTC offer indications of the number of consumer complaints per 100,000 population related to identity theft in each metropolitan statical area (MSA). Treated banks are then sorted into high and low risk exposure to identify theft groups if they are above or below the median values of the variable above in the year before cyberattacks (*Treated High (Low) Identity Theft Risk*). We finally estimate the following specification:

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

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) identify theft risk. In line with our baseline model, we estimate equation (3) with and without bank controls.

We employ a similar empirical framework to analyze how our results are affected by the degree of digital sophistication of depositors. We measure digital sophistication using information from Form 477 on internet access connections per thousands of households at the county level provided by the Federal Communication Commission[7].

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

The results reported in the first three columns of Panel A of Table 4 show that the relative decline in deposits in the treated group is stronger in counties where depositors show (plausibly) higher exposure to identity theft risk. We find that the coefficient of Treated High Identity Theft Risk × Post is negative and significant across all specifications and is statistically larger than the coefficient of Treated Low Identity Theft Risk × Post. Furthermore, in the last three columns of Panel A, we observe that decline in deposits for small banks after the cyberattacks is driven by depositors with lower levels of digital literacy. The fact that the negative consequences of cyberattacks on deposit growth are (primarily) driven by a higher exposure to identity theft risk and by digitally unsophisticated depositors is then consistent with concerns over private data losses as the main reason behind our findings.

3.3.2 Small Bank Risk and Depositor Response to Cyberattacks

Examining if there is heterogeneity in our results due to small bank risk can offer further insights regarding the observed slowdown in deposit growth. More precisely, if our results are due to concerns over the stability of hacked banks (and not due to concerns over private information loss), we should observe that after a successful cyberattack depositors should primarily withdraw deposits from riskier banks (e.g. Martinez Peria and Schmukler, 2001).

We assess the validity of this argument in Panel B of Table 4. In the first three columns of the Panel, we split treated small banks according to their riskiness (as measured by the log of Z-score) the year before the cyberattacks[8]. 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 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

We obtain similar results when we use estimates of the percentage of broadband subscriptions in a county provided by Tolbert and Mossberger (2020).

[8]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).

these two tests suggest that the observed depositor reactions are unlikely a reflection of concerns over bank risk.

In the Online Appendix we offer additional evidence against a risk interpretation by sequentially interacting our vector of bank controls (Size, ROA, NPL, Tier 1, Loan, Productivity) with Post. Our aim is to investigate if any change in bank characteristics post-shock is significant in dampening the economic significance of Treated $\times$ Post. If bank risk were crucial determinants of depositor reactions, we should observe significant decreases in the economic significance of the coefficient Treated $\times$ Post after we control for these variables. We do not find this to be the case: 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 hacked banks.

Therefore, our results are consistent with the observations of Kamiya et al. (2020), showing that the direct out-of-pocket costs (e.g., investigation and remediation costs, legal and regulatory penalties) resulting from cyberattacks only account for a small proportion (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.

3.3.3 Disentangling the two Explanations Using Household Borrowing Data

An alternative empirical strategy to understand the drivers behind our key finding is to focus on bank customers that are exposed to private information loss but not necessarily to funding losses. In this respect, besides deposit markets, banks engage in contractual relationships with households in mortgage markets. Indeed, cyberattacks should not generally expose borrowers to funding risk but as depositors they will share personal data with the lender. Accordingly, they are exposed to the loss of private information. It follows that if our findings for the deposit markets are driven by concerns about private data loss, we should also observe some negative effects on the lending relationships of the hacked bank in the mortgage market.

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 borrowers are concerned about the risk of losing personal data, 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 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)[9]. Each loan application in the HMDA dataset contains information on borrower demographics, loan characteristics, the decision to grant or not the mortgage, 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 the county where the mortgage was originated because they are likely to be loans that were submitted to independent mortgage brokers (Cortés, 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 to the analysis. We then aggregate HMDA loan-level variables to the bank-county-year and estimate the following difference-in-differences model:

$$\begin{aligned} \text{Lending}_{i,z,t} = \alpha + \beta_1 \text{Treated} \times \text{Post} + \mathbf{BANK} + \mathbf{COUNTY} \times \mathbf{TIME} \\ + \mathbf{CONTROLS}\varepsilon_{i,z,t}, \end{aligned} \tag{4}$$

[9]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).

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)[10]. The first two variables, therefore, take the borrowers' perspective while the remaining variables

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 Post $\times$ Treated and measures the change in one of the lending variables from the pre to the post shock period, as defined in equation (1), in the group of treated banks 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 definitions of these variables in Table A2 in the Online Appendix.

[TABLE 5 HERE]

Table 5 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 indicate, therefore, that small banks are more likely to attract riskier borrowers subsequent to a cyberattack.

[10]We winsorize the variables Applicant Income and Loan Amount at the 5% tails to minimize reporting errors.

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 evidence of an increase in the Loan-to-Income ratio of loans that have been approved. This increase indicates that hacked banks are forced to relax their lending standards to maintain approval rates. These results are consistent with the perspective that cyberattacks matter for bank customers that are not exposed to funding losses.

4 Spillover Effects in Local Deposit Markets

Up to this point, our analysis does not consider the possibility of spillover effects within local deposit markets. However, Kamiya et al. (2020) show negative spillovers at the industry level after successful cyberattacks on non-financial firms. More generally, in empirical settings involving companies operating in the same industry (and in the same geographic markets as in our analysis), the assumption of no spillovers is not entirely plausible. For instance, a broad theoretical banking literature highlights the presence of spillovers in deposit markets by identifying conditions under which deposits are withdrawn from banks affected by negative events and then reallocated towards other banking firms (see, for instance, Egan et al., 2017; Farhi and Tirole, 2012).

In line with the above banking literature, the negative spillover dynamics documented by Kamiya et al. (2020) for a sample of large listed firms are not necessarily indicative of what we should observe within deposit markets after a cyberattack. In particular, it is unlikely that depositors would perceive successful cyberattacks on small banks as a signal of a wider industry cybersecurity problem with negative spillovers for all other banks. Instead, under an "equilibrium framework" for local deposit markets, and consistent with the reallocation mechanism highlighted by theoretical banking studies, at least part of the withdrawn, or not-deposited funding, at hacked small banks should flow towards other banks operating in the same local market.

In this respect, since our results are driven by depositors' concerns about private data loss due to cybersecurity deficiencies, the positive spillovers should primarily favor banks that are seen as digitally and technologically safer by customers (e.g.

Chen et al., 2017). Large banks should then be an obvious choice for depositors given their financial capacity to comply with the continuous investments required by cybersecurity. Additionally, the presence of positive spillovers in favor of large banks would be consistent with the view that investment and innovation gaps between competing firms generate a business stealing effect in favor of firms with an innovation advantage and increase the risk of exit for less innovative firms (see, for instance, Bloom et al., 2013; Klette and Kortum, 2004). By contrast, if any negative spillovers would emerge, as in Kamiya et al. (2020), they are more likely to impact other (unaffected) small banks as their cybersecurity environment might be perceived as similar to that of hacked banks and as such equally vulnerable.

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 (banks below $10bln in terms of assets and not directly affected by the cyberattack). To conduct these tests, we employ an identification strategy that compares the evolution of deposits in the branches of untreated banks in counties where hacked banks operate with the branches of the same untreated banks operating in adjacent counties (where no branches of the hacked banks are present).

Our identification strategy is graphically presented in Figure 2, where we illustrate examples of how we categorize local markets. The local markets in which treated banks operate are illustrated in red and those markets in which hacked banks are not present are illustrated in blue. More precisely, Figure 2a graphically clarifies our empirical approach using the case of 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, therefore, the branches of large (small) banks in Middlessex and Norfolk are considered as indirectly treated whereas the branches in Worchester and Bristol are categorized as untreated.

[FIGURE 1 HERE]

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 empirical approach alleviates concerns over omitted bank characteristics driving our results since treated and untreated branches belong to the same group of banks.

4.1 Spillover Effects towards Large Banks

We start by focusing on spillovers towards larger banks. Panel A of Table 6 shows 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, reported in Panel B of Table 6, 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 rates between the two groups of branches is approximately equal to 15 percentage points. In summary, branches of large banks are able to attract more deposits in local markets where small banks suffered from a cyberattack. Therefore, our result indicates a "flight-to-reputation" effect in local deposit markets associated with cyberattacks on small banks and is consistent with the importance of depositor trust in relation to a bank's ability to protect private data (as established in our main analysis).

[TABLE 6 HERE]

To offer further support for the interpretation above, in Panel A of Table 7 we repeat the large bank spillover test by separating our sample of indirectly "treated" branches of large banks using the customer reputation score assigned to these banks by 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 and estimate the following model[11]:

$$\begin{aligned}\text{Ln(Deposits)}_{i,j,z,t} = \alpha &+ \beta_1(\text{Indirectly Treated High Reputation} \times \text{Post}) \\ &+ \beta_2(\text{Indirectly Treated Low Reputation} \times \text{Post}) \\ &+ \mathbf{BRANCH} + \mathbf{COUNTY} \times \mathbf{TIME} + \varepsilon_{i,t},\end{aligned} \tag{5}$$

If our results are driven by the search of large banks that offer better information safety to their customers, we should observe that the positive spillovers primarily benefit large banks with a higher customer reputation score. In line with this expectation, we observe that positive spillovers are significantly larger for large banks with exceptional reputation. The results further support an interpretation based on a "flight-to-reputation" effect.

[TABLE 7 HERE]

It might be, however, suggested, that the reputation score of large banks correlates with other bank characteristics that are driving our results. For instance, if large banks with better reputation are also those considered to offer more stability and safety, what we are capturing could be a "flight-to-safety" effect by depositors rather than a search for information security. Indeed, previous studies argue that depositors can look for a safe shelter to protect their savings, such as implicit government guarantees of large and systemically important banks (Farhi and Tirole, 2012). To test more specifically whether depositors seek out more systemically important banks, we identify globally systemically important US banks on the basis of the ranking published from 2011 by the Financial Stability Board. We then examine whether large banks being recognized as more systemically important show more positive spillovers after a cyberattack. The

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

results presented in Panel B of Table 7, however, do not support this alternative interpretation. We do not find this to be the case.

Ultimately, the positive spillovers we document in favor of large banks indicate a business stealing effect against smaller banking firms that are perceived as less competitive in terms of cybersecurity by depositors. This effect can be understood in the context of theoretical models highlighting the competitive effects caused by investment and innovation gaps among firms operating in the same market.

4.2 Spillover Effects towards Small Banks

Next, we examine spillovers to (untreated) small banks. We define the branches of small banks in our initial control group as indirectly treated because they are located in those counties where the hacked banks operate. The untreated group consists of the branches of the same banks that reside in unaffected adjacent counties. As before, we show in Panel A of Table 8 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 branches of small banks prior to the event.

[TABLE 8 HERE]

Panel B of Table 8 shows the regression results. Across all specifications, we find no evidence that the growth rate of deposits at (indirectly) treated branches differs significantly from the growth rate observed for untreated branches of the same small banks. Therefore, there is no evidence that a cyberattack on a small bank has any negative spillover effects on untreated small banks.

One potential explanation as to why we fail to detect a negative spillover in the full sample of untreated small banks is that some of the existing and potential depositors of these small banks are not sufficiently exposed to the shock affecting the hacked banks. For instance, within a given county, some untreated small banks might operate in local communities where hacked small banks do not operate branches. In these communities, therefore, the awareness of the cyberattack for depositors might not be

significant enough to raise widespread concerns over the cybersecurity of their own small banks as it would be required by a negative spillover. Indeed, Iyer et al. (2016) show that a depositor's network is a key determinant of the spread of bank runs within deposit markets. Additionally, Addoum et al. (2020) document that the bankruptcy of one firm negatively affects lending spreads for geographically proximate solvent firms. In a similar vein, Aharony and Swary (1996) show evidence of contagion for solvent banks with headquarters in closer geographic proximity to failing banks' headquarters, while Pino and Sharma (2019) show that contagion within the US banking system spreads locally[12].

To understand if the explanation above is plausible and there is a geographic proximity effect in terms of negative spillovers among small banks, we compute a measure of branch overlaps between the hacked bank and each untreated small bank based on ZIP codes. For each ZIP code, we count the overlaps in the branch network between two banks and scale the number of overlaps by the total number of branches of the untreated small bank. This measure, therefore, aims at capturing how closely the existing, and potential, depositors of these untreated banks have been exposed to the shock.

Next, indirectly treated small banks are sorted into high (low) degree of branch network overlaps if they are above (below) the median values of the variables in the year before a cyberattack. We then estimate a model similar to equation (4). We report the results in Panel C of Table 8. We find some evidence of negative spillovers for untreated small banks that share a high degree of branch overlaps with the hacked banks; that is, in hacked counties the branches of these banks show a slowdown in deposit growth after the cyberattacks compared to the branches operating in unhacked counties. For the group of banks with low branch network overlaps, there is no evidence of any spillover effect.

To summarize, while on average untreated small banks do not show any significant changes in their deposit growth rates after cyberattacks, we find that the hack

[12]A related stream of the literature documents that the geographic distance between a firm and its stakeholders decrease monitoring intensity and information collection incentives (Agarwal and Hauswald, 2010; Chen et al., 2020).

undermines the competitive positions of those small banks that have more geographic similarity with the hacked banks. The presence of a geographic dimension of negative spillovers from cyberattacks is in line with previous evidence on the local nature of contagion effects within the US banking sector and for corporate borrowers.

5 Aggregate and Real Effects

We find that a successful cyberattack on a small bank results in deposit flows to larger banks and, to some extent, a disadvantage for smaller unhacked banks that share more geographic similarity in the branch network with the hacked bank. In the next two sections, we provide evidence of the aggregate consequences that these market dynamics generate within local banking markets.

5.1 Market Structure Effects

The results reported in Section 4 hold important implications for the market structure of deposit markets. We find a shift in the competitiveness from small to large banks. As a result, we should expect an increase in the aggregate market share of large banks in these hacked local deposits markets.

[TABLE 9 HERE]

Along these lines, in Panel A of Table 9, we use our initial sample of treated and control banks to estimate a bank-county level regression where we employ as a dependent variable the deposit market share of each bank in a county. The coefficients on Treated x Post indicate a decrease of approximately 1 percentage point in the county market share of hacked small banks compared to small 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, therefore, the market share of treated banks decreases by approximately 14%.

In Panel B of Table 9, we assess the impact of the cyberattacks on the aggregate market shares of large banks. We compare the market shares of large banks in counties where hacked banks have branches and the market share in adjacent counties wherein hacked banks are absent. We find that the market share of large banks has shown a relative increase by about 3 percentage points.

5.2 Small Business Lending Effects

The growing influence of large banks in local deposit markets can also have implications in terms of access to credit for small firms. Since large banks tend to be less inclined to supply small business lending, especially in times of crises, local businesses could face increasing financial frictions (Bord et al., 2018; Chen et al., 2017). Therefore, the access to credit of (small) local businesses may deteriorate after a cyberattack, negatively impacting the development of local economies (Berger et al., 2017; Hakenes et al., 2015).

We test the validity of the argument above by relying on small business lending data based on the Community Reinvestment Act (CRA). This dataset offers the possibility to construct county-based measures of small business loans by aggregating information available at the bank level. Small business loans are defined as loans with a value lower than 1$ million.

To understand if the effects of cyberattacks have implications on small business lending, we compare the evolution of the log transformation of small business loans in the group of treated counties (that is, counties where we have at least one branch of small hacked banks) and the group of adjacent counties where these banks do not operate any branches. The matching between the treated counties and adjacent counties reduces concerns over omitted demand factors that can affect the evolution of small business lending. However, it is unlikely that it is sufficient to fully remove these concerns. We therefore take additional steps to remove the contamination effects coming from the demand side.

First, in all specifications, in addition to county fixed effects, we include the lag values of the number of small business loans granted in the previous year as a proxy for

the number of potential small customers. Second, in additional specifications we include lag values of county macro controls such as the GDP growth rate the log transformation of the unemployment rate and the growth rate of the number of establishments.

[TABLE 10 HERE]

The results reported in Panel A of Table 10 for the log transformation of total loans with a value lower than 1$ million offer some weak evidence of a decline in loan growth after a cyberattack. One possible explanation for this weak result is that our loan aggregate includes borrowers with heterogenous size that are not necessarily treated equally by large banks within the local banking market. We progress, therefore, by focusing on a tighter definition of small borrowers based on loans granted with a value lower than 250,000$. We then compare the results of this aggregate with the findings for larger loans (from 250,000$ up to 1$ million). The results reported in Panel B and C of Table 7 show some degree of heterogeneity across loan size groups, with only the very small borrowers suffering from a decline in the growth rate of the amount granted by local banks after the cyberattack.

Ultimately, our results are consistent with existing evidence on the importance of small banks in alleviating financial constraints of small firms (see, for instance, Berger et al., 2017). We find evidence of broad market-wide effects subsequent to cyberattacks on small banks, leading to a reduced competitiveness of these banks and a decrease in the growth of small business lending. This decrease is driven by the provision of loans to very small borrowers.

6 Conclusion

Cybersecurity is a rising concern for regulators and bankers. Unlike large banks with a wide range of human and financial resources to strengthen their IT infrastructure, small banks are more susceptible to cyberattacks. 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. We contrast two possible explanations for the effects we document. The first emphasizes the importance depositors assign to the ability of banks to protect private information. The second focuses on the potential funding loss for depositors due to broader concerns about bank stability highlighted by successful data breaches. Our findings validate the first interpretation. In particular, we show that the negative effects for hacked banks are more pronounced when depositors are likely to be more exposed to identity theft or have less digital knowledge, do not depend on small bank risk and extend to relationships with bank customers that are exposed to information loss but not the risk of loss of savings.

We next document that cyberattacks generate positive deposit spillovers to branches of unaffected large banks, operating in geographically proximate locales, as well as negative spillovers to small banks that have more branch overlaps with the hacked bank. Additionally, in response to the cyberattacks, depositors opt for large banks with high customer reputation but are not influenced by the implicit bailout guarantees of these banks. In other words, cyberattacks lead to a "flight-to-reputation" effect as larger, more reputable banks are likely to be seen by depositors as more secure against cyberattacks. The described effects result in a growing market share of large banks and a loss of competitive of small banks in local deposit markets. Not surprisingly, therefore, we document that the market-wide consequences of cyberattacks reduce access to credit for very small borrowers.

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

Addoum, J. M., A. Kumar, N. Le, and A. Niessen-Ruenzi (2020). Local bankruptcy and geographic contagion in the bank loan market. *Review of Finance 24*(5), 997–1037.

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

Aharony, J. and I. Swary (1996). Additional evidence on the information-based contagion effects of bank failures. *Journal of Banking & Finance 20*(1), 57–69.

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*.

Baker, A. C., D. F. Larcker, and C. C. Wang (2022). How much should we trust staggered difference-in-differences estimates? *Journal of Financial Economics 144*(2), 370–395.

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.

Behr, P., L. Norden, and F. Noth (2013). Financial constraints of private firms and bank lending behavior. *Journal of Banking & Finance 37*(9), 3472–3485.

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.

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.

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

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.

Brown, J. D. and J. S. Earle (2017). Finance and growth at the firm level: Evidence from sba loans. *The Journal of Finance 72*(3), 1039–1080.

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, 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.

Cortés, K. R. (2015). Did local lenders forecast the bust? Evidence from the real estate market. *Working Paper*.

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.

Degryse, H. and P. Van Cayseele (2000). Relationship lending within a bank-based system: Evidence from european small business data. *Journal of Financial Intermediation 9*(1), 90–109.

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.

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

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

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*.

Engels, C., B. Francis, and D. Philip (2021). The cost of privacy failures: Evidence from bank depositors' reactions to breaches. *SSRN Working Paper*.

Farhi, E. and J. Tirole (2012). Collective moral hazard, maturity mismatch, and systemic bailouts. *American Economic Review 102*(1), 60–93.

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.

Gordon, L. A., M. P. Loeb, and L. Zhou (2011). The impact of information security breaches: Has there been a downward shift in costs? *Journal of Computer Security 19*(1), 33–56.

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.

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.

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.

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*.

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

Klette, T. J. and S. Kortum (2004). Innovating firms and aggregate innovation. *Journal of Political Economy 112*(5), 986–1018.

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.

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.

Nicoletti, A. (2018). The effects of bank regulators and external auditors on loan loss provisions. *Journal of Accounting and Economics 66*(1), 244–265.

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.

Pino, G. and S. C. Sharma (2019). On the contagion effect in the us banking sector. *Journal of Money, Credit and Banking 51*(1), 261–280.

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).

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

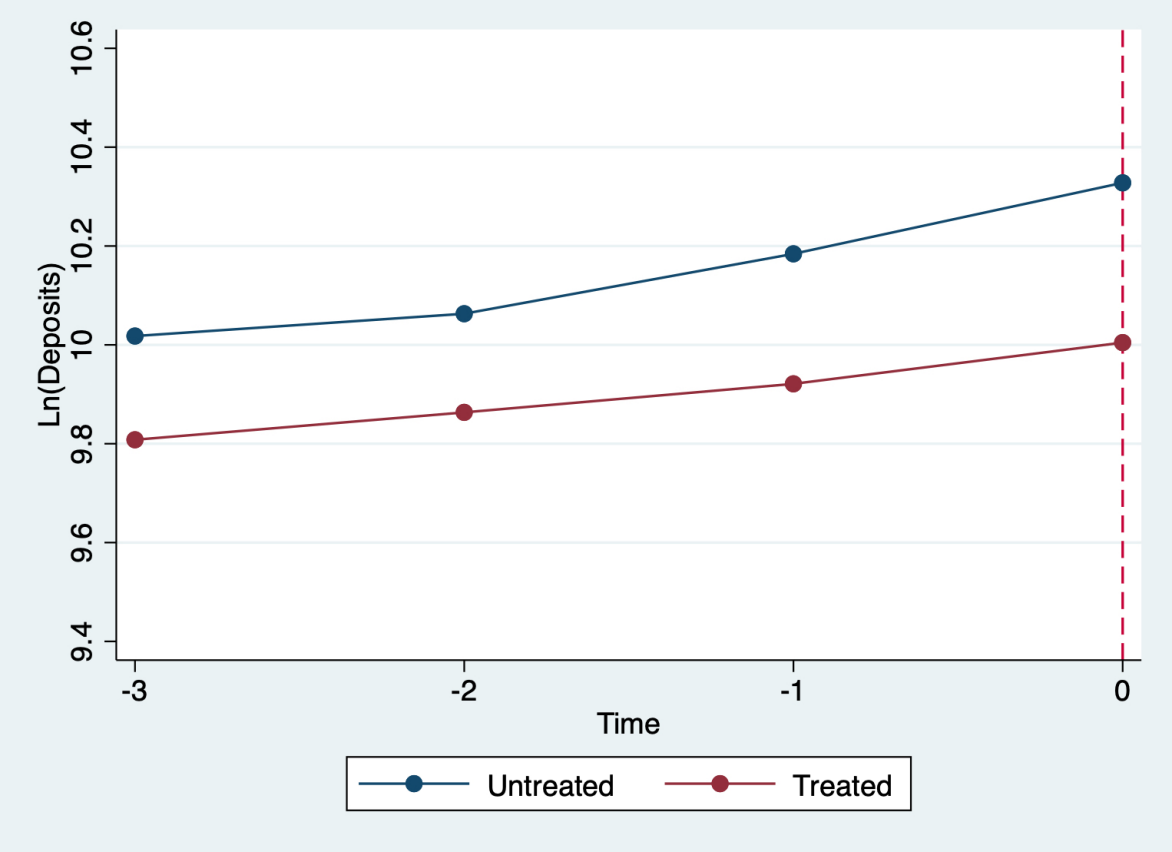
Skrastins, J. and V. Vig (2019). How organizational hierarchy affects information production. *The Review of Financial Studies 32*(2), 564–604.

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.

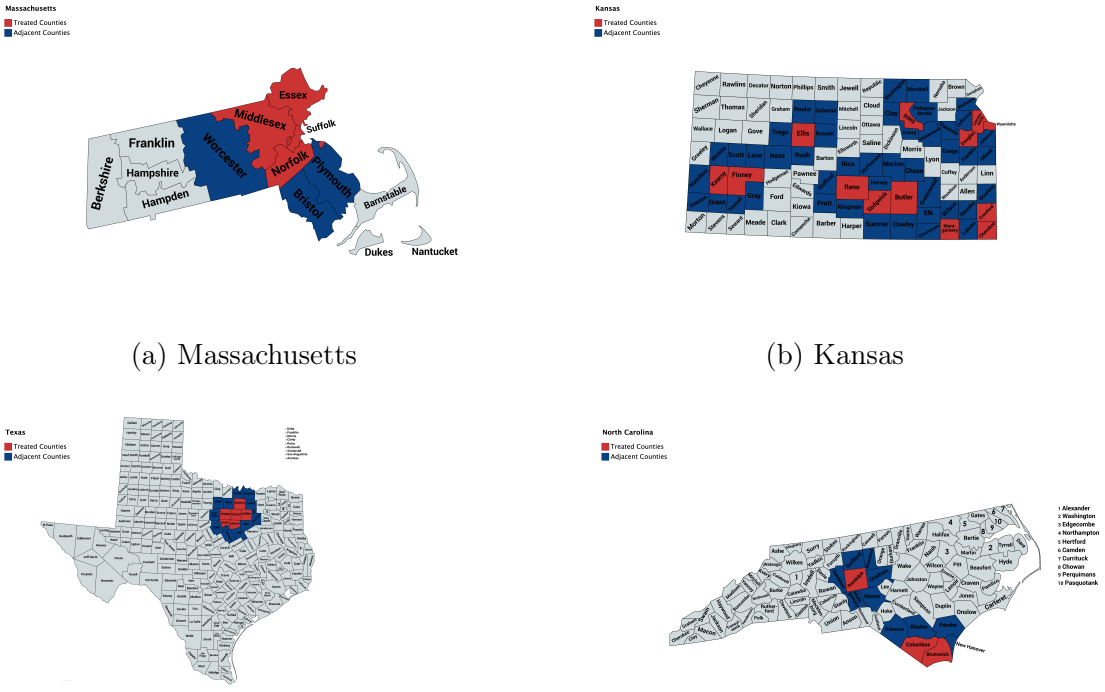
Tolbert, C. and K. Mossberger (2020). U.S. Current Population Survey & American Community Survey Geographic Estimates of Internet Use, 1997-2018.

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 results of a logit model for the pre-shock period. The dependent variable is an indicator variable equal to one if a bank was subsequnetly hacked; and zero otherwise. Panel D 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. Standard errors given in parentheses are corrected for heteroskedasticity and bank-level clustering. ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels.

Panel A	Descriptive Statistics					
	Obs.	Mean	Median	SD	25th	75th
	(1)	(2)	(3)	(4)	(5)	(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	Treated (A)	Untreated (B)	Normalized Diff. (A-B)	T-test (A-B)
	(1)	(2)	(3)	(4)	(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

Table 1 (cont.)
Descriptive Statistics and Parallel Trends

Panel C	Pre-Shock determinants of becoming a target		
	Ln(Deposits)		
	(1)	(2)	(3)
Ln(Deposits)	0.006	0.007	0.012
	(0.012)	(0.021)	(0.036)
ROA		-0.415	0.359
		(1.808)	(2.859)
NPL		-0.042	-0.198
		(0.584)	(0.564)
Tier 1		-0.093	-0.038
		(0.233)	(0.265)
Loan		-0.075	-0.027
		(0.136)	(0.134)
Productivity		-0.009	-0.010
		(0.005)	(0.009)
County FE	No	No	Yes
Observations	196	189	189
Adjusted R^2	0.00	0.01	0.01

Panel D	Parallel Trends			
	Treated (A)	Untreated (B)	Diff. (A-B)	T-value
	(1)	(2)	(3)	(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 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? Alternative Specifications

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 in column (1) to (4). Column (1) 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. Column (2) uses an alternative (-2;+2) years estimation window while column (3) employs a (-1;+1) year window. 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). In column (4), branch fixed effects (originally employed in our main results in Table 2 Panel B) are replaced with branch × cohort fixed effects. In column (5), the dependent variable Ln(Deposits) is the logarithmic transformation of the total amount of deposits of each bank in a given county and year. 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. In column (6), 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. 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.

	Baseline Regression: Alternative Specifications					
	(Bertrand et al., 2004)	(-2;+2)	(-1;+1)	Branch × Cohort FE	Bank-County	Falsification Test
	(1)	(2)	(3)	(4)	(5)	(6)
Treated × Post	-0.218**	-0.221***	-0.184**	-0.211***	-0.258***	-0.047
	(0.086)	(0.082)	(0.077)	(0.076)	(0.076)	(0.045)
Size Control	Yes	Yes	Yes	Yes	Yes	Yes
Other Bank Controls	Yes	Yes	Yes	Yes	Yes	Yes
Bank FE	No	No	No	No	Yes	No
Branch FE	Yes	Yes	Yes	Yes	No	Yes
County x Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4161	10679	6303	14382	2502	7887
Adjusted R^2	0.927	0.941	0.948	0.951	0.744	0.966

Table 4
What Drives 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 according to depositors' exposure to identify theft risk and digital sophistication and to small bank risk. In Panel A, we report results for identify theft and digital sophistication. To account for the exposure of depositors to the risk of identity theft, we rely on data from the Federal Trade Commission (FTC). The FTC offers data on number of consumer complaints per 100,000 population related to identity theft in each metropolitan statical area (MSA). Digital sophistication is measured using data from Form 477 on internet access connections per thousands of households at the county level, provided by the Federal Communication Commission. In Panel B, we report results for small bank risk. First, we identify riskier firms using the log of the Z-score in the year before the cyberattack. Those firms with Z-score above (below) the sample median are considered more (less) risky. In a second step, we define those banks as riskier (less risky) that jointly have NPL and Tier 1 rations above (below) the sample median. As indicated above, in both Panels banks are sorted into high (low) groups if they are above (below) the median in respect to each variable. 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 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	Identity Theft			Digital Sophistication		
	Ln(Deposits)			Ln(Deposits)		
	(1)	(2)	(3)	(4)	(5)	(6)
Treated High × Post	-0.446***	-0.438***	-0.402***	-0.063	-0.058	-0.050
	(0.103)	(0.102)	(0.095)	(0.039)	(0.041)	(0.044)
Treated Low × Post	-0.108**	-0.104**	-0.096**	-0.521***	-0.514***	-0.481***
	(0.045)	(0.046)	(0.048)	(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	0.936	0.936	0.936	0.937
Panel B	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	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	0.044	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

Table 5
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

Table 6
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 large bank spillovers. 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 and 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). Panels A provides 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 formally examines spillovers to large banks. 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 and small banks operating in unaffected adjacent counties) 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			
	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 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

Table 7
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 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 in adjacent counties) after the shock. More specifically, Panel A reports heterogeneous depositor results for a measure constructed to capture the reputation of 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). Panel B reports heterogeneous depositor results for a measure constructed to capture the systemic importance of large banks in relation to deposit flows. The TBTF measure is based on the List of Global Systemically Important Banks (G-SIBs) published by the Financial Stability Board (FSB). Treated banks are those that appear on the lists published by the FSB on an annual basis since 2011. 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	Large Bank Spillover Reputation Ln(Deposits)		
	(1)	(2)	(3)
Treated Top 5 Reputation × Post	0.374***	0.398***	0.417***
	(0.108)	(0.112)	(0.117)
Treated 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
Panel B	Large Bank Spillover Systemic Risk Ln(Deposits)		
	(1)	(2)	(3)
Treated TBTF × Post	0.158**	0.160**	0.137*
	(0.079)	(0.080)	(0.071)
Treated Non-TBTF × Post	0.143**	0.141**	0.170**
	(0.065)	(0.066)	(0.080)
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	0.015	0.019	-0.034
Observations	37603	37587	34696
Adjusted R^2	0.897	0.898	0.904

l

Table 8
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 small bank spillovers. 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 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 provides 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 formally examines spillovers to 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. To better understand the drivers of the depositor response, we test for heterogeneity in the spillover effect in Panel C. We compute a measure of branch overlaps between the hacked bank and each untreated small bank based on ZIP codes. For each ZIP code, we count the overlaps in the branch network between two banks and scale the number of overlaps by the total number of branches of the untreated small bank. We sort treated small banks into high (low) degree of branch network overlaps if they are above (below) the median values of the variables in the year before a cyberattack. The difference-in-differences estimate of the coefficient of Treated (High) Low 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. 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			
	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 B		Small Bank Spillover 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
Panel C		Small Bank Spillover Branch Overlap Ln(Deposits)		
		(1)	(2)	(3)
Treated High Branch Overlap × Post		-0.154***	-0.180***	-0.245***
		(0.039)	(0.044)	(0.065)
Treated Low Branch Overlap × Post		-0.024	-0.017	-0.026
		(0.086)	(0.087)	(0.096)
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		0.130	0.163*	0.219*
Observations		32165	31756	29539
Adjusted R^2		0.921	0.925	0.931

Table 9
Market share and local market concentration

The table below reports two sets of difference-in-differences analyses concerning cyberattacks on small banks at the bank-county level. 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 aggregated at the bank-county level. In Panel B, the aggregate market share is the deposit market share of each large 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 × 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. Panel A takes the perspective of the affected bank. Size is the logarithmic transformation of bank total assets in thousands of US$. Return on assets is the ratio between net income and total assets (ROA), tier 1 capital ratio is defined as total tier 1 capital divided by risk weighted assets (Tier 1), non-performing loans is defined as the fraction of non-performing loans with respect to total loans is a proxy for credit risk (NPL), loan is total loans divided by total assets (Loan) and productivity is defined as the ratio between total assets and the number of employees (Productivity). In Panel B, we explicitly consider economic drivers of deposit growth at the county level. In this specification we control for the following economic controls: GDP is the growth rate of the county-level gross domestic product, Unemployment is the natural logarithm of the unemployment rate at the county-level, Establishments is the growth rate in the number of establishments. All models include branch and county × year fixed effects. Standard errors given in parentheses are corrected for heteroskedasticity and county-level clustering. ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels.

Panel A	Bank-County Market Share		
	Deposit 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
Panel B	Aggregate Market Share (>10Bln)		
	Deposit market share		
	(1)	(2)	(3)
Treated × Post	0.030**	0.030**	0.032**
	(0.013)	(0.013)	(0.013)
GDP_{t-1}		0.021	0.016
		(0.021)	(0.024)
$Unemployment_{t-1}$			-0.045**
			(0.021)
$Establishments_{t-1}$			0.106
			(0.080)
County FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Observations	1357	1357	1200
Adjusted R^2	0.960	0.960	0.964

Table 10
Small Business Lending

The table below reports two sets of difference-in-differences analyses concerning cyberattacks on small banks at the bank-county level. Cyberattacks are identified using the breach classification "HACK" by Privacy Rights Clearinghouse (PRC). We test whether cyberattacks have an impact on the access to credit of (small) local businesses. We rely on small business lending data based on the Community Reinvestment Act (CRA). This dataset offers the possibility to construct county-based measures of small business loans by aggregating information available at the bank level. Small business loans are defined as loans with a value lower than 1$ million. In Panel A, we report results on the log transformation of total loans originated with a value lower than 1$ million. In Panel B and Panel C, we focus on a tighter definition of small borrowers based on loans granted with a value lower than 250,000$ and loans granted with a value from 250,000$ up to 1$ million respectively. Within all specification we control for the number of loans granted (*Ln(Number of Loans)*). We also control for several economic factors such as: GDP is the growth rate of the county-level gross domestic product, Unemployment is the natural logarithm of the unemployment rate at the county-level, Establishments is the growth rate in the number of establishments. All models include branch and county × year fixed effects. Standard errors given in parentheses are corrected for heteroskedasticity and county-level clustering. ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels.

Panel A	Ln(Loans Originated)		
	(1)	(2)	(3)
Treated × Post	-0.045*	-0.045*	-0.033
	(0.024)	(0.024)	(0.024)
Ln(Number of Loans)$_{t-1}$	1.177***	1.178***	1.185***
	(0.079)	(0.079)	(0.084)
GDP$_{t-1}$		-0.049	-0.083
		(0.092)	(0.100)
Unemployment$_{t-1}$			-0.109
			(0.067)
Establishments$_{t-1}$			0.275
			(0.262)
County FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Observations	1354	1354	1197
R^2	0.992	0.992	0.992
Panel B	Ln(Loans Originated <250k)		
	(1)	(2)	(3)
Treated × Post	-0.057***	-0.057***	-0.055***
	(0.019)	(0.019)	(0.019)
Ln(Number of Loans)$_{t-1}$	1.179***	1.180***	1.196***
	(0.058)	(0.058)	(0.062)
GDP$_{t-1}$		-0.041	-0.055
		(0.061)	(0.068)
Unemployment$_{t-1}$			-0.028
			(0.054)
Establishments$_{t-1}$			0.410***
			(0.151)
County FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Observations	1354	1354	1197
R^2	0.995	0.995	0.995
Panel C	Ln(Loans Originated >250k and <1M)		
Treated × Post	-0.043	-0.042	-0.011
	(0.039)	(0.039)	(0.038)
Ln(Number of Loans)$_{t-1}$	1.096***	1.099***	1.041***
	(0.139)	(0.141)	(0.140)
GDP$_{t-1}$		-0.142	-0.202
		(0.211)	(0.208)
Unemployment$_{t-1}$			-0.264**
			(0.106)
Establishments$_{t-1}$			-0.140
			(0.468)
County FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Observations	1322	1322	1168
R^2	0.973	0.973	0.975

Online Appendix

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
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	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