Data Breach Announcement Effect on Bank Operations and Performance

Isarin Durongkadej[*], Heng (Emily) Wang[†‡]

This version: September 2023

Abstract

A downside of the digital economy is that banks are prone to experiencing data breaches resulting from, for example, cyberattacks, system glitches, and employee negligence. We investigate how a data breach announcement affects bank operations and stock performance. Our findings show that banks have an outflow of insured and brokered deposits after a data breach announcement. Furthermore, we find that deposits transfer from banks with data breaches to banks with no history of experiencing data breaches. Data breaches negatively affect stock returns in both the short term and long term. However, we do not find a systematic long-term impact from the data breach announcement on bank operations. In terms of lending, banks with data breach tend to increase their lending after announcements, which is supported by the incentive of CEO compensations. Higher lending could be associated with CEOs' wealth preservation for banks with data breach.

JEL classification: G14, G21

Keywords: bank operations and performance, cybersecurity, data breach, bank stability.

Statements and Declarations
Partial financial support was received from Georgia College and State University as a summer research grant when Isarin Durongkadej was employed by the university.

[*]Lucas College and Graduate School of Business, Department of Accounting and Finance, San Jose State University, San Jose, CA, 95192; E-mail: isarin.durongkadej@sjsu.edu [Corresponding author]

[†]Martha and Spencer Love School of Business, Department of Finance, Elon University, Elon, NC 27244; Email: ewang3@elon.edu

[‡]We are grateful to Sara Yasar, Samuel Rosen, Chris Clark, Leng Ling, Brent Evans, and seminar participants at 2020 COBRA Conference, 2020 Southern Finance Association Annual Conference, 2021 Financial Management Association Annual Meeting, and 2022 Eastern Finance Association Annual Meeting for their helpful comments and suggestions. We thank Jonah Paul and Donavan Lingerfelt for research assistance. All errors are our own. Isarin Durongkadej received a research support grant from Georgia College and State University.

"In 2019, Capital One data breach compromised data of over 100 million people. The bank agreed to pay $190 million to settle claims"

The New York Times on December 23, 2021

1. Introduction

In the past few years, data breach incidents have increased sharply. Equifax, one of the three largest credit reporting agencies in the US, announced a data breach in 2017 which affected more than 150 million Americans. Equifax paid $650 million to settle claims from customers and investigations from federal and state governments. In 2019, Capital One revealed a data breach of 140,000 Social Security numbers and 80,000 bank accounts. The breach cost $190 million for Capital One to settle claims. More recent examples include the ransomware attacks on JBS, one of the Americas largest beef producers, and Colonial Pipeline, a company that delivers most of the gasoline in the East Coast. These cyberattacks cause firms' extra expenses such as expenses for settlements and inventing new preventative measures. However, an important indirect cost of a data breach is losing client confidence. Clients may lose trust in firms that have a data breach and decide to switch to another firm that offers similar products or services. If the same situation happens to a bank in which many depositors lose confidence and decide to withdraw their money from the data breach bank, will this affect banks' operations such as deposit and loan activities? The increasing frequency and size of data breaches means that the problem of data breaches will continue to get worse. Data breaches may not only influence banks' operations, but also affect banks' stock prices after a data breach is announced to the public. A decrease in stock price could imply a decrease in confidence that people have in their banks.

Given a potential impact on the economy through changes in banking activities, we believe that whether a data breach affects banks' operation and stock performance or not is an important question. On the one hand, we might expect that, after a data breach, people lose trust in their banks and withdraw their deposits. At the same time, investors sell banks' stocks because they expect higher expenses and lower revenues. On the other hand, people might trust their banks regardless of data breach risks and decide to take no action. Clients still trust their banks, because they believe that banks are able to handle the situation, or that banks will eventually compensate them for any damage that occurred from a data breach. In addition, literature on bank switching costs shows that clients have a switching cost, and it prevents them from conveniently moving their accounts away. For example, banks may tie up clients with other services such as mortgages, payroll, car loans, and investment portfolios. It would be inconvenient for clients to move their accounts away though they wanted to. In this case, there should not be much effect on bank operations or stock performance. Since there are two possible effects from a data breach on bank operations and performance, our paper strives to understand which effect is more important. The results in our paper could provide banks and regulators with more information to understand more about a consequence after a data breach announcement. The literature related to cyberattacks analyzes the effect of data breach incidents on stock markets (Lending, Minnick, and Schorno, 2018; Wang, Wang, and Wu, 2022), bond markets (Iyer, Simkins, and Wang, 2020), options markets (Piccotti and Wang, 2022), firm policies (Kamiya, Kang, Kim, Milidonis, and Stulz, 2021), and credit markets (Mikhed and Vogan, 2018). We analyze the effect of data breach incidents on banks' operations and their stock performance.

To address our research questions, we gather information about bank operations from call reports, data-breach announcements from Privacy Rights Clearing House[1], and stock data from CRSP. In terms of methodologies, we use difference-in-difference (DID) to analyze the impact on banks' operations before and after a data breach announcement. Specifically, we examine the effect of a data breach announcement on several key variables of bank operations such as deposits and loans. In addition to evaluating only on the quarter of a data breach announcement, we also evaluate the impact on the subsequent quarters. Since there could

[1]https://privacyrights.org/data-breaches?terms=&f%5B0%5D=years%3A2019

be a delay from the time that banks announce their data breach and to the time that clients receive the information, the impact might appear later in a subsequent period. To reduce a confounding factor, we apply propensity score matching on all observed variables to find a set of control banks. As common to the literature on the event study, we use Cumulative Abnormal Returns (CARs) to find the impact of the data breach on banks' performance surrounding the date that a data breach was announced to the public. We find that a data breach announcement has significant impacts on banks' operations and stock performance. However, we do not see evidence of a long-term impact on bank operations.

In terms of deposits, banks experience a decrease in insured deposits a quarter after a data breach announcement. Higher risk aversion of insured depositors could be the reason. For instance, insured depositors may try to reduce the future risk of their breached information being exploited by withdrawing money from the breached banks and depositing it in a bank with no data breach. For banks with multiple breach announcements, we find that clients start to stop withdrawing money from banks with a data breach after the fourth announcement. The reason could be that clients learn from the previous breaches that their accounts are handled well by their banks and there are no unauthorized charges on their accounts. Hence, they still leave their money in the account expecting that their accounts will be safe.

The breached records are divided into known and unknown records. On many occasions, banks did not know what information was lost in the breach. We test the effect separately between known and unknown breach records lost, because clients may perceive the two types of breach differently. Clients may feel that if banks do not know the number of records lost, it could imply that banks are unorganized (e.g., poor data storage procedure). For unknown records, the impacts of a data breach announcement on bank deposits are stronger in terms of timing and the size of the impact. After a data breach announcement, the insured deposits decrease by about one percent right at the quarter when it is announced. The impact occurs earlier and three times as much as the impacts shown in the pooled results. In addition, the results also show a decrease in time deposits a quarter after a data breach announcement. Time deposits have a higher switching cost relative to non-time deposits since withdrawing before the term is due could result in a penalty. The decrease in time deposits indicates that depositors trade off between the switching cost and future losses from a data breach. The results are consistent with our hypothesis that clients perceive unknown lost records breach to be worse than the breach with known lost records. In addition, we find that depositors withdraw money from a breach bank and deposit it to another bank with no history of data breach.

Data breaches also impact banks' stock returns. Our analysis shows that CARs are significantly negative within different windows surrounding the data breach announcement date. The data breach announcement affects banks' stock returns not only in the short term but also in the long term. We examine the long-term effect of a data breach announcement and find consistent results that CARs are continuously decreasing through the period from the announcement date to the third quarter after the announcement.

Since we find that a data breach announcement reduces the stock value, banks may increase their lending to enhance future financial performance. We find evidence of data breach banks increasing their lending after a data breach announcement. Our tests show that CEOs of data breached banks have higher deep in-the-money options than non-breached banks. The results support an argument that data breached banks lend more to preserve the stock value.

The contribution of our paper is, first, related to the literature on the liquidity of financial intermediaries as sufficient liquidity is important for banks to maintain the economic role of capital providers (e.g., Diamond and Rajan, 2001; Gatev, Schuermann, and Strahan, 2009; Imbierowicz and Rauch, 2014; Chen, Chen, and Huang, 2021). Direct or indirect costs of data breaches may negatively reduce economic activities. Our study gives implications to banks about preparation for the new digital age on data breaches.

Second, data breaches have different characteristics than other corporate events and are worth exploring in financial research. The financial losses of a data breach could be one of the most hard-hitting consequences for a company both directly and indirectly (e.g., Lending et al., 2018; Huang and Wang, 2021; Foerderer and Schuetz, 2022). The financial losses span from bank revenues to capital markets such as stock markets, bond markets, and options markets. In addition, a data breach induces other types of losses such as reputational damage, operational downtime, legal action, and loss of sensitive data. All of them would deteriorate the existing situation and lead to additional financial losses. According to 2020 McAfee's report, the annual worldwide cost associated with cyberattacks was estimated to be $1 trillion which is about one percent of the global GDP.

Third, our study contributes to the literature of emerging operational risks (e.g., Chernobai, Ozdagli, and Wang, 2021; Berger, Curti, Mihov, and Sedunov, 2022). Cyber risk is worth to analyze separately from other operational risks because of its unique characteristics and its growing threat to financial stability (Curti, Gerlach, Kazinnik, Lee, and Mihov, 2006). Data breach is unpredictable but does happen frequently. It is hard to avoid through internal improvements or corporate governance, but still can generate sizable losses. It could be a result of both external factors and internal control.

Lastly, our paper contributes to the bank financial crisis and contagion (e.g., Allen and Gale, 2000; Brusco and Castiglionesi, 2007; Baur, 2012; Cont and Schaanning, 2019). The scope of the damage from a data breach could go beyond the stakeholders of the banks to the entire economy. If our society can get benefits from moving financial transactions to a digital platform, banks should be proactive to prevent the violation of client privacy due to a data breach. This is to preserve the bank's confidence and the spillover effect on the financial system. Though, currently, we do not see a bank failing after a data breach, our paper shows that data breach incidents have an impact on bank operations and stock performance. The financial system has progressed to an online platform fast. The data breach issues are expected to be more, not less. To keep the data breach problem in check, banks, financial markets, and regulators should be aware of the impact of a data breach that could occur.

The rest of the paper is outlined as follows. Section 2 provides literature of the data breach in financial markets and hypothesis development. Data and variable construction are in Section 3. Section 4 describes methodologies. Section 5 provides results and discussions. Section 6 concludes.

2. Literature and Hypothesis Development

2.1. Banks and deposit levels after a data breach announcement

The global economy has been evolving into a digital platform. With much higher performance of smartphones and computers today, we can perform financial transactions more conveniently. However, the good and the bad usually come hand in hand. Concerning data breaches, companies have experienced system hacking, computer hardware stealing, and system glitches. In July 2019, Capital One revealed a data breach of 98 million US consumers. Capital One needs to pay the settlement of 190 million dollars.[2] In 2019, the Financial Crimes Enforcement Network (FINCEN) reported more than 12,500 cases of cyberattacks on banks.[3] Equifax was downgraded by Moody's in 2017 because of the cyberattack. *"We are treating this with more significance because it is the first time that cyberattack has been a named factor in an outlook change."* Joe Mielenhausen, a spokesperson for Moody's, told CNBC. [4]

Data breaches can be very costly. Equifax, a credit bureau company, announced in September 2017 that customers' information such as Social Security and driver's license numbers were compromised. Equifax paid around $650 million to settle claims from customers and investigations from federal and state governments. The incident exposed the personal information of more than 145 million people. After this incident, the stock price of

[2] https://www.capitalonesettlement.com/en
[3] https://www.fincen.gov/reports/sar-stats
[4] https://www.cnbc.com/2019/05/22/moodys-downgrades-equifax-outlook-to-negative-cites-cybersecurity.html

Equifax tumbled, and the CEO was forced to resign.[5] In Europe, policymakers adopted a new law, the General Data Protection Regulation, known as the GDPR. The GDPR allows regulators in each EU country to charge fines of up to 4 percent of revenue for a breach.[6] This regulation scopes the limit of fines that banks have to pay.

The consequences after a data breach in the banking industry could go far beyond the settlement that banks need to pay to their clients or federal regulators. Data breaches may reduce the deposit level of banks due to depositors losing confidence in bank security. Bank liquidity problems may arise because of the lower level of confidence (Diamond and Dybvig, 1983) or the level of uncertainty (Arifovic and Jiang, 2019). When there are more uncertainties, depositors may withdraw money from banks with a data breach and deposit it into another bank without a data breach; therefore, the deposit outflow after a data breach announcement may occur.[7]

Many articles have examined how financial crises affect banks' operations. For example, Peria, Soledad, and Schmukler (2001) find that depositors in Argentina, Chile, and Mexico discipline banks by withdrawing deposits during banking crises in the 1980s-1990s. Acharya and Mora (2015) examine how the financial crisis during 2007-2009 affects deposits and loans. However, the analysis of how a data breach affects banks' operations and performance has not been conducted.

Data breaches could affect a sense of trust within the financial system. People may move their accounts from a bank that experienced a data breach to another bank without a data breach. Clients may lose trust in a bank security system when their private data is compromised.[8] The worst-case scenario is that bank runs may occur, and further cause economy-wide damage.[9] The short-term stock performance of these banks should be neg-

[5]https://www.nytimes.com/2019/07/19/business/equifax-data-breach-settlement.html?searchResultPosition=1

[6]https://www.nytimes.com/2019/07/08/business/british-airways-data-breach-fine.html?searchResultPosition=9

[7]This is also consistent with Diamond and Rajan (2000). They show in their model that a bank run could occur when the level of uncertainty increases

[8]Campbell, Gordon, Loeb, and Zhou (2003) find that the stock market reacted negatively when the publicly traded US corporations reported their information security breaches in newspapers.

[9]Diamond and Dybvig (1983) and Bryant (1980) show in their models that when depositors lose confi-

atively affected for the same reason of lower confidence in the security system of the bank after a data breach. Other than the impact of data breaches on stock prices, researchers also explore how a data breach impacts firm value (Iyer et al., 2020). To our knowledge, however, the real economic cost of a data breach on bank operations and performance has not been examined.

Even so, clients may not move their accounts to another bank for several reasons, such as trusting that banks can handle the situation or banks having higher switching costs (Sharpe, 1990; Kim, Kliger, and Vale, 2003; Vesala, 2007). For example, a client may have many financial transactions with a bank such as mortgages, payroll, etc. Therefore, it could be inconvenient to move their accounts to another bank. Concerning switching costs, there may not be any effect on bank deposits after a data breach.[10] For the first hypothesis, we test whether banks experience deposit outflows after a data breach announcement.

2.2. *Effects of unknown and known breach records*

The effect of a data breach on banks' operations may also vary depending on whether the records of a breach are known or unknown. FDIC requires banks to notify clients when they have a data breach.[11] Banks may notify a larger group of potential clients whose privacy information may have been compromised. Banks' clients may feel less secure when banks are unable to tell the scope of the breach. In this case, data breaches with unknown records lost may affect the deposit outflows on a larger scale. From the perspective of the depositors, the benefit of moving their money to a safer bank may outweigh the switching cost. In terms of post-data breach managing costs, when the record number of a data breach is unknown, banks may not accurately estimate potential losses from the breach. In some cases, they reserve capital for the maximum possible losses. With higher capital reserves,

dence in their banks and withdraw their money at the same time, bank runs would occur and damage the economy

[10]Sharpe (1997) finds that banks have higher monopoly power when the switching cost is high. Kim et al. (2003) find that the switching costs also exist in banks' lending.

[11]https://www.fdic.gov/news/financial-institution-letters/2005/fil2705.html

banks' profits are expected to decrease, because they need to reduce their loan issuance. For banks, the potential costs resulting from a data breach include investigation and remediation costs, credit monitoring costs, legal fees, public relations fees, and costs of settling litigation or government investigations (Black, 2013). On the client side, clients could interpret the unknown record of a data breach as unorganized and reckless relative to the breach with known records. The breach of unknown records lost may reflect the overall operation of the banks and result in losing client trust. As a result, the impact of a data breach announcement on bank operations and performance could be larger for a data breach with unknown records lost.

2.3. Effects on bank stock returns

Additionally, the impact of data breach announcements could affect banks' stock returns as well. Spanos and Angelis (2016) conduct a survey analysis showing that more than 37 papers study the effect of data breaches on stock prices from 2003 to 2015. The stock market could negatively react to data-breach announcements because a data breach has a negative effect on the firm's profit such as potential financial losses from settlement claims or a bad reputation resulting in losing customers to another bank with no data breach. Some investors may choose to sell their stock holding when they feel uncertain about the future performance of banks with a data breach. In this case, stock investors perceive the data breach event as negative news that could negatively affect the bank's future revenues or costs. We employ CARs to estimate the stock market performance of banks surrounding the data breach announcement date.

In addition, banks with a larger number of breach records may experience a larger amount of negative CARs. For example, a bank that lost a large number of records of clients' information should have a larger impact on their stock price than a bank that lost a smaller number of records. Fang and Peress (2009) demonstrate that media coverage has an impact on stock performance. A larger number of breach records affect more clients, and they could potentially convey the message to others. After investors receive the information, they may sell the stock of a bank with a data breach to avoid future capital losses. Therefore, banks with a higher number of data breach records may experience a larger drop in CARs than the banks with a lower number of breach records.

2.4. Effects on bank loans

Bank managers and shareholders may have a conflicting goal (Allen and Saunders, 1992). Bank managers may focus on a short-term performance, but shareholders focus more on a long-term goal. In this context, bank managers may try to lend more after a data breach announcement to improve financial performance to compensate for a drop in bank stock values from the data breach announcement conditional on the level of liquidity after a data breach announcement. The literature on bank window dressing supports our conjecture. For example, a study by Bank for International Settlements by Garcia, Lewrick, and Sečnik (2021) shows that Banks in Europe suppress their year-end balance sheet to avoid being designated as Global Systemically Important Banks (G-SIBs), because G-SIBs will be subject to more scrutiny and higher capital requirement. In addition, Ho, Huang, Lin, and Yen (2016) find that over confident CEOs tend to lower lending standards and raise leverage before a financial crisis. Banks with over confident CEOs may increase lending after a data breach announcement if banks do not experience liquidity outflow after a data breach announcement. On the other hand, we may not find any effect on loan activities after a data breach announcement, since loan transactions have high switching costs. Sharpe (1990) finds that borrowers and banks have accumulated relationships and reduced the monitoring costs. The banks that have long-term relationships with borrowers know their clients better and provide an appropriate lending rate matching the client's risk. If another bank offers loans to the borrowers whom the bank has never had a relationship with, the borrowers might be charged a higher rate. The new bank charges a higher lending rate, because the bank has not developed enough relationships with the new clients to fully understand clients' risk profiles, and the bank compensates for their risk by charging a higher lending rate.

3. Data and Variable Construction

We obtain the data breach announcement dataset from Privacy Right Clearing House (PRC) for the period of 2005 to 2018. Other than the banking sector, the types of companies in the PRC database include other sectors such as retail, health care, and non-profit organizations. The data contains many interesting pieces of information: date made public, company name, city, state, type of breach, type of organization, total records, description of the incident, and information source. Appendix B provides examples of data breaches. The full sample of the data is publicly available on the PRC website.[12] The types of data breaches include CARD (Payment Card Fraud), HACK (Hacking or Malware), INSD (Insider), PHYS (Physical Loss), PORT (Portable Device), STAT (Stationary Device), DISC (Unintended Disclosure) and UNKN (Unknown).[13]

We obtain accounting data from Compustat, stock data from CRSP, and banking data from Bank Regulatory (call report). The identifiers (i.e. GVKEY, PERMNO, CIK, and CUSIP) of the firms in the data breach dataset from PRC are hand collected. The number of data breach announcements by financial institutions for the period of 2005 to 2018 is 209, with 124 unique institutions. After merging all the datasets, the number of data breach announcements by banks is 87 with 39 unique banks PERMCO-RSSDID linked. To merge all three datasets, first, the bank Call Report is merged with the CRSP data using CRSP-FED linking table. The Federal Reserve Bank of New York provides the PERMCO-RSSD

[12]https://privacyrights.org/data-breaches.

[13]CARD involves debit and credit cards that are not accomplished via hacking, such as skimming devices at point-of-service terminals. HACK refers to being hacked by an outside party or infected by malware. INSD is caused by insiders with legitimate access who intentionally breach information, such as an employee, contractor or customer. PHYS includes paper documents that are lost, discarded or stolen (non-electronic). PORT includes lost, discarded or stolen laptop, PDA, smartphone, memory stick, CDs, hard drive, data tape, etc. STAT refers to stationary computer loss (lost, inappropriately accessed, discarded or stolen computer or server not designed for mobility). DISC is unintended disclosure not involving hacking, intentional breach or physical loss, such as sensitive information posted publicly, mishandled or sent to the wrong party via publishing online, sending in an email, sending in a mailing, or sending via fax.

linking table starting from June 30, 1986.[14] PERMCO and RSSD are the unique ID for CRSP companies and banks, respectively. Our final dataset has 8,760 bank-quarter observations. The observations also include control banks or banks without a data breach. We use propensity score matching to match between the banks with and without a data breach announcement. We explain in more detail in the method section about our propensity score matching procedure.

We create bank variables in the same spirit as Acharya and Mora (2015). The description of variables is provided in Appendix A. Our sample period is quarterly from 2005 to 2018. We started the sample in 2005 which is the beginning year of PRC data. Appendix C provides specific details on how we construct banking variables.[15] Bank-level variables are from the bank's quarterly Call Reports. We merge the banking data at the bank and the bank holding level. For the bank holding, we aggregate banks under the same holding company to the top holder and we treat it as a single banking organization. In this paper, "banks" refer to banking organizations and individual banks. The standard errors in our analysis are clustered at the banking organization and individual bank levels. To remove the merger effect in the banking industry, we exclude samples with quarterly growth of total assets larger than 10 percent. All the growth rates are computed from the Call Reports and winsorized at the 1 percent tails. Our regression specification includes fixed effects for banks, bank district, and time (i.e., dummies of the quarter).

Our main dependent variables are the growths of deposits and loans. For the deposits, we create five different deposit accounts: total, core, insured, brokered, and time. For loans, there are three different types of loans: total, commercial and industrial (CI), and credit (loan and unused commitment). The control variables include other liquidity demand and bank solvency. A bank's exposure to liquidity demand is proxied by a bank's unused commitments ratio. The unused commitments ratio is computed as the ratio of unused loan commitments

[14] https://www.newyorkfed.org/research/banking_research/datasets.html

[15] We use a modified code provided by Professor Schnabl http://pages.stern.nyu.edu/ pschnabl/data/data_callreport.htm

to the sum of loans and unused commitments. The parts of the credit lines that have not been drawn down are unused commitments. We need to control for other liquidity demand because it could affect the level of deposit at banks when the demand is transferred from an off-balance sheet to an on-balance sheet. Other control variables related to a bank's liquidity and solvency are net wholesale funding, nonperforming loans (NPL), capital, real estate exposure, and size. Net wholesale funding is the liabilities net core deposit and liquid assets. The wholesale borrowing includes gross federal fund bought net gross federal fund sold and repos net reverse repos. Non-performing loans are loans that are past due for 90 days and nonaccruing. The capital ratio is the ratio of book capital to assets. Real estate exposure is controlled by loans backed by real estate to total loan outstanding. Table 1 shows summary statistics for the variables used in our analysis both for controls and dependent variables. Banks in our sample are healthy based on the bank capital ratio of 11.9 percent. A 0.7 percent mean of the net wholesale funding ratio indicates that the larger part of bank operations is backed by liquid funds. The proportion of NPL to loans is 1.7 percent whereas the proportion of real estate loans to loans is 48 percent. On average, core and insured deposits have a higher growth quarter to quarter than other types of deposits such as brokered deposits. Time deposits are the only type of deposit with a negative growth mean. Overall, loan growth has a similar growth rate as deposits.

[**Insert Table 1 near here**]

Table 2 shows summary statistics of the data breach announcement. Panel A shows the number of data breach announcements for each year. From 2005 to 2018, there are 87 breach announcements in total. The highest number of data breach announcements is in 2010 which has 12 data breach announcements. Panel B shows summary statistics of data breach types. The top two reasons for a data breach in banks are Portable Devices and Insider. Portable Devices reason is lost, discarded or stolen laptop, USB, hard drive, or any devices designed to be movable. The Insider reason is when an employee, contractor, or customer intentionally takes the data out or sells the data to a third party. Payment Card Fraud, Unintended Disclosure, and Hacking or Malware are tied for the third reason of data breach.[16] Panel C presents the summary statistics of data breach records lost. Of 87 data breach events, 44 events are known lost records while 43 events are unknown lost records. Known-record events mean banks know how many client records were lost, such as Social Security numbers, addresses, etc. On average, the records lost are 898,489 with the highest number of records of 17,000,000. Based on the median records lost of 6,000, the average of records lost is highly right skewed. The loss of 17,000,000 records was from Countrywide Financial Corp on August 2, 2008. A former employee of the company stole and sold sensitive personal information to outside parties.

[Insert Table 2 near here]

4. Testing Methods

Our first set of analyses is how a data breach affects deposit growths. We have two objectives to achieve. Our first goal is to compare the operations and performance of banks with and without a data breach. Second, we investigate when an effect starts after a data breach announcement, and whether the effect has any long-term impact. To test our hypotheses, we adapt the difference-in-difference (DID) model to handle different event dates. We create an interaction effect dummy to handle the event date in Equation (1) below.

$$Y_{i,t} = \alpha_t + c_i + k_d + \sum_{q=1}^{5} \beta_q D_{i,t+q} + \theta X_{i,t} + \epsilon_{i,t} \tag{1}$$

[16] For more explanation of other reasons, Payment Card Fraud is fraud relating to debit and credit cards such as skimming devices at point-of-service terminals. Unintended Disclosure is when the data were disclosed unintentionally. For example, sending private information to a wrong party falls into this category. Hacking or malware reason is when banks are hacked by an outside party or infected by malware. Unknown is when banks do not know the reason. Physical Loss is paper documents that are lost, discarded or stolen. Stationary Computer Loss is when the data breach is from losing, inappropriately accessing, discarding, stealing computer or server not designed for moving

where $Y_{i,t}$ is a dependent variable based on our hypotheses such as deposit and loan growths for bank i at quarter t, α_t is a time fixed effect, c_i is a bank fixed effect, k_d is a bank district fixed effect, $D_{i,t}$ are dummies of quarters at and after the data breach announcement. We consider five dummies to capture the timing effect of data breach announcement (the announcement quarter and four more quarters after). After banks announce a data breach, clients may not know right away after the data breach has been announced. For example, if banks announce the breach via email and clients miss it, clients will not know about the breach for some time. Therefore, analyzing the subsequent quarters after the announcement quarter could capture the delayed effect. X is the vector of all controls.

To test the stock market reaction of data breach announcements, we compute Cumulative Abnormal Returns (CARs) over the window (0,+1d), (0,+5d), (-1d,+1d), (-2d,+1d), (-1d,+2d) and (-1d, +3d) based on the CRSP value-weighted return. The number in the front (back) indicates the number of days before (after) a data breach announcement. For example, (-1d,+3d) is the window between a day before and three days after the date of data breach announcement. In this case, zero means the event date when a data breach was announced. The longer term is also tested under the window from ten days prior to the announcement date and three quarters after the announcement date to examine if there is any long-term effect of a data breach announcement. Apart from the CAR analysis for different windows surrounding the date of a data breach announcement, we also examine CAR for each type of breach and different levels of breach records lost.

4.1. Propensity Score Matching

To reduce a confounding effect and lessen an issue of randomness of a data breach bank, we match banks with and without a data breach announcement based on all the observable characteristics using propensity score matching. The observable characteristics are the control variables that reflect the use and source of bank capital, leverage capacity, asset size, and business characteristics. We use probit one-to-one matching without replacement, because this method satisfies the parallel trend assumption the most among all the methods. It is noted that, for a difference-in-difference model, the goal is to match sample and control banks so that dependent variables have the same trend prior to a data breach announcement, though their levels are not required to be the same. The final sample consists of 8,760 bank quarter observations. Table 3 shows the mean and median of dependent variables and controls between banks with and without a data breach. Banks with and without a data breach have similar capital ratio. The net wholesale funding ratio is negative for banks without any data breaches, but it is positive for banks with a data breach. This indicates that banks without a data breach financed their operation with more liquid funding. Banks with a data breach have higher NPL relative to overall loan and higher unused commitment. The higher portion of loans is real estate loans for banks with no data breach. Most of the growth in deposits and loans are similar between banks with and without a data breach. Some exceptions are that banks with a data breach have higher saving deposit growth, time deposit growth, and lower loan growth.

[**Insert Table 3 near here**]

4.2. Parallel Trend Test

A key assumption of the difference-in-difference (DID) approach is, without the treatment, the average change in the dependent variables would have been the same for both treatment and control groups (parallel trend assumption). In our case, if there is no data breach announcement, the change in the dependent variables should be the same for both banks with and without the announcement. We test the parallel trend hypothesis following the Equation (2) on the time period before the data breach announcement.

$$Y_{i,t} = \alpha + \beta_1 Time_t + \beta_2 Treat_i + \beta_3 Time_t \times Treat_i + \epsilon_{i,t} \tag{2}$$

$Y_{i,t}$ is a dependent variable. $Time_t$ is a time fixed effect. $Treat_i$ is a dummy equal to one if a bank announced its data breach. The parallel trend assumption is satisfied if β_3 is not statistically different from zero. If β_3 is not different from zero, it means that the changes in dependent variables prior to a data breach announcement are not different between banks with and without a data breach announcement. Since we test the hypothesis for each data breach announcement date, the results in Table 4 show average t-statistic values for each coefficient. Based on the average t-values of the β_3 column, most dependent variables satisfy the parallel trend assumption. None of the β_3 of any dependent variables are significant at 1 percent level. We have tried different propensity score matching methods and the method that is satisfied the parallel trend assumption the most is a probit model one-to-one matching without replacement.

[**Insert Table 4 near here**]

5. Empirical Results

5.1. ***Banks deposit flows after data-breach announcements.***

As we described in the hypothesis development section, a data breach might or might not affect the deposit growth of a bank. On the one hand, if the switching cost is high or clients trust that banks can handle the situation well, we would not see much effect of the data breach announcement on the deposit flows. On the other hand, the effect of a data breach announcement could affect the overall reputation and operation of a bank and result in deposit outflow after a breach. Columns (1) to (5) of Table 5 show different types of deposit accounts.

Our main explainable variables are the deposit growth rates in the announcement quarter (Qevent) and four subsequent quarters (Q1 after to Q4 after). Insured deposits show significant outflow of 0.5 percent at the data breach announcement quarter. Insured deposits are deposit amounts no more than \$100,000 before 2009Q3. After 2009Q3, the insured amount increased to \$250,000.[17] Since insured deposits are deposit accounts with a known limited amount of no more than \$250,000 (\$100,000 before 2009Q3), the decrease in insured deposits implies that depositors who have a deposit amount that does not exceed this cap are more sensitive to the data breach announcement. In addition, the result of core deposit growth shows that the deposit accounts with an amount less than \$100,000 are not sensitive to the data breach announcement. The results of both unchanged core deposit growth and an decrease in insured deposit growth imply that the deposit accounts with the amount between \$100,000 to \$250,000 are sensitive to a data breach announcement. Deposit insurance literature provides evidence that depositors are less sensitive to bank risk when deposits are insured by Federal Deposit Insurance Corporation (FDIC). Park and Peristiani (1998) show a positive relation between the probability of bank failure and the subsequent outflow of uninsured deposits. Moreover, Karas, Pyle, and Schoors (2013) find that depositors are less sensitive to bank risk reducing market discipline by depositors. However, it is noted that the FDIC insures deposits only when a bank fails. If depositors lose their money owing to a data breach, and the bank does not fail, the FDIC will not be responsible for the loss. In addition, the FDIC does not provide deposit insurance for deposit loss from identity theft. This could be a reason why depositors who are potentially more sensitive to banks' risk withdraw their funds after a data breach announcement. The FDIC states that unauthorized access to deposits can be covered by the Electronic Funds Transfer Act (EFT Act) and other consumer protections.[18] However, the EFT Act requires depositors to report an unauthorized electronic fund transfer that appears on a periodic statement within 60 days; otherwise, the depositors could be liable to the losses.[19] Though deposits could be covered by the EFT Act, depositors need to monitor their accounts themselves. Unlike the EFT Act, FDIC deposit insurance will automatically cover deposits when a bank fails without additional effort from the depositors to monitor their own accounts. As a result, it is possible

[17] The insured amount increased to \$250,000 in 2006Q2 for retirement account.

[18] https://www.fdic.gov/deposit/covered/notinsured.html

[19] FDIC Law, Regulations, Related Acts- 6500-ConsumerFinancial Protection Bureau-Part 1005 Electronic Fund Transfers (Regulation E) §1005.6 https://www.fdic.gov/regulations/laws/rules/

that insured depositors try to avoid the inconvenience of the EFT Act reporting requirement by moving their deposits to another bank with no data breach history.

Therefore, our findings are potentially not related to banks' likelihood of failure since their deposits are insured in that case. Our finding is more in line with the precautionary action to prevent losses from their information being exploited and depositors losing trust in the banks in handling their private information. In Table 5, we do not see a long-term effect after a data breach announcement. A sign of negative insured deposit growth only appears at the data breach announcement quarter without any subsequent impacts in later quarters.

We also find a 0.4 percent decrease in the brokered deposit growth almost a year after a data breach announcement. The delayed outflow of the brokered deposits could be explained by the types of inflow that created brokered deposits. Brokered deposits can be divided into Primary Purpose (PP) and Primary Purpose Exception (PPE). Not all third-party deposits are considered brokered deposits.[20] The Primary Purpose indicates that if the goal of the third-party deposits is to facilitate the deposit flows for the purpose of deposit, the deposit is considered a brokered deposit. A large proportion of Primary Purpose deposits are in the form of Certificate Deposits (CDs) which require customers to deposit their money for a certain time period. Typical CDs last between 12 to 36 months. In the case of Primary Purpose, some depositors may not want to lose the interest income from an early withdrawal though they may like to move to another bank after a data breach announcement. Primary Purpose Exception states that some deposits from the third-party may also qualify to be brokered deposits though their sole purpose is not to facilitate the deposit flow for customers. However, they need to satisfy either the rule of 1) No more than 25 percent of the assets are deposited 2) 100 percent of assets are in transactional accounts that have no interest or fees paid to depositors. For the Primary Purpose Exception, customers deposit their money through a third-party and may not even know which bank has their money. Unless the third party itself has a data breach, customers may not withdraw the money and, consequently,

[20]https://www.fdic.gov/news/board-matters/2020/2020-12-15-notice-dis-a-fr.pdf

there would be no brokered deposit decrease in banks with a data breach.

[Insert Table 5 near here]

5.2. ***Effects of unknown lost records of data breaches***

FDIC requires banks to notify clients impacted from a data breach.[21] If the affected clients are unknown, banks may need to notify a wider scope of clients to satisfy the FDIC requirement. Unknown lost records of a data breach may have a different effect from known lost records in this case, because more clients are notified than banks would have notified otherwise. Data breach banks that cannot define the scope of record losses may create an impression of recklessness and unorganized. Thus, clients might move their account somewhere else to make sure their assets are safe in case their stolen identity is used to access their accounts. Table 6 shows the effects of unknown lost records. The insured deposit growth reduces by 1 percent in the data breach announcement quarter. Compared to the main results, data breach announcements with unknown records lost have larger impacts in terms of the magnitude. The case of unknown records lost shows a two-times stronger deposit outflow compared to the main results. The main results show the outflow of 0.5 percent while the unknown records lost results show 1 percent outflow. Furthermore, the time deposit shows an outflow a quarter after the data breach announcement quarter. In the main results, a data breach announcement has no effect on the time deposit account. One could argue that the time deposit account has a higher switching cost. If depositors withdraw their money early from their time deposit accounts, they may lose accrued interest income. For the case of unknown lost records, the results indicate that the safety benefit of moving their account to a bank with no data breach outweighs the switching cost. Overall, for the case of unknown records lost, the impact of a data breach announcement on deposit flows seems to be more severe consistent with our hypothesis that clients may perceive a

[21] https://www.fdic.gov/news/financial-institution-letters/2005/fil2705.html

data breach with unknown lost records as a worse scenario than the case of a data breach with known lost records.

[Insert Table 6 near here]

5.3. ***Data breach announcement impact on bank stock returns***

Panel A of Table 7 shows that data breach announcements have a negative impact on banks' stock returns, which is consistent with our hypothesis. CARs are significantly negative within different windows, (0,+1d), (0,+5d), (-1d,+1d), (-2d,+1d). Among them, the CAR of (0,+5d) has the lowest value on average, -0.83%. The CAR within (-1d, +1d) is -0.4%, which has a smaller loss than -0.8% from Kamiya et al. (2021), possibly because Kamiya et al. (2021) examine attacks involving the loss of personal financial information only. However, we find a larger loss within (-1d, +1d) than the CAR, -0.31%, of the stocks taken advantage of by short sellers (Wang et al., 2022). Comparing with the CARs within (-1d, +1d), -0.1%, of all breaches in Lending et al. (2018), we find much larger losses, -0.4%, in the finance industry. Among several types of data breaches, Hacking or malware (HACK) has a more severe impact on the stock market, lower than -0.9% of CARs over the window (0,+1d) and window (-2d,+1d). This is consistent with the findings of Lending et al. (2018) that HACK induces larger losses in the stock market than other types of breaches. Large negative CARs also appear when the cause of a data breach is from losing physical documents (PHYS).

Panel B of Table 7 provides the CARs for different quantiles of total records. We divide financial firms into four groups from Q1 (highest number of breach records) to Q4 (lowest number of breach records). Unknowns are the banks with an unknown number of breach records. Table 7 shows that Q1 has the significantly largest loss of stock returns, which are, -1.76% over (0,+1d), -4.55% over (0,+5d) and -1.66% over (-1d,+1d). Overall, the losses are decreasing from Quantile 1 to Quantile 4 although most CARs are not significant in groups Q3 to Q4. The CARs of Unknown firms range from -0.3% to -0.46% yet they are not statistically significant. The results confirm our expectation that banks with a higher number of lost records experience a larger drop in CARs than banks with a lower number of lost records.

To eliminate the possibility that the data limitation may bias the overall results, we present CARs of the banks with available call reports and stock data in Panel C of Table 7. The results of banks with available call reports and stock data are qualitatively the same. Compared to all banks with data breach, banks in our sample after merging all the data have slightly smaller losses (-0.33% vs. -0.49%) over (0,+1) and slightly larger losses (-1.26% vs. -0.83%) over (0,+5). The reasons are that the banks with missing call reports and stock data are either relatively small banks or the data breach incidents have unknown records. Panel C also reports the CARs of large banks only (top five largest banks in a given quarter). Large banks experienced a loss of -0.39% over (0,+1), but did not retain significant losses 5 days after the data announcements. The insignificant losses over (0,+5) could be from investors' confidence in the large banks and large banks' quick and responsible reactions to data breaches. For example, Capital One disclosed a data breach on July 19, 2019. They immediately fixed the issue and promptly began working with federal law enforcement. The following clear reports are publicly available about who is responsible for the data breach incident, how the incident impacts customers, what Capital One did to protect clients, etc.[22]

[**Insert Table 7 near here**]

We also investigate the long-term effect of data breach announcements as shown in Figure 1. Consistent with the findings in the short-term effect, the CAR is continuously decreasing to around -5% through the period from the announcement date to three quarters post announcement. We acknowledge that the results can be noisy when we test for a longer period.

[**Insert Figure 1 near here**]

[22] https://www.capitalone.com/digital/facts2019/faq/

Overall, we find that a data breach announcement produces negative CARs surrounding the announcement date. Compared to all banks, large banks have lesser impact 5 days after the data breach announcement date. Hacking and physical document loss is the type of breaches with the most severe losses. In addition, data breaches are likely to have a long-term destructive impact on stock value.

5.4. ***Data breach effect on bank loan issuance***

Table 8 shows that the credit activities increase in a quarter after a data breach announcement. Credit in this case is the loans and unused commitments, such as home equity and credit card lines. Increases in Credit imply that banks issue more loans after a breach announcement. The result is consistent with Allen and Saunders (1992) that banks try to offset the bad reputation perceived by the public after a data breach announcement by issuing more loans in order to have better financial performance. Since the main revenue of banks is from net interest margin (NIM) or the difference between lending interest revenue and deposit interest expenses, increasing the lending volume should increase the banks' bottom line. Consistently, Column (4) of Table 8, shows higher NIM after the higher loan growth.

[Insert Table 8 near here]

Consistent with stronger incentive to take risk (DeYoung, Peng, and Yan, 2013), large banks may lend more after a data breach announcement than small banks. We also investigate further for large bank samples. To capture the effect of a data breach announcement on large banks, we use difference-in-difference-in-differences (DDD) as provided in Equation (3).

$$Y_{i,t} = \alpha_t + c_i + k_d + l_{i,t} + \sum_{q=1}^{5} \alpha_q D_{i,t+q} + \sum_{q=1}^{5} \beta_q D_{i,t+q} l_{i,t} + \theta X_{i,t} + \epsilon_{i,t} \tag{3}$$

$l_{i,t}$ is a dummy equal to one when the banks' asset size is in the top five largest banks in a given quarter, and other variables have the same definition as Equation (1). For all specifications, we control for the quarter, bank, and bank district fixed effects. The standard errors are clustered at the individual bank and bank holding levels.

Table 9 presents the impact of data breach announcement on large bank loans. It shows that the loan activities increase two quarters after the data breach announcement. Total loan increases 1.5 percent while commercial and industrial (CI) loan increases 0.4 percent two quarters after the data breach announcement quarter. The results are consistent with the findings in Table 8 that banks with a data breach potentially try to offset their bad reputation from a data breach with better future financial performance.

[Insert Table 9 near here]

We analyze further to understand the catalyst of higher lending after a data breach announcement. We examine the conflict of values between bank managers and shareholders stated in Allen and Saunders (1992). If bank managers' compensations tie to the stock value, bank managers might try to increase lending to improve company financial performance which affects the stock value. We apply a CEO's compensation proxy used in Ho et al. (2016) to examine whether CEOs of data breach banks have higher proportion of their compensation tie to bank stock value. Ho et al. (2016) use CEO's option moneyness to gauge CEO over confidence.[23] They find that banks with over confident CEO tend to lower lending standard and increase leverage before financial crises. From Table 10, we find that CEOs of banks with data breach have higher deep in-the-money options than non-data breach banks. The finding is consistent with higher lending after a data breach announcement, potentially, to preserve wealth for bank managers.

[Insert Table 10 near here]

[23]The option moneyness is based on the estimated strike price and per option realizable value following Core and Guay (2002) and Campbell, Gallmeyer, Johnson, Rutherford, and Stanley (2011). To find the estimated strike price, first, find total realizable value per share for exercisable option using data from ExecuComp (OPT_UNEX_EXER_EST_VAL/OPT_UNEX_EXER_NUM). Then, subtract the total realizable value per share for exercisable option from the stock price at the fiscal year end (PRCCF). The option moneyness is the per option realizable value divided by the estimated strike price.

5.5. ***Deposit flows between data-breach and non-data-breach banks***

Our main results show that banks with a data breach experience insured deposit outflow. An interesting question is whether those outflows become the inflows of banks with no data breach. We hypothesize that depositors would withdraw money from banks with a data breach and deposit it to a nearby bank without a data breach. Our hypothesis is based on the following ground. If depositors withdraw money from a data breach bank, they will put their money in an equivalent means of investment since the withdrawal would be considered an unplanned withdrawal. If depositors put money in a money market mutual fund or hold it as cash, there will not be any relationship between deposit growth between banks with and without a data breach. To test this conjecture, we follow Equation (4) below.

$$Y_{i,t} = \alpha_t + c_i + k_d + \beta_b DepositGrowth_{b,t} + \sum_{q=1}^{5} \alpha_q D_{i,t+q} + \sum_{q=1}^{5} \beta_q D_{i,t+q} DepositGrowth_{b,t} + \theta X_{i,t} + \epsilon_{i,t} \tag{4}$$

where $Y_{i,t}$ is insured deposit growth of no data-breach bank i at quarter t, α_t is a time fixed effect, c_i is a bank fixed effect, k_d is a bank district fixed effect. $DepositGrowth_{b,t}$ is the insured deposit growth of banks b with a data breach at quarter t. Since we assume an immediate withdrawal from a data breach bank and depositing into a no data breach bank, we match the quarter of deposit growths between banks with and without a data breach announcement. Therefore, quarter t of $Y_{i,t}$ always matches quarter $t+q$ of $D_{i,t+q}$. X is a vector of all controls. $D_{i,t}$ are dummies of quarters at and after the data breach announcement. A negative β_q is consistent with depositors withdrawing money from banks with a data breach and depositing it to banks without a data breach. To test Equation (4), we first match banks with and without a data breach based on their zip code. It is noted that we use all available non-breached banks in this step which is different from our main results that only use matched non-breached banks from the propensity score matching. We use all available non-breached banks in this matching step, because we assume that clients can move their deposits to any bank. The zip code (RSSD9220) and state (RSSD9200) information are retrieved from the Call Report. Then, we calculate the distance between the two banks in miles. We have three distance matching categories: less than 5 miles, less than 10 miles, and between 10 to 30 miles. It is less probable that depositors would withdraw money and redeposit it somewhere far away from the location of the breached bank in which the clients use, because it would be inconvenient to perform financial transactions in case clients need an in-person service. Another possibility is that, under a technological advancement, depositors may open an account online and transfer via an online channel without visiting the physical branch. In this case, the zip code of the newly opened bank account can be the bank headquarters, which could be far away from the bank with a data breach. If the latter argument is true, we should not find a significant relationship between the deposit growth of banks with and without a data breach that are located near each other. We also evaluate the hypothesis for banks within the same state. The motivation other than the distance for the state matching is that banks are regulated at both the federal and state levels.[24] Table 11 shows the results of the exchanged flow analysis between deposit growth of banks with and without a data breach. We find evidence that deposits outflow from the banks with a data breach to the banks without a data breach. For banks with and without a data breach located between 10 to 30 miles, we find a negative coefficient at the announcement quarter indicating that there is evidence of deposit flows between banks with and without a data breach announcement. A one percent drop in insured deposit growth in the data breach announcement quarter results in 0.347 percent increase in the insured deposit growth of banks without a data breach announcement that is located between 10 to 30 miles away from the breach bank. We also find evidence of the deposit flow between banks with and without a data breach announcement for other distance categories. For the category of less than 5 miles, the exchanged flow occurs later in the third quarter after a data breach announcement whereas, for the less than 10 miles, the exchanged flow occurs

[24] https://www.frbsf.org/education/publications/doctor-econ/2006/november/commercial-banks-regulation/

a quarter after a data breach announcement. For the same state category, we find strong evidence of exchanged flow a quarter after a data breach announcement.

[Insert Table 11 near here]

5.6. ***Multiple data breach effects***

Some banks have more than one data breach announcement. In this section, we analyze if responses from bank clients are different when banks have multiple breach announcements. On the support of stronger effect, bank clients may perceive a bank with multiple data breaches as having an unsafe security system. Consequently, clients could move their accounts away from a bank with multiple breaches. On the other hand, clients may learn from the previous breaches that their accounts are still safe and may decide to maintain the service with their bank. For the latter case, we may find less impact on the deposit flows for each announcement. Our approach is to first create a dummy variable for each data breach announcement. For example, on the first data breach announcement, we assign the first-breach dummy equal to one and zero otherwise. Then, on the second data breach announcement, the second-breach dummy is assigned a value of one and zero otherwise, and so on. Then, we interact the number of data breach announcement dummies with the announcement quarter and four quarters after. We perform the analysis on the insured deposit account, because we significantly find insured deposit outflow in our main results. In the results presented in Table 12, each column represents each announcement order. For example, column (1) is the results for the first announcement dummy. In this case, *BreachNumber* is equal to one when the announcement is the first time a data breach occurred. Table 12 shows that, after the fourth announcement, the results show no outflow from the insured deposit account. The last significant negative coefficient is three quarters after the fourth data breach announcement. The results imply that clients may feel it unnecessary to move their accounts away after a couple of data breaches, since their accounts were handled well by their banks based on the previous breaches.

[Insert Table 12 near here]

5.7. ***Robustness***

5.7.1. ***Breach disclosure law***

Romanosky, Telang, and Acquisti (2011) examine how disclosure laws on identity theft during 2002 to 2007 affect the number of identity thefts. They find that the law had marginal effect on the number of identity thefts. Between 2002 to 2007, many US states adopted data breach disclosure laws. The first state that adopted the laws was California in 2003 and no other states adopted the laws until 2005. Our data starts in 2005 which overlaps with the disclosure law adoption period. The overlapping period could impact our results, because some entities with data breach in some states were required to disclose their data breach incident while some other states were not. During such periods, our results could partly compare between the data breach entities that were required and were not required to disclose a data breach incident. Hence, we drop the year 2005 to 2007 to see if our results are still robust without the two-year period of disclosure law adoption. The results in Table 13 are qualitatively the same as our main results in Table 5.

[Insert Table 13 near here]

5.7.2. ***Spillover effect***

An earlier data breach announcement could create a psychological impact on clients overall confidence in their bank, even though their banks are not the one which announces the data breach incident (spillover effect). For example, when clients hear about a data breach in another bank and are afraid that their banks may also have a data breach in the future, they may decide to hold cash or move their deposits somewhere. Our results partly could be from the spillover effect of the earlier data breach announcement from another bank. To address the omitted variable concern, we control the earlier breach announcement effect by creating a cumulative number of data breach announcements as a control variable. Specifically, we create *Cumulative Breach* variable which is equal to one for the first breach announcement in our data. Then, in chronological order, the next announcement is assigned number two and so on. We add the *Cumulative Breach* variable as part of the control variables and rerun the regression. Table 14 shows that the results are qualitatively similar to the main results.

[**Insert Table 14 near here**]

5.7.3. ***Non-random shock***

One could argue that banks may be targeted by criminals in a non-random fashion. For example, large banks could be targeted by hackers more than small banks, because the benefit from breaching large banks is higher. However, hacking as a cause of a data breach in our sample is about 15 percent of the total. For other causes, such as insider, payment card fraud, and unintended disclosure, a counter argument would be large banks could have a better protocol to prevent a data breach from these types of breach than small banks could have. For instance, it could be easier for employees in smaller banks to take clients' information outside and sell it to a third party. The argument that some banks could be targeted more than other banks could be true when the cause is hacking. However, for other causes which are the majority of the data breach causes in our sample, the argument that some banks are prone to have more data breach is debatable. However, in the main results, we control for the large bank effect with large bank indicator as our control to lessen the issue of non-random target based on the hacking as the cause of breach. Large bank indicator is the dummy equal to one if a bank's asset size is the top five largest asset size in each quarter. In our propensity score matching, we also include asset size as our matching variable. Therefore, the issue could be mitigated by the size controls both as a control variable in the regressions and as a matching variable.

6. Conclusion

Data breaches could negatively affect banks' operations and performance when they announce that their customers' information such as their Social Security and account numbers are compromised. Customers may move their accounts (e.g., deposits) to another bank with no previous history of a data breach, because they lose confidence in the security system of the bank or to prevent their accounts from being exploited. Meanwhile, customers may not move their accounts elsewhere because of a high switching cost or trust that banks can handle the issue well. A consequence after a data breach is an important question for the banking industry and the economy, because the banking industry is a large part of our economy as seen in the impact from the subprime financial crisis. To evaluate the effect of a data breach announcement on banks' operations and stock performance, we perform a difference-in-difference approach and an event study, respectively. We find a sign of insured deposit outflows a quarter after a data breach announcement and brokered deposit outflows a year after. A typical US bank borrows short-term and lends long-term. Without stable deposits, banks may have a liquidity issue. Bank clients stop withdrawing money from their account after a data breach announcement when banks have multiple data breaches. Clients may learn from an earlier breach that their accounts are still handled well by their banks and see no reason to move their money away. The results are consistent with the real world that, so far, we have not seen any bank fails after a data breach and banks usually compensate for unauthorized transactions.

Banks may try to offset bad news from a data breach with better financial performance in a later period by reducing costs and increasing revenue. Supporting our argument that banks try to increase their revenue after a data breach announcement, we see evidence of banks increasing their lending amount after a data breach announcement. To support our argument, we show that CEOs of banks with data breach have significantly higher deep in-the-money options which give them incentives to maintain financial performance. Time deposits decrease when the breach lost records are unknown. The decrease in time deposits implies a stronger effect of a data breach on depositors' confidence, because time deposits have a higher switching cost compared to other types of deposit. We also find evidence that the deposits after a data breach announcement flow to near-by banks with no data breach. We explain that depositors would choose to open a new account with another bank close to the initial bank with the data breach for their convenience.

Our main findings show that data breach announcements affect banks' operations. However, so far, the effects have been short-term. Our paper is the first step to understand how a data breach could impact bank operations and performance. The contribution of our paper is to understand more about the new bank risk factors in the technology era such as data breaches and cybersecurity. In the future, if banks do not keep up with the new technology and security system, a data breach may trigger a liquidity problem and affect economic welfare. Areas of future research could be, first, how changes in banking activities, such as decreases in insured deposits, affect economic growth or different economic units' financial decisions. Second, one could further examine bank strategic actions in different situations. For example, some banks may expect a larger impact from a data breach announcement and decide to deploy strategic actions similar to what we find evidence in our paper that banks try to increase their lending after the announcement.

Fig. 1. Cumulative Abnormal Return: Mean & 95% Confidence Limits. This Figure provides CARs from pre-10 day to post-270 days relative to the data-breach announcement date. The estimation model is based on the CRSP value-weighted return.

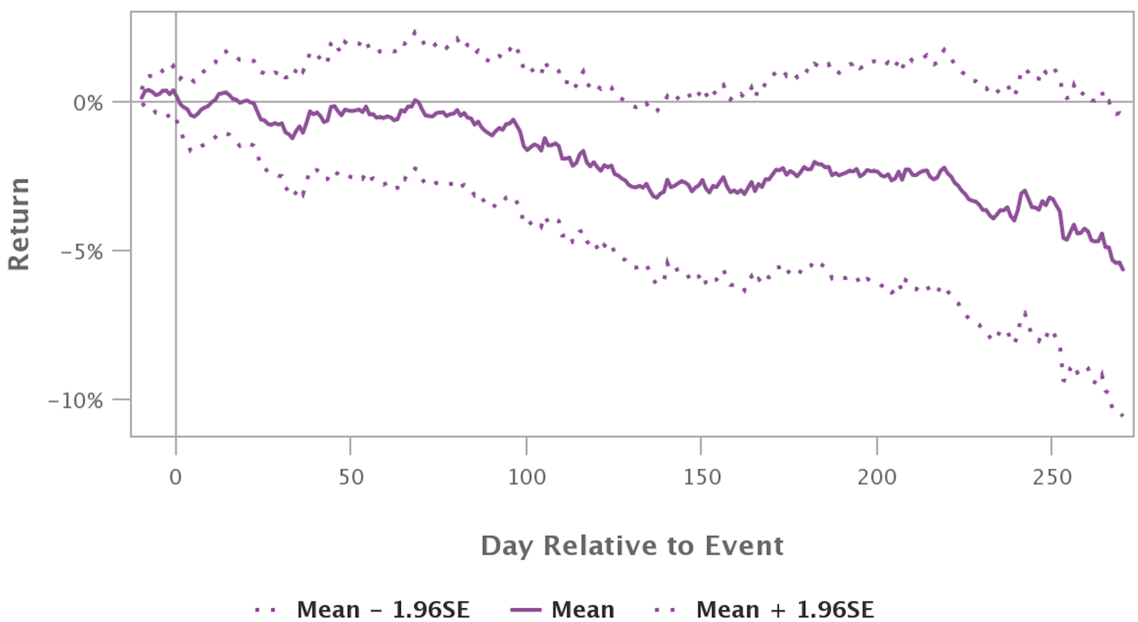

Table 1: Summary Statistics. This table provides the mean, standard deviation, 25 percentile, median, 75 percentile and N of control variables and dependent variables. The bank-level variables are quarterly data from 2005 to 2018. The banking data is obtained from Bank Regulatory (call report). The full description of each variable can be found in Appendix A.

Variables	Mean	StdDev	P25	Median	P75	N
Capital ratio (book capital to assets)	0.119	0.040	0.097	0.111	0.129	8,760
Net wholesale funding ratio (wholesale -liquid)	0.007	0.191	-0.096	-0.003	0.139	8,760
Nonperforming loans to loans	0.017	0.021	0.003	0.010	0.022	8,757
Real estate loan share	0.483	0.248	0.319	0.479	0.686	8,757
Unused commitment ratio	0.186	0.152	0.019	0.180	0.365	8,756
Quarterly growth of deposits	0.011	0.035	-0.007	0.009	0.027	8,586
Qaurterly growth of core deposits	0.009	0.031	-0.006	0.007	0.022	8,586
Quarterly growth of insured deposits	0.007	0.030	-0.004	0.002	0.011	8,586
Qaurterly growth of brokered deposits	0.001	0.013	-0.001	0.000	0.002	8,586
Quarterly growth of time deposits	-0.001	0.023	-0.004	-0.001	0.003	8,586
Qaurterly growth of loans	0.009	0.029	-0.005	0.006	0.021	8,586
Quarterly growth of C&I loans	0.002	0.010	-0.001	0.001	0.005	6,918
Quarterly growth of credit (loans+commitments)	0.007	0.036	-0.004	0.001	0.022	8,586
Qaurterly growth of net interest margin	0.001	0.014	0.000	0.007	0.009	8,586

Table 2: Panel A shows the number of data breach announcements for each year from 2005 to 2018. Panel B shows the number of data breach types. Panel C shows summary statistics of data breach records lost. Data breach data are retrieved from Privacy Rights Clearinghouse, https://privacyrights.org/data-breaches

Panel A: Data-Breach-Announcement Year

Year	Freq.	Percent
2005	9	10.34
2006	11	12.64
2007	5	5.75
2008	6	6.90
2009	3	3.45
2010	12	13.79
2011	9	10.34
2012	8	9.20
2013	7	8.05
2014	6	6.90
2015	2	2.30
2017	2	2.30
2018	7	8.05

Panel B: Type of Data Breach

Type of Breach	Freq.	Percent
Portable Devices	17	19.54
Insider	16	18.39
Payment Card Fraud	13	14.94
Unintended Dsiclosure	13	14.94
Hacking or Malware	13	14.94
Unknown	9	10.34
Physical Loss	4	4.60
Stationary Computer Loss	2	2.30

Panel C: Summary Statistics of Data Breach Records Lost

Mean	Std	P25	Median	P75	Min	Max	N
898,489	3,164,006	332	6,000	118,000	0	17,000,000	44

Table 3: This table shows the mean and median of bank variables with and without a data breach. First, banks with a data breach from Privacy Right Clearing House (PRC) are merged with the CRSP-FED linking table. CRSP-FED linking table links the bank entity and CRSP-PERMCO. Then, we match banks with a data breach with banks without a data breach using propensity score matching (PSM). We use the one-to-one PSM method of probit matching on all control variables without replacement. Banking data is from the Fed Call Report. Please see Appendix C for instructions on how to construct each variable.

	Banks without Data Breach		Banks with Data Breach	
	Mean	Median	Mean	Median
Net wholesale funding ratio (wholesale -liquid)	-0.042	-0.043	0.063	0.051
Nonperforming loans to loans	0.013	0.006	0.022	0.014
Unused commitment ratio	0.120	0.055	0.262	0.365
Real estate loan share	0.497	0.487	0.468	0.460
Quarterly growth of deposits	0.011	0.009	0.011	0.009
Qaurterly growth of core deposits	0.009	0.007	0.010	0.007
Quarterly growth of insured deposits	0.007	0.002	0.007	0.002
Qaurterly growth of brokered deposits	0.001	0.000	0.001	0.000
Quarterly growth of transaction deposits	0.003	0.002	0.002	0.001
Quarterly growth of saving deposits	0.006	0.003	0.008	0.005
Quarterly growth of time deposits	-0.001	-0.001	0.000	0.000
Qaurterly growth of loans	0.010	0.009	0.008	0.005
Quarterly growth of C&I loans	0.002	0.001	0.002	0.001
Quarterly growth of credit (loans+commitments)	0.008	0.005	0.005	0.000
Qaurterly growth of net interest margin	0.001	0.008	0.001	0.007

Table 4: This table shows the parallel trend test

$$Y_{i,t} = \alpha + \beta_1 Time_t + \beta_2 Treat_i + \beta_3 Time_t \times Treat_i + \epsilon_{i,t}$$

$Y_{i,t}$ is dependent variable. $Time_t$ is a time fixed effect. $Treat_i$ is a dummy equal to one if a bank announced its data breach. The parallel trend assumption is satisfied if β_3 is not statistically different from zero. We test the assumption for each data breach announcement date. The results in Table 4 show average t-statistic values for each coefficient. All dependent variables are quarterly growth. The banking data is from Call Report from 2005 to 2018. Please see the full definition for each variable in Appendix A. ***,**,* are significant at 1, 5, 10 percent, respectively.

Dependent Variables	α	$\beta 1$	$\beta 2$	$\beta 3$
Total deposit	7.09	1.51	1.29	1.54
Core deposit	3.84	2.26	1.15	0.95
Insured deposit	4.92	3.35	1.06	1.33
Brokered deposit	3.52	2.29	0.79	1.23
Time deposit	8.51	8.92	1.66	2.45**
Total loan	10.90	4.24	0.94	0.89
C&I loan	6.88	4.05	0.95	1.09
Credit (loan+commitment)	10.37	5.96	0.84	1.81*
Net interest margin	5.13	3.69	0.59	0.51

Table 5: The results show the data breach announcement effect on deposit growths. Dependent variables from column (1) to (5) are different types of quarterly deposit growth. *Deposits* quarterly growth of all types of deposits. *Core* is quarterly growth of transaction, saving, and time deposits less than $100,000. *Insured* is quarterly growth of insured deposits. *Brokered* is the quarterly growth of deposits from brokers. *Time* is the time deposit account. Please see the full definition for each type of deposit and other control variables in Appendix A. Qevent (Q1,2,3,4after) is a dummy variable equal to one if the interaction terms between a data breach bank dummy and a data breach announcement quarter (1,2,3,4 quarters after a data breach announcement quarter) dummy is equal to one. ***,**,* are significant at 1, 5, 10 percent, respectively.

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Deposits	Core	Insured	Brokered	Time
Qevent	-0.000	0.001	-0.005*	0.000	0.000
	(-0.13)	(0.37)	(-1.85)	(0.28)	(0.01)
Q1after	0.001	0.003	-0.001	-0.001	-0.002
	(0.32)	(1.00)	(-0.89)	(-0.57)	(-0.94)
Q2after	0.002	0.003	0.003	0.000	-0.001
	(1.00)	(1.17)	(0.82)	(0.40)	(-0.46)
Q3after	-0.003	-0.001	-0.002	-0.001	0.000
	(-1.16)	(-0.53)	(-0.91)	(-1.08)	(0.09)
Q4after	0.003	0.002	-0.002	-0.004**	-0.000
	(0.94)	(0.85)	(-0.69)	(-2.45)	(-0.14)
Unused Commitment Ratio	-0.008	-0.007	-0.006	-0.005	0.013*
	(-0.61)	(-0.63)	(-0.63)	(-0.62)	(1.75)
Net Wholesale Funding	-0.035***	-0.038***	-0.007*	-0.003	0.008
	(-3.62)	(-4.60)	(-1.66)	(-0.57)	(1.36)
NPL to Loans	-0.229***	-0.155***	-0.115***	-0.051***	-0.062**
	(-4.53)	(-2.84)	(-3.93)	(-2.63)	(-2.03)
Capital Ratio	-0.099*	-0.055	0.021	0.019	0.095***
	(-1.93)	(-1.35)	(0.53)	(1.57)	(3.37)
Large Bank Indicator	0.014***	0.011***	0.006***	0.001	0.009***
	(6.04)	(4.31)	(2.74)	(0.84)	(4.25)
Real Estate Loan Share	-0.011	-0.002	-0.011**	-0.001	0.004
	(-1.22)	(-0.25)	(-2.43)	(-0.14)	(1.17)
Constant	0.074***	0.043***	0.028***	0.005	-0.011**
	(5.68)	(4.59)	(3.20)	(0.86)	(-2.19)
Observations	8,585	8,585	8,585	8,585	8,585
R-squared	0.1343	0.1368	0.2617	0.0792	0.3779
Bank FE	YES	YES	YES	YES	YES
District FE	YES	YES	YES	YES	YES
Qtr FE	YES	YES	YES	YES	YES

Table 6: The results show the data breach announcement effects on deposit growths for a data breach with unknown breach records. Dependent variables from columns (1) to (5) are different types of quarterly deposit growth. *Deposits* quarterly growth of all types of deposits. *Core* is quarterly growth of transaction, saving, and time deposits less than \$100,000. *Insured* is quarterly growth of insured deposits. *Brokered* is the quarterly growth of deposits from brokers. *Time* is the time deposit account. Please see the full definition for each type of deposit and other control variables in Appendix A. Qevent (Q1,2,3,4after) is a dummy variable equal to one if the interaction terms between a data breach bank dummy and a data breach announcement quarter (1,2,3,4 quarters after a data breach announcement quarter) dummy is equal to one. ***,**,* are significant at 1, 5, 10 percent, respectively.

VARIABLES	(1) Deposits	(2) Core	(3) Insured	(4) Brokered	(5) Time
Qevent*Unknown	-0.006	0.001	-0.010**	-0.002	-0.003
	(-1.29)	(0.19)	(-2.22)	(-1.10)	(-0.86)
Q1after*Unknown	-0.011*	-0.007	-0.002	-0.003	-0.006**
	(-1.82)	(-1.32)	(-0.80)	(-1.01)	(-1.98)
Q2after*Unknown	-0.006	-0.003	-0.012	0.001	0.003
	(-0.82)	(-0.48)	(-1.52)	(0.36)	(0.55)
Q3after*Unknown	0.008	0.008*	0.008	0.003	-0.002
	(1.30)	(1.78)	(1.49)	(1.57)	(-0.39)
Q4after*Unknown	0.005	0.001	-0.009*	0.001	0.004
	(0.77)	(0.29)	(-1.87)	(0.29)	(0.82)
Constant	0.074***	0.043***	0.027***	0.005	-0.011**
	(5.74)	(4.59)	(3.40)	(0.86)	(-2.23)
Observations	8,585	8,585	8,585	8,585	8,585
R-squared	0.1348	0.1371	0.2626	0.0796	0.3782
Bank FE	YES	YES	YES	YES	YES
District FE	YES	YES	YES	YES	YES
Qtr FE	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES

Table 7: Cumulative Abnormal Returns (CARs) Results

Cumulative Abnormal Returns (CARs) following a data breach under the windows, (0,+1d), (0,+5d) and (-1,+1d). Panel A reports the CARs based on the CRSP value-weighted average returns for each data breach category. The categories of data breaches include CARD (Payment Card Fraud), HACK (Hacking or Malware), INSD (Insider), PHYS (Physical Loss), PORT (Portable Device), STAT (Stationary Device), DISC (Unintended Disclosure) and UNKN (Unknown). Panel B provides the Cumulative Abnormal Returns (CARs) based on quantiles of total records. The samples are divided into 4 groups based on the total records of data breaches. Q1 is the group with the highest breach records and Q4 is the group with the lowest breach records. Unknowns are the firms with an unknown number of breach records. Panel C shows the results for all available bank relative to the large bank samples. ***,**,* are significant at 1, 5, 10 percent, respectively.

Panel A: Full Sample

	(0,+1)			(0,+5)			(-1,+1)			(-2,+1)		
	Mean	Median	N	Mean	Median	N	Mean	Median	N	Mean	Median	N
Overall	-0.49%***	-0.28%	171	-0.83%*	-0.20%	171	-0.40%*	-0.16%	172	-0.52%**	-0.37%	172
	(-2.68)			(-1.95)			(-1.76)			(-2.06)		
CARD	-1.35%	-0.10%	14	-0.45%	0.42%	14	-1.32%	-0.12%	14	-1.64%	-0.41%	14
	(-1.26)			(-0.33)			(-1.14)			(-1.33)		
DISC	-0.19%	0.00%	27	-0.31%	-0.69%	27	-0.43%	-0.40%	27	0.04%	-0.06%	27
	(-0.47)			(-0.50)			(-1.01)			(0.09)		
HACK	-0.98%*	-0.37%	36	-2.35%	-0.18%	36	-0.97%	-0.70%	36	-1.33%*	-0.84%	36
	(-1.93)			(-1.40)			(-1.41)			(-1.90)		
INSD	0.35%	0.44%	23	0.21%	0.36%	23	0.40%	0.80%	23	0.40%	0.65%	23
	(1.08)			(0.28)			(0.85)			(1.01)		
PHYS	-0.89%	-0.18%	9	-1.09%	0.56%	9	-1.42%**	-0.26%	9	-1.51%**	-0.76%	9
	(-1.14)			(-1.09)			(-2.07)			(-2.16)		
PORT	-0.48%*	-0.39%	44	-0.72%	-0.35%	44	-0.28%	-0.36%	44	-0.48%	-0.67%	44
	(-1.89)			(-1.34)			(-1.07)			(-1.14)		
STAT	-0.67%**	-0.79%	4	-0.58%	-0.56%	4	1.19%	-0.06%	5	1.62%	0.39%	5
	(-2.05)			(-1.17)			(0.74)			(1.20)		
UNKN	-0.06%	-0.57%	14	-0.27%	-0.07%	14	0.42%	0.05%	14	-0.13%	-0.07%	14
	(-0.09)			(-0.48)			(0.42)			(-0.13)		

Panel B: Quantiles of Total Records

	(0,+1)			(0,+5)			(-1,+1)			(-2,+1)		
	Mean	Median	N	Mean	Median	N	Mean	Median	N	Mean	Median	N
Overall	-0.49%***	-0.28%	171	-0.83%*	-0.20%	171	-0.40%*	-0.16%	172	-0.52%**	-0.37%	172
	(-2.68)			(-1.95)			(-1.76)			(-2.06)		
Q1	-1.76%**	-0.10%	24	-4.55%*	-0.39%	24	-1.66%*	0.17%	24	-1.72%	-0.40%	24
	(-2.12)			(-1.93)			(-1.64)			(-1.54)		
Q2	-0.61%*	-0.37%	24	-1.14%*	-0.83%	24	-0.70%	-0.98%	24	-0.85%*	-1.04%	24
	(-1.67)			(-1.69)			(-1.55)			(-1.76)		
Q3	-0.29%	0.03%	25	0.75%	0.60%	25	0.06%	0.05%	25	0.13%	-0.23%	25
	(-0.61)			(0.75)			(0.10)			(0.21)		
Q4	0.27%	0.24%	24	0.00%	0.08%	24	0.66%	0.20%	24	0.16%	0.34%	24
	(0.72)			(-0.00)			(1.33)			(0.29)		
Unknown	-0.35%	-0.39%	74	-0.32%	0.13%	74	-0.40%	-0.34%	75	-0.46%	-0.48%	75
	(-1.62)			(-0.89)			(-1.38)			(-1.41)		

Panel C: Large Bank Sample

	(0,+1)			(0,+5)			(-1,+1)			(-2,+1)		
	Mean	Median	N	Mean	Median	N	Mean	Median	N	Mean	Median	N
All Banks	-0.33%**	-0.18%	91	-1.26%**	0.12%	91	-0.37%**	-0.13%	91	-0.36%*	-0.25%	91
	(-2.07)			(-2.13)			(-2.19)			(-1.85)		
Large Banks	-0.39%*	-0.18%	25	-0.30%	0.17%	25	-0.34%*	-0.47%	25	-0.38%	-0.28%	25
	(-1.86)			(-0.72)			(-1.84)			(-1.56)		

Table 8: The results show the data breach announcement's effect on loan growths. Columns (1) to (3) show results for different types of loan growth as dependent variables. *Loan* is quarterly growth of all types of loans. *CI* is the quarterly growth of commercial and industrial loans. *Credit* is quarterly growth of loans and commitments. Columns (4) is the result of the regression for Net Interest Margin (NIM) as dependent variables. Please see the full definition for each type of loan and other control variables in Appendix A. Qevent (Q1,2,3,4after) is a dummy variable equal to one if the interaction terms between a data breach bank dummy and a data breach announcement quarter (1,2,3,4 quarters after a data breach announcement quarter) dummy is equal to one. ***,**,* are significant at 1, 5, 10 percent, respectively.

	(1)	(2)	(3)	(4)
VARIABLES	Loans	CI	Credit	NIM
Qevent	-0.002	-0.000	0.002	-0.001
	(-0.92)	(-0.51)	(0.66)	(-0.88)
Q1after	0.003	-0.000	0.008**	-0.001
	(1.00)	(-0.34)	(2.39)	(-1.22)
Q2after	0.003	0.001	0.006	0.001*
	(0.88)	(1.14)	(1.38)	(1.86)
Q3after	-0.001	-0.000	0.000	0.000
	(-0.41)	(-0.06)	(0.20)	(0.37)
Q4after	0.004	-0.001	0.005	-0.001
	(1.11)	(-1.43)	(1.03)	(-1.58)
Unused Commitment Ratio	-0.000	-0.002	0.003	-0.001
	(-0.06)	(-0.89)	(0.30)	(-1.22)
Net Wholesale Funding	0.050***	0.010***	0.040***	0.001*
	(7.36)	(4.42)	(5.79)	(1.68)
NPL to Loans	-0.340***	-0.072***	-0.409***	-0.013***
	(-8.36)	(-4.46)	(-7.56)	(-3.73)
Capital Ratio	0.106**	0.034	0.145**	0.003
	(2.02)	(1.27)	(2.17)	(0.54)
Real Estate Loan Share	0.003	-0.011***	0.012	0.000
	(0.46)	(-3.91)	(1.47)	(0.44)
Constant	-0.004	0.004	-0.014	0.000
	(-0.34)	(0.97)	(-1.18)	(0.01)
Observations	8,585	6,917	8,585	8,585
R-squared	0.2088	0.1373	0.2907	0.8883
Bank FE	YES	YES	YES	YES
District FE	YES	YES	YES	YES
Qtr FE	YES	YES	YES	YES

Table 9: The results show the large bank data breach announcement effect on loan growths. Columns (1) to (3) show results for different types of loan growth as dependent variables. *Loan* is quarterly growth of all types of loans. *CI* is the quarterly growth of commercial and industrial loans. *Credit* is quarterly growth of loans and commitments. Columns (4) is the result of the regression for Net Interest Margin (NIM) as dependent variables. Please see the full definition for each type of loan and other control variables in Appendix A. Qevent (Q1,2,3,4after) is a dummy variable equal to one if the interaction terms between a data breach bank dummy and a data breach announcement quarter (1,2,3,4 quarters after a data breach announcement quarter) dummy is equal to one. *Largebank* is a dummy equal to one if banks' asset size is in the top five largest in a given quarter. ***,**,* are significant at 1, 5, 10 percent, respectively.

VARIABLES	(1) Loans	(2) CI	(3) Credit	(4) NIM
Qevent*Largebank	0.001	0.001	0.002	-0.000
	(0.23)	(0.83)	(0.28)	(-0.33)
Q1after*Largebank	-0.004	0.000	0.000	0.001
	(-0.79)	(0.22)	(0.06)	(0.75)
Q2after*Largebank	0.015***	0.004***	0.010	-0.000
	(2.91)	(2.67)	(1.28)	(-0.05)
Q3after*Largebank	-0.004	-0.000	-0.000	-0.001
	(-0.90)	(-0.06)	(-0.06)	(-0.85)
Q4after*Largebank	-0.010	0.002	-0.009	0.001
	(-1.56)	(1.04)	(-1.03)	(1.36)
Constant	-0.004	0.004	-0.014	0.000
	(-0.33)	(1.01)	(-1.15)	(0.02)
Observations	8,585	6,917	8,585	8,585
R-squared	0.2101	0.1387	0.2916	0.8884
Bank FE	YES	YES	YES	YES
District FE	YES	YES	YES	YES
Qtr FE	YES	YES	YES	YES
Controls	YES	YES	YES	YES

Table 10: Dependent variable is the option moneyness, *Moneyness*. The option moneyness is based on the estimated strike price and per option realizable value following Core and Guay (2002) and Campbell et al. (2011). To find the estimated strike price, first, find total realizable value per share for exercisable option using data from ExecuComp (OPT_UNEX_EXER_EST_VAL/OPT_UNEX_EXER_NUM). Then, subtract the total realizable value per share for exercisable option from the stock price at the fiscal year end (PRCCF). The option moneyness is the per option realizable value divided by the estimated strike price. Breached Bank Dummy is equal to one if banks have a data breach announcement and zero otherwise. Delta is the change in option value when the underlying stock price changes by 1 percent. Vega is the change in option value when the stock volatility changes by 1 percent. Insidedebt is the sum of present value of accumulated pension benefits and deferred compensation. Salary is the annual salary. Bonus is the dollar value of a bonus earned during the fiscal year. Stock is value of stock-related awards that do not have option like features. Total Compensation is all types of compensation in aggregate. ***,**,* are significant at 1, 5, 10 percent, respectively.

	(1)	(2)	(3)	(4)
VARIABLES	Moneyness	Moneyness	Moneyness	Moneyness
Breached Bank Dummy	91.4339***	79.0201***	79.9422***	101.7235***
	(12.41)	(5.26)	(5.50)	(9.05)
Delta		0.0166*	0.0416	0.0410
		(1.85)	(1.24)	(1.21)
Vega			-0.1400	-0.1615
			(-0.95)	(-1.09)
Insidedebt				0.0196
				(1.16)
Salary				-0.0106**
				(-2.20)
Bonus				-0.0005
				(-0.44)
Stock				0.0010
				(0.89)
Total Compensation		0.0002	0.0002	
		(0.74)	(0.71)	
Constant	53.8068***	56.1631***	55.9510***	20.0988
	(8.62)	(8.86)	(8.82)	(1.43)
Observations	5,882	5,712	5,712	5,323
R-squared	0.5174	0.5200	0.5210	0.5258
Bank FE	YES	YES	YES	YES
District FE	YES	YES	YES	YES
Qtr FE	YES	YES	YES	YES

Table 11: This table shows the results of deposit flows from breached banks to non-breached banks. The dependent variable is the insured deposit growth of non-data breached banks. *Breached Bank Insured Gr* is insured deposit growth of data breached banks. We first match the data breached banks with non breached banks by zip code. Then, calculate the distance in miles between breached and non-breached banks. Column (1) is for the distance of less than 5 miles between breached and non-breached banks. Column (2) is a distance of less than 10 miles. Column (3) is the distance between 10 to 30 miles. Column (4) is for the results if breached and non-breached banks are in the same state. The zip code (RSSD9220) and state (RSSD9200) information are retrieved from the Call Report. Please see the full definition for the control variables in Appendix A. ***,**,* are significant at 1, 5, 10 percent, respectively.

	(1)	(2)	(3)	(4)
VARIABLES	< 5 miles	< 10 miles	10 to 30	Same State
Qevent*Breached Bank Insured Gr	0.215	0.068	-0.347***	-0.096*
	(0.52)	(1.36)	(-2.74)	(-1.82)
Q1after*Breached Bank Insured Gr	-0.193	-0.363***	0.205	-0.078***
	(-1.53)	(-4.69)	(1.22)	(-2.67)
Q2after*Breached Bank Insured Gr	-0.006	0.060	-0.060*	-0.035
	(-0.06)	(0.76)	(-1.89)	(-1.57)
Q3after*Breached Bank Insured Gr	-0.188***	-0.148	0.053	0.106***
	(-4.04)	(-1.66)	(0.75)	(4.41)
Q4after*Breached Bank Insured Gr	0.012	0.101	-0.031	0.108**
	(0.11)	(0.84)	(-0.58)	(2.00)
Observations	7,587	16,082	33,165	88,896
R-squared	0.3828	0.3705	0.3579	0.3571
Bank FE	YES	YES	YES	YES
District FE	YES	YES	YES	YES
Qtr FE	YES	YES	YES	YES
Controls	YES	YES	YES	YES

Table 12: The results show the effect of multiple breach announcements on insured deposit growth. Each column represents the order of the data breach announcement. 1st is the first announcement, 2nd is the second announcement, and so on. *BreachNumber* is a dummy variable equal to one when the data breach incident is announced based on the number of breach announcement. For instance, BreachNumber for column (1) is the dummy variable equal to one when it is the first data breach announcement. The maximum number of data breach announcements is 10 times from the same bank. Please see the full definition for each type of deposit and other control variables in Appendix A. Qevent (Q1,2,3,4after) is a dummy variable equal to one if the interaction terms between a data breach bank dummy and a data breach announcement quarter (1,2,3,4 quarters after a data breach announcement quarter) dummy is equal to one. ***,**,* are significant at 1, 5, 10 percent, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
Qevent*BreachNumber	0.002	0.004	0.001	-0.008	-0.001	-0.007	0.005*	0.006	0.013***	-0.001
	(0.51)	(1.02)	(0.21)	(-1.13)	(-0.33)	(-0.96)	(1.81)	(1.26)	(3.51)	(-0.15)
Q1after*BreachNumber	0.003	0.005	-0.010**	0.005	0.000	-0.000	-0.002	-0.002	0.005**	-0.002
	(0.74)	(0.99)	(-2.49)	(0.99)	(0.04)	(-0.01)	(-0.67)	(-0.80)	(2.58)	(-0.50)
Q2after*BreachNumber	0.005	-0.011	0.008	0.008	-0.002	-0.004	0.000	0.001	-0.005	-0.001
	(0.63)	(-1.43)	(0.72)	(0.88)	(-0.51)	(-0.83)	(0.09)	(0.16)	(-1.14)	(-0.14)
Q3after*BreachNumber	-0.010*	0.021*	-0.002	-0.015**	0.005	-0.002	0.003	0.004	0.000	0.002
	(-1.90)	(1.97)	(-0.50)	(-2.01)	(1.48)	(-0.55)	(0.94)	(1.00)	(0.11)	(0.66)
Q4after*BreachNumber	0.002	-0.005	-0.003	0.020***	0.001	0.002	-0.002	0.002	0.004	0.006*
	(0.42)	(-0.79)	(-0.49)	(2.78)	(0.26)	(0.67)	(-0.44)	(0.56)	(1.27)	(1.88)
Observations	8,585	8,585	8,585	8,585	8,585	8,585	8,585	8,585	8,585	8,585
R-squared	0.2620	0.2626	0.2620	0.2627	0.2617	0.2620	0.2618	0.2618	0.2624	0.2617
Bank FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
District FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Qtr FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

Table 13: The results show the robustness test of data breach announcement effect on deposit growths by dropping years 2005 to 2007 during the adoption of disclosure laws. Dependent variables from columns (1) to (5) are different types of quarterly deposit growth. *Deposits* quarterly growth of all types of deposits. *Core* is quarterly growth of transaction, saving, and time deposits less than $100,000. *Insured* is quarterly growth of insured deposits. *Brokered* is the quarterly growth of deposits from brokers. *Time* is the time deposit account. Please see the full definition for each type of deposit and other control variables in Appendix A. Qevent (Q1,2,3,4after) is a dummy variable equal to one if the interaction terms between a data breach bank dummy and a data breach announcement quarter (1,2,3,4 quarters after a data breach announcement quarter) dummy is equal to one. ***,**,* are significant at 1, 5, 10 percent, respectively.

VARIABLES	(1) Deposits	(2) Core	(3) Insured	(4) Brokered	(5) Time
Qevent	0.000	0.003	-0.005*	0.000	0.001
	(0.04)	(0.98)	(-1.72)	(0.28)	(0.21)
Q1after	0.001	0.003	-0.001	0.001	-0.001
	(0.43)	(0.93)	(-0.66)	(1.15)	(-0.77)
Q2after	0.004	0.005*	0.004	0.001	0.001
	(1.43)	(1.68)	(0.89)	(0.91)	(0.44)
Q3after	-0.003	-0.000	-0.001	-0.001	0.002
	(-1.01)	(-0.00)	(-0.40)	(-0.36)	(0.93)
Q4after	0.002	0.002	-0.002	-0.004**	-0.000
	(0.72)	(0.72)	(-0.60)	(-2.33)	(-0.01)
Unused Commitment Ratio	-0.018	-0.015	-0.016	-0.009	0.007
	(-1.20)	(-1.22)	(-1.38)	(-1.05)	(0.90)
Net Wholesale Funding	-0.036***	-0.034***	-0.002	-0.005	0.011
	(-3.00)	(-3.35)	(-0.50)	(-1.03)	(1.53)
NPL to Loans	-0.246***	-0.177***	-0.117***	-0.070***	-0.081**
	(-4.25)	(-3.19)	(-3.05)	(-3.32)	(-2.23)
Capital Ratio	-0.192***	-0.143***	-0.077*	0.011	0.067*
	(-3.32)	(-3.58)	(-1.81)	(0.71)	(1.80)
Large Bank Indicator	0.011***	0.008**	0.006*	-0.001	0.007**
	(3.57)	(2.12)	(1.72)	(-0.38)	(2.20)
Real Estate Loan Share	-0.008	0.007	-0.015**	0.001	0.006*
	(-0.73)	(0.81)	(-2.45)	(0.17)	(1.68)
Constant	0.098***	0.059***	0.045***	0.006	-0.018***
	(8.06)	(6.11)	(6.24)	(1.05)	(-3.91)
Observations	7,098	7,098	7,098	7,098	7,098
R-squared	0.1497	0.1457	0.2850	0.0853	0.4194
Bank FE	YES	YES	YES	YES	YES
District FE	YES	YES	YES	YES	YES
Qtr FE	YES	YES	YES	YES	YES

Table 14: The results show the robustness test of data breach announcement effect on deposit growths by adding cumulative number of breaches, *Cumulative Breach*, as a control variable. Dependent variables from columns (1) to (5) are different types of quarterly deposit growth. *Deposits* quarterly growth of all types of deposits. *Core* is quarterly growth of transaction, saving, and time deposits less than \$100,000. *Insured* is quarterly growth of insured deposits. *Brokered* is the quarterly growth of deposits from brokers. *Time* is the time deposit account. Please see the full definition for each type of deposit and other control variables in Appendix A. Qevent (Q1,2,3,4after) is a dummy variable equal to one if the interaction terms between a data breach bank dummy and a data breach announcement quarter (1,2,3,4 quarters after a data breach announcement quarter) dummy is equal to one. ***,**,* are significant at 1, 5, 10 percent, respectively.

VARIABLES	(1) Deposits	(2) Core	(3) Insured	(4) Brokered	(5) Time
Qevent	-0.000	0.001	-0.005*	0.000	0.000
	(-0.19)	(0.26)	(-1.95)	(0.32)	(0.04)
Q1after	0.001	0.003	-0.001	-0.001	-0.002
	(0.30)	(0.96)	(-0.94)	(-0.55)	(-0.92)
Q2after	0.002	0.003	0.003	0.000	-0.001
	(1.03)	(1.20)	(0.84)	(0.39)	(-0.48)
Q3after	-0.003	-0.001	-0.002	-0.001	0.000
	(-1.13)	(-0.47)	(-0.87)	(-1.10)	(0.07)
Q4after	0.003	0.002	-0.002	-0.004**	-0.000
	(0.97)	(0.91)	(-0.63)	(-2.46)	(-0.16)
Cumulative Breach	-0.000	-0.000*	-0.000**	0.000	0.000
	(-0.84)	(-1.88)	(-2.40)	(0.82)	(0.56)
Unused Commitment Ratio	-0.007	-0.006	-0.006	-0.005	0.013*
	(-0.57)	(-0.57)	(-0.57)	(-0.64)	(1.74)
Net Wholesale Funding	-0.035***	-0.039***	-0.007*	-0.002	0.008
	(-3.63)	(-4.61)	(-1.71)	(-0.56)	(1.37)
NPL to Loans	-0.227***	-0.152***	-0.112***	-0.052***	-0.062**
	(-4.52)	(-2.77)	(-3.80)	(-2.68)	(-2.08)
Capital Ratio	-0.098*	-0.054	0.022	0.019	0.094***
	(-1.91)	(-1.30)	(0.56)	(1.54)	(3.35)
Large Bank Indicator	0.014***	0.011***	0.007***	0.001	0.009***
	(5.87)	(4.48)	(2.83)	(0.79)	(4.09)
Real Estate Loan Share	-0.011	-0.002	-0.011**	-0.000	0.004
	(-1.24)	(-0.28)	(-2.47)	(-0.13)	(1.19)
Constant	0.074***	0.043***	0.027***	0.005	-0.011**
	(5.68)	(4.58)	(3.18)	(0.87)	(-2.18)
Observations	8,585	8,585	8,585	8,585	8,585
R-squared	0.1344	0.1372	0.2621	0.0793	0.3779
Bank FE	YES	YES	YES	YES	YES
District FE	YES	YES	YES	YES	YES
Qtr FE	YES	YES	YES	YES	YES

References

Acharya, V. V., Mora, N., 2015. A crisis of banks as liquidity providers. The Journal of Finance 70, 1–43.

Allen, F., Gale, D., 2000. Financial contagion. Journal of Political Economy 108, 1–33.

Allen, L., Saunders, A., 1992. Bank window dressing: Theory and evidence. Journal of Banking & Finance 16, 585–623.

Arifovic, J., Jiang, J. H., 2019. Strategic uncertainty and the power of extrinsic signals evidence from an experimental study of bank runs. Journal of Economic Behavior & Organization 167, 1 – 17.

Baur, D. G., 2012. Financial contagion and the real economy. Journal of Banking & Finance 36, 2680–2692.

Berger, A. N., Curti, F., Mihov, A., Sedunov, J., 2022. Operational risk is more systemic than you think: Evidence from u.s. bank holding companies. Journal of Banking & Finance 143, 106619.

Black, J., 2013. Developments in data security breach liability. The Business Lawyer 69, 199–207.

Brusco, S., Castiglionesi, F., 2007. Liquidity coinsurance, moral hazard, and financial contagion. The Journal of Finance 62, 2275–2302.

Bryant, J., 1980. A model of reserves, bank runs, and deposit insurance. Journal of Banking & Finance 4, 335 – 344.

Campbell, K., Gordon, L., Loeb, M., Zhou, L., 2003. The economic cost of publicly announced information security breaches: Empirical evidence from the stock market. Journal of Computer Security 11, 431–448.

Campbell, T. C., Gallmeyer, M., Johnson, S. A., Rutherford, J., Stanley, B. W., 2011. Ceo optimism and forced turnover. Journal of Financial Economics 101, 695–712.

Chen, W.-D., Chen, Y., Huang, S.-C., 2021. Liquidity risk and bank performance during financial crises. Journal of Financial Stability 56, 100906.

Chernobai, A., Ozdagli, A., Wang, J., 2021. Business complexity and risk management: Evidence from operational risk events in u.s. bank holding companies. Journal of Monetary Economics 117, 418–440.

Cont, R., Schaanning, E., 2019. Monitoring indirect contagion. Journal of Banking & Finance 104, 85–102.

Core, J., Guay, W., 2002. Estimating the value of employee stock option portfolios and their sensitivities to price and volatility. Journal of Accounting Research 40, 613–630.

Curti, F., Gerlach, J., Kazinnik, S., Lee, M., Mihov, A., 2006. Cyber risk definition and classification for financial risk management. Journal of Operational Risk 1755, 2710.

DeYoung, R., Peng, E. Y., Yan, M., 2013. Executive compensation and business policy choices at us commercial banks. Journal of Financial and Quantitative Analysis 48, 165–196.

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

Diamond, D. W., Rajan, R. G., 2000. A theory of bank capital. The Journal of Finance 55, 2431–2465.

Diamond, D. W., Rajan, R. G., 2001. Liquidity risk, liquidity creation, and financial fragility: A theory of banking. Journal of Political Economy 109, 287–327.

Fang, L., Peress, J., 2009. Media coverage and the cross-section of stock returns. The Journal of Finance 64, 2023–2052.

Foerderer, J., Schuetz, S. W., 2022. Data breach announcements and stock market reactions: a matter of timing? Management Science 68, 7298–7322.

Garcia, L., Lewrick, U., Sečnik, T., 2021. Is window dressing by banks systemically important? BIS Working Papers .

Gatev, E., Schuermann, T., Strahan, P. E., 2009. Managing bank liquidity risk: How deposit-loan synergies vary with market conditions. The Review of Financial Studies 22, 995–1020.

Ho, P.-H., Huang, C.-W., Lin, C.-Y., Yen, J.-F., 2016. Ceo overconfidence and financial crisis: Evidence from bank lending and leverage. Journal of Financial Economics 120, 194–209.

Huang, H. H., Wang, C., 2021. Do banks price firms' data breaches? The Accounting Review 96, 261–286.

Imbierowicz, B., Rauch, C., 2014. The relationship between liquidity risk and credit risk in banks. Journal of Banking & Finance 40, 242–256.

Iyer, S. R., Simkins, B. J., Wang, H., 2020. Cyberattacks and impact on bond valuation. Finance Research Letters 33, 101215.

Kamiya, S., Kang, J.-K., Kim, J., Milidonis, A., Stulz, R. M., 2021. Risk management, firm reputation, and the impact of successful cyberattacks on target firms. Journal of Financial Economics 139, 719–749.

Karas, A., Pyle, W., Schoors, K., 2013. Deposit insurance, banking crises, and market discipline: Evidence from a natural experiment on deposit flows and rates. Journal of Money, Credit and banking 45, 179–200.

Kim, M., Kliger, D., Vale, B., 2003. Estimating switching costs: The case of banking. Journal of Financial Intermediation 12, 25–56.

Lending, C., Minnick, K., Schorno, P. J., 2018. Corporate Governance, Social Responsibility, and Data Breaches. Financial Review 53, 413–455.

Mikhed, V., Vogan, M., 2018. How data breaches affect consumer credit. Journal of Banking & Finance 88, 192 – 207.

Park, S., Peristiani, S., 1998. Market discipline by thrift depositors. Journal of Money, Credit and Banking pp. 347–364.

Peria, M., Soledad, M., Schmukler, S. L., 2001. Do depositors punish banks for bad behavior? market discipline, deposit insurance, and banking crises. The Journal of Finance 56, 1029–1051.

Piccotti, L. R., Wang, H., 2022. Informed trading in the options market surrounding data breaches. Global Finance Journal p. 100774.

Romanosky, S., Telang, R., Acquisti, A., 2011. Do data breach disclosure laws reduce identity theft? Journal of Policy Analysis and Management 30, 256–286.

Sharpe, S. A., 1990. Asymmetric information, bank lending and implicit contracts: A stylized model of customer relationships. The Journal of Finance 45, 1069–1087.

Sharpe, S. A., 1997. The effect of consumer switching costs on prices: A theory and its application to the bank deposit market. Review of Industrial Organization 12, 79–94.

Spanos, G., Angelis, L., 2016. The impact of information security events to the stock market: A systematic literature review. Computers & Security 58, 216–229.

Vesala, T., 2007. Switching costs and relationship profits in bank lending. Journal of Banking & Finance 31, 477–493.

Wang, H. E., Wang, Q. E., Wu, W., 2022. Short selling surrounding data breach announcements. Finance Research Letters p. 102690.

Appendix A. Dependent Variables and Controls

Equations	Variables	Description
$\frac{\Delta Deposits_t}{Assets_{t-1}}$	Quarterly growth of deposits	Quarterly change in deposits divided by assets
$\frac{\Delta CoreDeposits_t}{Assets_{t-1}}$	Quarterly growth of core deposits	Quarterly change in core deposits divided by assets. Core deposits are the sum of transaction deposits, saving deposits, and time deposits less than \$100,000
$\frac{\Delta InsuredDeposits_t}{Assets_{t-1}}$	Quarterly growth of insured deposits	Quarterly change in insured deposits divided by assets. Insured deposits are calculated from the accounts of \$100,000 or less before 2009Q3. After 2009Q3, the insured amount increased to \$250,000.
$\frac{\Delta BrokeredDeposits_t}{Assets_{t-1}}$	Quarterly growth of brokered deposits	Quarterly change in brokered deposits divided by assets.
$\frac{\Delta TimeDeposits_t}{Assets_{t-1}}$	Quarterly growth of time deposits	Quarterly change in time deposits divided by assets.

Equations	Variables	Description
$\frac{\Delta Loans_t}{Assets_{t-1}}$	Quarterly growth of loans	Quarterly change in loans divided by assets.
$\frac{\Delta Credit_t}{(Assets+Committment)_{t-1}}$	Quarterly growth of credit (loans+commitments)	Quarterly change in credit divided by asset size and unused commitment. Credit is the sum of loans and unused commitments.
$\frac{\Delta CILoans_t}{Assets_{t-1}}$	Quarterly growth of CI loans	Quarterly change in commercial and industrial loans divided by asset size
$\frac{Committment_t}{(Loans+Committment)_t}$	Unused commitment ratio	Unused commitment divided by the sum of unused commitments and loans

Equations	Variables	Description
$\frac{(Wholesalefunds-Liquidassets)_t}{(Assets)_t}$	Net wholesale funding	Wholesale funds less liquid assets to total assets. Wholesale funds are the sum of large time deposits, deposits booked in foreign offices, subordinated debt and debentures, gross federal funds purchased, repos, and other borrowed money. Liquid assets are cash, federal funds sold and reverse repos, and securities excluding MBS/ABS securities.
$\frac{Bookcapital_t}{(Assets)_t}$	Capital Ratio	Book capital to asset size
$\frac{Realestateloans_t}{(Loans)_t}$	Real estate loan share	Real estate loans divided by total loans
Large bank dummy	Large bank indicator	Dummy variable equal to one for the top five largest banks by asset size in each quarter

Appendix B. Data Breach Announcement Examples

Date Made Public	Company	City	State	Type of breach	Total Records	Description of incident
17-Jan-18	Ameriprise Financial, Inc.	Minneapolis	Minnesota	DISC	56	Ameriprise Financial suffered an inadvertent disclosure of 56 records, including SS numbers and names
5-Feb-18	1st Mariner Bank	Baltimore	Maryland	HACK	1500	1st Mariner Bank experienced a phishing attack that resulted in the exposure of the records of 1500 persons. Information exposed included Social Security Numbers, as well as names in combination with credit card or financial account information.
26-Feb-18	Southern National Bancorp of Virginia, Inc	Glen Allen	Virginia	HACK	24999	Southern National Bancorp of Virginia suffered a breach affecting 24,999 records, including social security numbers, driver's license number or non-driver identification card numbers, as well as financial account numbers or credit card numbers, in combination with the security code, access code, password or PIN for the account
20-Apr-18	SunTrust Banks, Inc.	Atlanta	Georgia	HACK	1500000	SunTrust Banks Inc. said an employee may have stolen the information of about 1.5 million customers and provided it to a criminal third party, the latest example of a potential breach that underscores the vulnerability of consumers, private data. The Atlanta-based bank on Friday said the employee, who no longer works at SunTrust, attempted to access client information, although it has not identified significant fraudulent activity around the accounts involved

Appendix C. Variable Description (call report)

Bank level data are from the quarterly Call Report. We aggregate banks to top holder level (RSSD9348) when RSSD9348 is available. If RSSD9348 is not available for any bank, all variables are calculated at the bank level. The sample excludes non-US banks. To control the merger, we also exclude banks with more than 10% growth in assets from the previous quarter. All the growth rates are the quarterly change divided by the beginning period assets. For the assets, we use RCFD2170 for bank holding level and RCON2170+RCFN2170 for bank level. All growth rates are winsorized at 1% tails.

When we download call report variables in the balance sheet focusing on bank holding level (RCFD), many variables are missing after 2011. To recover the missing information, we first need to understand how a variable is calculated. First, we need to understand that there are three types of call report:[25]

1. FFIEC 031: Banks with domestic and foreign offices
2. FFIEC 041: Banks with domestic offices only
3. FFIEC 051: Banks with domestic offices only and total assets less than $5 billion

For example, if we need *total assets* variable from call report, we can call RCFD2170. If RCFD2170 is missing, we can also call RCON2170. RCFD2170 refers to information from FFIEC 031. To access the information from FFIEC 041 and 051, we will call RCON 2170 for total assets. The prefix RCFD is specifically for FFIEC 031. It refers to banks with domestic and foreign offices. If the variable is still missing, we trace back how the variable is calculated. For example, total assets (RCFD2170) is calculated from the sum of domestic assets (RCON2170) and foreign assets (RCFN2170).

Dependent variables

[25] https://www.ffiec.gov/ffiec_report_forms.htm

- Quarterly growth of deposits: RCFD2200. If RCFD2200 is missing, deposits are the sum of interest-bearing deposits (RCON6636) and non-interest bearing deposits (RCON6631).
- Quarterly growth of core deposits

  Core deposits are the sum of transaction deposits, saving deposits, and time deposits less than $100,000. RCON2215+RCON6810+RCON0352+RCON6648
- Quarterly growth of insured deposits

  The threshold for insured deposit was $100,000 or less for any account until 2009Q3 when the amount was increased to $250,000. Note retirement account increased the amount to $250,000 earlier in 2006Q2. Insured deposits before 2006Q2: RCONF049. Since 2006Q2, insured deposits: RCONF049+RCONF045
- Quarterly growth of brokered deposits

  Deposits received from brokers and dealers: RCON2365
- Quarterly growth of time deposits

  Deposits from time deposit accounts less than $100,000 and more than $100,000: RCON6648+RCON2604
- Quarterly growth of loans

  Loans are RCFD1400. If RCFD1400 is missing, we use RCON1400. If RCON1400 is missing, we use RCFD2122 - RCFD3123 which is total loans and leases held for investment and held for sale less allowance for loan and lease losses. If RCFD3123 is missing, we use RCFD2122. Next, if Loans variable is still missing, we use RCON2122. If RCON2122 is missing, we then use RCFDB528+RCFD5369 which are loans and leases held for investment and held for sale.

  Commercial and industrial (CI) loans are RCFD1766. If RCFD1766 is missing, we use RCON1766. If RCON1766 is missing, we use the sum of RCFD1763 and RCFD1764 which are CI loans to US addresses (domicile) and non-US addresses (domicile). Lastly, if CI loans variable is still missing, we use the sum of RCON1763 and RCON1764.

- Quarterly growth of credit

  Credits are the sum of loans (RCFD1400) and unused commitments (RCFD3814+RCFD3816+RCFD3817+RCFD3818+RCFD6550+RCFD3411). In this case, the denominator of the growth rate is the sum of beginning period assets and commitment.

Controls

- Unused commitment ratio

  Unused commitment consists of RCFD3814 +RCFD3816 +RCFD3817 +RCFD3818 +RCFD6550 +RCFD3411. If the variable is missing, we use RCON3814 +RCON3816 +RCON3817 +RCON3818 +RCON6550 +RCON3411

  Unused commitment ratio is unused commitments divided by the sum of unused commitments and loans.

- Net wholesale funding to asset ratio

  The ratio of wholesale funds (excludes liquid assets) to total assets

  Wholesale funds are the sum of deposits booked in foreign offices, large time deposits, subordinated debt and debentures, gross federal funds purchased, repos, and other borrowed money: RCFN2200 + RCON2604 + RCFD3200 + RCFD2800 (From 2002Q1: RCONB993+RCFDB995) + RCFD3190.

  If RCFD3200 is missing, we set it as RCON3200. If RCFD2800 is missing, we use RCON2800. If RCON2800 is missing, we use RCONB993+RCONB995. If RCFD3190 is missing, we use RCON3190.

  Liquid assets are the sum of cash (RCFD0010), federal funds sold and reverse repos (Before 2002Q1: RCFD1350, From 2002Q1: RCONB987 +RCFDB989), and securities excluding MBS/ABS securities (Before 2009Q2: RCFD1754 +RCFD1773 (RCFD8500 +RCFD8504 +RCFDC026 +RCFD8503 +RCFD8507 +RCFDC027), From 2009Q2: RCFD1754 +RCFD1773 (RCFDG300 +RCFDG304 +RCFDG308 +RCFDG312 + RCFDG316 + RCFDG320 +RCFDG324 +RCFDG328 +RCFDC026 +RCFDG336

+RCFDG340 +RCFDG344 +RCFDG303 +RCFDG307 +RCFDG311 +RCFDG315 +RCFDG319 +RCFDG323 +RCFDG327 +RCFDG331 +RCFDC027 +RCFDG339 +RCFDG343 +RCFDG347).

To maximize the availability of liquid assets, we use alternative IDs for each variable used to calculate liquid assets. For cash, if RCFD0010 is missing, we use RCON0010. For securities excluding MBS/ABS securities before 2009Q2, we use (RCON1754 + RCON1773) - (RCON8500 + RCON8504 + RCONC026 + RCON8503 + RCON8507 + RCONC027).

For 2009Q2 and after, we use (RCON1754 + RCON1773) - (RCONG300 + RCONG304 + RCONG308 + RCONG312 + RCONG316 + RCONG320 + RCONG324 + RCONG328 + RCONC026 + RCONG336 + RCONG340 + RCONG344 + RCONG303 + RCONG307 + RCONG311 + RCONG315 + RCONG319 + RCONG323 + RCONG327 + RCONG331 + RCONC027 + RCONG339 + RCONG343 + RCONG347)

- Nonperforming loans to loans

  Nonperforming loans (NPL) are loans past due 90 days or more and non-accruals: RCFD1407 + RCFD1403. If NPL is missing, we use RCON1407+ RCON1403. If NPL is still missing, we use RCONF174 + RCONF175 + RCON3494 + RCON5399 + RCONC237 + RCONC239 + RCON3500 + RCONF180 + RCONF181 + RCFNB573 + RCFD5378 + RCFD5381 + RCFD1597 + RCFD1252 + RCFD1255 + RCFDB576 + RCFDK214 + RCFDK217 + RCFD5390 + RCFD5460 + RCFDF167 + RCFDF170 + RCONF176 + RCONF177 + RCON3495 + RCON5400 + RCONC229 + RCONC230 + RCON3501 + RCONF182 + RCONF183 + RCFNB574 + RCFD5379 + RCFD5382 + RCFD1583 + RCFD1253 + RCFD1256 + RCFDB577 + RCFDK215 + RCFDK218 + RCFD5391+ RCFD5461 + RCFDF168 + RCFDF171

  If NPL is still missing, we use RCONF174 + RCONF175 + RCON3494 + RCON5399 + RCONC237 + RCONC239 + RCON3500 + RCONF180 + RCONF181 + RCONB835 + RCON1607 + RCONB576 + RCONK214 + RCONK217 + RCON5460 + RCON1227

+ RCONF176 + RCONF177 + RCON3495 + RCON5400 + RCONC229 + RCONC230 + RCON3501 + RCONF182 + RCONF183 + RCONB836 + RCON1608 + RCONB577 + RCONK215 + RCONK218 + RCON5461 + RCON1228

- Capital ratio

  Book capital (RCFD3210) to asset ratio. If RCFD3210 is missing, we use RCON3210.

- Indicator for large banks

  If a bank organization is in the top 5 largest bank organization by assets, the indicator is equal to 1; 0 otherwise.

- Real estate loan share

  Loans backed by real estate (RCFD1410) divided by total loans

  If RCFD1410 is missing, then we use RCON1410. If RCON1410 is missing, we use RCFDF158 + RCFDF159 + RCFD1420 + RCFD1797 + RCFD5367 + RCFD5368 + RCFD1460 + RCFDF160 + RCFDF161. if the real estate variable is still missing, we use RCONF158 + RCONF159 + RCON1420 + RCON1797 + RCON5367 + RCON5368 + RCON1460 + RCONF160 + +RCONF161

- District time trends

  Federal Reserve district (RSSD9170)