Customer Satisfaction and the Durability of Retail Banking Relationships*

Joel F. Houston[†], Hongyu Shan[‡], and Yu Shan[§]

April 3, 2021

Abstract

Building on 699,008 reviews of 75,903 U.S. bank branches, this paper proposes the first detailed empirical framework to study how banks engage with customers at the branch level. In a within bank-county estimation, we exploit reputation damage as exogenous negative shocks to deposit-taking and find that branches receiving higher ratings mitigate deposit outflows more effectively following these shocks. These results are stronger in neighborhoods with higher income and lower population mobility, and for branches of community banks. Overall, our work highlights the value of customer interactions as a novel and important non-price factor influencing the durability of retail banking relationships.

JEL classification: G20, G21, G28, G41

Keywords: retail banking, customer satisfaction, deposit, mortgage, textual analysis

*We thank Linda Allen, Svenja Dube, Mark Flannery, Chris James, Sehoon Kim, Jay Ritter, Benjamin Segal, Maya Waisman, Yihui Wang, An Yan, and seminar participants at Fordham University and the University of Florida for helpful comments and suggestions. Shuai Hao provided excellent research assistance. We are responsible for any remaining errors.

[†]Corresponding author. Eugene Brigham Chair in Finance, Warrington College of Business, University of Florida, Gainesville, FL 32611. Email: joel.houston@warrington.ufl.edu.

[‡]Assistant Professor of Finance, Gabelli School of Business, Fordham University, New York, NY 10023. Email: hshan6@fordham.edu.

[§]Assistant Professor of Finance, John Molson School of Business, Concordia University, Montreal, QC, Canada H3H 0A1. Email: yu.shan@concordia.ca.

1 Introduction

Banks are relationship driven and the durability of the relationships depends critically on customer satisfaction. In recent years, the importance of customer engagement has drawn tremendous interest from academics and practitioners. Notably, Drechsler, Savov, and Schnabl (2017) model the "percentage of non-switching customers" as a major determinant of the branch's local market power. McKinsey (2019) also highlights the value of customer interactions as branches transform into a more consultative role featuring personalized services. However, researchers have struggled to empirically examine how banking experiences vary across branches, and whether these inter-branch differences contribute to the durability of retail banking relationships.

In this paper, we quantify, for the first time, the quality of customer interactions at the branch level. We utilize the Google Map Platform, which provides a staggering amount of information including customer ratings and detailed reviews for over 150 million physical locations around the world. As the most popular search engine in the world, Google has built trust and credibility among global customers. A recent survey shows that 77% of consumers are willing to leave a review when asked, and 81% of consumers visited Google reviews in the past year.[1] In this study, we focus on retail banking locations and obtain 699,008 reviews on 75,903 unique bank branches in the United States. We demonstrate that while individual branch ratings are fairly consistent over time, only 30% of the variations in ratings across branches are explained by bank-county-specific characteristics, suggesting that there remain considerable variations in customer satisfaction ratings across branches operated by the same bank in a given county.

Utilizing this novel database, we explore how customer satisfaction influences a branch's ability to retain deposits in the wake of adverse reputation shocks.[2] We identify these shocks based on the extant literature on depositor discipline that document that deposit flows are affected by banks' negative signals in ESG news coverage (Homanen 2018; Chen, Hung, and Wang 2019) or financial markets (Chavaz and Slutzky 2019). We first validate this premise using the RepRisk database, which tracks negative news stories related to the bank's business conduct and its treatment of

[1] "Podium 2017 State of Online Reviews," `https://learn.podium.com/rs/841-BRM-380/images/Podium-2017-State-of-Online-Reviews.pdf`.

[2] Deposit-taking is one of the most critical retail banking functions. According to the FDIC, deposits represent 81% of banks' total liabilities (Q3 2020, *Quarterly Banking Profile* Report).

various stakeholders including customers, employees, and the surrounding environments and communities. We confirm that branch deposit growth subsides following negative news coverage, and the magnitude of the impact increases with the severity of the incidence. We include branch fixed effects (FEs, hereafter) to preclude the impacts of time-invariant branch-level omitted variables (e.g., geographical locations or proximity to businesses), and county by year FEs to exclude the effects of time-varying county-specific events (e.g., political elections or natural disasters). Economically, depending on the severity of the incidence, the deposit growth rate is 0.7% to 2.6% lower following reputation shocks, equivalent to 7.45% to 27.66% of the average deposit growth rate. Our results suggest that bank customers penalize banks for behaviors that are perceived to be detrimental to stakeholders, and demonstrate the validity of using negative ESG incidents as a unique, quasi-exogenous empirical setting to examine the durability of banking relationship as a function of customer interactions.[3]

Building on the negative relationship between deposit growth and adverse reputation shocks, we examine the durability of retail banking relationships when these shocks ripple through branch networks. Specifically, we explore whether branches with higher customer ratings are more resistant to bank-level reputation shocks. Arguably, the theoretical links between customer ratings and relationship durability remain unclear. One possibility is that customer depositing decision is independent of their in-branch banking experiences, given the rise of mobile and internet channels of deposit-taking. However, observations from practitioners show that customer visits to physical branches remain the most significant channel for deposit-taking. According to McKinsey (2019), among customers that opened a core banking product in the past two years (e.g., checking account), only 13% acquired them digitally. Deloitte (2020) echoes the findings by showing that most customers prefer branches over digital channels when opening new accounts for both simple (e.g., savings accounts and debit cards) and complex products (e.g., loans).

Even if customer satisfaction affects a branch's ability to withstand negative shocks, the direction of the impact remains an empirical question. One possibility is that higher-rated branches have "more to lose" and would therefore observe a relatively larger drop in deposits following

[3]We focus on negative ESG news as negative shock to deposit taking, but have not included positive ESG news as positive shocks for two reasons: First, positive ESG news are more likely to be strategic and subject to green-washing bias. Second, Chen, Hung, and Wang (2019) has documented that depositors do not respond to positive ESG shock.

an adverse reputation shock. Alternatively, higher-rated branches may have successfully accumulated sufficient goodwill to resist the negative impact. In this scenario, investments in customer banking experiences would help protect banks against subsequent hits to their reputation. Our results are consistent with this latter expectation. We find a positive and significant cross-sectional relationship between branch ratings and deposit growth. Conditional on a reputation shock, a one-standard-deviation increase in the branch's customer rating attenuates the negative impact on deposit growth by 0.3% to 0.4%. Consequently, we offer strong evidence that banks with stronger levels of customer satisfaction are better positioned to withstand the shocks to their reputation.

Admittedly, an empirical examination of the connections between customer ratings and the durability of banking relationships is subject to two identification challenges: measurement errors and omitted variables. Measurement errors arise because the utility functions of the retail clientele for different branches may be different. In those cases, the ratings and reviews are not comparable across branches. Omitted variables are also a valid concern when time-varying bank-county unobserved variables are associated with both branch ratings and branch deposit growth. For example, when a bank is planning to retreat from a local market, the decision may affect both its commitment to customer services and its efforts for keeping deposit growth.[4]

We alleviate these endogeneity concerns by conducting a within bank-county-shock estimation. This extremely restrictive setup allows us to observe how different branches of the same bank operating in the same geographic region differentially respond to common reputation shocks, and how these responses depend on the branches' level of customer satisfaction. The identifying assumption for this estimation is that customers with similar preferences and utility functions cluster by bank and location, and therefore the ratings by reviewers connected to the same bank in the same county are more comparable. Furthermore, high dimensional bank by county by shock year FEs help us preclude the impact of time-varying bank-county specific omitted variables. Admittedly, there remains some heterogeneity across different neighborhoods within the same county. With this concern in mind, we also check the robustness of our results by using even more granular bank by ZIP Code by shock year FEs to preclude the impact of time-varying neighborhood-specific

[4] Note that reserve causality is unlikely a major concern in this study. The empirical design utilizes bank reputation shocks - the exact timing of the news reporting by outsiders is arguably quasi-exogenous and out of the control of corporate insiders. The likelihood of changes in branch deposit growth reversely leading to negative news coverage initiated by outsiders is extremely low.

variables, and confirm that our results still persist.

To shed some light on the potential drivers of our results on customer satisfaction, we document the heterogeneous effects of customer satisfaction across counties with differential socioeconomic conditions. First, we find a stronger effect of customer satisfaction in counties with higher income per capita, suggesting that wealthier customers place a higher economic value on non-price factors, such as customer satisfaction. Second, we find a stronger effect of customer satisfaction in counties with lower population mobility, which is measured by the share of the population that migrates into and out of the county during our sample period. Given that deposit growth can be established through retaining existing customers and/or attracting new customers, our results indicate that customer satisfaction may have a larger impact on the stickiness of existing customers following adverse shocks.

We further investigate on the heterogeneous effects of customer satisfaction by comparing the results between community banks and large regional or super-regional banks. We posit that effective customer interactions are more crucial for community banks that heavily rely on relationship banking using soft information collected through years of experience with local customers and the business community. Indeed, we find that the effects of customer satisfaction are stronger among branches that are part of a community bank, and the results are robust to alternative asset cutoffs.

Our analysis of depositor decisions suggests that banks with higher levels of customer satisfaction build more durable relationships and are better positioned to confront the effects of adverse reputation shocks. However, the multi-faceted customer banking experience remains a black-box that is only partially captured by the aggregate rating. It remains unclear what customers truly value and what dimensions of customer interactions consequently contribute to the durability of retail banking relationships. Thus, we decompose the customer banking experience by analyzing the detailed textual reviews that form the basis of Google ratings. In this analysis, we follow Li et al. (2020) to employ a semi-supervised machine learning-based approach to extract key topics from the underlying textual reviews. We highlight four main dimensions that drive customer satisfaction: 1) accessibility of the services; 2) quality of the products, 3) hospitality of the staff; and 4) quality of facilities. For each branch, we construct indicators that measure the extent to which customers care about each dimension.

We employ the topic-specific rating indicators to examine the services that depositors care most about in the wake of reputation damage. We find that the accessibility dimension, which captures location, hours and efficiency of branch operations, is the key factor that drives customer retention. These results suggest that the relative accessibility of branch services determine the customer's effective switching costs. Notably, pricing and fee-related features of products do not play a role in retaining customers after adverse shocks.

While our results show that higher customer satisfaction helps retain depositors following adverse reputation shocks, the impact of customer satisfaction can go above and beyond depositor retention to borrower growth. As a final robustness test, we also consider an alternative set of tests where we focus on the demand for residential mortgages. We employ natural disasters as positive shocks to local residential mortgage demand (see Cortés and Strahan 2017 and Dlugosz et al. 2019), and explore whether higher levels of customer satisfaction enable banks to capture additional mortgage business in the aftermath of these natural disasters.

We find that, although natural disasters generate an average increase in mortgage demand, the increase is significantly larger for banks with higher levels of customer satisfaction. For banks with local Google ratings at the top (bottom) tercile among all bank-county observations within the state, the number of mortgage applications increased by 55.8% (35.5%) in the disaster year. The result is robust to loans of different purposes, which arguably contain different proportions of new and returning customers. Our results indicate that banks with higher levels of customer satisfaction are better positioned to capture stronger mortgage growth in the face of demand shocks.[5] Altogether, these results provide further evidence that Google ratings offer useful insights into customer attitudes, and that these attitudes meaningfully affect a bank's ability to retain and attract customers in the aftermath of reputation shocks and natural disasters.

Our results contribute to several distinct, but interconnected literature. First, we provide the first micro-level measure of customer satisfaction among banking customers that is based on observed Google ratings. Despite the voluminous literature estimating the relationship between customer satisfaction and loyalty (see reviews in Shankar, Smith, and Rangaswamy 2003 and Ku-

[5]With respect to topic-specific rating, we find that hospitality of the staff and the quality of facility are the two main determinants that drive new business. Arguably, these findings highlight the important influence that the "human element" has on the quality of banking relationships, especially during uncertain periods following natural disasters.

mar, Dalla Pozza, and Ganesh 2013), whether increasing service satisfaction in the banking sector leads to better loyalty or performance is still unclear, arguably because bank customer satisfaction is difficult to define or measure.[6] The existing literature typically uses survey-based data (Loveman 1998; Huang and Sudhir 2021) whose time window, customer base, or the number of banks covered are relatively limited. Our measures of branch customer satisfaction are more extensive and granular covering the vast majority of branches in the U.S. These measures provide insights into the factors that drive customer satisfaction in banking, and also provide further support for papers in other areas that have focused on the efficacy of web-based ratings. In this regard, we also argue that the observed Google ratings provide not only an important measure of the satisfaction of existing customers, but also an important signal to potential customers.

Second, our paper contributes to the evolving literature on the social responsibility of banking institutions. There are ongoing concerns about the responsiveness of companies to consumer and community needs. These concerns are particularly pronounced in key regulated areas such as banking. Different banks likely have different views on the relative importance of these issues, and these views may ultimately influence a bank's willingness to invest in its "social capital." (Chava 2014; Homanen 2018; Houston and Shan 2019). In this regard, our study explores the connections between investments in social capital and customer satisfaction, and yields particular insights into the specific factors that matter most to customers and that lead to more durable relationships. In particular, our results demonstrate that these investments create valuable customer loyalty, which helps banks maintain more durable relationships amid significant internal and external shocks.

Third, our paper highlights another important factor that influences the strength and durability of banking relationships. In this vein, our work is related to the long-standing theories of relationship lending (e.g., Sharpe 1990; Berger and Udell 1995; Puri and Rocholl 2008) and depositor behavior (e.g., Iyer and Puri 2012; Iyer, Puri, and Ryan 2016). Relatedly, our results also indicate that customer satisfaction is an important non-price factor influencing the establishment and durability of banking relationships. In this regard, our paper contributes to the literature that has documented that other non-price factors affect banking relationships, such as political orientation (Khwaja and Mian 2005), reputation (Ross 2010; Gopalan, Nanda, and Yerramilli 2011; Chava

[6] Huang (2020) presents the first evidence on the financial consequences of customer satisfaction. The paper documents that small business with higher ratings are more likely to be approved for small business loans.

2014), and cultural and legal origins (Mian 2006; Giannetti and Yafeh 2012).

Fourth, our paper sheds light on the role of traditional bank branches in the digital era. Branches, as the traditional banking channel, play a key role in bank-customer interactions. Despite the continued pressure of digitization, local bank branches still have inherent advantages in facilitating access to credit (Nguyen 2019), building brand presence (Jacques et al. 2018), and maintaining customer relationships (Larsson and Viitaoja 2017). Moreover, customers appear to place more trust in businesses with whom they have a good rapport, and arguably these links are even more important when the business specializes in complex financial products (Buttle and Maklan 2019). Indeed, survey data show that the effects of customer satisfaction with branches on overall satisfaction are at least twice as large as satisfaction with online or mobile channels (Srinivas and Wadhwani 2019). To this extent, our results highlight the apparent benefits of human interaction in maintaining durable customer relationships and suggest that there may be risk associated with a quick overreliance on technology. Moreover, our findings may indicate an ongoing niche role for smaller community banks that utilize human interactions to build and maintain customer relationships.

2 Data

2.1 Google Reviews on U.S. Bank Branches

We obtain 699,008 reviews on 75,903 unique bank branches in the United States from Google Map. The Google Map Platform is built with the most comprehensive, global points of interests data. With its online review and photosphere systems, Google Map provides real-world insights and immersive location experiences for over 150 million physical locations around the world. The service was first offered to Android and iOS users in September 2008. In 2013, it has grown into the most popular App with 54% of global smartphone users using it at least once. In this study, we focus on the Google profiles observed since 2014, when the service expanded to a sizable user base to ensure that banking experiences are widely shared on the platform.

To construct the database, we first extract the list of bank branches from the FDIC Summary of Deposits (SOD) as of June 30, 2019. We locate the Google profile of each branch by searching the combination of bank name and branch address as the key words in the Google search engine. This automated process is programmed in Python and enabled by the Wextractor Google Map API. After identifying the Google profile of each branch in the Google Map Platform, we download the rating, time and the textual content of all reviews. The downloading process was performed during March 2020, so our sample covers all reviews left as of February 2020.

We further tokenize and clean the textual content of the reviews by removing 1) punctuation, 2) numbers, 3) non-English words, and 4) stop words using the NTLK list of English stop words. Finally, we only keep reviews with more than three words after these cleaning steps. We record the number of words in each review as the length of the review. The yearly branch-level Google rating (*Rating*) is calculated as the average of all ratings of the branch as of every June 30. We also control for the yearly branch-level number of Google ratings of the branch ($Num\_Reviews$) as of every June 30.[7]

2.2 Reputation Shocks

This study measures a bank's perceived reputation related to ESG and business conduct issues using the RepRisk database. The database tracks negative news incidents of firms from January 2007 to June 2019. A dedicated team of analysts leverage a combination of artificial intelligence and curated human analysis to track a universe of over 95,000 firms globally. Over 80,000 public sources and stakeholders in 20 languages are screened on a daily basis. Once an incident is identified, analysts conduct additional analysis to (1) confirm that the incident is indeed related to the firm's ESG activities or business conduct, (2) remove possible duplicate media coverage on the same incident to make sure each risk event only enters once into the RepRisk Platform, and (3) identify the specific nature of the incident, by mapping it to 28 issues and 45 topics including "discrimination in employment", "controversial products", and "tax evasions", etc. Each incident is assigned three proprietary scores based on severity (harshness), reach (influence), and novelty (newness). Finally, the monthly RepRisk Index is updated, reflecting the ensuing impact of the news incident on the firm's perceived reputation.

[7] In the robustness test section, we also calculate the yearly bank-county-level Google rating as the average of all ratings of the branches within the same bank-county as of every December 31.

We capture the reputation shock to a bank by exploiting the jumps in its RepRisk Index. According to RepRisk, the magnitude of the jump increases in the severity of negative news incidents. The RepRisk Index is a non-broken, time-series variable ranging from 0 to 100, with 100 representing the worst perceived reputation. A yearly increase in the index is driven by the intensity of negative news coverage on the bank's business conduct and lending practice (Houston and Shan 2019). Instead of arguing for a "one-size-fits-all" threshold, we create four indicator variables that are associated with increasing levels of damage to the bank's reputation - *Rep_10*, *Rep_15*, *Rep_20*, and *Rep_25*, which equal one if the RepRisk Index increases by more than the corresponding magnitudes over the past year, and zero if the jump is less than 10.[8] In addition to indicator variables, we also define a continuous variable, *Rep_Chg*, to capture the reputation damage. It is defined as the maximum jump in the RepRisk Index over the past year. To merge with other databases, we identify the banks' RSSD ID by cross-checking the banks' ISIN and name in the SNL database.

2.3 Branch Deposits

We obtain branch office deposits data from the FDIC Summary of Deposits (SOD) database. SOD reports the annual survey of branch office deposits as of June 30 for all FDIC-insured depository institutions.[9] All insured institutions with branch offices are required to submit the survey. Besides branch deposits, the survey also reports comprehensive data including branch location, date of establishment, institution type, and name of the top holding company, etc. We construct the following variables using SOD data: the amount of deposits (*Deposits*) in the branch; the annual growth in branch deposits (*Depositgrowth*); a local bank dummy (*Local*) that equals one if a bank obtains more than 65 percent of its deposits from a single county, and zero otherwise; an important market dummy (*Important*) that equals one if a county is in the top quartile of

[8] In Table 5, we examine the validity of the measures by regressing deposit growth on the proxies of reputation shock. In a within-branch setting, the economic magnitudes of the indicator variables increase in the magnitude of jumps in RepRisk Index, even after controlling for branch and county-year FEs. Details of the validity tests are reported in Section 4.1.

[9] In our study, we assume that the branch that the depositor visits is the same branch where the deposits are assigned. In reality, FDIC allows banks to assign deposits consistent with their existing internal record-keeping practices. The general guidelines are the following: 1) deposits assigned to the office in closest proximity to the account holder's address, 2) deposits assigned to the office where the account is most active, 3) deposits assigned to the office where the account was opened, or 4) deposits assigned to offices for branch manager compensation or similar purposes (*FDIC Summary of Deposits Reporting Instructions*, June 30, 2020). Our assumption is consistent with the first and second guidelines. If banks choose to comply with the 3) or 4) then it creates attenuation bias. In those cases, our results are the underestimation of the true economic effects.

deposits among all the counties in which a bank has branches and zero otherwise; a market share variable ($Countyshare$) that measures the market share of a given bank branch in a given county by deposits; a branch location dummy ($Samestate$) that equals one if a branch is in the same state with the headquarter of the bank, and zero otherwise; a new branch dummy (New) that equals one if a branch was established within the past five years, and zero otherwise.

We obtain the branch-level deposit rates data from the RateWatch database. The deposit rates are available for a wide variety of deposit products such as CDs, checking/saving accounts, and money market accounts with different minimum account sizes and maturities. Following Drechsler, Savov, and Schnabl (2017) and Lin (2020), we focus on one of the most popular products - the 12-month CD product with a minimum account size of $10,000. We take the average weekly rates at the branch level during June and July, when the level of deposits in SOD is reported. Then, we construct the following variables: a rate setter dummy ($Ratesetter$) that equals one if a branch is a local rate setter and zero otherwise; a better rate dummy ($Betterrate$) that equals one if the average rate of 12-month CD products is higher than the county median, and zero otherwise.

2.4 Bank Balance Sheet and Local Characteristics

A bank's financial conditions such as lending opportunities, profitability, liquidity, and sensitivity to interest rate risk can affect how depositors perceive the bank. Thus, we control for a rich set of bank financial variables using the data from the Consolidated Report of Condition and Income (Call Report), and calculate the following variables: a small bank dummy ($Small$) that equals one for banks with assets less than two billion dollars, and zero otherwise; the share of loans in total assets ($Loan$); return on assets (ROA) measured as the ratio of annualized net income to gross total assets; liquidity ($Liquidity$) measured as bank cash divided by total deposits; sensitivity to interest rate risk ($Sensitivity$) measured as the ratio of the absolute difference between short-term assets and short-term liabilities to gross total assets. Lastly, we collect the following county-level variables from the American Community Survey (ACS) conducted by the U.S. Census Bureau: income per capita ($Income$) and mobility ($Mobility$) measured as the number of people who migrate into and out of the county between 2014 and 2018. Detailed variable definitions are available in Appendix A1. We merge the Google rating data, SOD data, and RateWatch data using the bank branches'

RSSD ID, and then merge with the parent banks' reputation shock data and financial data using the parent banks' RSSD ID. Our final merged sample includes 216,241 reviews of 35,978 distinct branches owned by 177 distinct banks.

2.5 Summary of Statistics

In Figure 1A, we visualize the average ratings of bank branches by county. We show that the average ratings of bank branches in the Midwest are higher than those in the Northeast and the West. Also note that branches in suburban and rural areas enjoy higher ratings than those in urban locations. In Figure 1B, we map out the ratings for the largest and second-largest US banks by deposits - JP Morgan Chase Bank and Bank of America. Notably, there is significant overlap in their branching networks, and JP Morgan Chase is rated higher in most of the counties. This observation is consistent with the findings in the U.S. National Banking Satisfaction Study (J.D. Power 2019). While the geographic heterogeneity in customer experience is interesting, we apply branch or bank-county FEs in the regression analysis to preclude the impact of time-invariant bank-location-level omitted variables.

Figure 2A documents the number of reviews by quarter. The growing popularity of Google Map over the past few years has brought the quarterly number of reviews on bank branches from 2,142 in 2014Q1 to 53,301 in 2019Q3. However, the influx of new users and new reviews do not significantly change how people rate their banking experiences over time. Figure 2B shows that the national average branch ratings are largely stationary. It is not surprising given that how banks interact with customers is unlikely to change significantly during our sample period. In light of this observation, we exploit the cross-sectional heterogeneity in customer ratings of banking experience, rather than the time-series variation, in our empirical analysis.

Table 1 reports the summary statistics for the Google rating-RepRisk-SOD-RateWatch-merged sample which we employ to examine the reaction of bank deposit-taking to reputation shock and customer satisfaction. The sample is constructed at the year-branch level. The table shows that an average branch-year has 3.178 reviews with an average rating of 3.389. The variation of rating is also sizable. The relative standard deviation of rating is 0.287 (0.972/3.389=0.287). The mean value for the four reputation shock dummies varies from 18.7% to 55%. The average annual deposit growth of branches in our sample is at 9.4%.

3 Decomposing Customer Satisfaction

Our study presents a novel empirical framework to understand how banks interact with customers at the branch level. In Section 3.1, we estimate branch-level Google ratings in additional specifications and summarize all of their R^2 values to capture the exact portion of the variation in ratings that can be explained by observable characteristics. Section 3.2 examines the determinants of branch-level Google ratings. Section 3.3 presents external validity tests on this novel measure and show that variations in ratings significantly correlate with how customers feel. In Section 3.4, we decompose the ratings along four important and interpretable dimensions - two that center on the interpersonal aspects of the banking experience (accessibility and hospitality), and another two that focus on customer perception about in-branch product offerings and facility.

3.1 Variance Decomposition

We regress ratings on a rich set of fixed effects to examine if geographic or bank-specific characteristics explain the variations. In Panel A of Table 2, the dependent variable is the Google ratings of bank branches observed on December 31, 2019. The bank, county, and bank by county FEs are included in columns 1, 2 and 3 respectively. Results show that bank FEs explain 22.7% of the variations and county FEs explain only 9%. Column 3 shows that 70.4% of the variations remain unexplained for branches owned by the same bank and operating in the same county. Note that we drop singletons when applying high dimensional fixed effects (Correia 2016), which leads to a drop in the sample size when a greater number of FEs are included.

In Panel B of Table 2, the dependent variable is the Google ratings of bank branches observed in every year-end from 2014 to 2019 (if non-missing). Time-varying bank and county FEs (column 4 and 5) only explain 13.8% and 6% of the variations. Branch FEs (column 3, 7 and 8) explain more than 67% of variations, leaving around 30% of variations attributable to within-branch time-series characteristics.

There are two key takeaways from the variance decomposition analysis. First, banking experiences vary significantly among branches, even across those operated by the same bank in the same county. The Google profiles provide a unique opportunity for us to examine the banking experiences at each individual branch. Second, we confirm that branch-level google ratings are largely stationary during our sample period (i.e., 2014-2019). Guided by these findings, we focus on exploiting the cross-sectional heterogeneity, rather than time-series variations, following exogenous shocks to bank deposit-taking.

3.2 Determinants of Branch-Level Ratings

In this section, we examine the determinants of branch-level ratings by estimating the following cross-sectional regression:

$$Rating_i = \beta_0 + \beta_1 X_k + \beta_2 D_{k,j} + \beta_3 T_i + FE + \varepsilon_i, \tag{1}$$

where i indexes branches, k banks and j counties. The regression is performed at the branch level. The dependent variable is the Google rating of branch i observed on December 31, 2019.[10] The independent variables consist of bank (X_k), bank-county ($D_{k,j}$), and branch (T_i) level characteristics. Bank characteristics include the *Small* and *Local* dummies. Bank-county characteristics include the *Important* dummy. Branch characteristics include *Countyshare*, *Ratesetter*, *Betterrate*, *Samestate*, and *New* dummies. We include county FEs to capture any unobservable time-invariant county-specific determinants of customer satisfaction or service quality. Standard errors are clustered at the county level. [11]

Regression results are reported in Table 3. Within the same county, branches owned by local and small banks have higher ratings than those owned by larger and/or national banks. *Countyshare* and *Important* are negatively and significantly correlated with ratings, which indicate that branches facing less local competition have lower ratings. *Betterrate* is positively and significantly related to ratings, suggesting that pricing factors are key determinants of customer satisfaction. Finally,

[10] Our analysis is based on the cross-sectional regression of branch ratings on observable characteristics. We didn't regress yearly ratings using the panel data because both the ratings and independent variables are quite stationary over time (see Table 2 Panel B for details), which inevitably inflate the t-statistics of our coefficient estimation.

[11] The reason that we use a cross-sectional regression instead of a panel regression is because both the dependent and independent variables are highly stable across time. Using a panel regression in this case would lead to inflated and spurious t-statistics. Thus, we employ a more conservative cross-sectional regression.

we show that recently established branches, and branches located in the headquarter state have higher ratings. Newer branches generally feature better location and upgraded facilities, which are preferred by customers. Consistent with Deng and Elyasiani (2008) and Goetz, Laeven, and Levine (2016), branches located in the home state are more intensively monitored by headquarter offices, leading to fewer agency problems and a higher quality of customer service.

3.3 External Validation

A valid measure of customer engagement should quantify how customers perceive the banking services at the micro level. This section presents external validation of our measure. Our analysis confirms that branch-level ratings effectively capture how customers perceive during their banking experiences.

Using the sentiment word list developed by Loughran and McDonald (2011), we calculate the weights of positive and negative words in each review (frequency of mentions over the length of the review). We then obtain the average weight of positive (*Positive*) and negative (*Negative*) words for the reviews received by the same branch as of December 31, 2019, and use them as the proxies for the overall sentiment of customer interactions at the branch.

We also employ the NRC Emotion Lexicon to detect eight emotions of the reviewers: anger, anticipation, disgust, fear, joy, sadness, surprise and trust. Similar to the sentiment analysis above, we calculate the weights of words related to each emotion and then average the weights across all reviews received by the branch as of December 31, 2019 to obtain the eight emotion variables: *Anger*, *Anticipation*, *Disgust*, *Fear*, *Joy*, *Sadness*, *Surprise*, and *Trust*.

In Table 4, we regress the branch-level ratings on the proxies for sentiment and emotions. The coefficient on *Positive* is positive and significant at the 1% level, while the coefficient on *Negative* is negative and significant at the 1% level. The coefficients on emotion metrics including *Trust* and *Joy* are positive and significant at the 1% level, while those on *Sadness*, *Disgust*, *Fear*, and *Anger* are negative and significant at the 1% level. The results collectively demonstrate that branch-level ratings are an effective measure of how customers perceive their banking experiences.

3.4 Banking Experience: Accessibility, Hospitality, Product and Facility

In this section, we disentangle customer experience along four dimensions. Following Li et al. (2020), we apply Word2Vector, a semi-supervised approach, to extract topics from unstructured textual documents. The Word2Vector algorithm is a word-embedding model that identifies, for each seed word provided by the user, an expanded set of synonyms (Mikolov et al. 2013). It is based on a simple, time-tested concept in linguistics: Words tend to co-occur with neighboring words with similar meanings (Harris 1954). For a given word, the algorithm searches for the neighboring words in the textual document and creates a vector matrix consisting of the frequencies of each neighboring word. The model then applies a neural network to reduce the dimension of the matrix to a fixed number. The similarity between any two words in the document can be calculated as the cosine product between the two corresponding vector representations. Lastly, the algorithm performs the bootstrapping process to iteratively associate words gleaned from the document to each seed word. The most similar words (i.e., those with the highest cosine similarity) are considered as an expanded dictionary to the original set of seed words provided by the user.

Specifically, we first sort all words in the reviews by their corresponding frequencies of mentions and manually screen the top 1,000 words through many iterations to understand what customers value. We select five seed words with similar meanings along each dimension from this high-frequency word list. The seed words for Topic 1 consist of "location", "call", "line", "wait", and "time", which we label as topic *Accessibility*. The seed words for Topic 2 consist of "communicate", "assist", "experience", "greet", and "solve", which we label as topic *Hospitality*. The seed words for Topic 3 consist of "checking", "mortgage", "investment", "rate", and "fee", which we label as topic *Product*. The seed words for Topic 4 consist of "building", "lobby", "parking", "facility", and "atm", which we label as topic *Facility*. This step leaves us 20 seed words associated with four topics.

For each seed word, we follow Li et al. (2020) to obtain its unique neighboring words in the reviews and their corresponding frequencies of mentions. We define neighboring words as the five words before and after the position of the seed word. The information about neighboring words is condensed into a 100-dimensional vector. We associate the seed words with the words gleaned from the reviews to calculate their cosine products – the top 20 closest synonyms of each seed word from the same topic are pooled together as the expanded dictionary for the topic. In Appendix A3, we report the top words in the expanded dictionary for each of the four topics.[12]

We define the branch-level sub-component ratings on each of the topics as follows:

$$Topic\ Rating_{i,t,d} = \frac{\sum_{n \in N_{i,t}} \left(\frac{SR_{i,t,n} \times \sum_{b \in B_{i,t,n}} (\P[b \in TD])}{B_{i,t,n}} \right)}{N_{i,t}}, \tag{2}$$

where i indexes branches, t refers to the year of observation of Google profiles, n refers to the review left in the Google profiles, and d refers to the specific topic (*Accessibility*, *Hospitality*, *Product*, or *Facility*). For each branch, we first obtain a list of reviews observed as of the year, $n = 1, 2, ..., N_{i,t}$. For each review, we further break it into a list of words (unigrams), $b = 1, 2, ..., B_{i,t,n}$. TD is the expanded dictionary of the topic. $\P[\cdot]$ is the indicator function. We adjust the ratings by subtracting the constant number three, re-centering the scale of ratings as [-2, 2] instead of [1, 5], which we denote as $SR_{i,t,n}$. In essence, we calculate the topic loadings in each review (i.e., the frequency of mentions of words in the topic-specific expanded dictionary over the total number of words in the review), and then multiple it with the scaled rating ($SR_{i,t,n}$) assigned to the review. Using the ratings with a re-centered scale, we ensure that this product will be negative if customers rate the branch below the median of the original rating scale (less than 3 out of a [1,5] scale), and positive if otherwise.[13] Lastly, we aggregate the product and divide the sum by the total number of reviews observed as of year t. By construction, the branch's rating on *Accessibility* is highest if all ratings are highest, and the reviews only consist of the words from the *Accessibility* expanded dictionary.

In the robustness section, we also construct the following bank-county level sub-component

[12] The words in each expanded dictionary are ranked by their $Fweight$, which is reported in column 2 of Appendix A3. For each occurrence of the word, we scale it by the length of the review. $Fweight$ is the sum of all scaled occurrences of the word in the reviews. We also provide the un-scaled number of occurrences, namely the total frequency of mentions $Freq$ in the third column.

[13] Note that subtracting all ratings by a constant number does not change the time-series or cross-sectional variations in branch ratings. This step, however, reduces the measurement errors in the product which is used to construct the decomposed topic ratings.

ratings which are used in regressions of mortgage lending on natural disasters:

$$Topic\ Rating_{k,j,t,d} = \frac{\sum_{n\in N_{k,j,t}} \left(\frac{SR_{k,j,t,n} \times \sum_{b\in B_{k,j,t,n}} (\P[b\in TD])}{B_{k,j,t,n}} \right)}{N_{k,j,t}}, \tag{3}$$

where k indexes bank, and j refers to county. Other variables are defined in the same way as described in the branch-level measure.

4 Empirical Strategies and Results

4.1 Reputation Shock and Deposit-Taking

We first validate the premise of using reputation damage as a shock to branch deposit growth. As described in Section 2.2, we use the RepRisk data to create four indicator variables (*Rep_10*, *Rep_15*, *Rep_20*, and *Rep_25*) that correlate with greater intensity of negative news coverage of the bank. We posit that a branch's deposit flow is a function of its perceived reputation among customers. We validate the relationship between these novel proxies of reputation shocks and branch-level deposit growth using the following specification:

$$\begin{aligned} Depositgrowth_{i,t} =& \beta_0 + \beta_1 Rep\_J_{k,t} + \beta_2 Deposits_{i,t-1} + \beta_3 Betterrate_{i,t-1} \\ &+ \beta_4 Countyshare_{j,k,t-1} + \beta_5 Loan_{k,t-1} + \beta_6 ROA_{k,t-1} \\ &+ \beta_7 Liquidity_{k,t-1} + \beta_8 Sensitivity_{k,t-1} + FE + \varepsilon_{i,j,k,t} \end{aligned} \tag{4}$$

where i indexes branches, k banks, j counties, and t years. The regression is performed at the branch-year level. The dependent variable is the growth in deposits at branch i in year t. The independent variable, *Rep_J*, is an indicator variable that equals one if the RepRisk Index of the bank which owns the branch increases more than J (J=10, 15, 20, or 25) over the past year, and zero if the increase is less than 10. To alleviate concerns about the choice of cutoffs to create the dummy variables, we also replace the dummies with *Rep_Chg*, the bank's maximum jump in the RepRisk Index over the past year. In our regressions, we also include several lagged time-varying branch-level and bank-level control variables: *Deposits*, *Betterrate*, *Countyshare*, *Loan*, *ROA*, *Liquidity*, and *Sensitivity*.[14]

Literal use of FEs in the regression further mitigates the confounding effects from other omitted variables. Branch FEs preclude the impacts of time-invariant branch-level characteristics, such as its location/proximity to area businesses. County by year FEs rule out the possibility that any county-specific events, such as political elections or natural disasters, bias our estimation. Standard errors are clustered at the branch level.

In constructing the sample, we only include the branches that have non-missing reviews in a given year. Also note that the sample size decreases in J because we exclude the observations whose RepRisk Index jumps between 10 and J, from column 2 to column 4. In this way, we keep the control group constant, which consists of observations whose RepRisk Index changes less than 10. This allows us to sensibly compare the economic magnitudes of coefficient estimates while holding the benchmark largely constant across the four regressions.[15]

Regression results are presented in Table 5. The coefficients of Rep_J are negative and statistically significant at the 1% level in all columns. Economically, according to column 4, the deposit growth at branches of banks with a reputation shock (Rep_25=1) is lower than other branches by 2.6%, equivalent to 27.66% of the mean deposit growth (9.4%). We present the magnitudes of the coefficients and their corresponding 95% confidence interval in Figure 3. We show that the economic magnitudes of the coefficient estimates increase in J, which is consistent with our hypothesis that greater intensity of negative news coverage is associated with more deposit outflows. Our results confirm the impact of negative ESG performance on depositors' funding decisions, consistent with Homanen (2018) and Chen, Hung, and Wang (2019).[16]

[14] We exclude *Samestate*, *New*, and *Ratesetter* from control variables because there is no (or little) variation in these variables within the the fixed effects cluster (branch).

[15] In Appendix A4, we keep the sample size constant across all regressions. Our main results are robust to these variations - the direction of association is the same, and the statistical significance also holds in all columns except column 3.

[16] Despite the restrictive fixed effects in our model, it is likely that the results are predominantly driven by a few major banking scandals in recent years, e.g., Wells Fargo account fraud scandal. To check the robustness of our results, we drop the Wells Fargo reputation shocks from the RepRisk sample and rerun the regressions. The results are quantitatively and qualitatively similar.

4.2 Customer Satisfaction and Deposit-Taking

Building on the findings from Homanen (2018) and the results in Section 4.1, we adopt an event study approach to further examine how customer satisfaction mitigates deposit outflows. We focus on the banks that have experienced a reputation shock and exploit the heterogeneous Google ratings among branches operated by the same affected bank in a given area. The specification is as follows:

$$\begin{aligned} Depositgrowth_{i,t} =& \beta_0 + \beta_1 Rating_{i,t} + \beta_2 Num\_Reviews_{i,t} + \beta_3 Deposits_{i,t-1} \\ &+ \beta_4 Betterrate_{i,t-1} + \beta_5 Countyshare_{k,t-1} + \beta_6 New_{i,t} \\ &+ \beta_7 Ratesetter_{i,t} + FE + \varepsilon_{i,k,t} \end{aligned} \tag{5}$$

where i indexes branches, k banks, j counties, and t years. The regression is performed at the branch-year level, where branches are the branch offices of banks that has experienced a reputation shock in the past 12 months (an increase in RepRisk index of at least 10, 15, 20, or 25). We employ a restrictive specification using bank by county by shock year fixed effects, which compare the deposit flows of branches within a same bank and within a same county during the year of the reputation shock. The analysis precludes the impacts of any time-varying bank and county characteristics on deposit flows. Standard errors are clustered at the bank by county by shock year level.

Regression results are reported in Table 6. Each column focuses on a subsample of banks that experienced reputation shocks of different intensities (J=10, 15, 20, or 25, respectively). We find that deposit growth increases with the branch-level Google rating. The coefficient estimates on *Rating* are positive and statistically significant at the 1% to 10% levels in columns 1 to 4. The economic magnitudes are sizable. Take column 2 for example, conditional on the reputation shock, a one-standard deviation increase (0.972) in the branch's ratings is associated with a 3.89% increase in deposit growth ($0.972 \times 0.004 = 0.03888$). Chen, Hung, and Wang (2019) show that negative bank social performance reduces depositors' willingness to finance the bank by decreasing their trust in banks. Our results suggest that, better customer relationship built at the branch-level may help mitigate the negative impact of loss of trust due to reputation damage.

Also look at the coefficients of the time-varying branch-level control variables. Deposit growth decreases with *Deposits* and increases with *New*, consistent with the expectation that new and smaller branches may experience a higher percentage growth in deposits. Deposit growth increases with *Betterrate*, indicating that branches that offer higher deposits rates attract more deposits. We don't observe a definite relationship between the number of Google reviews and deposit growth.

4.3 Sub-sample Analysis by County Characteristics

In this sub-section, we explore the heterogeneous effects of customer satisfaction across counties with different socioeconomic conditions. Specifically, we follow Equation 5 to conduct sub-sample analysis based on two crucial county-level characteristics: income per capita and population mobility.[17] Local income per capita serves as a proxy for customer wealth. Population mobility is indicative of the compositions of existing and new customers. Arguably, customer satisfaction may be more relevant in low-mobility markets where there are longer-standing relationships. Alternatively, to the extent that newer customers rely more heavily on Google ratings to select banks, then we might instead find stronger effects in high-mobility counties. In Panel A of Table 8, we conduct the regressions among branches that are located in counties where the income per capita is in the top 25% and bottom 25% nationally. In Panel B of Table 8, we conduct the regressions among branches that are located in counties where the population mobility is in the top 25% and bottom 25% nationally.

We find that the relationship between branch rating and deposit growth is only significant in high income per capita and low mobility counties. The findings are consistent with our hypothesis that the economic value of non-price factors, such as customer satisfaction, increases in customer wealth. Furthermore, given that a resilient retail banking relationship builds on retaining existing customers and/or attracting new customers, our results indicate that customer satisfaction may have a larger impact on the behaviors of existing customers.

[17] In order to fully control for unobservable bank-county level omitted variables, we are unable to conduct the cross-sectional analysis by including an interaction term. County-level dummy variables will be fully absorbed by high dimensional bank-county-shock FEs.

4.4 Sub-sample Analysis by Bank Size

We further posit that the links between branch Google ratings and deposit growths vary between community banks and large regional or super-regional banks. Specifically, we examine how the links between ratings and deposit growth vary among banks with domestic assets in the bottom 25% (maximum assets $0.8 billion) and those with domestic assets in the top 25% (minimum assets $3.36 billion) of all banks in a given year. In Table 9, we show that the relationship is stronger for smaller banks, but only when we consider cases with severe reputation damage (Rep_20 and Rep_25). An untabulated test shows that our results are robust to alternative cutoffs of $1 billion for community banks and $10 billion for large banks. Our results suggest that customer satisfactions may be more crucial for community banks that heavily rely on relationship banking and soft information accumulated through effective interactions with local customers and businesses.

4.5 What do Customers Really Value: The Heterogeneous Effects of Review Topics

In this sub-section, we employ customer ratings on topics of *Accessibility*, *Hospitality*, *Product*, and *Facility*, and examine what aspects of the banking experience are most likely to help branches mitigate deposit outflows in the wake of reputation shocks. We investigate this question using the following specification:

$$\begin{aligned} Depositgrowth_{i,t} = &\beta_0 + \beta_1 Topic\ Rating_{i,t,d} + \beta_2 Num\_Reviews_{i,t} + \beta_3 Deposits_{i,t-1} \\ &+ \beta_4 Betterrate_{i,t-1} + \beta_5 Countyshare_{k,t-1} + \beta_6 New_{i,t} \\ &+ \beta_7 Ratesetter_{i,t} + FE + \varepsilon_{i,k,t} \end{aligned} \tag{6}$$

where i indexes branches, k banks, d topics, and t years. The regression is performed at the branch-year level, where branches are the branch offices of banks that has experienced a reputation shock in the past 12 months (an increase in the RepRisk Index by at least 10, 15, 20, or 25) and years represent the years in which the reputation shocks occurred. $Rating_{i,t,d}$ is the average google rating of branch i on topic d as of year t, whose calculation is presented in Equation 2. Other variables and fixed effects are defined in the same way as those in Equation 5. Standard errors are clustered at the bank by county by shock year level.

The results are presented in Table 7. To conserve space, we compress the control variables but we do present the full table in Appendix A5. We find that the coefficient on *Accessibility* is positive and statistically significant in all four specifications, while the coefficient on *Facility* is positive and statistically significant when the reputation shock is relatively weaker (*Rep*_10=1 or *Rep*_15=1). The effects are also economically sizable. Take column 2 as an example: for each one standard deviation increase in *Accessibility* (0.114), the deposit growth increases by 0.425% (3.731 $\times$ 0.114 = 0.425), equivalent to 4.52% of the mean deposit growth (9.4); for each one standard deviation increase in *Facility* (0.068), the deposit growth increases by 0.173% (2.550 $\times$ 0.068 = 0.173), equivalent to 1.84% of the mean deposit growth (9.4). The coefficients on other topics are not statistically significant. Berger, Kravitz, and Shibut (2021) argue that depositors, especially the uninsured ones, are responsive to the condition of their local banks because they wish to minimize the potential for convenience losses in the event of bank failure. Our results suggest that the additional convenience offered by local branch offices, through accessible services and superior facilities, is helpful to mitigate the concerns about convenience losses and retain depositors in the wake of reputation shocks.

5 Robustness and Discussions

In this section, we discuss and extend our main results. Section 5.1 discusses a potential selection issue and address it using an alternative empirical framework. In Section 5.2, we use the most granular data to control for remaining heterogeneity across neighborhoods in the same county. Section 5.3 analyzes if deposits flow out of the banking sector following reputation shocks. Lastly, Section 5.4 provides further robustness on the relationship between customer satisfaction and durability of retail relationship, by 1) examining the other important retail banking function - mortgage lending, and 2) exploiting an alternative positive shock on business expansion instead of a negative shock on deposit retention.

5.1 A Difference-in-Differences Approach for Google Ratings and Deposit Growth

In Section 4.1, we utilized an event study approach to investigate whether branches with stronger levels of customer satisfaction are better able to withstand negative shocks to reputation. The event study approach allows us to exploit the heterogeneous Google ratings among branches of the same bank in the same county and investigate the effect of customer satisfaction on deposit growth. However, the bank by county by shock year fixed effects require each bank to have at least two branches within a county. This can introduce a sample selection bias in that for those bank-county pairs that have only one branch, customers might have limited choices of where to move their deposits (if they want to move their deposits across branches within the same bank and county). In this case, the effects of reputation and customer satisfaction on deposit growth may be limited, such that our previous results may overestimate their mean effects.

In this sub-section, we use a difference-in-differences approach to resolve the sample selection issue. Instead of comparing the deposit growth rates between branches of the same bank, we compare branches of banks that are subject to reputation shocks with those that are not. This allows us to loosen the restrictions on the number of branches per county-bank pair, and examine the joint effect of customer satisfaction and reputation shock on deposit growth.[18] The treatment group includes all branches of banks that experienced a reputation shock in the past 12 months. For each treatment branch in each treatment year, the control branches include all branches located in the same county that did not experience any reputation shocks in the past 12 months. We consider a treatment branch and its control branches as a cohort. We estimate the joint effect of customer satisfaction and reputation shock using the following specifications:

$$\begin{aligned} Depositgrowth_{i,j,k,t} =& \beta_0 + \beta_1 Rep\_J_{k,t} \times Rating_{i,j,k,t} + \beta_2 Rep\_J_{k,t} + \beta_3 Rating_{i,j,k,t} \\ &+ \Lambda \times C + FE + \varepsilon_{i,j,k,t}, \end{aligned} \tag{7}$$

[18]The reasons that we do not use the difference-in-differences approach in the main regression are two folds: i) It cannot effectively rule out the effects of time-varying bank-county-specific omitted variables. It is impossible to include fixed effects at the finest level given the treatment dummy is constructed at the bank-year level; ii) In counties with very few banks or are monopolized by a large bank, it can be hard to find sensible control group that are comparable with the treatment banks. Unlike experiments that build on natural disaster where adjacent counties are sensibly considered as controls that could, but wasn't affected, identifying branches that could/should, but wasn't affected by a reputation event is extremely hard. Thus, we are just using a difference-in-differences approach as a robustness check to our results.

where C is a vector of control variables of branch and bank characteristics including *Num_Reviews*, *Deposits*, *Samestate*, *New*, *Ratesetter*, *Ln*(*Loan*), *Betterrate*, *Countyshare*, *ROA*, *Liquidity*, and *Sensitivity*. We employ the cohort fixed effects to eliminate any county-year-specific determinants of deposit growth. The standard errors are clustered by cohort.

Appendix A6 presents the results based on the estimation of Equation 7. The results reconfirm our main findings. The coefficients on *Rep_J* are negatively significant at the 1% level (except for J=25), suggesting a negative impact of reputation shock on deposit growth. This is consistent with the results in Table 5. The interaction term between *Rep_J* and *Rating* is positively significant at the 1% or 10% levels, reconfirming that higher levels of customer satisfaction help mitigate the negative impact of reputation shocks on deposit growth. This confirms the results we find in Table 6. In sum, both the event study method in Section 4 and the difference-in-differences approach in this sub-section show that customer satisfaction plays a crucial role in helping branches withstand shocks to bank reputation.

5.2 Heterogeneity in Within-county Branch Resources Allocations

The restrictive bank by county by shock year fixed effects in Equation 5 preclude the impacts of any time-varying bank and county specific characteristics on deposit flows. However, even within the same county, there could still be large variations in neighborhood characteristics (e.g., customer wealth). For example, within New York County (Manhattan), the income average for the top 1% is more than 110 times that of the bottom 99%.[19] If banks allocate more internal resources to enhance the customer experience (i.e., branch ratings) of those branches in more affluent neighborhoods compared to other branches, and if this internal resource allocation is correlated with deposit growth in the wake of reputation shocks, then our results may be confounded.

Although it is difficult to fully control for all the neighborhood characteristics that can affect both deposit growth and customer experience, we try to alleviate this concern by analyzing within even more granular geographical areas. Specifically, we apply bank by ZIP Code by shock year fixed effects to Equation 5. This compares the deposit flows of branches within a same bank and

[19]See https://www.mytwintiers.com/news-cat/infographic-shows-new-york-has-the-highest-income-inequality/.

within a same ZIP Code in the year of reputation shock. Appendix A7 reports the results. As expected, the more restrictive set of fixed effects significantly reduces the sample size, as they require a bank to have at least two branches within a same ZIP Code. The coefficient on *Rating* remains statistically significant at the 1% to 10% levels across reputation shocks of various severity, and the economic magnitude still remains sizable. This result indicates that our previous results in Table 6 are less likely to be confounded by geographical variations in neighborhood characteristics and internal resource allocations.

5.3 Do Deposits Flow out of the Banking Sector Following Reputation Shocks?

As deposits flow out of the banks that are subject to reputation shocks, it remains unknown where the deposits flow to. The depositors can either move their deposits out of the county to other branches of the same bank or other banks, or within the county to neighboring or competitor banks, or both. Directly measuring the change in the deposit growth at each individual neighboring or competitor bank branch is difficult because of the possibility of multiple destinations of deposit flow. Instead, we can test whether the deposits still stay within the same county, or flow to other areas, by examining the response of county-level deposits to reputation shocks occurring on the banks in the county. If the county-level deposit growth significantly drops after reputation shocks, then it is evident that deposits flow out of the focal county to other areas. If there is no significant change in the county-level deposit growth, then it is more likely that deposits flow within the county to neighboring or competitor banks. We construct a set of county-year variables, $Share\_Rep\_J$, to measure the share of branch deposits in a county that are subject to a jump in the RepRisk Index by more than J (J=10, 15, 20, or 25), and run the following panel OLS regression of county-level deposit growth on these variables individually:

$$Depositgrowth\_County_{j,t} = \beta_0 + \beta_1 Share\_Rep\_J_{j,t} + \Lambda \times C + FE + \varepsilon_{j,t}, \tag{8}$$

where $Depositgrowth\_County_{j,t}$ is the county-level deposit growth. C is a vector of control variables including lagged total deposits of all bank branches in the county ($Deposits\_County$), the number of bank branches per 1,000 population in the county ($Branch$), and the Herfindahl-Hirschmann index of deposits at all bank branches in the county (HHI). We employ county fixed effects and year fixed effects to control for any unobservable county-specific or time-specific determinants of deposit growth. Standard errors are clustered at the county level.

Appendix A8 presents the results of Equation 8 estimation. We find no significant change in the county-level deposit growth following reputation shocks of various severity, indicating that it is less likely that deposits flow out of the focal county following reputation shocks. Instead, deposits may flow within the county to neighboring or competitor banks.

5.4 The Impact of Customer Satisfaction on Mortgage Demand

While we have shown that higher customer satisfaction helps retain depositors in the wake of reputation shocks, the effect of customer satisfaction on the durability of retail banking relationships can go above and beyond depositor retention to borrower growth. In this sub-section, we examine another important retail banking product on the asset side - residential mortgage. We investigate whether customer satisfaction helps attract more mortgage businesses using natural disasters as positive shocks to local residential mortgage demand.[20]

We collect annual mortgage application data for the period between 2014 and 2018 from the Home Mortgage Disclosure Act (HMDA) database and Presidential Disaster Declaration data from the Federal Emergency Management Agency (FEMA).[21] As the HMDA does not disclose which specific branch a mortgage application is submitted to, we collapse the Google rating data and HMDA data into bank-county-year level observations.

Our empirical design follows Dlugosz et al. (2019) and involves a triple difference-in-differences approach. A county is considered a treatment county if it was hit by at least one natural disaster in a given year (treatment year) and didn't experience any disaster in the preceding year (control year). A county is considered a control county if it is located in the same state but didn't experience any disaster during the two-year window. We consider all the disasters occurred in a given state in a given year as one disaster event, and the treatment counties and their control counties in the

[20] Natural disasters can generate positive shocks to local mortgage demand because disaster-affected residents must rebuild or replace their damaged homes and businesses (Cortés and Strahan 2017). For example, a household may take a refinancing loan which converts home equity into cash in order to pay for home repair.

[21] FEMA reports all natural disasters declared by the President of the United States and provides associated information including incidence dates, disaster types, and affected counties, etc. The property damages caused by the President-declared disasters are generally severe, so that they are likely to create shocks to mortgage demand.

disaster event as a cohort.[22] We compare mortgage applications to higher-rated banks and lower-rated banks in treatment and control counties during the two-year event window. Specifically, we estimate the following regression:

$$
\begin{aligned}
Ln(Applications)_{j,k,c,t} = & \beta_1 Treat_{j,c} \times Post_{c,t} \times High\ Rating_{j,k,t} + \beta_2 Treat_{j,c} \times Post_{c,t} \\
& + \beta_3 Treat_{j,c} \times High\ Rating_{j,k,t} + \beta_4 Post_{c,t} \times High\ Rating_{j,k,t} \\
& + \beta_5 Treat_{j,c} + \beta_6 Post_{c,t} + \beta_7 High\ Rating_{j,k,t} + \Lambda \times C \\
& + FE + \varepsilon_{j,k,c,t},
\end{aligned}
\tag{9}
$$

where j indexes counties, k banks, c cohorts, and t years. $High\ Rating_{j,k,t}$ equals one if bank k's average Google rating in county j in year t is at the top tercile among all bank-county-year observations within the state, and zero if it is at the bottom tercile. $Treat_{j,c}$ is a dummy variable that equals one for treatment counties and zero for control counties within a cohort c. $Post_{j,c}$ is a dummy variable that equals one for the disaster incidence year and zero for the prior year within a cohort. We include a list of control variables in C to ensure that $High\ Rating$ captures a distinct feature of banks not subsumed by other bank or local factors. We control for bank balance sheet variables, local bank density and concentration variables, socioeconomic and demographic characteristics, as well as the one-year lagged dependent variable (see a full list of control variables in Appendix A9 and their definition in Appendix A1).

We employ several alternative fixed effects to control for unobservable factors. First, we employ disaster year by state fixed effects (cohort fixed effects), which ensures that we compare mortgage demand across banks within economically and socially similar areas and closer time windows. Second, we employ disaster year by state by bank fixed effects. This allows us to do a within bank comparison of the mortgage demand within economically and socially similar areas and closer time windows. Third, we use bank by county fixed effects to control for any unobservable time-invariant local and bank characteristics. Finally, we add bank by county fixed effects to disaster year by state fixed effects. We cluster the standard errors at the county by bank level.

Appendix A2 reports the summary statistics for the Google rating-FEMA-HMDA-merged sam-

[22] The control counties are matched with treatment counties with replacement. For example, for a control county, a same county-year observation can show up in one cohort as a treatment year observation, and show up again in another cohort as a control year observation.

ple. Panel A shows the frequency of natural disasters and the number of affected counties. Our sample includes 138 unique natural disasters and 1,485 affected counties. On average, 10.76 counties were hit by each disaster. The three most frequent types of disaster are severe storms, floods, and hurricanes, followed by snow and tornado. Panel B shows that for all banks in our sample, 24.8% of them have a Google rating that is at the top tercile in its state, while 75.2% are in the bottom tercile. The average number of reviews for an average bank in an average county in a given year is 8.120.[23] The annual number of mortgage applications for an average bank in an average county is 82 applications.

Table 10 reports the results of Equation 9.[24] The results show that mortgage demand increases following natural disasters. The coefficient on $Treat \times Post$ is positive and statistically significant at the 1% level and is robust across alternative choices of fixed effects. Economically, taking column 1 as an example, the number of applications surged by 35.5% for lower-rated banks in treatment counties. This result validates the premise of using natural disasters as demand shocks. We also find that following natural disasters, the increase in loan applications to highly-rated banks is higher than that to lower-rated banks. The coefficient on $Treat \times Post \times High\ Rating$ is positive and statistically significant at the 5% or 1% level and is robust across alternative fixed effects. The estimate is also economically significant. Taking column 1 as an example, the increase in the number of applications to highly-rated banks exceeded that to lower-rated banks by 20.3%. In sum, the results in Table 10 suggest that in cases of demand shocks, mortgage borrowers strongly prefer those branches with higher customer satisfaction to seek funding.[25]

We then investigate what aspects of branch services are more crucial to mortgage borrowers' choices of lending bank, we utilize the topic rating measures as defined in Equation 3 to explain borrowers' mortgage applications to each bank after natural disasters. To allow easier interpretation

[23] Note that the average number of reviews at the branch level is 3.178 (See Table 1). While too few number of reviews at the branch level may subject the results to measurement errors, this concern is less relevant in this test, given that we have a higher number of reviews at the bank-county-year level.

[24] We compress the control variables to conserve space, and show the complete table in Appendix A9.

[25] A concern about our result could be that it is driven by households who seek to refinance their mortgages after natural disasters. They may just go to their original bank with whom they have an existing mortgage lending relationship, instead of going to a high customer satisfaction bank for financing. Since those highly-rated banks originally had more mortgage businesses (the significant coefficient on the stand-alone term *High Rating* in Table 10), the coefficient on the triple-interaction term may just reflect those returning households who originally had mortgage contracts with the bank. To alleviate the concern, we re-estimate Equation 9 separately on new purchases loans and refinance loans, because new home purchase borrowers are less likely to be returning customers. We report the results in Appendix A11. Our results are robust to the choices of different loan purposes.

of the triple difference-in-differences results, we construct a $High\ Topic\ Rating$ dummy that equals one if a banks average Google rating in a given county in a given year on given topic is at the top tercile among all bank-county-year observations of that topic in the state, and zero if it is at the bottom tercile. Then, we estimate a similar difference-in-differences regression as Equation 9 but replace $High\ Rating$ with $High\ Topic\ Rating$.

Appendix A10 presents the results. We find that loan demand is higher for banks who have high ratings on hospitality and facility. The coefficients on the triple interaction term are statistically significant at the 1% level for hospitality and facility. The economic magnitude is also sizeable. Taking column 6 as an example, the increase in the number of applications to banks with high hospitality rating exceeded low hospitality rating banks by 33.9%. Similarly, taking column 8, the increase in the number of applications to banks with high facility rating exceeded that to low facility rating banks by 38.3%. The results indicate that human interactions with branch employees and facilities of the branch are crucial criteria for borrowing when they are choosing the lending bank. We also find that the role of accessibility and product are less significant. The coefficient on the triple interaction term for accessibility and product is still positive, but less significant statistically (significant at the 5 or 10 percent levels) and economically. The results in Appendix A10 indicate that human interactions and bank facility play crucial roles in affecting the decisions of mortgage borrowers in the wake of natural disasters. Banks with better customer interactions and facilities gain popularity because they better resisted the disasters and responded to customers' urgent needs in the face of disasters.[26]

6 Conclusion

This paper provides a detailed exploration of customer satisfaction for a broad range of U.S. bank branches over an extended time period. We use Google ratings to measure the overall level of customer satisfaction, and exploit machine learning techniques to capture the key determinants of customer satisfaction from the information embedded in the specific comments that accompany these ratings. We find that branch Google ratings are correlated with a number of factors related

[26] Indeed, following natural disasters, banking regulators (e.g., the Federal Reserve and FDIC) usually expedite any request to operate temporary banking facilities to provide more convenient availability of services to affected borrowers.

to bank size, organizational structure, and pricing behavior. We also find that while these ratings are generally consistent over time, there are considerable variations in ratings among branches operated by the same bank in a given county. Consistent with the notion that customer attitudes drive business, we also find a strong association between customer satisfaction and deposit growth after incorporating a wide range of controls and fixed effects.

The key question we explore is whether branches with higher levels of customer satisfaction are better able to withstand exogenous shocks to their overall bank's reputation as captured by the RepRisk database. While we find strong evidence that there is a significant decline in deposit growth in the aftermath of these negative reputation shocks, we show that the branches with higher levels of customer satisfaction are significantly more able to mitigate the impact of these shocks. Overall, these results provide compelling evidence that banks with stronger customer satisfaction have more durable retail relationships. We also find that the durability of retail relationships is greater in areas with higher income levels and lower population mobility. Notably, we also find that these customer links are significantly more important for smaller community-oriented banks than they are for branches of larger bank holding companies.

Taking a closer look at the specific customer comments, we consider four key dimensions influencing customer satisfaction: 1) accessibility of the services; 2) the quality of the products, 3) hospitality of the staff; and 4) the quality of facilities. We find that the accessibility dimension is the key factor driving depositor retention.

Altogether, our results provide compelling evidence that the existence and durability of retail banking relationships are significantly driven by non-price factors. These findings confirm the value of the continued role of community-oriented banking despite the ongoing consolidation within the industry, and provides an interesting backdrop to consider the impact of continued technological changes that have dramatically transformed the various ways in which banks interact with their customers.

References

Berger, A. N., T. Kravitz, and L. Shibut. 2021. The many facets of bank competition: Evidence from an extraordinary dataset. Working Paper, Working Paper.

Berger, A. N., and G. F. Udell. 1995. Relationship lending and lines of credit in small firm finance. *Journal of Business* 351–81.

Buttle, F., and S. Maklan. 2019. Customer relationship management: concepts and technologies. Working Paper.

Chava, S. 2014. Environmental externalities and cost of capital. *Management Science* 60:2223–47.

Chavaz, M., and P. Slutzky. 2019. Worried depositors. Working Paper.

Chen, Y.-C., M. Hung, and L. L. Wang. 2019. Depositors' responses to public nonfinancial disclosure. Working Paper.

Correia, S. 2016. Linear models with high-dimensional fixed effects: An efficient and feasible estimator. Working Paper.

Cortés, K. R., and P. E. Strahan. 2017. Tracing out capital flows: How financially integrated banks respond to natural disasters. *Journal of Financial Economics* 125:182–99.

Deloitte. 2020. Recognizing the value of bank branches in a digital world.

Deng, S., and E. Elyasiani. 2008. Geographic diversification, bank holding company value, and risk. *Journal of Money, Credit and Banking* 40:1217–38.

Dlugosz, J., Y. K. Gam, R. Gopalan, and J. Skrastins. 2019. Decision-making delegation in banks. Working Paper.

Drechsler, I., A. Savov, and P. Schnabl. 2017. The deposits channel of monetary policy. *Quarterly Journal of Economics* 132:1819–76.

Giannetti, M., and Y. Yafeh. 2012. Do cultural differences between contracting parties matter? evidence from syndicated bank loans. *Management Science* 58:365–83.

Goetz, M. R., L. Laeven, and R. Levine. 2016. Does the geographic expansion of banks reduce risk? *Journal of Financial Economics* 120:346–62.

Gopalan, R., V. Nanda, and V. Yerramilli. 2011. Does poor performance damage the reputation of financial intermediaries? evidence from the loan syndication market. *Journal of Finance* 66:2083–120.

Harris, Z. S. 1954. Distributional structure. *Word* 10:146–62.

Homanen, M. 2018. Depositors disciplining banks: The impact of scandals. Working Paper.

Houston, J. F., and H. Shan. 2019. Corporate esg profiles and banking relationships. Working Paper.

Huang, G., and K. Sudhir. 2021. The causal effect of service satisfaction on customer loyalty. *Management Science* 67:317–41.

Huang, R. 2020. The financial consequences of customer satisfaction: Evidence from yelp ratings and SBA loans. Working Paper.

Iyer, R., and M. Puri. 2012. Understanding bank runs: The importance of depositor-bank relationships and networks. *American Economic Review* 102:1414–45.

Iyer, R., M. Puri, and N. Ryan. 2016. A tale of two runs: Depositor responses to bank solvency risk. *Journal of Finance* 71:2687–726.

Jacques, F., M. Nunez Maxwell, P. Patiath, and D. Stephens. 2018. Reimagining the bank branch for the digital era. Working Paper.

J.D. Power. 2019. U.S. National Banking Satisfaction Study.

Khwaja, A. I., and A. Mian. 2005. Do lenders favor politically connected firms? rent provision in an emerging financial market. *Quarterly Journal of Economics* 120:1371–411.

Kumar, V., I. Dalla Pozza, and J. Ganesh. 2013. Revisiting the satisfaction–loyalty relationship: empirical generalizations and directions for future research. *Journal of Retailing* 89:246–62.

Larsson, A., and Y. Viitaoja. 2017. Building customer loyalty in digital banking. *International Journal of Bank Marketing* 35:858–77.

Li, K., F. Mai, R. Shen, and X. Yan. 2020. Measuring corporate culture using machine learning. *Review of Financial Studies* Forthcoming.

Lin, L. 2020. Bank deposits and the stock market. *Review of Financial Studies* 33:2622–58.

Loughran, T., and B. McDonald. 2011. When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *Journal of Finance* 66:35–65.

Loveman, G. W. 1998. Employee satisfaction, customer loyalty, and financial performance: an empirical examination of the service profit chain in retail banking. *Journal of Service Research* 1:18–31.

McKinsey. 2019. Rewriting the rules: Succeeding in the new retail banking landscape.

Mian, A. 2006. Distance constraints: The limits of foreign lending in poor economies. *Journal of Finance* 61:1465–505.

Mikolov, T., I. Sutskever, K. Chen, G. S. Corrado, and J. Dean. 2013. Distributed representations of words and phrases and their compositionality. In *Advances in Neural Information Processing Systems*, 3111–9.

Nguyen, H.-L. Q. 2019. Are credit markets still local? evidence from bank branch closings. *American Economic Journal: Applied Economics* 11:1–32.

Puri, M., and J. Rocholl. 2008. On the importance of retail banking relationships. *Journal of Financial Economics* 89:253–67.

Ross, D. G. 2010. The "dominant bank effect:" how high lender reputation affects the information content and terms of bank loans. *Review of Financial Studies* 23:2730–56.

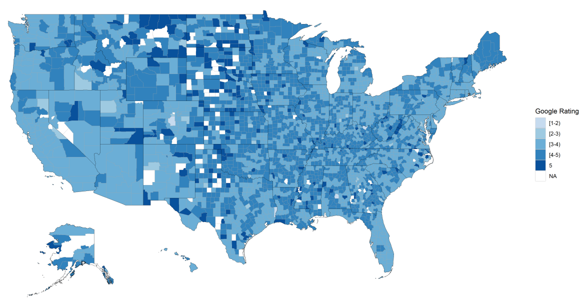
Shankar, V., A. K. Smith, and A. Rangaswamy. 2003. Customer satisfaction and loyalty in online and offline environments. *International Journal of Research in Marketing* 20:153–75.

Sharpe, S. A. 1990. Asymmetric information, bank lending, and implicit contracts: A stylized model of customer relationships. *Journal of Finance* 45:1069–87.

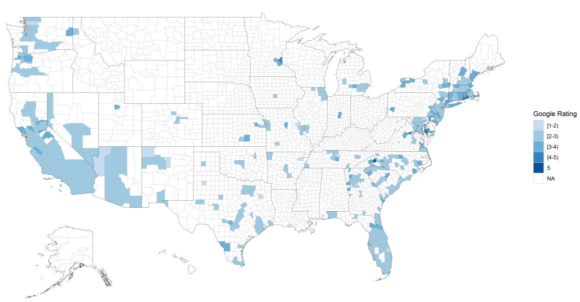
Srinivas, V., and R. Wadhwani. 2019. Recognizing the value of bank branches in a digital world: Findings from the global digital banking survey.

Figure 1. Average Google Rating Scores by County

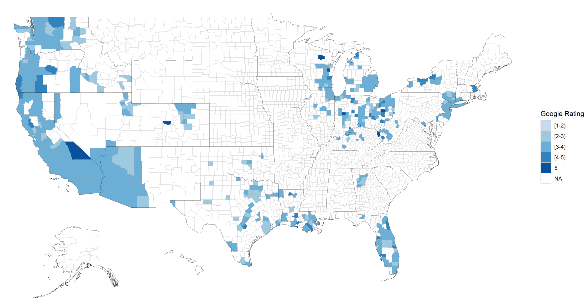
These figures show the county-level average Google rating scores for branches of all banks (Figure 1A), branches of Bank of America, N.A. (Figure 1B), and branches of JPMorgan Chase Bank, N.A. (Figure 1C), as observed on December 31, 2019.



Panel A All Banks



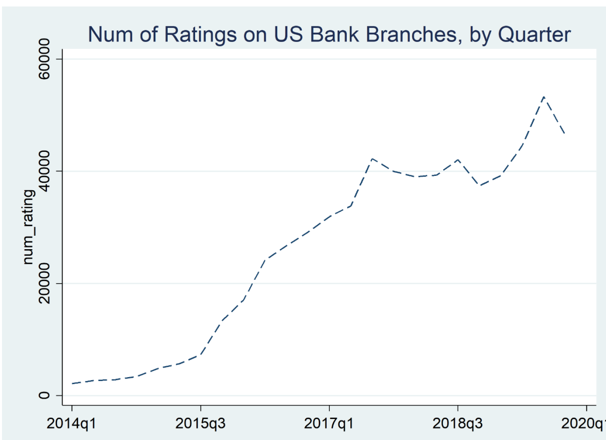
Panel B Bank of America, N.A.



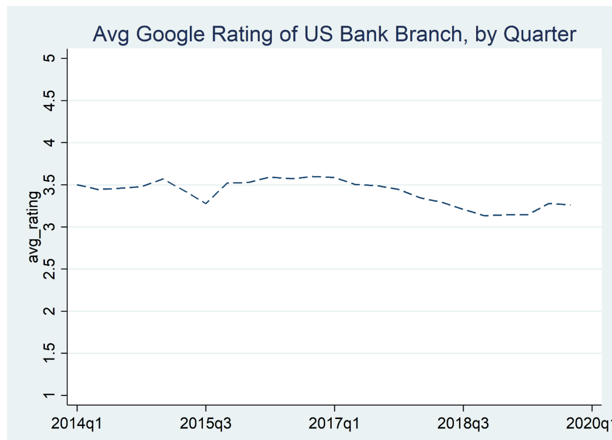
Panel C JPMorgan Chase Bank, N.A.

Figure 2. Number and Average Scores of Google Rating by Quarter

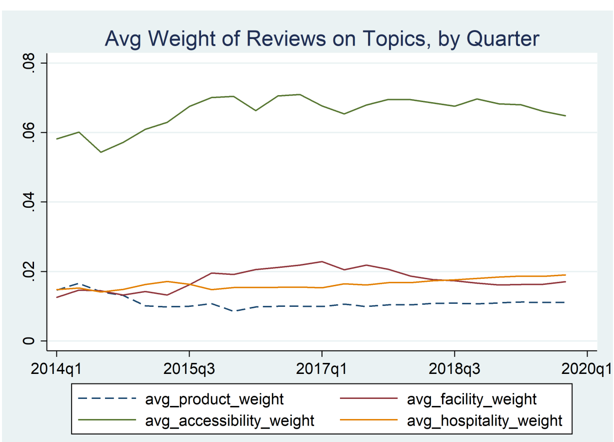
The following figures summarize Google ratings by quarter from 2014 to 2019. Figure 2A reports the number of new Google reviews on U.S. bank branches in each quarter. Figure 2B reports the average Google rating scores of U.S. bank branches by quarter. Figure 2C reports the average weight (frequency of mentions of topic-specific words over the length of the review) of reviews for four topics: product, facility, accessibility, and hospitality. Figure 2D reports the average rating scores by topics.



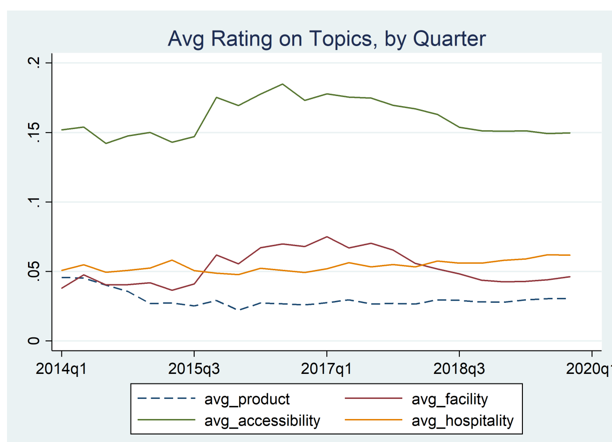
Panel A



Panel B



Panel C



Panel D

Figure 3. Effects of Reputation Shocks and Google Ratings

This figure reports the coefficient estimates and the corresponding 95% confidence intervals of panel OLS regressions of branch deposit growth on the reputation shock dummies.

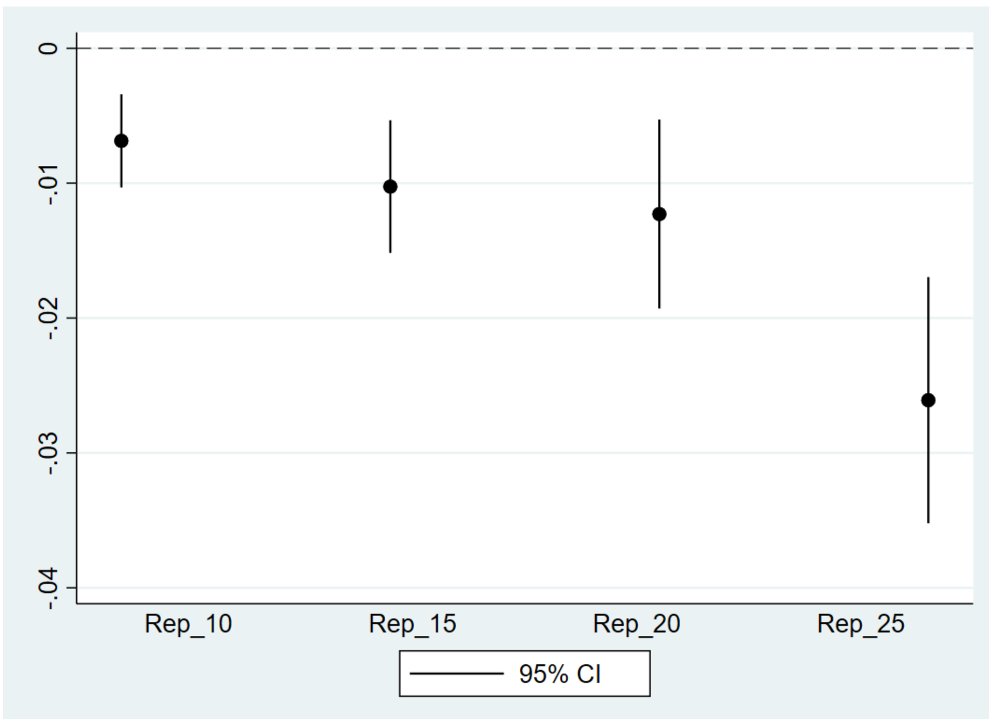


Table 1. Summary Statistics
This table reports the summary of statistics for the key variables in the Google Rating-RepRisk-SOD-merged sample.

	N	Mean	Median	S.D.
Branch Characteristics				
Depositgrowth	135,834	9.404	5.989	21.171
Deposits	135,834	91.705	60.374	120.207
Betterrate	135,834	0.688	1.000	0.463
Samestate	135,834	0.205	0.000	0.404
New	135,834	0.009	0.000	0.094
Ratesetter	135,834	0.027	0.000	0.161
Google Rating				
Rating	135,834	3.389	3.100	0.972
Num_Reviews	135,834	3.178	1.000	4.817
Accessibility Rating	135,834	-0.014	0.000	0.114
Hospitality Rating	135,834	0.004	0.000	0.052
Product Rating	135,834	-0.004	0.000	0.043
Facility Rating	135,834	0.002	0.000	0.068
Reputation Shocks				
Rep_10	135,834	0.550	1.000	0.497
Rep_15	98,785	0.382	0.000	0.486
Rep_20	81,092	0.247	0.000	0.431
Rep_25	75,121	0.187	0.000	0.390
Rep_Chg	135,834	11.888	11.000	8.904
Bank Characteristics				
Loan	135,834	0.562	0.593	0.153
ROA	135,834	0.005	0.005	0.003
Liquidity	135,834	0.085	0.062	0.099
Sensitivity	135,834	-0.095	-0.108	0.109
Countyshare	135,834	0.039	0.007	0.096
County Characteristics				
Income	135830	32.536	30.732	8.848
Mobility	135822	110.542	105.234	37.436

Table 2. Variance Decomposition Analysis

This table examines how bank branches' Google ratings are explained by a variety of fixed effects. Panel A reports the results on the cross-sectional regressions of Google ratings as of December 31, 2019 on a variety of fixed effects. Panel B reports the results on the panel regressions of Google ratings on a variety of fixed effects. Adjusted R^2s are reported for each model.

Panel A - Cross-sectional

	Rating		
	(1)	(2)	(3)
Bank FE	+		
County FE		+	
Bank × County FE			+
Observations	73,941	74,906	63,844
Adjusted R^2	0.227	0.090	0.296

Panel B - Panel

	Rating							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Bank FE	+							
County FE		+						
Branch FE			+				+	+
Year × Bank FE				+		+	+	
Year × County FE					+	+	+	
Year × Bank × County FE								+
Observations	299,248	299,364	294,479	291,958	296,679	289,055	284,045	245,211
Adjusted R^2	0.149	0.068	0.672	0.138	0.060	0.160	0.677	0.670

Table 3. Determinants of Branch-level Ratings

This table reports the deterministic regression of Google rating. The dependent variable is the Google rating of each bank branch as of December 31, 2019. The independent variables consist of bank, bank-county, and branch level characteristics. Bank characteristics include *Small* and *Local* dummies. *Small* equals one if a bank has less than two billion in assets. *Local* equals one if a bank obtains more than 65% of its deposits from a single county. Bank-county characteristics include the *Important* dummy. *Important* equals one if a county is in the top quartile of deposits among all the counties in which a bank has branches. Branch characteristics include *Countyshare*, *Ratesetter*, *Betterrate*, *Samestate*, and *New*. *Countyshare* is the branch's market share of deposit in the county. *Ratesetter* equals one if the branch is rate-setting branch. *Betterrate* equals one if the average rate of 12-month CD products is higher than the county median. *Samestate* equals one if a branch is in the same state with the headquarter of the bank. *New* equals one if a branch was established within the past five years. Standard errors are clustered at the county level and t-statistics are shown in parentheses. Statistical significance at the 10%, 5% and 1% levels is denoted by *, ** and ***, respectively.

	Rating	
	(1)	(2)
Small	0.454***	0.263***
	(9.55)	(7.31)
Local	0.204***	0.117***
	(4.15)	(3.68)
Important	-0.082***	-0.058**
	(-3.56)	(-2.17)
Countyshare		-0.790***
		(-4.02)
Ratesetter		-0.004
		(-0.14)
Betterrate		0.183***
		(4.26)
Samestate		0.257***
		(4.51)
New		0.245***
		(6.34)
County FE	+	+
Observations	67,986	64,905
Adjusted R^2	0.129	0.152

Table 4. Branch Rating and Customer Sentiment

This table reports how branch-level ratings are correlated with the metrics of sentiments and emotions. The dependent variable is the branch-level Google rating as of December 31, 2019. The independent variables are the weights of words associated with each sentiment and emotion in all of the branch's reviews (i.e., frequency of mentions over the total length of the reviews). Standard errors are clustered at the county level and t-statistics are shown in parentheses. Statistical significance at the 10%, 5% and 1% levels is denoted by *, ** and ***, respectively.

	Rating	
	(1)	(2)
Positive	2.584***	
	(15.24)	
Negative	-5.516***	
	(-19.78)	
Trust		0.824***
		(5.88)
Sadness		-0.838***
		(-2.60)
Surprise		-0.236***
		(-2.65)
Joy		2.387***
		(14.63)
Disgust		-0.292
		(-0.78)
Fear		-1.489***
		(-3.79)
Anger		-4.125***
		(-14.36)
Anticipation		-1.452***
		(-6.48)
County FE	+	+
Observations	67,481	67,481
Adjusted R^2	0.259	0.232

Table 5. Reputation Shock and Deposit Growth

This table reports the results on the OLS regressions of branch deposit growth on the bank's reputation shock proxies. *Depositgrowth* is the annual growth in branch deposits. *Rep_J* is a dummy variable that turn on if the RepRisk Index jumps more than J ($J=$ 10, 15, 20, or 25) over the past year, and zero if the jump is less than 10. *Rep_Chg* is the bank's maximum jump in RepRisk Index over the past year. *Deposits* is the amount of branch deposits. *Betterrate* is a dummy variable that equals one if the average rate of 12-month CD products at the branch is higher than the county median, and zero otherwise. *Countyshare* is the market share of the branch in the county by deposits. *Loan* is the share of loans in total assets. *ROA* is the bank's return on assets. *Liquidity* is cash divided by total deposits. *Sensitivity* is the sensitivity to interest rate risk. Standard errors are clustered at the branch level and t-statistics are shown in parentheses. Statistical significance at the 10%, 5% and 1% levels is denoted by *, ** and ***, respectively.

	Depositgrowth				
	(1)	(2)	(3)	(4)	(5)
Rep_10	-0.687***				
	(-3.90)				
Rep_15		-1.026***			
		(-4.09)			
Rep_20			-1.229***		
			(-3.44)		
Rep_25				-2.609***	
				(-5.60)	
Rep_Chg					-0.034***
					(-3.72)
Lag Deposits	-0.152***	-0.175***	-0.177***	-0.172***	-0.152***
	(-12.34)	(-13.13)	(-11.64)	(-10.50)	(-12.34)
Lag Betterrate	0.015	1.271**	1.896***	2.064***	-0.048
	(0.04)	(2.54)	(2.68)	(2.61)	(-0.13)
Lag Countyshare	-225.506***	-200.464***	-229.612***	-226.832***	-225.369***
	(-14.95)	(-12.77)	(-13.23)	(-12.00)	(-14.93)
Lag Loan	3.893***	4.174***	6.015***	6.979***	4.433***
	(3.79)	(2.97)	(3.05)	(3.36)	(4.27)
Lag ROA	34.706	65.075**	39.148	96.593*	40.329
	(1.32)	(2.10)	(0.87)	(1.96)	(1.52)
Lag Liquidity	2.543*	2.076	8.875**	13.498***	3.523**
	(1.78)	(0.73)	(2.18)	(2.99)	(2.51)
Lag Sensitivity	5.637***	4.772***	8.743***	11.809***	5.774***
	(4.41)	(3.34)	(4.49)	(5.15)	(4.52)
Fixed Effects	Branch + County × Year				
Observations	135,834	93,510	70,696	63,829	135,834
Adjusted R^2	0.157	0.127	0.106	0.109	0.157

Table 6. Reputation Shock, Google Rating, and Deposit Growth

This table reports the OLS regression of branch deposit growth on the branch-level Google ratings for banks that experienced reputation shocks. The observations are at the branch-year level, where branches are the branch offices of banks that experienced a reputation shock in the past 12 months (an increase in RepRisk index of at least 10, 15, 20, or 25 corresponds to columns 1 to 4, respectively) and years are the years of reputation shocks. *Depositgrowth* is the annual growth in branch deposits. *Rating* is the average Google rating of all the existing Google reviews on the branch. *Num_Reviews* is the number of Google ratings. *Deposits* is the amount of branch deposits. *Betterrate* is a dummy variable that equals one if the average rate of 12-month CD products is higher than the county median, and zero otherwise. *Countyshare* is the market share of the branch in the county by deposits. *New* is a dummy variable that equals one if a branch was established within the past five years, and zero otherwise. *Ratesetter* is a dummy variable that equals one if a branch is a local rate setter, and zero otherwise. Standard errors are clustered at the bank by county by shock year level and t-statistics are shown in parentheses. Statistical significance at the 10%, 5% and 1% levels is denoted by *, ** and ***, respectively.

	Depositgrowth			
	Rep_10=1	Rep_15=1	Rep_20=1	Rep_25=1
	(1)	(2)	(3)	(4)
Rating	0.270***	0.385***	0.338**	0.355*
	(3.59)	(3.67)	(2.09)	(1.87)
Num_Reviews	0.037**	0.035	-0.074	-0.202*
	(1.99)	(1.01)	(-0.78)	(-1.84)
Lag Deposits	-0.019***	-0.015***	-0.011***	-0.010***
	(-18.81)	(-9.03)	(-4.11)	(-2.70)
Lag Betterrate	4.978***	6.773***	13.457***	13.167***
	(3.40)	(2.93)	(4.26)	(3.80)
Lag Countyshare	-14.027***	-18.490***	-22.629**	-20.531*
	(-4.10)	(-3.71)	(-2.48)	(-1.87)
New	31.666***	28.884***	36.148***	37.762***
	(19.01)	(10.53)	(8.86)	(8.50)
Ratesetter	-16.685***	-12.350***	-11.049***	-13.329***
	(-12.52)	(-6.95)	(-4.31)	(-4.45)
Fixed Effects	Bank × County × Shock Year			
Observations	68,153	31,929	14,483	10,576
Adjusted R^2	0.151	0.130	0.150	0.160

Table 7. The Differential Effects of Topic Ratings during Reputation Shocks

This table reports the effects of topic ratings on deposit growth. The independent variable (*Accessibility*, *Product*, *Hospitality*, *Facility*) is the average topic rating on the branch up to the current year, where topic rating is defined as the relative word mentioning on a topic (frequency of key words associated with a specific topic scaled by the total number of words in the review) times the rating of the review (scaled to -2 to +2). The key words for topics are listed in Appendix A3. Columns (1) to (4) correspond to banks that experienced a reputation shock (an increase in RepRisk index of 10, 15, 20, and 25, respectively, since last June) and years of reputation shocks. Control variables include *Num_Reviews*, *Lag Log Deposits*, *Lag Betterrate*, *Lag Countyshare*, *New*, and *Ratesetter*. Standard errors are clustered at the bank-county-shock year level and t-statistics are shown in parentheses. Statistical significance at the 10%, 5% and 1% levels is denoted by *, ** and ***, respectively.

Panel A

	Depositgrowth			
	Rep_10=1	Rep_15=1	Rep_20=1	Rep_25=1
	(1)	(2)	(3)	(4)
Accessibility	2.234***	3.731***	4.372***	5.299***
	(3.41)	(3.98)	(2.93)	(3.08)
Controls	+	+	+	+
Fixed Effects	Bank × County × Shock Year			
Observations	68,153	31,929	14,483	10,576
Adjusted R^2	0.151	0.130	0.150	0.161

Panel B

	Depositgrowth			
	Rep_10=1	Rep_15=1	Rep_20=1	Rep_25=1
	(1)	(2)	(3)	(4)
Hospitality	1.215	1.575	4.039	6.149
	(0.88)	(0.79)	(1.13)	(1.47)
Controls	+	+	+	+
Fixed Effects	Bank × County × Shock Year			
Observations	68,153	31,929	14,483	10,576
Adjusted R^2	0.151	0.130	0.149	0.160

Panel C

	Depositgrowth			
	Rep_10=1	Rep_15=1	Rep_20=1	Rep_25=1
	(1)	(2)	(3)	(4)
Product	1.981	0.297	-2.101	-2.384
	(1.14)	(0.12)	(-0.59)	(-0.48)
Controls	+	+	+	+
Fixed Effects	Bank × County × Shock Year			
Observations	68,153	31,929	14,483	10,576
Adjusted R^2	0.151	0.130	0.149	0.160

Panel D

	Depositgrowth			
	Rep_10=1	Rep_15=1	Rep_20=1	Rep_25=1
	(1)	(2)	(3)	(4)
Facility	1.765*	2.550*	2.794	2.900
	(1.79)	(1.69)	(1.30)	(1.12)
Controls	+	+	+	+
Fixed Effects	Bank × County × Shock Year			
Observations	68,153	31,929	14,483	10,576
Adjusted R^2	0.151	0.130	0.149	0.160

Table 8. Cross-Sectional Analysis by County Income and Mobility

This table reports the sub-sample OLS regression of branch deposit growth on the branch-level Google ratings for banks that experienced reputation shocks. The observations are on the branch-year level. In Panel A, we conduct the regressions among branches that are located in counties whose income per capita are in the top 25% and bottom 25% nationally. In Panel B, we conduct the regressions among branches that are located in counties whose population mobility are in the top 25% and bottom 25% nationally. *Depositgrowth* is the annual growth in branch deposits. *Rating* is the average Google rating of all the existing Google reviews on the branch. *Num_Reviews* is the number of Google ratings. *Deposits* is the amount of branch deposits. *Betterrate* is a dummy variable that equals one if the average rate of 12-month CD products is higher than the county median, and zero otherwise. *Countyshare* is the market share of the branch in the county by deposits. *New* is a dummy variable that equals one if a branch was established within the past five years, and zero otherwise. *Ratesetter* is a dummy variable that equals one if a branch is a local rate setter, and zero otherwise. Standard errors are clustered at the bank by county by shock year level and t-statistics are shown in parentheses. Statistical significance at the 10%, 5% and 1% levels is denoted by *, ** and ***, respectively.

Panel A: Income Per Capita

	Depositgrowth							
	Rep_10=1		Rep_15=1		Rep_20=1		Rep_25=1	
	Bottom25%	Top25%	Bottom25%	Top25%	Bottom25%	Top25%	Bottom25%	Top25%
Rating	0.185	0.334**	0.099	0.849***	-0.042	0.615*	0.116	0.706*
	(1.25)	(2.29)	(0.51)	(3.97)	(-0.15)	(1.84)	(0.33)	(1.91)
Num_Reviews	0.038	0.067**	-0.017	0.081*	-0.136	-0.101	-0.018	-0.217
	(0.91)	(2.33)	(-0.22)	(1.77)	(-0.75)	(-0.77)	(-0.10)	(-0.90)
Lag Deposits	-0.018***	-0.020***	-0.011***	-0.015***	-0.008	-0.008	-0.004	-0.008
	(-7.41)	(-11.69)	(-2.92)	(-5.18)	(-1.32)	(-1.32)	(-0.49)	(-0.92)
Lag Betterrate	6.505***	-2.016	9.041**	-1.886	15.857***	7.878	15.672***	7.324
	(2.65)	(-0.76)	(2.29)	(-0.44)	(2.75)	(1.55)	(2.97)	(1.32)
Lag Countyshare	-18.068***	-18.258*	-19.515***	-36.815**	-26.248**	-36.978	-32.007*	-40.913
	(-3.83)	(-1.75)	(-2.59)	(-2.07)	(-2.01)	(-0.55)	(-1.91)	(-0.48)
New	24.891***	37.716***	27.527***	32.961***	37.634***	37.283***	35.914***	39.819***
	(7.35)	(12.32)	(4.79)	(6.17)	(5.07)	(4.51)	(4.96)	(4.03)
Ratesetter	-6.734***	-28.784***	-7.906***	-24.105***	-8.186**	-22.888***	-6.652*	-26.930***
	(-2.99)	(-10.38)	(-2.59)	(-5.45)	(-2.12)	(-3.83)	(-1.69)	(-3.58)
Fixed Effects	Bank × County × Shock Year							
Observations	14,961	17,976	6,805	8,399	3,087	3,874	2,222	2,842
Adjusted R^2	0.129	0.187	0.107	0.171	0.114	0.175	0.152	0.177

Panel B: Mobility

	Depositgrowth							
	Rep_10=1		Rep_15=1		Rep_20=1		Rep_25=1	
	Bottom25%	Top25%	Bottom25%	Top25%	Bottom25%	Top25%	Bottom25%	Top25%
Rating	0.363**	0.052	0.351*	0.128	0.619**	0.086	0.581*	0.153
	(2.24)	(0.34)	(1.76)	(0.58)	(2.02)	(0.28)	(1.70)	(0.41)
$\text{Num}_{R}eviews$	0.014	0.026	0.006	0.012	-0.002	-0.089	-0.119	-0.218
	(0.38)	(0.83)	(0.09)	(0.19)	(-0.02)	(-0.42)	(-0.79)	(-0.66)
Lag Deposits	-0.022***	-0.021***	-0.017***	-0.014***	-0.016***	-0.010	-0.016***	-0.010
	(-11.26)	(-9.29)	(-5.82)	(-3.94)	(-4.03)	(-1.46)	(-3.15)	(-1.11)
Lag Betterrate	2.585	1.462	6.514	3.914	13.626**	11.623**	15.233**	12.692**
	(0.95)	(0.59)	(1.53)	(1.02)	(2.21)	(2.22)	(2.40)	(2.16)
$\text{Lag}_{C}ountyshare$	-21.795	-14.240***	-31.351*	-21.271***	-9.456	-14.830	-13.261	-11.154
	(-1.34)	(-2.88)	(-1.73)	(-3.04)	(-0.56)	(-0.90)	(-0.67)	(-0.56)
New	39.751***	29.355***	40.268***	28.602***	49.099***	37.374***	49.591***	41.006***
	(10.47)	(8.84)	(6.72)	(5.55)	(5.28)	(4.94)	(4.81)	(4.78)
Ratesetter	-26.123***	-15.466***	-21.866***	-12.786***	-20.367***	-12.753***	-23.615***	-15.511***
	(-7.36)	(-6.46)	(-4.83)	(-4.34)	(-2.88)	(-2.87)	(-2.74)	(-2.92)
Fixed Effects	Bank × County × Shock Year							
Observations	17,957	15,885	8,506	7,375	3,914	3,275	2,877	2,397
Adjusted R^2	0.148	0.140	0.138	0.123	0.149	0.161	0.154	0.164

Table 9. Cross-Sectional Analysis by Bank Size

This table reports the OLS sub-sample regression of branch deposit growth on the branch-level Google ratings for banks that experienced reputation shocks. The observations are on the branch-year level, where branches are the branch offices of banks that experienced a reputation shock in the past 12 months (an increase in RepRisk index of at least 10, 15, 20, or 25 respectively) and years are the years of reputation shocks. We report the results using the sub-sample banks with domestic assets in the bottom 25% (maximum assets $0.8 billion) and those with domestic assets in the top 25% (minimum assets $3.36 billion) among all banks each year. *Depositgrowth* is the annual growth in branch deposits. *Rating* is the average Google rating of all the existing Google reviews on the branch. *Num_Reviews* is the number of Google ratings. *Deposits* is the amount of branch deposits. *Betterrate* is a dummy variable that equals one if the average rate of 12-month CD products is higher than the county median, and zero otherwise. *Countyshare* is the market share of the branch in the county by deposits. *New* is a dummy variable that equals one if a branch was established within the past five years, and zero otherwise. *Ratesetter* is a dummy variable that equals one if a branch is a local rate setter, and zero otherwise. Standard errors are clustered at the bank by county by shock year level and t-statistics are shown in parentheses. Statistical significance at the 10%, 5% and 1% levels is denoted by *, ** and ***, respectively.

	Depositgrowth							
	Rep_10=1		Rep_15=1		Rep_20=1		Rep_25=1	
	Bottom25%	Top25%	Bottom25%	Top25%	Bottom25%	Top25%	Bottom25%	Top25%
Rating	0.293***	0.281**	0.437***	0.388*	0.594**	0.258	0.573*	0.241
	(2.820)	(2.307)	(3.072)	(1.958)	(2.158)	(0.896)	(1.925)	(0.633)
Num_Reviews	0.077	0.033*	0.130	0.012	0.068	-0.136*	-0.457	-0.255*
	(1.131)	(1.676)	(1.350)	(0.334)	(0.127)	(-1.649)	(-1.548)	(-1.921)
Lag Deposits	-0.015***	-0.020***	-0.008***	-0.018***	-0.000	-0.012***	0.002	-0.010**
	(-6.741)	(-17.709)	(-2.896)	(-9.004)	(-0.010)	(-3.837)	(0.269)	(-2.448)
Lag Betterrate	5.340*	5.359***	4.348	7.569**	11.677**	16.248***	14.952***	16.038***
	(1.869)	(2.740)	(0.904)	(2.391)	(2.034)	(3.121)	(2.603)	(2.848)
Lag Countyshare	-11.319*	-19.304***	-11.172	-23.574***	15.545	-44.077***	17.997	-54.233***
	(-1.725)	(-4.446)	(-0.965)	(-3.645)	(0.492)	(-2.730)	(0.505)	(-2.734)
New	33.907***	32.348***	36.740***	26.639***	40.580***	40.166***	43.620***	41.434***
	(7.822)	(17.466)	(5.878)	(8.179)	(3.317)	(7.550)	(3.511)	(7.234)
Ratesetter	-15.899***	-18.667***	-16.796***	-11.634***	-13.496	-10.099***	-17.015*	-12.179***
	(-5.059)	(-11.621)	(-3.933)	(-5.312)	(-1.390)	(-2.824)	(-1.782)	(-2.684)
Fixed Effects	Bank × County × Shock Year							
Observations	22,696	39,691	12,867	14,495	4,667	5,668	4,133	3,360
Adjusted R^2	0.126	0.163	0.119	0.118	0.106	0.169	0.104	0.196

Table 10. Google Rating and Mortgage Demand During Natural Disasters
This table reports the results on the triple difference-in-differences regressions of mortgage applications around natural disasters. *Treat* is a dummy variable that equals one for the disaster affected counties, and zero for control counties. *Post* is a dummy variable that equals one for the disaster incidence year, and zero for the preceding year. *High Rating* is a bank-county-year-level dummy variable that equals one if the bank's average Google rating in the county is at the top tercile among all banks within the same state, and zero if it is at the bottom tercile. *Ln(Applications)* is the natural logarithm of the annual number of mortgage applications to the bank in the a county. Control variables include *Lag Ln(Applications)*, *Ln(Loan)*, *ROA*, *Liquidity*, *Sensitivity*, *Has Missing*, *Small*, *Local*, *Important*, *Branch*, *Deposit per Capita*, *Unemployment*, *Population*, *White*, *Female*, *Education*, *Income*, *Senior*, *Manufacturing Labor*, *Information Labor*. Standard errors are clustered at the county by bank level and t-statistics are shown in parentheses. Statistical significance at the 10%, 5% and 1% levels is denoted by *, ** and ***, respectively.

	Ln(Applications)			
	(1)	(2)	(3)	(4)
Treat × Post × High Rating	0.203***	0.199***	0.149**	0.184***
	(3.57)	(3.15)	(2.53)	(3.05)
Treat × Post	0.355***	0.343***	0.246***	0.269***
	(13.43)	(11.95)	(9.79)	(10.82)
Treat × High Rating	-0.196***	-0.176***	-0.243***	-0.200***
	(-5.43)	(-4.17)	(-5.42)	(-4.14)
Post × High Rating	-0.722***	-0.756***	-0.568***	-0.710***
	(-20.96)	(-19.70)	(-17.26)	(-20.42)
Treat	-0.150***	-0.147***	0.038*	-0.131***
	(-9.72)	(-8.90)	(1.66)	(-5.23)
Post	-0.849***	-0.899***	-0.271***	-0.668***
	(-48.98)	(-45.42)	(-18.23)	(-31.04)
High Rating	0.621***	0.649***	0.720***	0.778***
	(26.75)	(23.86)	(23.09)	(23.97)
Controls	+	+	+	+
Disaster Year × State FE	+			+
Disaster Year × State × Bank FE		+		
Bank × County FE			+	+
Observations	37,173	37,170	37,170	37,170
Adjusted R^2	0.691	0.690	0.606	0.655

Appendix

Table A1. Variable Definitions

Variable	Definition	Source
Google Rating		
Rating	Branch-level cumulative average Google rating.	Google Rating
HighRating	A dummy variable that equals one if the bank-county-year average Google rating belongs to the top tercile among all bank-county observations in the state in the year, and zero if it belongs to the bottom tercile.	Google Rating
HighTopicRating	A dummy variable that equals one if the bank-county-year average Google rating on a specific topic (weight of words mentioning times rating of the review) belongs to the top tercile among all bank-county observations in the state in the year, and zero if it belongs to the bottom tercile.	Google Rating
Num_Reviews	Cumulative number of Google ratings.	Google Rating
Missing	A bank-county-year level dummy variable that equals one if a bank has at least one branch with missing Google rating in a year, and zero otherwise.	Google Rating
ReputationShock		
Rep_J	A dummy variable that equals one if the bank has an increase in the RepRisk Index by at least J (J=10, 15, 20, or 25) in the past 12 months, and zero if the increase is below 10.	RepRisk
Share_Rep_J	The share of branch deposits in a county that is subject to a jump in the RepRisk Index by more than J (J=10, 15, 20, or 25).	SOD, RepRisk
Branch Characteristics		
Deposits	Branch deposits.	SOD
Countyshare	The market share of the branch in the county by deposits.	SOD
Betterrate	A dummy variable that equals one if the average rate of 12-month CD products at the branch is higher than the county median, and zero otherwise.	RateWatch
Ratesetter	A dummy variable that equals one if a branch is a local rate setter, and zero otherwise.	RateWatch
Samestate	A dummy variable that equals one if a branch is in the same state with the headquarter of the bank, and zero otherwise.	SOD
New	A dummy variable that equals one if a branch was established within the past five years, and zero otherwise.	SOD
Bank Characteristics		
Loan	The share of loans in total assets.	Call Report
ROA	Return on assets, measured as the ratio of the annualized net income to gross total assets.	Call Report
Liquidity	Cash divided by bank total deposits.	Call Report
Sensitivity	The sensitivity to interest rate risk, defined as the ratio of the absolute difference between short-term assets and short-term liabilities to gross total assets.	Call Report
Small	A dummy variable that equals one if a bank has less than two billion dollars in assets, and zero otherwise.	Call Report
Local	A dummy variable that equals one if a bank obtains more than 65% of its deposits from a single county, and zero otherwise.	SOD
Important	A dummy variable that equals one if a county is in the top quartile of deposits among all the counties in which a bank has branches, and zero otherwise.	SOD
County Characteristics		
Branch	Number of bank branches per 1,000 population in the county.	SOD
Depositgrowth_County	Total deposits of all bank branches in the county.	SOD
Deposits per capita	Amount of bank deposits per capita.	SOD
HHI	The Herfindahl-Hirschmann index of deposits at all bank branches in the county.	SOD
Unemployment	Local unemployment rate.	ACS
Population	Local total population size in millions.	ACS
White	Share of white people in local population.	ACS
Female	Share of female in local population.	ACS
Education	The population that are over 25 years and with high school education (or higher) divided by the total population older than 25.	ACS
Income	Income (in thousand dollars) per capita.	ACS
Mobility	The number of people who migrate into and out of the county between 2014 and 2018, per thousand of county population.	ACS
Senior	The share of population that is over 65 years old.	ACS
Manufacturing Labor	The share of labor working in manufacturing industry.	ACS
Information Labor	The share of labor working in information industry.	ACS

Table A2. Summary Statistics of Natural Disasters

This table presents the summary statistics for the Google Rating-FEMA-HMDA-merged sample. Panel A shows the frequency and severity of the natural disasters used by the sample. Panel B shows the summary statistics of the key variables in the sample.

Panel A: Frequency and Severity of Natural Disasters			
Disaster Type	**Frequency**	**Average Number of Counties Affected**	**Total Number of Counties Affected**
Severe Storm	55	9.58	527
Flood	48	10.19	489
Hurricane	15	17.00	255
Snow	10	15.40	154
Tornado	4	1.25	5
Severe Ice Storm	3	15.00	45
Coastal Storm	1	1.00	1
Mud/Landslide	1	8.00	8
Volcano	1	1.00	1
Total	138	10.76	1,485

Panel B: Google Rating-FEMA-HMDA-merged Sample				
	N	**Mean**	**Median**	**S.D.**
Mortgage Demand				
Mortgage Application	37,178	82.100	31.000	153.000
Log Applications	37,178	3.280	3.470	1.670
Google Rating				
High Rating	37,178	0.248	0.000	0.432
Num Reviews	37,178	8.120	3.000	19.3
Bank Characteristics				
Ln(Loan)	37,178	14.400	13.700	2.800
ROA	37,178	0.005	0.005	0.002
Liquidity	37,178	0.060	0.041	0.061
Sensitivity	37,178	-0.130	-0.126	0.124
Small	37,173	0.465	0.000	0.499
Local	37,178	0.206	0.000	0.404
Important	37,178	0.244	0.000	0.430
County Characteristics				
Branch	37,178	382	338	182
Deposits per capita	37,178	22.500	18.800	13.000
Unemployment	37,178	0.067	0.064	0.025
Population	37,178	0.273	0.095	0.451
White	37,178	0.823	0.866	0.140
Female	37,178	0.504	0.507	0.014
Education	37,178	0.876	0.889	0.058
Income	37,178	28.100	27.100	6.960
Senior	37,178	0.156	0.151	0.042
Manufacturing Labor	37,178	0.119	0.106	0.063
Information Labor	37,178	0.017	0.016	0.008

Table A3. Topic Word List

Accessibility			Hospitality			Product			Facility		
Word	**Fweight**	**Frequency**	**Word**	**Fweight**	**Frequency**	**Word**	**Fweight**	**Frequency**	**Word**	**Fweight**	**Frequency**
time	3,113.58	68,943	experience	1,900.42	30,401	fees	633.23	14,270	atm	2,367.55	31,162
place	2,125.39	22,202	see	530.17	14,604	loan	518.99	14,531	inside	598.20	11,617
wait	1,840.16	31,216	walk	355.75	7,971	checking	395.19	11,206	parking	569.96	7,413
waiting	1,023.72	20,039	deal	343.90	7,115	charge	359.31	9,210	clean	508.73	4,249
phone	1,020.38	20,817	helping	329.62	7,012	fee	353.70	10,384	atms	339.60	4,343
take	878.39	21,484	smile	327.72	5,410	mortgage	233.95	6,596	building	297.68	4,166
come	852.07	20,060	attitude	316.62	6,862	savings	204.18	6,217	front	274.11	7,110
times	767.44	17,536	visit	309.83	6,100	charged	164.66	5,242	debit	250.02	8,352
person	710.55	18,391	talk	295.87	7,471	rates	148.81	2,617	window	248.78	5,772
day	694.66	20,878	welcoming	221.17	2,603	interest	140.27	4,055	machine	221.33	5,239
lines	670.36	7,828	speak	218.95	6,512	charges	135.25	3,936	lobby	201.68	4,266
hour	594.33	11,487	reviews	190.96	5,077	rate	130.42	3,460	outside	196.41	3,960
busy	577.94	8,886	greeted	180.04	4,258	loans	123.09	2,860	app	187.20	4,339
took	517.35	14,331	review	171.17	5,835	acct	66.74	2,067	atmosphere	174.44	1,723
called	479.12	17,972	experiences	168.18	3,150	saving	37.28	1,014	desk	129.55	3,583
waited	446.21	9,727	dealt	165.92	3,416	refinance	34.00	1,047	windows	121.02	2,138
area	363.73	7,147	assist	162.99	3,688	score	31.81	1,140	machines	119.63	2,065
usually	309.11	5,551	experienced	150.14	3,096	cd	30.20	768	view	106.04	908
takes	297.79	5,796	fix	139.85	3,394	lower	28.39	837	environment	103.18	1,078
branches	297.62	6,580	welcome	129.16	2,271	terms	27.24	802	broken	99.37	1,451
front	274.11	7,110	greet	125.23	2,454	lending	26.92	640	doors	72.27	1,493
least	269.76	7,000	remember	125.00	2,624	lender	26.13	796	park	71.67	1,583
week	268.12	7,959	assistance	109.63	2,484	added	25.61	919	facility	67.89	973
mins	263.81	5,446	smiling	99.10	1,391	investment	25.03	641	space	60.52	961
store	256.75	3,743	handle	96.52	2,613	financing	24.10	622	views	58.34	355
window	248.78	5,772	explain	88.53	2,992	finance	23.36	657	accessible	56.07	693
hold	242.13	7,505	resolve	81.28	2,216	bucks	23.02	541	cars	52.92	1,284
locations	218.70	4,174	meet	79.12	1,905	product	22.78	562	traffic	50.81	1,002
counter	205.77	5,039	serve	76.36	1,483	mortgages	21.23	542	floor	46.36	887

Table A4. Reputation Shock and Deposit Growth

This table re-estimates Equation 4 using the same sample across all columns. *Depositgrowth* is the annual growth in branch deposits. *Rep_J* is a dummy variable that equals one if the bank has an increase in the RepRisk Index by at least J (J=10, 15, 20, or 25) in the past 12 months, and zero if the increase is below J (instead of 10 as in Table 5). *Deposits* is the amount of branch deposits. *Betterrate* is a dummy variable that equals one if the average rate of 12-month CD products at the branch is higher than the county median, and zero otherwise. *Countyshare* is the market share of the branch in the county by deposits. *Loan* is the share of loans in total assets. *ROA* is the bank's return on assets. *Liquidity* is cash divided by total deposits. *Sensitivity* is the sensitivity to interest rate risk. Standard errors are clustered at the branch level and t-statistics are shown in parentheses. Statistical significance at the 10%, 5% and 1% levels is denoted by *, ** and ***, respectively.

	Depositgrowth				
	(1)	(2)	(3)	(4)	(5)
Rep_10	-0.687***				
	(-3.90)				
Rep_15		-1.064***			
		(-6.10)			
Rep_20			-0.242		
			(-1.07)		
Rep_25				-0.734***	
				(-2.89)	
Rep_Chg					-0.034***
					(-3.72)
Lag Deposits	-0.152***	-0.152***	-0.152***	-0.152***	-0.152***
	(-12.34)	(-12.34)	(-12.33)	(-12.33)	(-12.34)
Lag Betterrate	0.015	0.026	-0.021	-0.087	-0.048
	(0.04)	(0.07)	(-0.06)	(-0.23)	(-0.13)
Lag Countyshare	-225.506***	-225.226***	-225.371***	-225.320***	-225.369***
	(-14.95)	(-14.92)	(-14.93)	(-14.92)	(-14.93)
Lag Loan	3.893***	5.390***	4.026***	4.143***	4.433***
	(3.79)	(5.07)	(3.92)	(4.05)	(4.27)
Lag ROA	34.706	44.061*	29.485	35.466	40.329
	(1.32)	(1.65)	(1.10)	(1.33)	(1.52)
Lag Liquidity	2.543*	3.433**	3.652***	3.786***	3.523**
	(1.78)	(2.44)	(2.58)	(2.69)	(2.51)
Lag Sensitivity	5.637***	5.969***	5.625***	5.682***	5.774***
	(4.41)	(4.67)	(4.40)	(4.45)	(4.52)
Fixed Effects	Branch + County × Year				
Observations	135,834	135,834	135,834	135,834	135,834
Adjusted R^2	0.157	0.157	0.156	0.157	0.157

Table A5. The Effects of Topic Ratings

The following tables reports the complete tables of Table 7.

Panel A

	Depositgrowth			
	Rep_10=1	Rep_15=1	Rep_20=1	Rep_25=1
	(1)	(2)	(3)	(4)
Accessibility	2.234***	3.731***	4.372***	5.299***
	(3.41)	(3.98)	(2.93)	(3.08)
Num_Reviews	0.037**	0.035	-0.069	-0.188*
	(1.97)	(1.02)	(-0.72)	(-1.72)
Lag Deposits	-0.019***	-0.015***	-0.011***	-0.010***
	(-18.82)	(-9.05)	(-4.14)	(-2.74)
Betterrate	4.962***	6.730***	13.425***	13.151***
	(3.39)	(2.91)	(4.25)	(3.80)
Countyshare	-14.030***	-18.441***	-22.368**	-19.968*
	(-4.10)	(-3.70)	(-2.45)	(-1.82)
New	31.688***	28.928***	36.206***	37.842***
	(19.01)	(10.53)	(8.85)	(8.50)
Ratesetter	-16.715***	-12.405***	-11.126***	-13.420***
	(-12.54)	(-6.98)	(-4.33)	(-4.47)
Fixed Effects	Bank × County × Shock Year			
Observations	68,153	31,929	14,483	10,576
Adjusted R^2	0.151	0.130	0.150	0.161

Panel B

	Depositgrowth			
	Rep_10=1	Rep_15=1	Rep_20=1	Rep_25=1
	(1)	(2)	(3)	(4)
Hospitality	1.215	1.575	4.039	6.149
	(0.88)	(0.79)	(1.13)	(1.47)
Num_Reviews	0.028	0.022	-0.089	-0.216**
	(1.52)	(0.64)	(-0.94)	(-1.98)
Lag Deposits	-0.019***	-0.015***	-0.011***	-0.010***
	(-18.79)	(-9.04)	(-4.12)	(-2.72)
Betterrate	4.977***	6.748***	13.453***	13.147***
	(3.40)	(2.92)	(4.25)	(3.79)
Countyshare	-14.097***	-18.542***	-22.492**	-20.118*
	(-4.12)	(-3.72)	(-2.47)	(-1.84)
New	31.724***	28.954***	36.248***	37.883***
	(19.03)	(10.53)	(8.87)	(8.52)
Ratesetter	-16.727***	-12.413***	-11.115***	-13.396***
	(-12.54)	(-6.98)	(-4.33)	(-4.47)
Fixed Effects	Bank × County × Shock Year			
Observations	68,153	31,929	14,483	10,576
Adjusted R^2	0.151	0.130	0.149	0.160

Panel C

	Depositgrowth			
	Rep_10=1	Rep_15=1	Rep_20=1	Rep_25=1
	(1)	(2)	(3)	(4)
Product	1.981	0.297	-2.101	-2.384
	(1.14)	(0.12)	(-0.59)	(-0.48)
Num_Reviews	0.028	0.021	-0.092	-0.222**
	(1.50)	(0.61)	(-0.97)	(-2.03)
Lag Deposits	-0.019***	-0.015***	-0.011***	-0.010***
	(-18.78)	(-9.03)	(-4.12)	(-2.71)
Betterrate	4.979***	6.748***	13.451***	13.176***
	(3.40)	(2.92)	(4.25)	(3.80)
Countyshare	-14.106***	-18.603***	-22.752**	-20.644*
	(-4.12)	(-3.73)	(-2.50)	(-1.88)
New	31.723***	28.961***	36.267***	37.897***
	(19.02)	(10.54)	(8.87)	(8.51)
Ratesetter	-16.730***	-12.424***	-11.151***	-13.435***
	(-12.55)	(-6.99)	(-4.34)	(-4.48)
Fixed Effects	Bank $\times$ County $\times$ Shock Year			
Observations	68,153	31,929	14,483	10,576
Adjusted R^2	0.151	0.130	0.149	0.160

Panel D

	Depositgrowth			
	Rep_10=1	Rep_15=1	Rep_20=1	Rep_25=1
	(1)	(2)	(3)	(4)
Facility	1.765*	2.550*	2.794	2.900
	(1.79)	(1.69)	(1.30)	(1.12)
Num_Reviews	0.029	0.022	-0.089	-0.220**
	(1.57)	(0.65)	(-0.95)	(-2.01)
Lag Deposits	-0.019***	-0.015***	-0.011***	-0.010***
	(-18.79)	(-9.04)	(-4.12)	(-2.71)
Betterrate	4.972***	6.751***	13.460***	13.193***
	(3.39)	(2.92)	(4.25)	(3.80)
Countyshare	-14.136***	-18.601***	-22.711**	-20.589*
	(-4.13)	(-3.73)	(-2.49)	(-1.88)
New	31.709***	28.935***	36.227***	37.844***
	(19.01)	(10.53)	(8.86)	(8.50)
Ratesetter	-16.729***	-12.411***	-11.122***	-13.397***
	(-12.54)	(-6.98)	(-4.33)	(-4.47)
Fixed Effects	Bank $\times$ County $\times$ Shock Year			
Observations	68,153	31,929	14,483	10,576
Adjusted R^2	0.151	0.130	0.149	0.160

Table A6. Reputation Shock, Google Rating, and Deposit Growth - A Difference-in-differences Approach

This table reports the results of a difference-in-differences approach that re-estimates the effects of branch-level Google ratings on mitigating the negative effects of reputation shock on deposit growth. The treatment group include all branches of banks that experienced a reputation shock in the past 12 months (an increase in RepRisk index of at least 10, 15, 20, or 25). For each treatment branch in each treatment year, the control branches for it include all branches in the same county but did not experience any reputation shocks in the past 12 months (the increase in the RepRisk index was less than 10). The treatment branch and its control branches are considered a cohort. *Depositgrowth* is the annual growth in branch deposits. *Rep_J* is a dummy variable that equals one if there is an increase in the RepRisk Index by at least J (J=10, 15, 20, or 25) in the past 12 months, and zero if the increase is below 10. *Rating* is the average Google rating of all the existing Google reviews on the branch. Control variables include *Num_Reviews*, lagged *Deposits*, *Samestate*, *New*, *Ratesetter*, *Loan*, *Betterrate*, *Countyshare*, *ROA*, *Liquidity*, and *Sensitivity*. Standard errors are clustered at the cohort level and t-statistics are shown in parentheses. Statistical significance at the 10%, 5% and 1% levels is denoted by *, ** and ***, respectively.

	Depositgrowth			
	(1)	(2)	(3)	(4)
Rep_10 × Rating	0.554***			
	(2.95)			
Rep_15 × Rating		0.596***		
		(3.15)		
Rep_20 × Rating			0.624***	
			(3.29)	
Rep_25 × Rating				0.397*
				(1.86)
Rep_10	-1.058***			
	(-4.29)			
Rep_15		-1.002***		
		(-4.02)		
Rep_20			-0.962***	
			(-3.85)	
Rep_25				-0.285
				(-0.93)
Rating	0.080**	0.047	0.047	0.144***
	(2.54)	(1.47)	(1.46)	(4.17)
Controls	+	+	+	+
Cohort Fixed Effects	+	+	+	+
Observations	350,862	345,413	342,484	231,965
Adjusted R-squared	0.096	0.096	0.097	0.083

Table A7. Reputation Shock, Google Rating, and Deposit Growth: ZIP Code Fixed Effects

This table re-estimates Equation 5 using bank by ZIP Code by shock year fixed effects. The observations are at the branch-year level, where branches are the branch offices of banks that experienced a reputation shock in the past 12 months (an increase in RepRisk index of at least 10, 15, 20, or 25 corresponds to columns 1 to 4, respectively) and years are the years of reputation shocks. *Depositgrowth* is the annual growth in branch deposits. *Rating* is the average Google rating of all the existing Google reviews on the branch. *Num_Reviews* is the number of Google ratings. *Deposits* is the amount of branch deposits. *Betterrate* is a dummy variable that equals one if the average rate of 12-month CD products is higher than the county median, and zero otherwise. *Countyshare* is the market share of the branch in the county by deposits. *New* is a dummy variable that equals one if a branch was established within the past five years, and zero otherwise. *Ratesetter* is a dummy variable that equals one if a branch is a local rate setter, and zero otherwise. Standard errors are clustered at the bank by county by shock year level and t-statistics are shown in parentheses. Statistical significance at the 10%, 5% and 1% levels is denoted by *, ** and ***, respectively.

	Depositgrowth			
	Rep_10=1	Rep_15=1	Rep_20=1	Rep_25=1
	(1)	(2)	(3)	(4)
Rating	0.393**	0.668***	0.776*	0.829*
	(2.31)	(2.71)	(1.93)	(1.73)
Num_Reviews	-0.019	-0.084	-0.369	-0.563
	(-0.41)	(-0.93)	(-1.28)	(-1.44)
Lag Deposits	-0.024***	-0.021***	-0.012***	-0.011**
	(-13.84)	(-7.22)	(-2.74)	(-2.01)
Lag Betterrate	7.772***	10.903***	25.333***	29.979***
	(3.61)	(3.37)	(4.87)	(5.18)
Lag Countyshare	-16.335***	-25.085***	-39.657**	-48.819**
	(-2.97)	(-3.06)	(-2.40)	(-2.03)
New	33.207***	32.392***	43.546***	45.733***
	(13.37)	(8.10)	(7.10)	(6.70)
Ratesetter	-18.528***	-14.011***	-10.171**	-8.069
	(-8.36)	(-4.62)	(-2.04)	(-1.28)
Fixed Effects	Bank $\times$ ZIP Code $\times$ Shock Year			
Observations	23,957	10,009	3,513	2,598
Adjusted R^2	0.147	0.132	0.186	0.218

Table A8. The Effects of Reputation Shock on County-level Deposit Growth

This table reports the results on the OLS panel regressions of yearly county-level deposit growth on the share of branch deposits that is subject to reputation damage. *Depositgrowth_County* is the county-level deposit growth. *Share_Rep_J* is the share of branch deposits in a county that is subject to a jump in the RepRisk Index by more than J (J=10, 15, 20, or 25). *Deposits_County* is the total deposits of all bank branches in the county. HHI is the is the Herfindahl-Hirschmann index of deposits at all bank branches in the county. Standard errors are clustered at the county by bank level and t-statistics are shown in parentheses. Statistical significance at the 10%, 5% and 1% levels is denoted by *, ** and ***, respectively.

	Depositgrowth_County			
	(1)	(2)	(3)	(4)
Share_Rep_10	1.063			
	(0.71)			
Share_Rep_15		-0.148		
		(-0.09)		
Share_Rep_20			-1.878	
			(-1.56)	
Share_Rep_25				-2.061
				(-1.47)
Lag Deposits_County	-0.064	-0.063	-0.064	-0.064
	(-1.42)	(-1.41)	(-1.41)	(-1.42)
Lag Branch	8.055**	7.939**	7.968**	7.982**
	(2.10)	(2.10)	(2.11)	(2.11)
Lag HHI	34.776***	34.647***	34.699***	34.696***
	(4.16)	(4.17)	(4.17)	(4.17)
Fixed Effects	County + Year			
Observations	15,923	15,923	15,923	15,923
Adjusted R^2	0.158	0.157	0.158	0.157

Table A9. Google Rating and Mortgage Demand During Natural Disasters
The following table reports the complete Table 10.

	Ln(Applications)			
	(1)	(2)	(3)	(4)
Treat × Post × High Rating	0.203***	0.199***	0.149**	0.184***
	(3.57)	(3.15)	(2.53)	(3.05)
Treat × Post	0.355***	0.343***	0.246***	0.269***
	(13.43)	(11.95)	(9.79)	(10.82)
Treat × High Rating	-0.196***	-0.176***	-0.243***	-0.200***
	(-5.43)	(-4.17)	(-5.42)	(-4.14)
Post × High Rating	-0.722***	-0.756***	-0.568***	-0.710***
	(-20.96)	(-19.70)	(-17.26)	(-20.42)
Treat	-0.150***	-0.147***	0.038*	-0.131***
	(-9.72)	(-8.90)	(1.66)	(-5.23)
Post	-0.849***	-0.899***	-0.271***	-0.668***
	(-48.98)	(-45.42)	(-18.23)	(-31.04)
High Rating	0.621***	0.649***	0.720***	0.778***
	(26.75)	(23.86)	(23.09)	(23.97)
Lag Ln(Applications)	0.757***	0.652***	0.112***	0.092***
	(172.97)	(93.75)	(7.79)	(7.00)
Ln(Loan)	0.029***	0.084***	0.068***	0.050***
	(17.27)	(18.13)	(15.33)	(13.09)
ROA	-6.647***	-91.712***	-67.331***	-50.719***
	(-4.29)	(-15.03)	(-13.74)	(-12.01)
Liquidity	-0.589***	-0.275	0.032	-0.045
	(-8.47)	(-0.93)	(0.13)	(-0.21)
Sensitivity	0.400***	2.542***	1.691***	2.270***
	(11.51)	(18.53)	(14.09)	(21.27)
Has Missing	0.067***	0.110***	0.235***	0.095***
	(6.12)	(7.85)	(7.39)	(3.41)
Small	0.077***	0.245	0.072	-0.011
	(8.28)	(1.63)	(0.67)	(-0.12)
Local	-0.024**	-0.252***	-0.278***	-0.226***
	(-2.43)	(-3.32)	(-4.75)	(-4.28)
Important	0.114***	0.237***	0.072	0.059
	(11.78)	(19.59)	(1.14)	(1.08)
Branch Per Capita	-0.000***	-0.000***	0.002***	0.000
	(-8.22)	(-12.46)	(4.15)	(0.73)
Deposits per capita	0.002***	0.001***	0.015**	0.023***
	(4.44)	(3.20)	(2.19)	(3.66)
Unemployment	-1.995***	-2.739***	-5.609***	-13.045***
	(-8.23)	(-10.59)	(-3.90)	(-11.79)
Population	0.069***	0.214***	22.946***	17.777***
	(5.56)	(15.85)	(14.84)	(12.36)
Share of White People	0.154***	0.117***	9.481***	6.267***
	(3.86)	(2.71)	(4.78)	(4.77)
Share of Female	0.897***	1.624***	-2.991	-6.690**
	(3.41)	(5.75)	(-0.67)	(-2.01)
Education	0.829***	1.268***	6.882***	10.437***
	(8.40)	(11.41)	(4.63)	(8.46)
Income	-0.013***	-0.015***	-0.677***	-0.438***
	(-14.07)	(-14.45)	(-41.82)	(-28.26)
Share of Senior People	-0.241**	-0.395***	-33.008***	1.204
	(-2.13)	(-3.53)	(-9.65)	(0.43)
Share of Manufacturing Labor	0.041	0.096	3.592***	3.343***
	(0.59)	(1.23)	(2.69)	(3.12)
Share of Information Labor	4.049***	3.585***	-1.523	3.551
	(6.35)	(5.24)	(-0.38)	(1.13)
Disaster Year × State FEs	+			+
Disaster Year × State × Bank FEs		+		
Bank × County FEs			+	+
Observations	37,173	37,170	37,170	37,170
Adjusted R^2	0.691	0.690	0.606	0.655

Table A10. The Differential Effects of Topic Ratings During Natural Disasters

This table reports the results on triple difference-in-differences regressions that exam the differential effects of topic ratings on mortgage demand following natural disasters. *High Topic Rating* is a dummy variable that equals one if the topic rating of a bank in a county is at the top tercile in the state in the year, and zero if it is at the bottom tercile. The topic rating (*Accessibility*, *Product*, *Hospitality*, *Facility*) is the simple average topic rating of all the reviews of the bank in the county up to the current year, where topic rating is defined as the relative word mentioning on a topic (frequency of key words associated with a specific topic scaled by the total number of words in the review) times the rating of the review. The key words for topics are listed in Appendix A3. *Treat* is a dummy variable that equals one for the disaster affected counties, and zero for control counties. *Post* is a dummy variable that equals one for the disaster incidence year, and zero for the preceding year. Control variables include *Lag Ln(Applications)*, *Ln(Loan)*, *ROA*, *Liquidity*, *Sensitivity*, *Missing*, *Small*, *Local*, *Important*, *Branch*, *Deposit per Capita*, *Unemployment*, *Population*, *White*, *Female*, *Education*, *Income*, *Senior*, *Manufacturing Labor*, *Information Labor*. Standard errors are clustered at the county by bank level and t-statistics are shown in parentheses. Statistical significance at the 10%, 5% and 1% levels is denoted by *, ** and ***, respectively.

	Ln(Applications)							
	Accessibility		Product		Hospitality		Facility	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treat × Post × High Topic Rating	0.206**	0.188**	0.193*	0.193*	0.296***	0.339***	0.319***	0.383***
	(2.55)	(2.37)	(1.74)	(1.79)	(3.46)	(4.08)	(3.17)	(3.86)
Treat × Post	0.425***	0.358***	0.434***	0.363***	0.396***	0.323***	0.411***	0.333***
	(16.53)	(16.11)	(17.10)	(16.66)	(15.61)	(14.84)	(16.21)	(15.29)
Treat × High Topic Rating	-0.091*	-0.172***	-0.109	-0.262***	-0.102**	-0.282***	-0.115*	-0.259***
	(-1.80)	(-2.63)	(-1.59)	(-2.99)	(-2.00)	(-4.21)	(-1.89)	(-3.07)
Post × High Topic Rating	-1.103***	-1.074***	-1.314***	-1.279***	-1.274***	-1.253***	-1.378***	-1.412***
	(-23.94)	(-24.86)	(-21.23)	(-22.15)	(-26.34)	(-28.00)	(-24.42)	(-26.81)
Treat	-0.209***	-0.194***	-0.207***	-0.192***	-0.197***	-0.179***	-0.203***	-0.193***
	(-13.67)	(-8.27)	(-13.93)	(-8.37)	(-12.94)	(-7.68)	(-13.49)	(-8.36)
Post	-1.130***	-0.852***	-1.201***	-0.904***	-1.140***	-0.852***	-1.188***	-0.884***
	(-65.56)	(-41.31)	(-71.63)	(-44.06)	(-68.63)	(-41.61)	(-71.46)	(-42.97)
High Topic Rating	0.680***	0.769***	0.763***	0.823***	0.711***	0.709***	0.816***	0.952***
	(22.58)	(17.72)	(19.46)	(14.19)	(23.28)	(15.39)	(22.27)	(17.61)
Lag Ln(Applications)	0.639***	0.102***	0.642***	0.108***	0.640***	0.104***	0.639***	0.103***
	(95.83)	(7.47)	(97.54)	(7.95)	(96.79)	(7.75)	(96.38)	(7.63)
Ln(Loan)	0.108***	0.059***	0.110***	0.059***	0.119***	0.065***	0.116***	0.062***
	(21.89)	(14.99)	(22.39)	(14.83)	(23.61)	(15.97)	(23.39)	(15.50)
ROA	-116.946***	-66.487***	-119.745***	-67.582***	-121.133***	-68.506***	-120.631***	-68.899***
	(-17.76)	(-15.14)	(-17.93)	(-15.20)	(-18.08)	(-15.40)	(-18.06)	(-15.45)
Liquidity	-0.147	0.206	-0.130	0.188	-0.058	0.203	-0.115	0.204
	(-0.47)	(0.92)	(-0.41)	(0.83)	(-0.18)	(0.90)	(-0.36)	(0.90)
Sensitivity	3.175***	2.751***	3.169***	2.761***	3.083***	2.706***	3.095***	2.734***
	(24.99)	(27.75)	(25.20)	(27.92)	(24.59)	(27.67)	(24.83)	(27.95)
Missing	0.077***	0.064**	0.069***	0.053**	0.064***	0.043*	0.069***	0.053**
	(6.35)	(2.48)	(5.76)	(2.07)	(5.40)	(1.68)	(5.77)	(2.05)

Small	0.351**	0.046	0.316**	0.030	0.348**	0.046	0.299**	0.001
	(2.27)	(0.52)	(2.13)	(0.34)	(2.33)	(0.55)	(1.96)	(0.01)
Local	-0.290***	-0.242***	-0.254***	-0.222***	-0.270***	-0.236***	-0.258***	-0.228***
	(-3.78)	(-4.74)	(-3.33)	(-4.40)	(-3.49)	(-4.63)	(-3.26)	(-4.40)
Important	0.214***	0.066	0.217***	0.087*	0.221***	0.089*	0.218***	0.091*
	(18.81)	(1.29)	(19.21)	(1.71)	(19.44)	(1.74)	(19.09)	(1.79)
Branch Per Capita	-0.000***	0.000	-0.000***	0.000	-0.000***	-0.000	-0.000***	0.000
	(-12.94)	(0.14)	(-13.56)	(0.16)	(-13.27)	(-0.25)	(-13.35)	(0.04)
Deposit Per Capita	0.001***	0.029***	0.002***	0.027***	0.002***	0.029***	0.002***	0.026***
	(4.06)	(4.80)	(5.11)	(4.45)	(5.16)	(4.81)	(4.75)	(4.40)
Unemployment	-2.717***	-13.931***	-2.715***	-14.215***	-2.640***	-14.117***	-2.670***	-14.094***
	(-10.87)	(-12.69)	(-10.95)	(-12.71)	(-10.69)	(-12.75)	(-10.72)	(-12.67)
Population	0.214***	19.741***	0.203***	20.836***	0.200***	21.702***	0.198***	21.247***
	(16.66)	(15.47)	(15.73)	(16.35)	(15.83)	(16.93)	(15.27)	(16.61)
White	0.170***	5.722***	0.176***	5.619***	0.185***	5.665***	0.191***	5.622***
	(4.07)	(4.34)	(4.24)	(4.20)	(4.48)	(4.30)	(4.58)	(4.19)
Female	1.534***	-5.557*	1.526***	-6.010*	1.568***	-5.948*	1.591***	-6.213*
	(5.71)	(-1.65)	(5.72)	(-1.77)	(5.87)	(-1.77)	(5.89)	(-1.84)
Education	1.353***	10.679***	1.412***	11.381***	1.408***	10.799***	1.440***	11.101***
	(12.82)	(8.77)	(13.50)	(9.15)	(13.40)	(8.79)	(13.67)	(8.99)
Income	-0.017***	-0.455***	-0.017***	-0.476***	-0.017***	-0.465***	-0.017***	-0.477***
	(-16.60)	(-31.92)	(-17.07)	(-32.97)	(-16.87)	(-32.58)	(-17.34)	(-33.06)
Senior	-0.462***	-1.795	-0.506***	-1.192	-0.533***	-1.363	-0.552***	-1.254
	(-4.29)	(-0.64)	(-4.71)	(-0.43)	(-4.96)	(-0.48)	(-5.10)	(-0.44)
Manufacturing Labor	0.172**	1.927*	0.191***	1.885*	0.172**	2.042*	0.155**	1.944*
	(2.39)	(1.87)	(2.68)	(1.80)	(2.42)	(1.95)	(2.15)	(1.87)
Information Labor	3.772***	5.084	3.871***	5.357*	3.692***	4.339	3.710***	4.864
	(5.63)	(1.61)	(5.80)	(1.68)	(5.57)	(1.38)	(5.54)	(1.52)
DisasterYear × State FEs		+		+		+		+
Disaster Year × State × Bank FEs	+		+		+		+	
Bank × CountyFEs		+		+		+		+
Observations	42,476	42,478	42,826	42,828	42,922	42,924	42,948	42,950
Adjusted R-squared	0.684	0.64	0.682	0.638	0.685	0.642	0.683	0.64

Table A11. Google Rating and Mortgage Demand by Loan Purpose

This table re-estimates Table 10 by examining mortgage applications by loan purpose (new purpose or refinance). *Treat* is a dummy variable that equals one for the disaster affected counties, and zero for control counties. *Post* is a dummy variable that equals one for the disaster incidence year, and zero for the preceding year. *High Rating* is a bank-county-year-level dummy variable that equals one if the bank's average Google rating in the county is at the top tercile among all banks within the same state, and zero if it is at the bottom tercile. *Ln*(*Applications*) is the natural logarithm of the annual number of mortgage applications to the bank in the a county. Control variables include *Lag Ln*(*Applications*), *Ln*(*Loan*), *ROA*, *Liquidity*, *Sensitivity*, *Missing*, *Small*, *Local*, *Important*, *Branch*, *Deposit per Capita*, *Unemployment*, *Population*, *White*, *Female*, *Education*, *Income*, *Senior*, *Manufacturing Labor*, *Information Labor*. Standard errors are clustered at the county by bank level and t-statistics are shown in parentheses. Statistical significance at the 10%, 5% and 1% levels is denoted by *, ** and ***, respectively.

	Log Applications							
	New Purchase				Refinance			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treat × Post × High Rating	0.182***	0.166***	0.112**	0.138**	0.132***	0.134**	0.096*	0.129**
	(3.47)	(2.82)	(2.10)	(2.55)	(2.60)	(2.38)	(1.83)	(2.42)
Treat × Post	0.285***	0.283***	0.193***	0.212***	0.309***	0.293***	0.214***	0.235***
	(11.46)	(10.48)	(8.36)	(9.26)	(12.26)	(10.78)	(9.15)	(10.11)
Treat × High Rating	-0.177***	-0.163***	-0.197***	-0.157***	-0.154***	-0.124***	-0.197***	-0.162***
	(-5.36)	(-4.12)	(-4.86)	(-3.61)	(-4.70)	(-3.29)	(-4.87)	(-3.78)
Post × High Rating	-0.561***	-0.599***	-0.460***	-0.569***	-0.577***	-0.592***	-0.426***	-0.546***
	(-18.13)	(-17.30)	(-15.71)	(-18.48)	(-18.78)	(-17.33)	(-14.60)	(-17.78)
Treat	-0.133***	-0.132***	0.019	-0.111***	-0.122***	-0.118***	0.035	-0.108***
	(-9.02)	(-8.20)	(0.90)	(-4.70)	(-8.32)	(-7.42)	(1.62)	(-4.61)
Post	-0.817***	-0.852***	-0.244***	-0.568***	-0.742***	-0.789***	-0.241***	-0.581***
	(-52.19)	(-47.45)	(-18.18)	(-27.83)	(-45.16)	(-42.36)	(-17.52)	(-28.68)
High Rating	0.493***	0.531***	0.572***	0.615***	0.526***	0.541***	0.613***	0.659***
	(23.54)	(21.38)	(20.42)	(21.02)	(25.09)	(22.21)	(21.83)	(22.85)
Controls	+	+	+	+	+	+	+	+
Disaster Year × State FEs	+			+	+			+
Disaster Year × State × Bank FEs		+				+		
Bank × County FEs			+	+			+	+
Observations	33,613	33,610	33,610	33,610	35433.000	35,430	35,430	35,430
Adjusted R^2	0.663	0.667	0.614	0.655	0.703	0.706	0.644	0.685