Bank Competition amid Digital Disruption: Implications for Financial Inclusion

Erica Xuewei Jiang, Gloria Yang Yu, and Jinyuan Zhang*

Abstract

This paper studies *how* banks compete amid digital disruption and the resulting distributional effect across consumers. Digital disruption increases the geographic coverage of banking services, bringing new entrants to local markets. However, as digital customers shift from branches to digital services, banks close branches, and the remaining branching banks gain market power among non-digital customers that rely on branches. Consequently, digital customers benefit from the intensified bank competition at the cost of non-digital customers who pay higher prices for branch services and face the risk of financial exclusion. We provide empirical evidence by exploiting the staggered expansion of 3G networks, instrumented by regional distribution of lightning strike frequency. Using a structural model, we further quantitatively decompose the benefit and costs of digital disruption resulting from banks' pricing, branching, and entry decisions. The results highlight the role of banks' endogenous responses to digital disruption in widening the gap in access to banking services.

*Jiang is from the University of Southern California, Marshall School of Business. erica.jiang@marshall.usc.edu. Yu is from Singapore Management University. gloriayu@smu.edu.sg. Zhang is from UCLA Anderson School of Management. jinyuan.zhang@anderson.ucla.edu. Runtong Cai provided outstanding research assistance. We are grateful to Kenneth Ahern, Greg Buchak, Matteo Benetton, Philip Bond, Marco Di Maggio, Andrea Eisfeldt, Valentin Haddad, Barney Hartman-Glaser, Gerard Hoberg, Rawley Heimer, Gregor Matvos, Shaun McRae, Clemens Otto, Amit Seru, Anthony Lee Zhang and seminar and conference participants at UCLA Anderson, USC Marshall, Singapore Management University, University of Washington at Seattle, University of Technology Sydney, IIOC 2022 for comments. Yu thanks the support from Sim Kee Boon Institute for Financial Economics at SMU. *First version: Feb, 2022.*

1 Introduction

The impact of technological innovation—e.g., mobile banking and online lending—on the banking industry and financial inclusion is central to policy discussions.[1] The widely accepted view is that fintech can democratize access to financial services, increase the competition of financial intermediaries, and improve financial inclusion (Philippon, 2016, 2019). However, not everyone has equal access to digital services, either due to a lack of capability to understand technology or unable to afford a device.[2] Survey data shows a previously overlooked sharp divergence in how consumers access banking services over the past decade: digital consumers—i.e., younger, more-educated, and higher-income—adapt to digital platforms quickly, while non-digital—older, less-educated, and lower-income—still heavily rely on branches.

Motivated by this observation, we study *how* banks compete amid digital disruption and the resulting distributional effect across consumers. Although non-digital consumers still rely on branches, digital disruption has shifted digital consumers' preference from branch services to digital services. As the average preference for branches declines, banks close costly branches, and more digital banks enter the market. The intensified competition from digital banks forces incumbents with branches to specialize in the market segment in which they have a comparative advantage: branching banks target non-digital customers and exploit market power on them by charging higher prices. Consequently, although digital consumers benefit from intensified bank competition, non-digital consumers bear higher costs to access banking services and thus face the risk of financial exclusion. This distributional effect is important with the rising concern about the aging society and digital inequality.

To empirically examine the above story, we exploit the staggered introduction of the third-generation wireless mobile telecommunications (3G) networks. The 3G technology is the critical infrastructure that allows users to freely browse the internet anywhere and to access banking services without going to the physical branches. As the 3G infrastructure was slowly constructed in different regions across the U.S., this setting provides us with substantial variations in both the time series and the cross-section.

We begin by verifying the digital divide premise using the Federal Deposit Insurance

[1] See, for example, digital divide, digital inequality and United Nation's discussion.

[2] Such a social phenomenon is called the "digital divide" by policymakers and academics.

Corporation (FDIC) surveys. We show that younger, richer, and more educated consumers shift from branches to mobile banking as 3G expands to their residential areas, while older, poorer, and less educated consumers have barely adopted mobile banking. The diverging preference is robust to different specifications of fixed effects.

Our empirical analysis unfolds in two steps. We first examine how the expansion of 3G networks affects banks' competition dynamics regarding branching decisions, geographic expansion, and pricing in a staggered difference-in-difference (DiD) framework. We establish a significantly positive relationship between the expansion of 3G networks and branch closures at the bank-county level: banks shut down more branches in regions with higher 3G coverage. Such effect is robust to the inclusion of bank-county and bank-state-year fixed effects. At county-level, the total number of branches decreases by 1.3% when 3G networks fully cover the county. Moreover, the impact of digital disruption on branch closures is much more salient in counties with more young consumers who have a lower preference for branches.

Moreover, digital services empowered by 3G networks allow banks to expand geographically with fewer or no branches. As 3G enters and covers the entire region, the number of branches a bank has in regions where it originates mortgages declines by 2%, and new entrants have 0.6% fewer branches.

Our conceptual framework also uncovers a novel effect of digital disruption on banks' pricing strategies. On the one hand, the intensified competition forces banks to charge lower prices. On the other hand, as banks close branches after digital disruption, non-digital consumers, who are less adaptive to digital services, are left with limited choices. To some degree, digital disruption transforms a market from a pooling equilibrium, where all serve both types of consumers, to a separating equilibrium, where some banks retain more branches and serve most non-digital customers, while others develop digital services and target digital customers. As a result, the consumer pool of branching banks contains more non-digital customers, enabling them to exploit market power on the latter. With more inelastic demand, branching banks optimally charge higher prices.

We take this hypothesis to the data and confirm that banks with more local branches charge higher prices relative to banks with fewer or no local branches as 3G coverage increases. In the deposit markets, we find that the deposit spreads charged by banks with branches increase with 3G coverage. In the lending market, banks with more branches charge higher loan origination fees as the 3G networks expand, relative to banks with fewer branches.

To address the endogeneity concern that omitted factors drive both 3G network expansion and banking decisions, we exploit an instrumental variable (IV) strategy, following Manacorda and Tesei (2020) and Guriev et al. (2021). We use the population-weighted frequency of lightning strikes per area to predict the expansion of 3G networks. Frequent lightning strikes substantially increase the costs of providing service and maintaining the infrastructure, hence slowing down the rollout of 3G construction. The IV regressions confirm the causal impact of the expansion of 3G networks on banks' branching, entry, and pricing decisions.

The above findings collectively suggest that a new banking market structure emerges amid the digital disruption after the 3G expansion: banks *without* a competitive advantage in operating branches compete on prices and serve consumers that prefer digital services, whereas banks *with* a competitive advantage in operating branches invest in branches, charge higher prices, and serve consumers that rely on branch services.

After establishing how digital disruption changes banks' branching and pricing decisions, we discuss the subsequent benefits and costs to consumers. On the benefit side, we have witnessed the expansion of the geographic scope of competition from the local market to the national market over the past decade. Take the mortgage market as an example, from 2009 to 2017, the entire distribution of the number of counties covered by each lender shifted rightward. The geographic expansion of banks increases local competition. We confirm that the expansion of 3G networks and the resulting changes in banks' competing strategies partially contribute to these trends. The expansion of 3G networks leads more banks to serve a region. Quantitatively, as 3G coverage increases from 0 to 100%, a region is served by 3.2% more banks. Hence, a full 3G coverage reduces local lending market concentration by 47.3 bps. Relative to the average county concentration, the effect translates into an economically meaningful 5.2% reduction in concentration.

However, the benefits are not equally shared among consumers, and non-digital consumers sometimes can even suffer. We show that non-digital consumers pay higher prices to access banking services—higher mortgage origination fees and deposit spreads—after 3G expands in a region. Moreover, survey data shows that non-digital consumers, including those who are more senior, lower-income, and less-educated, are more likely to be excluded from banking services after 3G arrives in town. Importantly, the 3G expansion increases the unbanked rate of non-digital consumers partially by causing those banked individuals to lose banking access. According to reported reasons for leaving banks of these previously banked individuals, *high fees* is a critical aspect associated with their opt-out decision after the 3G networks arrive.

The reduced form results highlight the overall benefits and costs of digital disruption through banks' endogenous changes in their branching, pricing, and entry decisions. In reality, these changes can occur simultaneously, and thus, the reduced-form analyses can not distinguish the impact of each decision on the costs of digital disruption. To provide more insights along this line, we build a structural model to formalize our conceptual framework and quantitatively decompose the effects of various banking decisions.

The basic model framework is in the spirit of Berry et al. (1995).[3] Our main innovation is to model banks' endogenous pricing, branching, and entry decisions simultaneously as they face heterogeneous consumer preferences for branch services. On the demand side, consumers obtain banking services characterized by prices, the number of branches, and digital banking quality bundles. Motivated by the stylized facts, we model two groups of consumers, young and old, who differ in their preference for branching and digital services. On the supply side, banks compete to set prices for their services and choose the number of branches that incur costs. As consumers value branch services, banks optimally choose to operate branches and charge markups for their banking services.

We estimate the model using bank deposit market data for pre- and post-digital disruption as measured by the local 3G coverage ratio. We use the estimated model to decompose the impacts of banks' responses on different consumers. Digital disruption allows more digital consumers to access banking services via cheaper options. On average, digital consumers pay lower prices and are more likely to be financially included. However, for non-digital consumers, when banks only adjust prices, they pay 0.8% higher prices, which arises from a 7% increase in branching banks' prices as they understand their customer pools have come more captive after the digital disruption. As banks shut down branches, 11% of non-digital consumers leave the banking system. Lastly, as new banks enter the market, digital consumers benefit significantly from the fiercer competition, while non-digital consumers start to be better off from the digital disruption. Meanwhile, however, we see the largest difference in financial inclusion between the two groups of consumers. In other words, banks' endogenous

[3]The discrete choice model by Berry et al. (1995) has been applied to the banking sector by Buchak et al. (2018a); Jiang (2019); Xiao (2020); Benetton (2021); Wang (2020); Robles-Garcia (2019).

responses to digital disruption may turn the digital divide into digital inequality.

Our paper contributes to the growing literature on the costs and benefits of financial technology. The literature argues that digital disruption will likely bring in new players, increase competition in the banking industry, and enhance consumers' welfare (Philippon, 2016; Vives, 2019). In terms of the benefits of financial technology, Buchak et al. (2018b) and Fuster et al. (2019) highlight that technology increases the speed of mortgage origination without causing higher defaults. Bartlett et al. (2019) find that algorithmic scoring, compared to face-to-face assessment, reduces price discrimination in the lending market. Di Maggio et al. (2021) find that the use of alternative data can better assess borrowers' creditworthiness. In terms of the costs, Fuster et al. (2020) and Blattner and Nelson (2021) suggest that minorities benefit less from machine learning models due to their noisy hard information, and Jiang et al. (2022) finds that the use of machine limits the incorporation of same-race loan officers' soft information. Our paper extends both views by pointing out that while digital disruption benefits digital consumers by bringing in new banks, the enhanced competition does not benefit non-digital consumers equally, even if digital services are accessible to everyone.

Along this line, we also contribute to the literature on financial inclusion. Traditionally, branches play an essential role in providing banking services to local residents in both developing countries (Fonseca and Matray, 2022) and developed countries (Brown et al., 2019). The recent digital disruption provides new opportunities for financial inclusion, reaching income and racial groups with low financial representation (Philippon, 2019; Yogo et al., 2021). Our paper connects these two strings of literature and highlights the impacts on financial inclusion from the interplay between digital disruption and branches. The externality of digital disruption through changing banks' pricing and branching decisions can reversely undermine some under-representative households.

Amid digital disruption, a rising concern among policymakers is that disparate access to digital services can contribute to persistent social inequality.[4] Digital divide can be caused by the availability of advanced infrastructures (Saka et al., 2021; Lee et al., 2021), and high transaction costs (Pierre et al., 2018; Jack and Suri, 2014). WorldBank (2016) emphasizes that developing regions do not benefit from new digital technologies owing to the lack of

[4]Many policy discussions about unbanked population. For example, https://www.fdic.gov/analysis/household-survey/index.html

highspeed internet. Our results highlight a new force that can cause digital inequality even in economically developed regions: the endogenous bank responses to consumers' heterogeneous preferences for digital services.

This paper also contributes to the literature on banking competition (Cetorelli and Strahan, 2006; Garmaise and Moskowitz, 2006; Drechsler et al., 2017; Jiang, 2019; Buchak et al., 2018a; Buchak and Jørring, 2021; Benetton, 2021; Robles-Garcia, 2019). Most of the existing papers focus on banks' price competition; see, for example, Egan et al. (2017); Xiao (2020). Our paper adds to this literature by showing how banks' branching decisions interact with pricing decisions when consumers have heterogeneous preferences for branch services. In this regard, we also contribute to the literature that studies the real effect of banks' branch networks (Jayaratne and Strahan, 1996; Huang, 2008; Jayaratne and Strahan, 1997; Beck et al., 2010). Ménard and Ghertman (2009) and Hubbard and Hubbard (1994) show that branching facilitates diversification of bank portfolios and hence stabilizes banking systems. Carlson and Mitchener (2006) and Kuehn (2018) point out branch banking increases competition, causing the exit of weak banks.

The rest of this paper proceeds as follows. Section 2 presents a conceptual framework on how digital disruption affects bank competition. Section 3 describes the data. Section 4 presents all empirical results. In Section 5, we build a structural model to formalize our conceptual framework and quantitatively decompose the effects of various banking decisions Section 6 concludes.

2 Conceptual Framework

This section introduces a conceptual framework about how digital disruption affects bank competition, which is based on the intuition of the model that we are going to develop in Section 5.

Consider an economy with two groups of consumers that differ in their preferences for banking services. Non-digital consumers prefer branch services to digital services, while digital consumers are the other way around. They obtain banking services from two types of banks, which have different comparative advantages. Traditional banks (T-banks) have lower marginal costs of operating branches, while FinTech banks (F-banks) offer better digital services. Facing consumers who value lower prices and better services, banks compete on both dimensions.

As shown in Figure 1, technological innovation over the past decade has been accompanied by a shift in the primary way to access banking services from branches to mobile apps. From figure 2, we see that the shift is mainly driven by digital consumers, who are younger, richer, and more-educated. Motivated by these facts, we model digital disruption as a preference shock, which simultaneously reduces digital consumers' preference for branch services and increases their preference for digital banking. Meanwhile, we assume that non-digital consumers still heavily rely on branches.

As the average preference for branches declines after the digital disruption, banks shut down costly branches, and digital banks enter the market. The intensified competition from F-banks forces incumbents to specialize in the market segment in which they have a comparative advantage: T-banks target non-digital customers and exploit market power on them by charging higher prices, while F-banks target digital customers and compete on prices. Consequently, although digital customers benefit from intensified bank competition, non-digital customers suffer from having fewer branches serving the market segment, paying higher costs to access banking services, and eventually facing a higher risk of financial exclusion.

According to this framework, we derive the following testable predictions related to banks' endogeneous responses to digital disruption:

PREDICTION 1. *Digital disruption induces banks to shut down branches, especially in regions with more young consumers.*

Before the digital disruption, consumers rely on branches, which limits the geographic expansion of banks as operating branches is costly. As digital disruption lowers the value of branches, banks with a higher marginal cost of operating branches are able to attract customers without having a branch, which increases their profitability. Moreover, as more young consumers shift to F-banks for better digital services, the aggregate market share of F-banks increases. Both the increased profit margin and the enlarged market share invite more F-banks to enter the market.

PREDICTION 2. *Digital disruption induces entries of F-banks.*

Unlike the traditional view, our model shows that these new entries do not lead to a uniform reduction of the prices charged by banks on their services. On the one hand, the new entry of F-bans intensifies competition and forces F-banks to lower prices. On the other hand, as banks close branches after digital disruption, non-digital consumers, who are less adaptive to digital services, are left with limited choices. In some sense, digital disruption transforms a market from a pooling equilibrium, where both types of banks serve both types of consumers, to a separating equilibrium, where T-banks serve most non-digital customers, and F-banks serve most digital customers. As a result, banks with a competitive advantage in operating branches (i.e., T-banks) strategically shift their focus toward non-digital consumers. As T-banks' demand curve becomes more inelastic, they optimally charge higher prices.

PREDICTION 3. *Digital disruption leads to diverging pricing strategies of the two types of banks: T-banks charge higher prices while F-banks charge lower prices.*

Banks' endogenous responses result in the following effects of digital disruption:

PREDICTION 4. *Digital disruption increases the number of banks that serve each region and lowers the market concentration as measured by the HHI.*[5]

PREDICTION 5. *Following digital disruption, digital consumers pay lower prices to access banking services, while non-digital consumers pay higher prices.*

PREDICTION 6. *Following digital disruption, the unbanked rate of digital consumers declines, while the unbanked rate of non-digital consumers rises.*

3 Data and Descriptive Statistics

3.1 Data Sources

FDIC Survey of Household Use of Banking and Financial Services The survey has been conducted by the FDIC biennially since 2009. Each survey collects responses from around 33,000 consumers, including their bank account ownership, like whether they are

[5] Both digital consumers and non-digital consumers benefit from reduced concentration as long as the market is not completely segmented, i.e., non-digital consumers can only access banking services through branches. We elaborate this point when we formally introduce the model in Section 5.

bank or unbanked, the primary methods they access their bank accounts if they are banked, why they are unbanked if they don't have a bank account, and saturated set of demographic information. Like other survey data, it puts weights on each response to help reflect the full population.

3G Coverage We use digital maps of 3G network coverage from 2007 to 2018 provided by Collins Bartholomew's Mobile Coverage Explorer (Guriev et al. 2021). These maps gather coverage data that mobile network operators submit to the GSM Association, and essentially provide an indicator variable identifying the availability of 3G for each 1×1-kilometer binary grid cell. To combine data on mobile network coverage with the county-level banking data, we calculate 3G coverage in each county-year as the weighted average of the value of 3G availability weighted by the population density in each grid cell across one county's polygon. This measure captures the proportion of population in one county with access to the 3G networks.

Frequency of Lightning Strikes The World Wide Lightning Location Network (WWLLN) provides the exact coordinate and timestamps of all detected lightning strikes in the US. To measure the degree to which a county is affected by lightning strikes, we average the annual number of lightning strikes during our sample period, weighting each strike by the population density in each grid cell across one county's polygon. The measures reflect the amount of population potentially affected by lightning strikes.

Bank Branch Information The bank branch-level information is extracted from the Federal Deposit Insurance Corporation (FDIC), which is the annual survey of branch offices as of June 30 each year for all FDIC-insured institutions. Note that FDIC only insures deposits in banks, so this data does not include FDIC-insured entities, such as credit unions.

Deposit Rate RateWatch provides the interest rates paid on the branch-level deposits. The interest rates paid on branch-level deposits are obtained from RateWatch, which provides weekly deposit rates on products that include certificates of deposit (CDs), money market accounts, etc. The data are aggregated to the quarterly frequency by averaging the deposit rates for each product of each branch. We focus on $10,000 12-month and 36-month CDs, which are the most popular time deposit product offered across all U.S. branches.

Lending We obtain loan-level mortgage origination data from Home Mortgage Disclosure Act (HMDA) database. HMDA includes the vast majority of residential mortgage applications in the United States.

County Demographics County-level demographic features, including GDP, population, employment, and per capita income, are collected from the BEA. Variables pertaining to real economic outcomes are obtained from the Quarterly Census of Employment and Wages; these variables include the annual average of quarterly business establishment counts and of monthly employment.

3.2 Descriptive Statistics

Household use of banking services Table 1 presents summary statistics for the key variables. Panel A reports respondents' characteristics from FDIC surveys. The banked population increased from 92.4% in 2009 to 94.6% in 2019. Compared to unbanked households, banked households are older (50.4 v.s. 43.9 years old), higher-income ($57,915 v.s. $21,277), better-educated, more likely to be white individuals, and higher chance to have phones.

The primary way to access banking services of banked individuals was through online banking in 2013 and turned to through mobile banking in 2019. We also see a significant drop in the usage of branches as the main way to access banks, from 32.5% in 2013 to 21.0% in 2019. On average, individuals who prefer branches are older, poorer, and under-educated than those who use online and mobile banking.

Until 2019, there is still 5.4% of the unbanked population, which represents approximately 7.1 million U.S. households. Strikingly, 49.2% of these unbanked population once had a bank account, but lost their access to banking services. These previously banked individuals have a slightly higher proportion of white ethnics but similar to other unbanked populations in other aspects. Panel B examines the reasons for households being unbanked. While the lack of wealth is the major reason, unbanked households are also not concerned about high fees and lack of trust in banks.

Other variables Panel C and D of Table 1 describe county and bank characteristics in our sample. To avoid the influence of outliers, all variables (expect for county characteristics) are winsorized at the 1% and 99% level.

4 Reduced Form Evidence

We present empirical evidence for predictions in Section 2. We begin by introducing our empirical design and identification. We then provide evidence for the digital divide and how banks respond to digital disruption before studying the resulting benefits and costs.

4.1 Empirical Design and Identification

4.1.1 3G Expansion

3G technology drives most of the growth in individual broadband subscriptions over the past decade.[6] 3G mobile service allows users to freely browse the internet anywhere and access banking services without going to the physical branches. 3G coverage affects consumers' reliance on bank branches (i) at the extensive margin by affecting the probability of getting banking services via digital channels rather than branches, (ii) at the intensive margin by affecting the frequency of using bank branches, and (iii) qualitatively by changing what transactions people do with banks through branches. The qualitative advantages that a mobile broadband connection brings for a number of banking transactions, such as bank account management and transfer, are particularly well-suited for digital access. The ease of connection also makes a qualitative difference by engaging users in digital banking (Rainie and Wellman 2012).

We exploit the timing of 3G expansion in the US as US banks' exposure to technology disruption. Figure 3 illustrates the expansion of 3G networks at the county level over the sample period. It presents maps of 3G coverage in 2007, 2012, and 2018. Evidently, the expansion of 3G is staggered across regions and over time.

[6] According to OECD data in 2013, mobile broadband subscriptions in OECD regions saw year-over-year growth of 16.63 percent, and the total subscriptions were more than double those of Fixed wired broadband.

4.1.2 Main Specifications and Identification

We examine the effect of 3G expansion on banks' branching and pricing decisions in difference-in-differences (DiD) settings:

$$\text{County level: } Y_{c,t} = \beta 3\text{G Coverage}_{c,t} + \lambda X_{c,t} + \mu_c + \nu_{s,t} + \epsilon_{c,t}, \tag{1}$$

$$\text{Bank-county level: } Y_{b,c,t} = \beta 3\text{G Coverage}_{c,t} + \lambda X_{c,t} + \mu_{b,s,t} + \nu_{b,c} + \epsilon_{b,c,t}, \tag{2}$$

where b, s, c and t index bank, state, county and year respectively. Specification (1) is at county-level, while specification (2) is at bank-county level. In both specifications, the key variable of interest is $3GCoverage_{c,t}$, which is the share of the population with potential access to 3G in county c in year t. $X_{c,t}$ is a vector of county controls, including GDP, per capita income, and the population at the county level. We include these variables to capture the economic development in the areas, which potentially relates to a faster expansion of the 3G networks.

The inclusion of fixed effects makes both specifications feature a DiD design. In Specification (1), μ_c and $\nu_{s,t}$ are county fixed effects and state-year fixed effects, respectively, which absorb county-specific time-invariant characteristics and state-level time-specific shocks and eliminates relevant variation. The inclusion of μ_c allows us to exploit the time-series variation within a county. Similarly, in Specification (2), $\mu_{b,c}$ is bank-county fixed effects, which allows us to exploit the time series variation in bank b's decisions in county c. $\nu_{b,s,t}$ is bank-state-year fixed effects, which controls for things like shocks to bank b's business activities in state s in year t, idiosyncratic exposure to changes in state regulation, and other bank-specific or state-specific shocks.

Including these fixed effects provides a stringent identification. In Specification (1), we exploit the cross-sectional variations in changes in 3G coverage within a state-year, while in Specification (2), we examine a given bank's responses across counties that experience different levels of 3G expansion within a state-year. The main identification assumptions are that the 3G expansion is uncorrelated with unobservable factors that also trigger banks' branching, pricing, and entry decisions, and that 3G coverage itself is not driven by changes in banks' behaviors or the aforementioned omitted variables. Although these assumptions are not directly testable, we present several robustness checks such as pre-trend analysis in support of the causal interpretation.

Bartik instrument for 3G expansion To address the remaining concerns about the identification assumptions, we adopt the IV approach proposed by Manacorda and Tesei (2020). We construct the population-weighted frequency of lightning strikes per square kilometer, and use it to instrument the speed of 3G expansion, following Guriev et al. (2021); Manacorda and Tesei (2020). The relevance condition between lightning spikes and the speed of 3G network expansion has been verified by multiple studies (Andersen et al., 2012). Frequent lightning and the resulting electrostatic discharges can damage the infrastructure for mobile coverage and negatively affect the transmission of signals. These negative impacts reduce the profits of service providers as power protection and maintenance are costly and increase the risk of intermittent communications. Therefore, we expect that areas with more lightning incidents have lower supply and slower adoption of 3G networks.

Specifically, the first stage of the 2SLS is specified as follows:

$$\text{3G Coverage}_{c,t} = \beta_1 \text{High Lightning}_c \times t + \beta_2 \text{High Lightning}_c \times t^2 \quad (3)$$

$$+ \gamma X_{c,t} + \mu_c + \nu_{s,t} + \epsilon_{c,t}. \quad (4)$$

High Lightning_c is an indicator variable that equals 1 if county c's average population-weighted frequency of lightning strikes from 2007 to 2018 is higher than the state median, and 0 otherwise. We interact lightning strikes with time trend t and t^2 to capture the growth feature of 3G coverage, similar to Guriev et al. (2021). Moreover, to take into consideration that the initial status of 3G networks in our sample may affect the speed of expansion, we add the interaction term of time trends and county-level 3G coverage in 2007. We then estimate the second stage using predicted county-level 3G coverage.

Column 1 of Table 2 presents the first-stage relationship. The significantly negative coefficient confirms the tight relationship between the frequency of lightning strikes and 3G coverage. The estimated Cragg-Donald Wald F statistic is 20.68, higher than the 5% significance critical value for a Stock-Yogo weak IV test.

The identification assumption is that the frequency of lightning strikes affects banking decisions in the local region only through its impact on the expansion of 3G networks conditional on other covariates. The exclusion restriction is likely to be valid in our context because banks' decisions to close branches are unlikely to be driven by weather conditions.[7]

[7]Notice that extreme weathers asymmetrically affect banks' branch opening and closure decisions. To the extent

One potential concern is that lightning strikings may concentrate in certain areas (e.g., southeastern states) whose common characteristics other than lighting frequency can generate spurious correlations between lightning strikes, 3G coverage, and bank decisions. To mitigate this concern, we define high lighting frequency counties *within a state* and thus account for regional factors such as geographical locations. In addition, this approach ensures that the estimation is not given by a handful of states with extreme weather conditions.

Figure 4 plots the distribution of high-lightning-frequency counties across the US. Within each state, we still see a clustering of counties on a certain side of a state. Table IA.1 shows a balance of characteristics (growth trend of GDP, population, unemployment rate, and the share of young population) between counties with high and low lightning strikes frequencies within the same state. This evidence alleviates the concern that lightning strikes may affect 3G coverage through economic conditions.

4.2 Premise: Digital Divide

We begin by presenting evidence for consumers' heterogeneous preferences for digital banking and how their preferences diverge over time. Figure 2 shows a sharp increase in the usage of mobile banking as the primary way to access banking services for younger, more-educated, and higher-income consumers. In 2019, mobile banking became the dominant way to access banking services for these groups of consumers, whereas older, less-educated, and lower-income households still heavily rely on branches. We confirm in Table IA.2 that such divergence coincides with the staggered 3G expansion: younger, richer, and more educated consumers shift away from branches and towards online and mobile banking after 3G expands to their residential areas. The diverging preference is robust to different specifications of fixed effects.

that weather is relatively persistent, it shall only affect banks' initial decisions with a county—branch opening. For example, if perennial bad weathers force local residents to rely on vehicles going out, banks may choose to locate their branches sparsely when entering the area. However, the same weather condition would not then affect whether banks close these branches.

4.3 Bank Endogeneous Responses

4.3.1 Branch Closure

Bank-county level results As Prediction 1 states, banks optimally close branches as digital disruption reduces average consumers' preference for branches. We examine whether banks close more branches in regions with higher 3G coverage by estimating Specification (2) with the log of the number of branches as the outcome variable.[8] Table 2 presents the results. The coefficient estimates on 3G coverage are significantly negative in column 2 with a less saturated specification as well as in column 3, which corresponds to Specification (2), suggesting that the expansion of 3G networks is indeed associated with banks' decisions to close branches. The number of branches of an average bank in a county drops by 6.6% after the area is fully covered by 3G networks (column 3). We confirm the qualitative result in the IV setting introduced in Section 4.1.2. The magnitude of the IV estimate in column 4 of Table 2 is larger than that of OLS estimates. One potential explanation is that consumers in regions with frequent lightning strikes may favor benefits brought by 3G networks more, and in response, banks close branches more aggressively in these regions.[9]

County level results The above bank-county analysis focuses on individual banks' decisions. Local consumers' access to branch services may not be affected if more banks enter the local market, opening new branches to replace the closed branches. We then study the impact of 3G coverage on county-level branch closures by estimating Specification (1).

On average, an increase of 3G coverage from 0 to 100% leads to a 1.3% reduction in the total number of branches (Table 2 column 6). Roughly speaking, the total number of branches dropped by 9% during our sample period, and the 3G expansion contributed 15% to this decrease. The result holds in the IV setting (Table 2 column 7). Again, the IV estimate is bigger than the OLS estimate, suggesting that the effect of 3G expansion on branch closure is more prominent in high-lighting strikes regions.

We further examine the heterogeneous effects on regions with more and fewer digital

[8] In robustness checks, we use the number of branch closures scaled by last year's total branch numbers as the outcome variable. This alternative measure ensures that the closed branches are not replaced by new ones.

[9] In such a case, IV regressions effectively estimate the *local* average treatment effect (ATE) in regions with frequent lightning strikes, whereas OLS regressions estimate the ATE over the entire sample. Similar magnitude differences are shown in Guriev et al. (2021).

consumers, which are approximated by the median age of a county. We consider counties with a median age below 40 to have more digital customers. The significantly negative coefficient on the interaction term between 3G coverage and the young county dummy shows that the 3G-induced branch closures mainly concentrate in counties with more digital consumers. This result is again in line with Prediction 1.

Parallel Trend Dynamic panel analyses may be subject to a concern that regions with early 3G introduction experience different trends in bank branch closures from areas with late 3G introduction. We alleviate such a concern by conducting an event study in Appendix C by exploiting sharp increases in counties' 3G coverage. We define a treatment event as a county's 3G coverage increasing by more than 50% from the previous year. Given the monotonic increasing feature of 3G coverage, such an event can happen at most once for one county. For each treated county, we construct a control county if a county has the closest matching score based on county characteristics but did not experience a sharp increase in 3G coverage ever or reach 30% 3G coverage within three years upon the treatment event. Figure 6 presents the dynamic differences between the two groups around years with sharp increases in 3G coverage. We see that the treated counties start to experience significant branch closures at the time of sharp 3G networks expansion, and this effect gets stronger over the years. In contrast, the differences during the 3-year pre-event window are small in magnitude and statistically indistinguishable from zero. Hence, there exist no pretends.

4.3.2 Entry and Geographic Expansion

As Prediction 2 states, digital disruption induces new entries with fewer or no branches. We examine whether banks are more likely to enter new markets without opening a branch after the introduction of 3G. To this end, we estimate Specification (2) with the outcome variable being the log of the number of branches a bank has in a county-year using a sample of entrants that did not originate loans in the county before.[10]

As 3G expands to a region, banks enter a new market with 0.6% fewer branches (Table 3 column 1). The IV regression in Table 3 column 2 confirms that 3G coverage causally reduces banks' propensity to rely on branches when entering a new county. We further confirm that

[10]We need to observe banks' activities even if they do not have branches in a region. Hence, we use the HMDA data, which collects banks' mortgage origination activities in all counties based on borrowers' location.

banks serve the mortgage market in a region with 2% fewer branches after 3G coverage increases from 0 to 100% (Table 3 column 3). The result holds in the IV setting in column 4. These results confirm Prediction 2 that the emergence of digital disruption induces entries of banks with fewer branches.

4.3.3 Pricing

Our model uncovers a novel effect of digital disruption on banks' pricing strategies, as stated in Prediction 3. We next take it to the data and study whether Banks with local branches charge higher prices relative to banks with fewer or no local branches as 3G coverage increases. We examine both the deposit pricing and the loan pricing.

Deposit Pricing. Due to data availability, we only observe the deposit rates charged by banks in counties where they have local branches. Therefore, our test for deposit pricing focuses on the pricing strategies of banks with local branches and examines whether these banks charge higher prices after 3G coverage increases. We estimate Specification (2) with the outcome variable being $DepositSpread_{b,c,q}$, the spread between federal fund rates and deposit rates charged by bank b in county c in quarter q.

Table 4 reports the results. Comparing two counties in the same state, branching banks charge 2 bps higher spread on its 12-month CDs in counties with a 100% increase in 3G coverage than in counties without 3G coverage (columns 1 and 2).[11] The result holds in the IV setting where we instrument 3G coverage with lightning strikes frequency (column 3). The effect is relatively weaker for 36-month CDs, but the signs in columns 4-6 are consistent with our prediction.

Loan Pricing. To further examine the pricing divergence of branching banks and non-branching banks, we turn to the loan pricing data. Consumers pay upfront origination fees when getting a mortgage loan. Since 2018, HMDA has started collecting loan-level origination fees, along with other information about loan and borrower characteristics. We examine the impact of 3G expansion on loan origination fees by estimating the following

[11] We do no include the most stringent fixed effects bank×state×quarter because bank have limited rate-setting branches in one state.

specification:

$$\begin{aligned} Price_{b,j,c} =& \beta_1 \text{3G Coverage}_c + \beta_2 \text{3G Coverage}_c \times \text{Branch}_{b,c} + \beta_3 \text{Branch}_{b,c} \\ & + \gamma X_j + \mu_c + \nu_b + \epsilon_{b,j,c}, \end{aligned} \tag{5}$$

where b, j, and c index bank, borrower, county, and state, respectively. Since the test sample focuses on 2018 only, we drop the time subscript t in this loan-level specification. $Price_{b,j,c}$ is origination fee charged by bank b in county c to borrower j, $Branch_{b,c}$ is an indicator for whether bank b has a branch in county c or log of the number of bank b's branches in county c. μ_c and ζ_b are county, state, and bank fixed effects, respectively. The inclusion of county and bank fixed effects allows comparison of two banks' pricing strategies within a county after taking out their average pricing power across the US. X_j are borrower-loan controls, including loan size, loan type (i.e., conventional, FHA, VA, or RHS), loan purpose (home purchases, refinancing, or others), loan-to-value ratios, debt-to-income ratios, and borrowers' income, gender, age, and race.

3G coverage has a sizeable impact on origination fees. A full 3G penetration leads to a 0.797 percentage-point reduction in non-branching banks' origination fees (Table 5 column 1). In contrast, banks with branches charge higher prices in response to 3G penetration. Banks with 3% more branches in a local county charge 0.17 ($0.322 \times 3 - 0.797$) percentage-point higher origination fees. The intuition is in line with our prediction that banks with branches are more differentiated following 3G expansion and hence do not decrease fees as much as non-branch banks do.

Overall, we find consistent empirical evidence that 3G reduces deposit and loan pricing because it lowers the operating costs of different types of banks. However, the diverging consumer preference for branches gives branching banks an edge, which allows them to charge relatively higher deposit and loan prices than non-branching banks after the 3G expansion.

4.4 Benefits and Costs

After establishing the impact of 3G expansion on how banks change their branching and pricing strategies, we discuss the subsequent benefits and costs in this section.

4.4.1 Benefits: Increased Competition

Over the past decade, we have witnessed the expansion of the geographic scope of competition from the local market to the national market. Take the mortgage market as an example, panel (a) and (b) of Figure 7 show that lenders have become more geographically dispersed: from 2009 to 2017, the entire distribution of the number of counties covered by each lender shifts rightward, and the distribution mass of lender geographic concentration has moved closer to zero (the most geographically dispersed).[12] Meanwhile, local competition goes up: panel (c) and (d) of Figure 7 show that county-level mortgage market HHI indices decrease, especially in the largest 500 counties.[13]

We argue that the expansion of 3G networks and the resulting changes in banks' competing strategies partially contribute to these trends. We examine whether 3G penetration increases local competition as stated in model Prediction 4. We estimate Specification (1), where the outcome variables are two county-level competition measures: Herfindahl-Hirschman index (HHI) and the log number of lenders serving the region.[14]

Table 6 reports the results. In this table, we construct the measures using all types of lenders to better estimate market competition. Local competition increases as the 3G networks penetrate a region: a full 3G coverage reduces HHI by 47.3 bps (column 1). Relative to the average county HHI, the effect translates into an economically meaningful 5.2% reduction in concentration. Also, the expansion of 3G is associated with more banks serving a region. Quantitatively, as 3G coverage increases from 0 to 100%, a region is served by 3.2% more banks, amounting to 3 additional banks. IV results are reported in Columns 2 and 4, confirming the findings.

[12] Geographic concentration of a lender is calculated as the sum of squared share of mortgage origination activity in each county, i.e., $\sum_k \in \mathbb{K}_i \frac{Volume_{ik}}{\sum_k \in \mathbb{K}_i Volume_{ik}}$. The average number of counties covered by each lender increased from 24 to 40, amounting to a 67% increase relative to the 2009 average, while the bottom quartile, the median, and the top quartile have increased by 75% (from 4 to 7), 50% (from 8 to 12), and 50% (from 16 to 24), respectively. The average geographic concentration has declined by 26% since 2009. In 2017, there were 896/3128 (29%) lenders with geographic concentration below 0.2, for example, whereas there were only 592/4282 (14%) lenders with geographic concentration below 0.2 in 2009.

[13] Panel (c) of Figure 7 presents the entire distribution of county-level HHI index in 2009 and the distribution in 2017. The distribution shifts left, and the median HHI index dropped by 20% from 5% in 2009 to 4% in 2017.

[14]

$$HHI_{ct} = 10000 \times \sum_{l \in L_{(c,t)}} S_{lct}^2$$

where $S_{l,c,t}$ denotes the market share of lender l in county c and year t, and $L_{(c,t)}$ is the set of lenders that originated loans in county c and year t.

This section and section 4.3.2 collectively illustrate that 3G expansion induces incumbent banks to shut down branches and induces new banks to enter a market without (with fewer) branches. Consequently, competition is intensified.

4.4.2 Costs: Unequal Banking Access

Although the competition is enhanced after the 3G expansion, the benefits are not equally shared by all consumers. This section studies the distributional effect between digital and non-digital consumers, as stated in Predictions 5 and 6, and highlights that non-digital customers could be even worse off as a result of banks' responses to digital disruption.

Intensive Margin: Banking Service Cost We begin by examining the effect of 3G expansion on the costs paid by digital and non-digital consumers to access banking services (Prediction 5). Digital consumers (i.e., young consumers) who can freely switch away from branches can benefit from intensified competition brought by non-branching banks, whereas non-digital customers who are captivated by branches suffer from worse pricing charged by branching banks. To this end, we exploit within county variation by comparing the average loan origination fees and interest rates paid by different age groups following the expansion of the 3G networks.[15] Formally, we estimate the following specification:

$$\begin{aligned} Price_{b,j,c} = & \beta_1 \text{3G Coverage}_c + \beta_2 \text{3G Coverage}_c \times \text{Borrower Age}_j + \beta_3 \text{Borrower Age}_j \\ & + \gamma X_j + \mu_c + \zeta_b + \epsilon_{b,j,c}, \end{aligned} \tag{6}$$

where b, j, and c index bank, borrower, county, and state, respectively. $BorrowerAge_j$ are a set of indicator variables for borrowers' age range, and other variables are defined in Equation (5). The independent variable of interest is the interaction term between 3G coverage and indicator variables for borrowers' age range. The inclusion of loan characteristics and fixed effects allows us to identify the differential effect of 3G penetration on old versus young borrowers while accounting for loan, borrower, bank, and county characteristics.

Table 7 reports the results.[16] On average, borrowers below age 35 pay lower origination fees in counties with higher 3G coverage. However, the fees (columns 1 and 2) and interest

[15] We do not have data on depositor characteristics and hence focus on borrowers in this section.

[16] This table includes only bank lenders, and Table IA.9 includes loans originated by all types of lenders.

rates (columns 3 and 4) paid by borrowers above age 35 relative to the fees paid by borrowers below age 35 increase as 3G coverage increases. Comparing two counties, one with no 3G coverage and one with 100%, borrowers with age above 34 and below 55 pay 0.213 percentage-point higher origination fees, and borrowers with age above 55 pay 0.543 percentage-point higher origination fees than borrowers with age below 35. We reach similar qualitative results about interest rates.

Overall, our findings are consistent that non-digital borrowers pay more origination fees and loan interest rates amid the 3G digital disruption.

Extensive Margin: Unbanked Lastly, Prediction 6 states that the unbanked rate declines among digital consumers but rises among non-digital consumers upon digital disruption. We analyze how digital disruption influences financial inclusion.

We unmask the differential implication of 3G coverage on banking access across demographic groups using the FDIC Survey of Consumers Use of Banking and Financial Services. Panel (a) of Figure 8 shows that the unbanked rate of the population below age 55 dropped by about 4% from 2009 to 2019, while an additional 1% population above age 55 became unbanked. Panel (b) shows similar patterns: the unbanked rate of the population with phones dropped by about 2% from 2009 to 2019, while an additional 2% population without phones became unbanked.

Table 8 relates consumers' bank account ownership to 3G expansion. The outcome variable in columns 1 and 2 is an indicator variable that equals one if the respondent does *not* have any bank account.[17] 3G introduction lowers the unbanked rate by a larger degree for younger and lower-income respondents. Interestingly, the unbanked rate declines more for non-white consumers. The reason for racial minorities is possibly due to their mistrust and perceived discrimination during in-person interactions with bankers, which is mitigated by digital services. Subsample analysis in Table IA.7 suggests that 3G expansion induces older, poor, less-educated, and no-phone consumers to become more unbanked, suggesting that these consumers get strictly worse off after digital disruption.

We then examine whether digital disruption affects the likelihood of existing bank account

[17]Since the survey only records the MSA location of a respondent, we aggregate 3G coverage to the MSA level (the weighted average of the value of 3G availability weighted by the population density in each MSA's polygon). All columns control for MSA-year fixed effects. In addition, even columns also include MSA-year fixed effects to account for local economic development confounders.

holders becoming unbanked in columns 3-4 of Table 8. The outcome variable is an indicator of whether the survey participant loses banking access conditional on having banking access previously. Across various demographics, age is the key distinguishing feature driving the heterogeneity: a 100% increase in 3G coverage increases old consumers' likelihood of losing banking access by 1.9% (in unreported table). This magnitude is close to the coefficient estimate in column 1 of Table IA.7, suggesting that the 3G expansion negatively affects old consumers primarily by turning them from banked to unbanked. This 1.9% magnitude represents a 50% effect relative to the average rate of losing banking access. In short, 3G expansion induces some old consumers to opt-out of banking services.

We then dig into what drives individuals to quit banking access in Table 9. Consistent with our predictions, more individuals report high fees charged by banks as the reason for leaving banks after 3G expansion. Collectively, our findings suggest that the financial inclusion gap between digital and non-digital customers enlarges after the 3G expansion.

The above evidence highlights the essential role of banks' endogenous response to digital disruption in affecting digital inequality. Our new channel sheds light on the puzzling fact that some banked households can become unbanked amid digital disruption. That is, the digital disruption renders non-digital consumers at the risk of being excluded from banking services. This result, in particular, deserves regulators' attention.

5 Quantifying the Bank Competition Channel

The previous section identifies the overall benefits and costs of digital disruption through banks' endogenous changes in their branching, pricing, and entry decisions. In reality, these changes can occur simultaneously, and thus, the reduced-form analyses can not distinguish the impact of each decision on the costs of digital disruptions. To provide more insights along this line, we build a structural model to formalize our conceptual framework and quantitatively decompose the effects of various banking decisions.

5.1 Setup

5.1.1 Consumer Demand for Banking Services

There are two groups of consumers, young (representing digital) and old (representing non-digital), with a measure μ_y and μ_o and $\mu_y + \mu_o = 1$. Consumers, indexed by i, are looking to obtain one dollar worth of banking services, which can be seen as either one dollar of deposit or mortgage. They choose among J_T traditional banks (T-type) and J_F FinTechs (F-banks) or stay unbanked. Each option is indexed by j and is characterized by a price, the number of operating branches, and the digital banking quality bundle, $\{r_j, b_j, d_j\}$. We denote the unbank decision as choice 0.

The utility consumer type i derives from choosing bank j is

$$u_{i,j} = -\alpha_i r_j + \beta_i b_j + \gamma_i d_j + \epsilon_{i,j} \tag{7}$$

where α_i is rate sensitivity, β_i is preference for branch services, and γ_i is preference for digital banking services. $\epsilon_{i,j}$ is a mean-zero idiosyncratic utility shock, which follows the generalized extreme value distribution with correlation coefficient $\lambda_J \in \{\lambda_T, \lambda_F\}$.[18] λ_T and λ_F are the nested logit coefficients which specify the correlation among bank options within T-type and F-banks, respectively. The nested setup assumes that, from consumers' perspectives, T-type and F-banks are independent choices while banks within each type have a correlation of $1 - \lambda_T$ (or 1-λ_F).

The differences between young and old consumers are threefold. First, old consumers derive more utility from branching services (that is, $\beta_o > \beta_y \geq 0$). This is justified by the fact in Figure 2 that old people are much more likely to access banking services via a branch than young people. The same fact justifies the second difference between the young and the old in our model: young consumers value digital services, but the old do not (that is, $\gamma_y > \gamma_0 \geq 0$). Lastly, old consumers are less price-sensitive than young consumers (that is, $\alpha_y > \alpha_o \geq 0$).

Overall, the utility function features (1) consumers' preference for a lower service price,

[18]The generalized extreme value distribution has the following cumulative distribution function $F(\epsilon_{i,1}, ..., \epsilon_{i,J}) = \exp\left(-\sum_{t\in\{T,F\}} \left(\sum_{j=1}^{J_t} e^{-\epsilon_{i,j}/\lambda_t} \right)^{\lambda_t} \right)$.

a higher number of bank branches, and better digital banking quality, (2) heterogeneous preference where young people value digital banking services, whereas old people value more branch services, and (3) banks compete more aggressively within each group than across groups due to the nested structure.

Consumer i chooses bank j if it delivers the highest utility and its utility is also higher than the utility of being unbanked, which is normalized to be 0:

$$u(j;\alpha_i,\beta_i,\gamma_i,\epsilon_{i,j}) \geq u(k;\alpha_i,\beta_i,\gamma_i,\epsilon_{i,j}), \quad \forall k \in 0,1,...,J. \tag{8}$$

Given the assumed distribution of $\epsilon_{i,j}$, we can derive a probability that consumer i chooses bank j, which we denote as $s_{i,j}$. Then, the overall demand for bank j's service is characterized as follows

$$D_j = \sum_i \mu_i s_{i,j}, \quad i \in \{y,o\}. \tag{9}$$

5.1.2 Banks

Banks, indexed by j, provide differentiated banking services (i.e., lending or deposit taking) to consumers. They earn revenue from offering banking services and pay to run branches which are valued by consumers.

There are J_T traditional banks (T-type) and J_F FinTech banks (F-banks). Banks within each type are symmetric. T-banks and F-banks mainly differ in 1) their marginal cost of operating branches and 2) their digital service quality. Traditional banks have a lower cost to operate branches than FinTechs (that is, $\kappa_T < \kappa_F$), while FinTech banks have better digital service quality than traditional banks (that is, $d_T < d_F$). These assumptions are motivated by the empirical fact that traditional banks have sophisticated branch networks, whereas FinTech banks tend to provide services remotely.[19] Apart from these two dimensions, we do not impose restrictions on their marginal cost c_j and entry cost FC_j (to be defined). These parameters are not essential to our model predictions.

Conditional on serving a region, bank j sets the price of its banking services, r_j, and

[19]FinTech banks are broadly defined as lenders with fewer or no local branches.

decides the number of branches, b_j, to maximize their profits:

$$\max_{r_j, b_j} (r_j - c_j)D_j - \frac{1}{2}\kappa_j b_j^2, \tag{10}$$

where D_j is the demand for bank j's banking service, and the second term is the total cost of operating branches.[20] The total bank profit, with the optimal decisions $\{r_j^*, b_j^*\}$, net of entry cost FC_j is

$$\pi_j = (r_j^* - c_j)D_j - \frac{1}{2}\kappa_j (b_j^*)^2 - FC_j. \tag{11}$$

A bank serves a region as long as $\pi_j \geq 0$.

5.1.3 Equilibrium

An equilibrium is a market structure comprising the number of banks of each type, $\{J_T, J_F\}$; the pricing decisions , $\{r_T, r_F\}$; the branching decisions, $\{b_T, b_F\}$; and the market shares, $\{D_T, D_F\}$, such that

1. Consumers maximize utility, taking market structure, branching, and pricing as given (Equation (8) holds for all consumers);
2. Banks set prices and choose the number of branches to maximize profits, taking market structure and the pricing decisions of other lenders as given (Equation (10) holds for all banks);
3. The number of banks of each type $\{J_T, J_F\}$ is set such that the least profitable bank has a positive π_j and no new bank wants to enter the market (Equation (11) holds true for the marginal bank).

In equilibrium, the likelihood that consumer i chooses bank j with the following probability:[21]

$$s_{i,j} = \frac{A_{i,j}}{Z_{i,t}} \frac{Z_{i,t}^{\lambda_t}}{1 + \sum_{t \in \{T,F\}} Z_{i,t}^{\lambda_t}}, \quad t \in \{T, F\}, \tag{12}$$

[20] We choose the quadratic cost function such that we can derive the interior solution for the number of operating branches.

[21] We derive the equilibrium in the Appendix D.

where $Z_{i,t} = \sum_{j=1}^{J_t} \exp(\frac{1}{\lambda_t}(-\alpha_i r_j + \beta_i b_j + \gamma_i d_j))$. The proportion of depositor i stays unbnaked is

$$s_{i,0} = \frac{1}{1 + \sum_{t \in \{T,F\}} Z_{i,t}^{\lambda_t}}. \tag{13}$$

Given Equation (12), banks' optimal pricing and branching decisions are

$$r_j = c_j + \frac{\sum_{i \in y,o} \mu_i s_{i,j}}{\sum_{i \in y,o} \mu_i \frac{\alpha_i}{\lambda_t} s_{i,j} \Big(1 + (\lambda_t - 1)\frac{A_{i,j}}{Z_{i,t}} - \lambda_t s_{i,j}\Big)}, \tag{14}$$

$$b_j = \frac{1}{\kappa_j}(r_j - c_j) \sum_{i \in y,o} \mu_i \frac{\beta_i}{\lambda_t} s_{i,j} \Big(1 + (\lambda_t - 1)\frac{A_{i,j}}{Z_{i,t}} - \lambda_t s_{i,j}\Big), \tag{15}$$

With these expressions, banks' market shares and profits can be derived. Lastly, the entry condition along with the equilibrium profit function yields the number of banks.

5.2 Estimation

We estimate the model using bank deposit data and the FDIC survey data. The estimation sets the stage for the ensuing counterfactual analysis.

We define markets at the MSA level to match the most granular location in the FDIC surveys. As we do not have individual depositors' characteristics, we define young and old depositors at the zip-code level: a young zip-code has a median age below the MSA median.

We classify banks into T-banks and F-banks based on the number of branches per dollar of domestic deposits. When this ratio is small, it suggests that the bank can serve a large amount of deposits with fewer branches and hence is more likely to be a F-bank, and vice versa. We calculate the average number of branches per dollar of domestic deposits from 2007 to 2018. Banks below the median are classified as F-banks and T-banks otherwise.

We calculate the total deposits, $D_{y,j}$ ($D_{o,j}$), that bank j takes from young (old) consumers in a given MSA by summing up all of its deposits in young (old) zip-codes within the MSA. To calculate market shares, we first find the share of the unbanked population in an MSA ($s_{i,0}, i \in \{y, o\}$), from the FDIC survey data as in our reduced form analysis. We then calculate each lender's market share as

$$s_{i,j} = \frac{D_{i,j}}{\sum_j D_{i,j}} \times \frac{1}{s_{i,0}}, \quad i \in \{y, o\}. \tag{16}$$

We define that a region has experienced digital disruption when it is covered by more than 50% 3G networks. For simplicity, we assume that only young consumers change preferences after digital disruption: their preferences shift from branches to digital services. In other words, young consumers have β_y^{pre} and γ_y^{pre} as the preferences for b_j and d_j before disruption, and β_y^{post} and γ_y^{post} afterwards. Apart from it, other parameters stay the same across the pre- and post- disruption periods.

We begin by estimating the demand:

$$ln(s_{ij}) - ln(s_{i0}) = -\frac{\alpha_i}{\lambda_t} r_j + \frac{\beta_i}{\lambda_t} b_j + \frac{\gamma_i}{\lambda_t} d_j - (\lambda_t - 1) log(Z_{i,t}). \tag{17}$$

We find r_j as the average deposit rate spread charged by each bank and b_j as its total number of branches in any given MSA. We do not have measures on digital services offered by banks, and for simplicity, we set $d_F = 1$ and $d_T = 0$. $(\lambda_t - 1)log(Z_{i,t})$ can be treated as fixed effects at the consumer group and bank-type levels. Because of fixed effects, we can only estimate the difference $\Delta\gamma_y = \gamma_y^{post} - \gamma_y^{pre}$. With the estimated demand-side parameters, we can derive the marginal cost parameters c_j and k_j using Equations (14) and (15), and derive FC_j from setting $\pi_j = 0$ in Equation (11).

Table 10 presents estimated parameters. Young consumers are more price-elastic than old consumers. λ_T and λ_F are strictly smaller than 1, suggesting that banks are more substitutable within types. Lastly, young people's preference for branches drops significantly after the digital disruption. Regarding supply side parameters, we find the marginal cost of operating branches is higher for F-banks than T-banks, consistent with our assumption.[22]

[22]The marginal costs c_j are negative because deposit spreads, which are differences between federal fund rates and deposit rates, are negative for more than half of the banks in the sample. Since we are interested in the relative change before and after the digital disruption, we argue the negative level of deposit spreads is not an issue.

5.3 Counterfactual Analysis

We use the estimated model to decompose the impacts of banks' responses on heterogeneous consumers and shed light on their benefits and costs when banks 1) only change prices (we call short-run), or 2) change branching and pricing decisions (we call medium-run), or 3) decide to enter or exit a market with endogenous decisions on branching and pricing decisions (we call long-run).

We first solve our full model with $\gamma_i = 0$ and $\beta_y = \beta_y^{pre}$ together with other parameters. This yields the equilibrium benchmark before digital disruption. After digital disruption, we set $\gamma_y = \Delta\gamma_y$ and $\beta_y = \beta_y^{post}$, and re-solve the model for three different scenarios. In case 1), we fix b_t and J_t as in the benchmark case, and only allow banks to decide r_j in Equation (10). In case 2), we fix J_t as in the benchmark case, and allow banks to optimally choose r_j and b_j. In case 3), we re-solve the full model with the new set of parameters. Then we compare three new equilibrium outcomes to the benchmark case, respectively.

Figure 9 plots percentage changes in the average rate paid by young and old groups, as well as the respect banked proportion for three scenarios. The digital disruption allows more digital customers to shift to cheap services provided by F-banks, and hence, intuitively, digital customers, on average, pay lower service fees and get more financially included. What is intriguing is the impact on non-digital customers. In the short run, when banks only adjust rates, we see old customers pay 0.8% higher service fees. This effect arises from a 7% increase in services fees of T-banks as they understand their customer pools become more captive after the digital disruption. Banks shut down branches in the medium run, limiting branch access for non-digital customers. This, together with higher service fees charged by T-banks, results in 11% of non-digital customers leaving banks. In the long run, more F-banks enter the market, intensifying the competition, especially among F-banks. Digital customer benefits dramatically from the fiercer competition: an average customer pays 26.1% lower service fees, and the banked population increases by 15%. When F-banks' price becomes very appealing, some non-digital customers also shift from T-banks to F-banks. This effect lowers the average price paid by non-digital customers and incentivize more of them to use bank services. However, in this scenario, we see the largest difference in financial inclusion between the two groups of consumers. In other words, banks' endogenous response to digital disruption could turn the digital divide into digital inequality in the long run.

6 Conclusion and Discussion

In this paper, we study banks' branching, pricing, and entry decisions in response to digital disruption and the resulting the benefits and costs to consumers. While non-digital consumers still rely on branches, digital disruption has shifted digital consumers' preference from branch services to digital services. As the average preference for branches declines, banks close costly branches and serve consumers remotely, and more digital banks enter the market. The intensified competition from digital banks forces incumbents with branches to specialize in the market segment in which they have a comparative advantage: branching banks target non-digital customers and exploit market power on them by charging higher prices. Consequently, digital customers benefit from the intensified bank competition at the cost of non-digital customers who pay higher prices for branch services and face the risk of financial exclusion.

This paper speaks to several prevailing policy discussions. Our findings highlight that the benefit of digital disruption may come at the cost of non-digital consumers, which receives less attention in the current discussion of how technology affects the economy. We also bring in a new perspective of diverging customer preference and product differentiation in analyzing how technology affects bank competition, which is missed in the current discussion. Importantly, we also calibrate the consequences of digital disruption in terms of financial inclusion and potential price discrimination. By unraveling the heterogeneous consumers and banks, we hope this paper can provoke new insights into the interaction between technology and financial intermediaries.

References

Andersen, Thomas Barnebeck, Jeanet Bentzen, Carl-Johan Dalgaard, and Pablo Selaya, 2012, Lightning, it diffusion, and economic growth across us states, *Review of Economics and Statistics* 94, 903–924.

Bartlett, Robert, Adair Morse, Richard Stanton, and Nancy Wallace, 2019, Consumer-lending discrimination in the fintech era, *National Bureau of Economic Research* .

Beck, Thorsten, Ross Levine, and Alexey Levkov, 2010, Big bad banks? the winners and losers from bank deregulation in the united states, *The Journal of Finance* 65, 1637–1667.

Benetton, Matteo, 2021, Leverage regulation and market structure: A structural model of the uk mortgage market, *The Journal of Finance* 76, 2997–3053.

Berry, Steven, James Levinsohn, and Ariel Pakes, 1995, Automobile prices in market equilibrium, *Econometrica* 841–890.

Blattner, Laura, and Scott Nelson, 2021, How costly is noise? data and disparities in consumer credit, *arXiv preprint arXiv:2105.07554* .

Brown, James R, J Anthony Cookson, and Rawley Z Heimer, 2019, Growing up without finance, *Journal of Financial Economics* 134, 591–616.

Buchak, Greg, and Adam Jørring, 2021, Do mortgage lenders compete locally? implications for credit access .

Buchak, Greg, Gregor Matvos, Tomasz Piskorski, and Amit Seru, 2018a, Beyond the balance sheet model of banking: Implications for bank regulation and monetary policy .

Buchak, Greg, Gregor Matvos, Tomasz Piskorski, and Amit Seru, 2018b, Fintech, regulatory arbitrage, and the rise of shadow banks, *Journal of Financial Economics* 130, 453–483.

Carlson, Mark, and Kris James Mitchener, 2006, Branch banking, bank competition, and financial stability, *Journal of Money, Credit and Banking* 38, 1293–1328.

Cetorelli, Nicola, and Philip E Strahan, 2006, Finance as a barrier to entry: Bank competition and industry structure in local us markets, *The Journal of Finance* 61, 437–461.

Di Maggio, Marco, Dimuthu Ratnadiwakara, and Don Carmichael, 2021, Invisible primes: Fintech lending with alternative data, *Available at SSRN 3937438* .

Drechsler, Itamar, Alexi Savov, and Philipp Schnabl, 2017, The deposits channel of monetary policy, *The Quarterly Journal of Economics* 132, 1819–1876.

Egan, Mark, Ali Hortaçsu, and Gregor Matvos, 2017, Deposit competition and financial fragility: Evidence from the us banking sector, *American Economic Review* 107, 169–216.

Fonseca, Julia, and Adrien Matray, 2022, The real effects of banking the poor: Evidence from brazil, Technical report, National Bureau of Economic Research.

Fuster, Andreas, Paul Goldsmith-Pinkham, Tarun Ramadorai, and Ansgar Walther, 2020, Predictably unequal? the effects of machine learning on credit markets, *The Effects of Machine Learning on Credit Markets (October 1, 2020)* .

Fuster, Andreas, Matthew Plosser, Philipp Schnabl, and James Vickery, 2019, The role of technology in mortgage lending, *The Review of Financial Studies* 32, 1854–1899.

Garmaise, Mark J, and Tobias J Moskowitz, 2006, Bank mergers and crime: The real and social effects of credit market competition, *the Journal of Finance* 61, 495–538.

Guriev, Sergei, Nikita Melnikov, and Ekaterina Zhuravskaya, 2021, 3g internet and confidence in government, *The Quarterly Journal of Economics* 136, 2533–2613.

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

Hubbard, R Glenn, and R Glenn Hubbard, 1994, *Money, the financial system, and the economy* (Addison-Wesley Reading, MA).

Jack, William, and Tavneet Suri, 2014, Risk sharing and transactions costs: Evidence from kenya's mobile money revolution, *American Economic Review* 104, 183–223.

Jayaratne, Jith, and Philip E Strahan, 1996, The finance-growth nexus: Evidence from bank branch deregulation, *The Quarterly Journal of Economics* 111, 639–670.

Jayaratne, Jith, and Philip E Strahan, 1997, The benefits of branching deregulation, *Economic Policy Review* 3.

Jiang, Erica Xuewei, 2019, Financing competitors: Shadow banks' funding and mortgage market competition, *Available at SSRN 3556917* .

Jiang, Erica Xuewei, Yeonjoon Lee, and Will Shuo Liu, 2022, Disparities in consumer credit: The role of loan officers in the fintech era, *Available at SSRN 4035764* .

Kuehn, Joseph, 2018, Spillovers from entry: the impact of bank branch network expansion, *The RAND Journal of Economics* 49, 964–994.

Lee, Jean N, Jonathan Morduch, Saravana Ravindran, Abu Shonchoy, and Hassan Zaman, 2021, Poverty and migration in the digital age: Experimental evidence on mobile banking in bangladesh, *American Economic Journal: Applied Economics* 13, 38–71.

Manacorda, Marco, and Andrea Tesei, 2020, Liberation technology: Mobile phones and political mobilization in africa, *Econometrica* 88, 533–567.

Ménard, Claude, and Michel Ghertman, 2009, *Regulation, deregulation, reregulation: institutional perspectives* (Edward Elgar Publishing).

Philippon, Thomas, 2016, The fintech opportunity .

Philippon, Thomas, 2019, On fintech and financial inclusion .

Pierre, Bachas, Paul Gertler, Sean Higgins, and Enrique Seira, 2018, Digital financial services go a long way: transaction costs and financial inclusion, *AEA Papers and Proceedings* 108, 444–448.

Rainie, Harrison, and Barry Wellman, 2012, *Networked: The new social operating system*, volume 10 (Mit Press Cambridge, MA).

Robles-Garcia, Claudia, 2019, Competition and incentives in mortgage markets: The role of brokers, *Unpublished working paper* .

Saka, Orkun, Barry Eichengreen, and Cevat Giray Aksoy, 2021, Epidemic exposure, fintech adoption, and the digital divide .

Vives, Xavier, 2019, Digital disruption in banking, *Annual Review of Financial Economics* 11, 243–272.

Wang, Olivier, 2020, Banks, low interest rates, and monetary policy transmission, *NYU Stern School of Business* .

WorldBank, Group, 2016, *World development report 2016: Digital dividends* (World Bank Publications).

Xiao, Kairong, 2020, Monetary transmission through shadow banks, *The Review of Financial Studies* 33, 2379–2420.

Yogo, Motohiro, Andrew Whitten, and Natalie Cox, 2021, Financial inclusion across the united states, *Available at SSRN 3934498* .

Figures

Figure 1. Change of Ways to Access Banking Services

The bar chart shows time series of the primary ways consumers access banking services from 2013 to 2019. Source: FDIC Survey of Household Use of Banking and Financial Services.

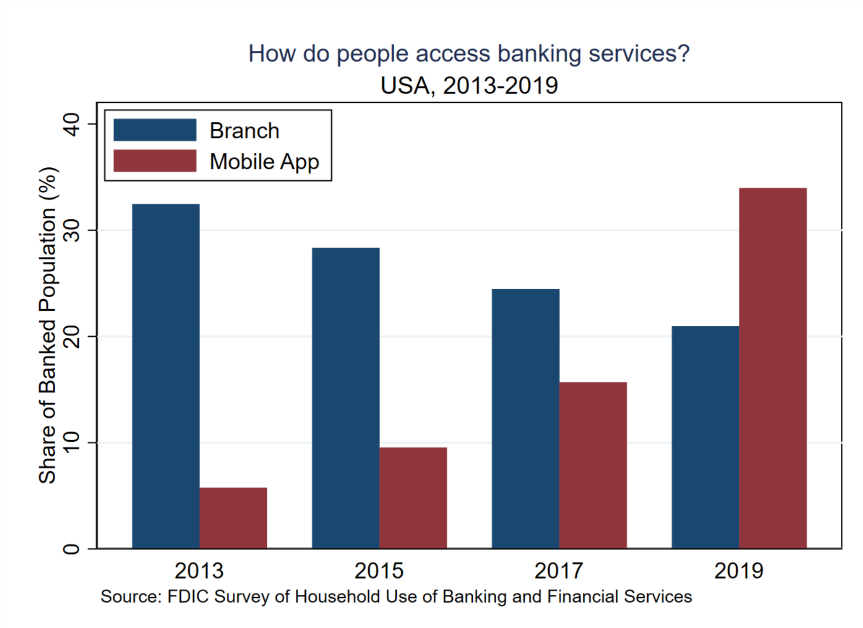
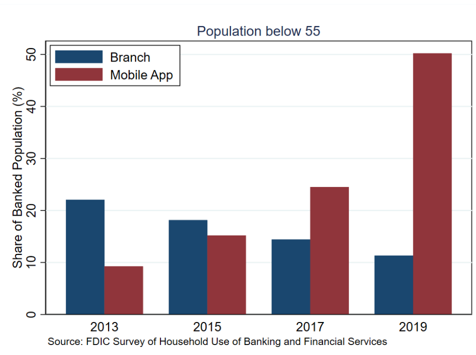


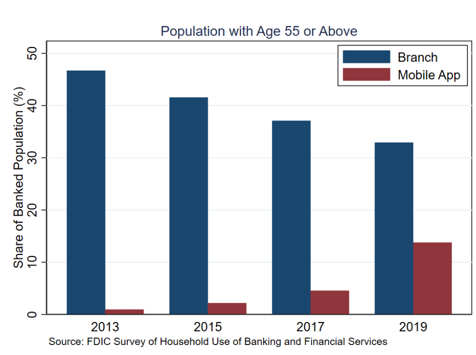
Figure 2. Change of Ways to Access Banking Services

The bar charts show the ways consumers in different age buckets to access banking services. Panel (a) and (b) plot the share of survey participants that access banking services via branch and via mobile app for young and old consumers, defined as below or above 55-year old, respectively. Panel (c) and (d) plot the same time series for consumers with college or higher degrees and those without. Panel (e) and (f) compares consumers with more than $50,000 income to those without. Source: FDIC Survey of Household Use of Banking and Financial Services.

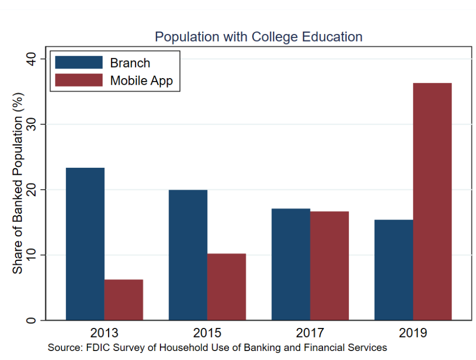
(a) Young



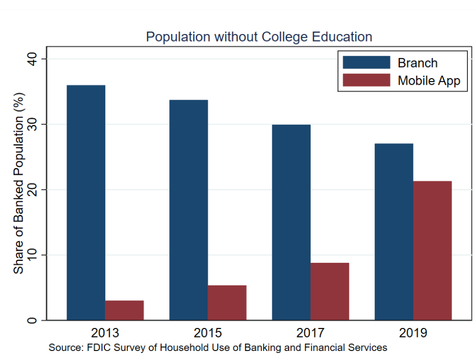
(b) Old



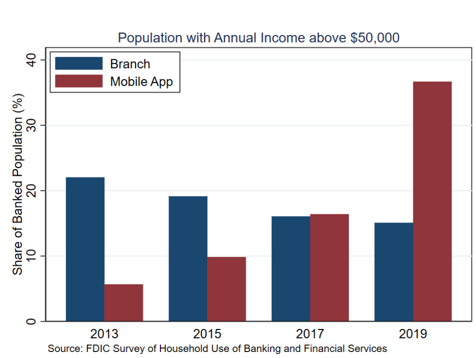
(c) More-educated



(d) Less-educated



(e) High-income



(f) Low-income

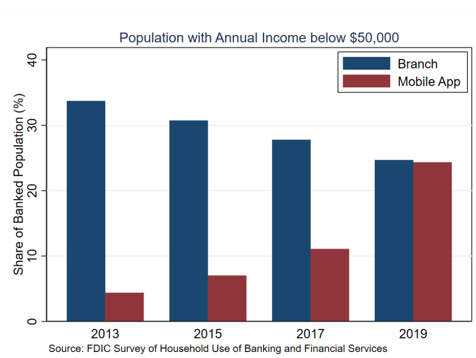
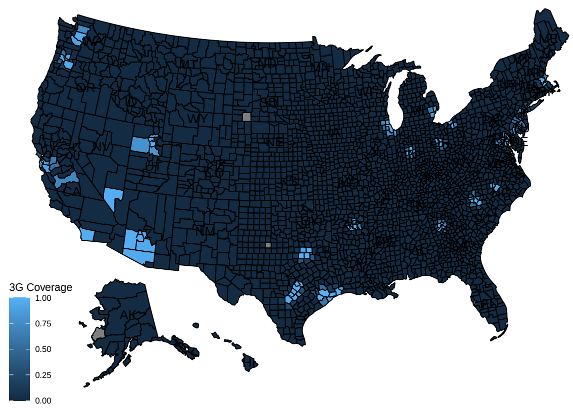


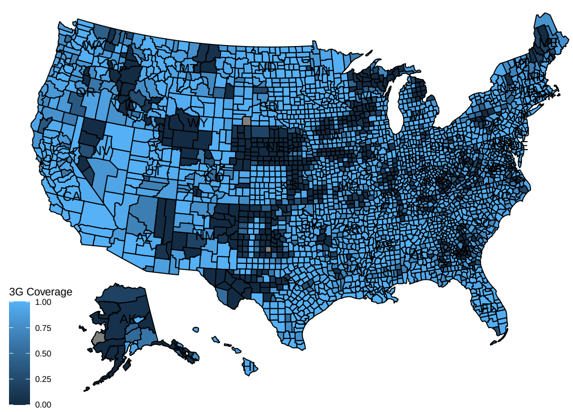
Figure 3. Maps of 3G Coverage

This figure plots 3G coverage at the county level in 2007, 2012, and 2018. 3G coverage is calculated as the average of the value of 3G availability weighted by the population density in each grid cell across all grid-cells in each county's polygon. Source: Collins Bartholomew's Mobile Coverage Explorer

(a) 2007



(b) 2012



(c) 2018

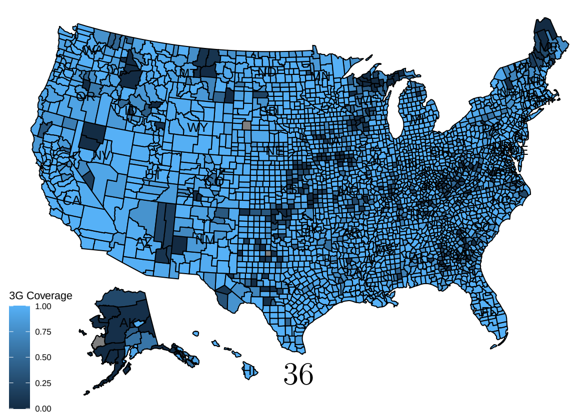


Figure 4. Maps of Counties with High Lightning Strikes Within Each State

This figure plots counties with higher-than-median lightening strike frequency within each state. Lightening strike frequency is calculated as the sum of strikes in each country from 2007 to 2018, weighting each by the population density in the 0.1×0.1 decimal degree grid cell of the lightning strike location. Source: World Wide Lightning Location Network

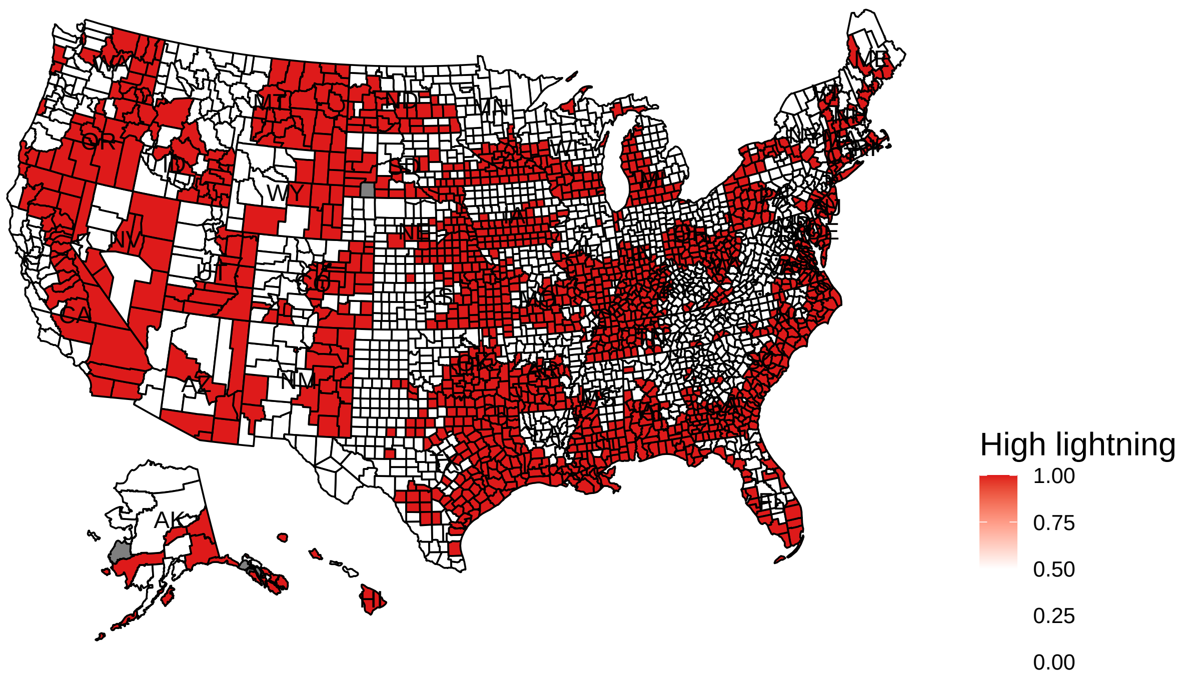
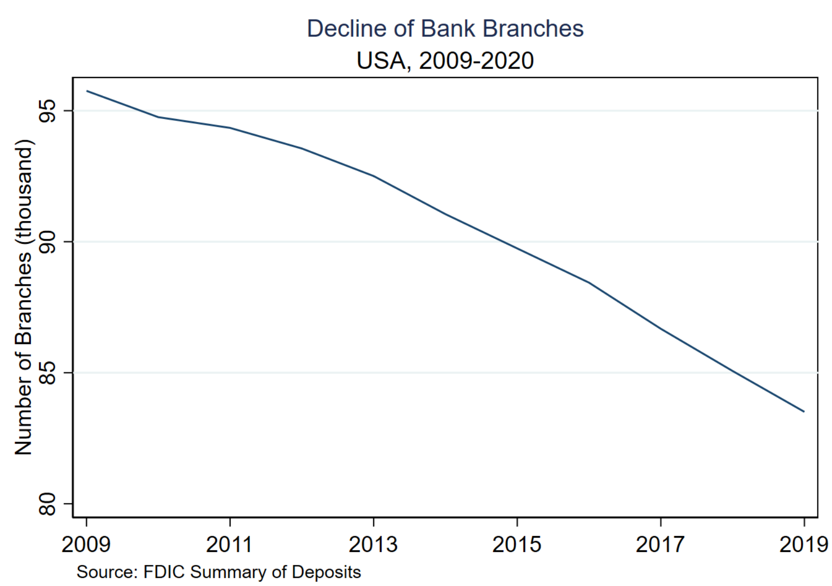


Figure 5. Bank Branch Closure

Panel (a) of this figure plots the time series of total number of branches from 2009 to 2020. Panel (b) plots county-level per-bank branch closure rate from 2009 to 2018 against the county share of population below 55. Source: FDIC Summary of Deposits and the American Community Survey.

(a) Total



(b) Cross-Section

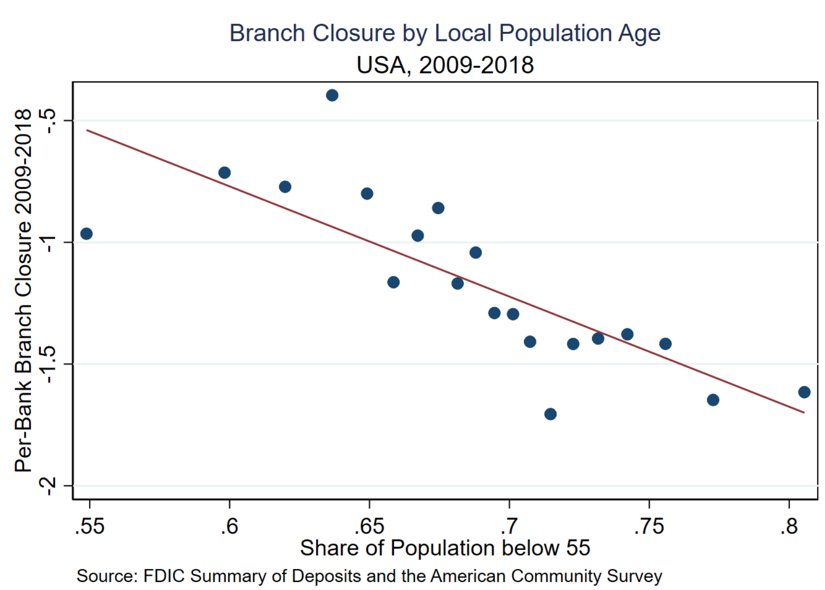


Figure 6. Event Study for Bank Closure

This figure plots dynamic DiD results for branch closure at the county-year level. The treatment group includes counties whose 3G coverage increased more than 50% in one year. The control group is constructed using matching methodology, as described by Table IA.4.

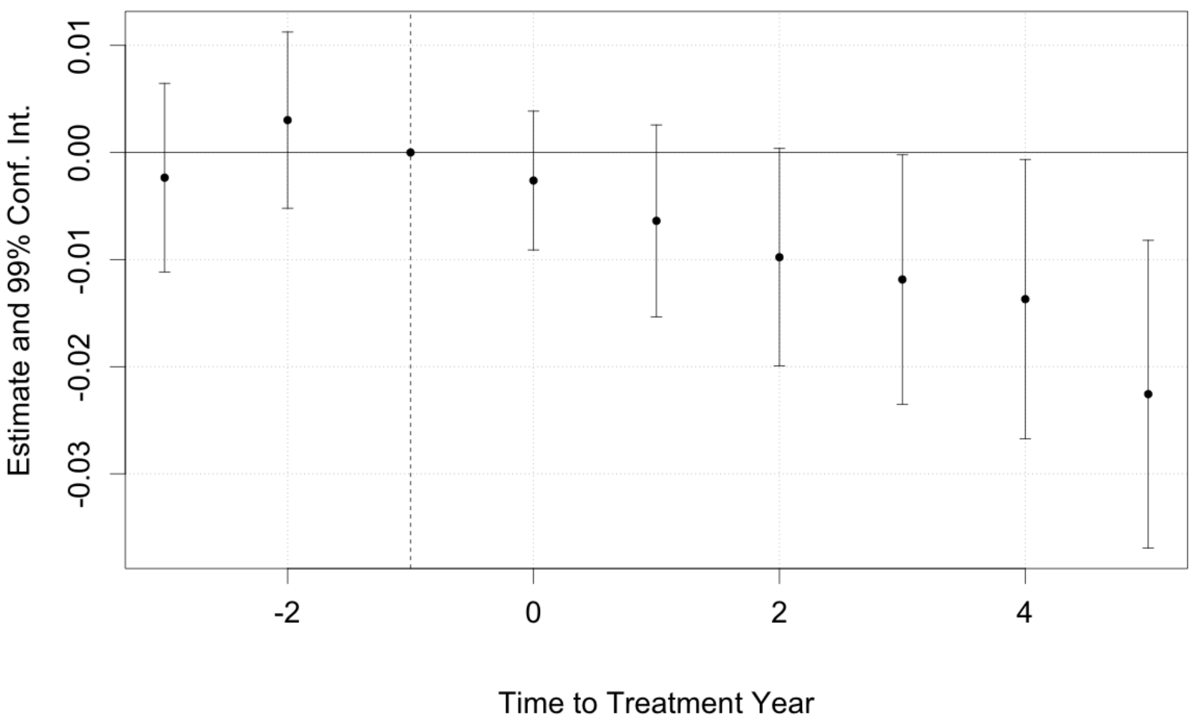
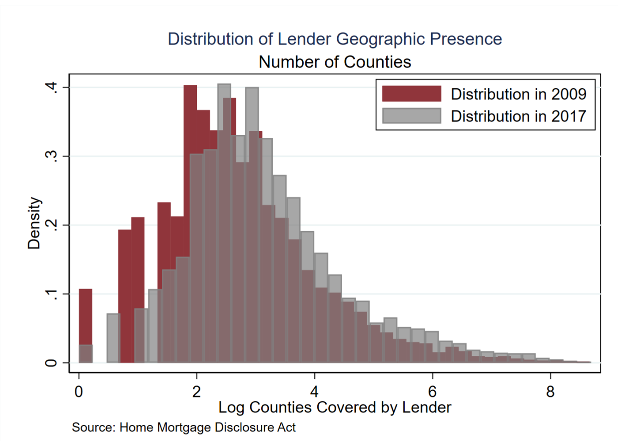


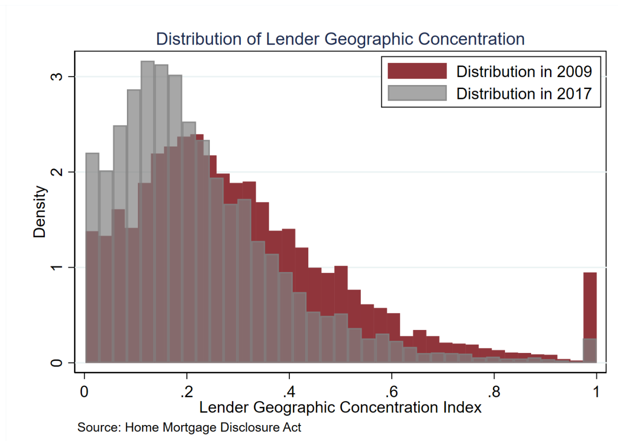
Figure 7. Geographic Expansion and Increased Bank Competition

This figure plots the distributions of geographic expansion of lenders in 2009 versus 2017. Panel (a) plots the histogram of log number of counties covered by each mortgage originator in 2009 and in 2017. Panel (b) plots the histogram of the geographic concentration. Geographic concentration of a lender is calculated as the sum of squared share of mortgage origination activity in each county, i.e., $\sum_k \in \mathbb{K}_i \frac{Volume_{ik}}{\sum_k \in \mathbb{K}_i Volume_{ik}}$. Panel (c) and (d) plot the histograms of county HHI index, where (c) is full sample, and (d) focuses on the largest 500 counties in the US. Source: HMDA.

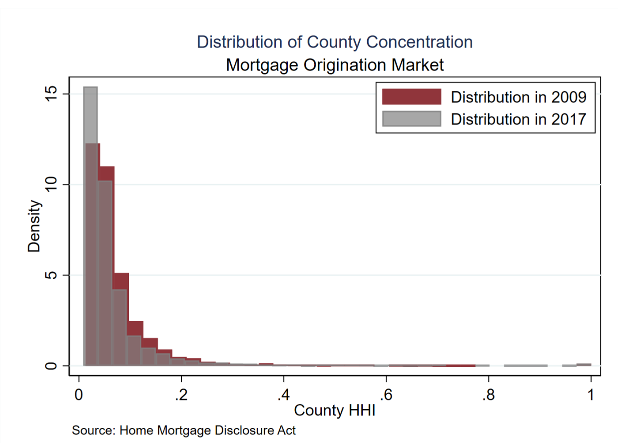
(a) Number of Counties



(b) Lender Geographic Concentration



(c) County HHI - Full



(d) County HHI - Big

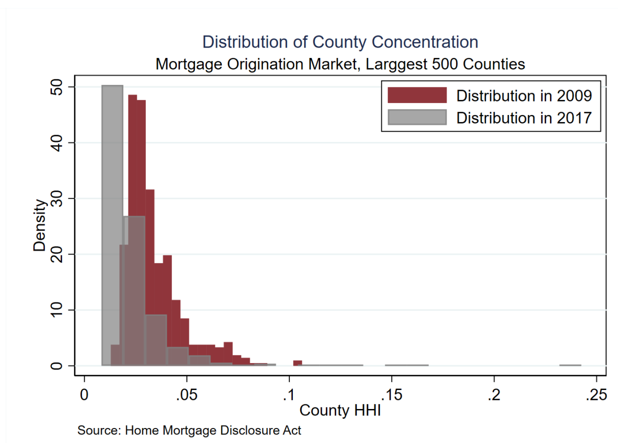
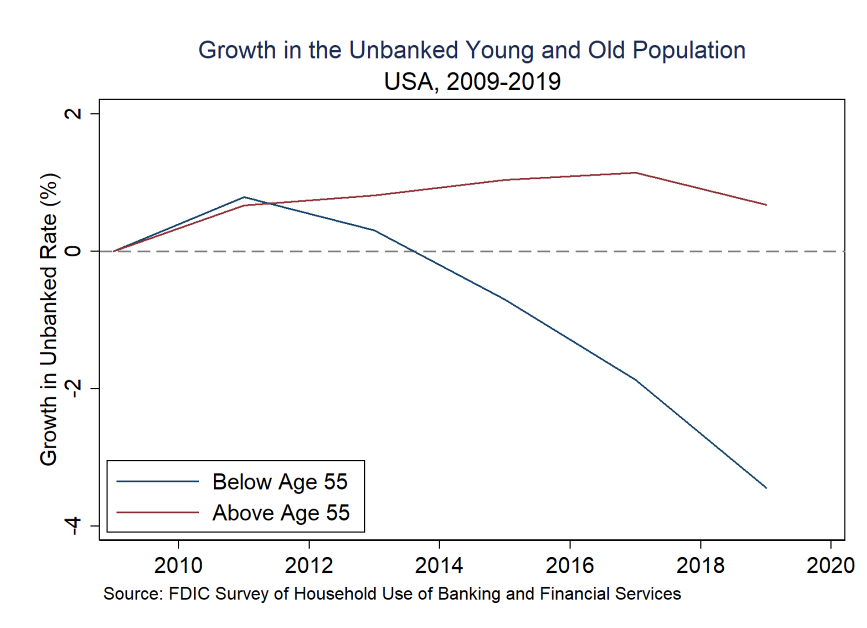


Figure 8. Growth in the Unbanked Young and Old Population

This figure plots the growth rate of unbanked consumers under 55 versus above 55 over years (Panel (a)) and with versus without phones (Panel (b)). Source: FDIC Survey of Household Use of Banking and Financial Services.

(a) Age



(b) Phone

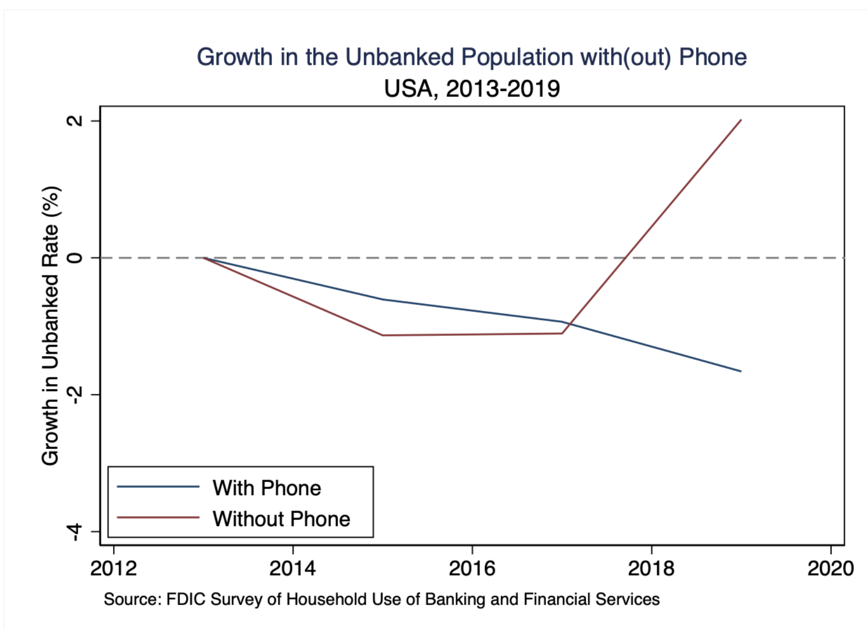
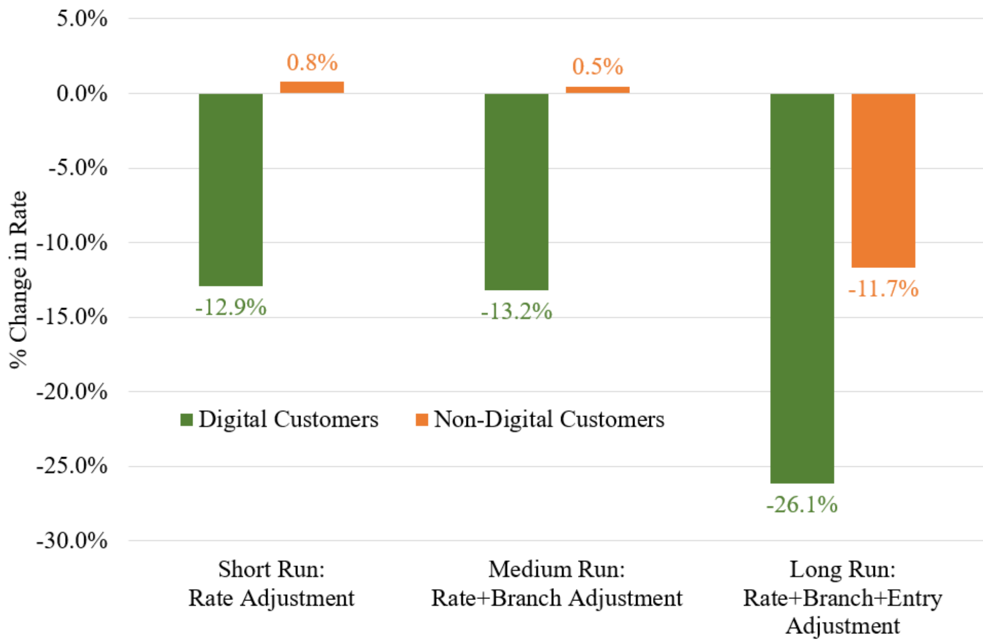
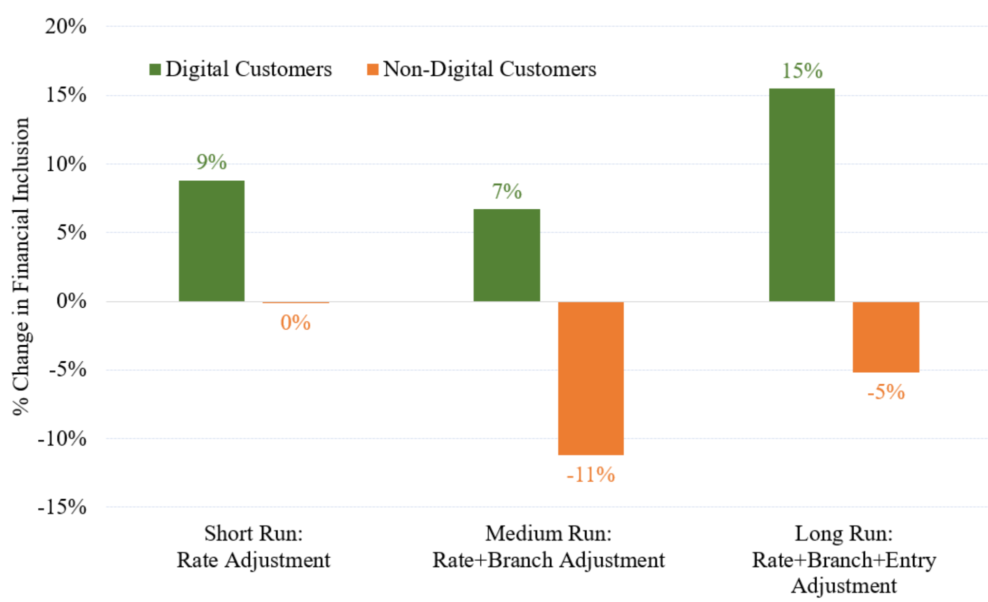


Figure 9. Decomposition of Impacts from Structural Estimation

The figures plot impacts on bank service fee and financial inclusion when banks are only allowed to 1) change pricing decision, 2) change pricing and branching decisions, and 3) all previous decisions and entry decisions.





7 Tables

Table 1 Summary Statistics

Panel A presents characteristics of banked and unbanked households and the sub-categories using FDIC surveys. "Bank Teller", "Online Banking", and "Mobile Banking" refer to the primary way that banked individuals use to access the bank services. Panel B presents reasons of being unbanked given by unbanked consumers using FDIC surveys. Panels C and D report the summary statistics of bank and county data used. Variables are defined in Section 3.1.

Panel A: FDIC survey (characteristics)						
	Banked	Bank Teller	Online Banking	Mobile Banking	Unbanked	Losing bank access
	2009 → 2019		2013 → 2019		2009 → 2019	
Proportion	92.4% → 94.6%	32.5% → 21.0%	33.1% → 22.8%	5.8% →34.0%	7.6% →5.4%	50.2% → 49.2%
Age	50.4	57.6	49.4	39.8	43.9	44.6
Income	57,915	47,092	71,623	65,249	21,277	20,875
College education	65.1%	51.2%	81.2%	77.1%	26.4%	33.1%
White	70.6%	69.8%	76.8%	63.2%	31.2%	40.9%
Phone	92.1%	83.5%	97.1%	99.0%	78.5%	83.9%

Panel B: FDIC survey (reasons of being unbanked)							
	NoMoney	NoTrust	HighFee	NoPrivacy	Inconvenience	AccountProblem	NoProducts
Unbanked	47.2%	16.8%	13.5%	4.4%	5.5%	10.8%	1.7%
Losing bank access	45.7%	15.6%	16.4%	4.1%	6.0%	10.7%	1.5%

Panel C: County characteristics								
	Count	Mean	St. Dev.	$q_{5\%}$	$q_{25\%}$	Median	$q_{75\%}$	$q_{95\%}$
3G Coverage	34,081	0.593	0.440	0.000	0.000	0.865	0.990	1.000
Per capita income, in $K	33,586	38.143	11.376	25.346	30.987	36.067	42.645	57.162
GDP, in $B	33,586	5.419	23.537	0.121	0.356	0.909	2.634	21.598
Population, in K	34,070	101.471	323.015	3.088	11.445	26.265	68.296	433.952
$\#Branch_{c,t}$	34,081	29.056	73.955	2.000	5.000	10.000	22.000	119.000
#Lenders	31,226	106.179	101.995	11.000	38.000	74.000	136.000	321.000
HHI (in bps)	31,226	910.684	893.637	269.791	436.381	648.843	1045.187	2395.384
Share^{wBranch}, in %	31,226	47.859	24.590	0.000	31.710	51.876	66.806	82.010
Lightning	3,220	1.701	1.307	0.041	0.616	1.406	2.671	4.222

Panel D: Bank-county characteristics								
	Count	Mean	St. Dev.	$q_{5\%}$	$q_{25\%}$	Median	$q_{75\%}$	$q_{95\%}$
$\#Branch_{b,c,t}$	668,019	2.007	5.522	0.000	0.000	1.000	2.000	7.000
Spread$_{b,t,qt}^{12MCD10K}$, in%	445,830	0.009	0.909	-1.475	-0.475	-0.100	0.535	1.665
Spread$_{b,t,qt}^{36MCD10K}$, in %	407,209	-0.447	1.015	-2.125	-1.065	-0.520	0.220	1.310

Table 2 Effect of 3G Coverage on Branch Closure

This table reports how 3G coverage affect branch closures at the bank-county-year and county-year level. $\#Branches_{b,c,t}$ is the number of branches of bank b in county c of year t whereas $\#Branches_{c,t}$ is the aggregated number of branches in a county. 3G coverage$_{c,t}$ is the proportion of population with access to 3G networks in county c at year t. Log(PerCapitaIncome), log(CountyGDP), log(TotalPop) are per ca capital income, GDP and population at the county level. 'MidAge" is the median age in a county. High lightening strikes represent counties whose average population-weighted frequency of lightning strikes across 2007 to 2018 is higher than the state median. Column (1) presents the results of the first stage of IV regression, and columns (4)(7)(8) are results for the second stage. Standard errors are clustered at the county level. Numbers in parentheses are t-statistics. *, **, *** represent statistical significance at 10%, 5% and 1% level, respectively.

	First stage 3G Coverage	Bank-county Level $\log(1+\#Branches_{b,c,t})$			County Level $\log(\#Branches_{c,t})$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\mathbb{1}$(High Lightning) $\times$ $Year$	-0.002**							
	(-2.437)							
3G Coverage		-0.069***	-0.066***		-0.008**	-0.013***		
		(-9.823)	(-10.233)		(-2.556)	(-3.748)		
3G $\widehat{\text{Coverage}}$				-0.722***			-0.212***	-0.016
				(-5.476)			(-2.906)	(-0.520)
3G $\widehat{\text{Coverage}}$ $\times$ $\mathbb{1}(MidAge <= 40)$								-0.012**
								(-2.093)
log(PerCapitaIncome)	-0.246***	-0.144***	-0.076***	-0.158***	0.057***	0.056***	0.009	0.053***
	(-6.106)	(-4.584)	(-2.720)	(-7.228)	(3.973)	(3.495)	(0.450)	(4.801)
log(CountyGDP)	0.067***	-0.008	0.004	0.013*	-0.008	-0.009	0.003	-0.009**
	(3.796)	(-0.623)	(0.336)	(1.808)	(-1.052)	(-1.273)	(0.535)	(-2.168)
log(TotalPop)	0.118*	0.775***	0.656***	0.357***	0.081**	0.144***	0.138***	0.132***
	(1.842)	(12.297)	(11.766)	(9.815)	(2.516)	(4.131)	(7.345)	(7.906)
log(# Bank)	0.071***	0.213***	0.256***	0.328***	0.460***	0.443***	0.459***	0.446***
	(4.077)	(10.986)	(16.600)	(26.387)	(28.453)	(27.641)	(62.261)	(87.765)
3G Coverage$_{2007}$$\times$ Year	-0.081***			-0.035***			-0.011*	0.004*
	(-49.884)			(-3.350)			(-1.827)	(1.815)
$\mathbb{1}(MidAge <= 40)$								0.009*
								(1.936)
County FE	✓				✓	✓	✓	✓
Bank$\times$County FE		✓	✓	✓				
Year FE					✓			
State$\times$Year FE	✓					✓	✓	✓
Bank $\times$ Year Fe		✓						
Bank$\times$State$\times$Year FE			✓	✓				
Adjusted R^2	0.825	0.856	0.900		0.996	0.996		
Observations	33,575	344,211	333,420	333,420	33,575	33,575	33,575	33,575

Table 3 Effect of 3G Coverage on Banks' Branching Decisions in the Lending Market

This table reports how 3G coverage affects the branching decisions of banks in the lending market. The analysis unit is bank-county-year level. Columns (1)(2) include both incumbent and new entry banks (which did not originate any loans in previous year) for a given county. Columns (3)(4) include new entry banks only. $\#Branches_{b,c,t}$ is the number of branches of bank b in county c of year t. Columns (2)(4) report results with lighting strikes frequency as an instrument to 3G coverage. Standard errors are clustered at the county level. Numbers in parentheses are t-statistics. *, **, *** represent statistical significance at 10%, 5% and 1% level, respectively.

	$log(1 + \#Branches_{b,c,t})$			
	Entry		All Banks	
	(1)	(2)	(3)	(4)
3G Coverage	-0.006***		-0.022***	
	(-2.993)		(-8.819)	
3G $\widehat{\text{Coverage}}$		-0.004**		-0.099***
		(-2.178)		(-3.351)
log(PerCapitaIncome)	0.008**	0.008**	-0.053***	-0.064***
	(2.443)	(2.472)	(-4.711)	(-9.562)
log(CountyGDP)	0.007***	0.007***	0.005	0.005**
	(5.664)	(5.603)	(1.367)	(2.394)
log(TotalPop)	-0.005	-0.005	0.426***	0.364***
	(-1.483)	(-1.577)	(13.159)	(47.586)
3G $\text{Coverage}_{2007}\times$ Year		-0.000		0.002
		(-1.067)		(0.719)
Bank×County FE			✓	✓
State×Bank×Year FE	✓	✓	✓	✓
Observations	207,665	207,354	981,232	980,331
Adjusted R^2	0.183		0.946	

Table 4 Impact of 3G on Deposit Pricing

This table reports the impact of 3G on deposit pricing for banks with branches. The dependent variable is the deposit spread. All columns include bank, county and year-quarter fixed effects. Columns (2) and (5) further include bank×county fixed effects. We use instrumented 3G coverage by lightening strikes as independent variables in columns (3) and (6). Standard errors are clustered at the county level. Numbers in parentheses are t-statistics. *, **, *** represent statistical significance at 10%, 5% and 1% level, respectively.

	$\text{Spread}^{12MCD10K}$			$\text{Spread}^{36MCD10K}$		
	(1)	(2)	(3)	(4)	(5)	(6)
3G Coverage	0.019**	0.018**		0.014*	0.012	
	(2.281)	(2.102)		(1.662)	(1.506)	
3G $\widehat{\text{Coverage}}$			0.562***			0.523***
			(13.265)			(10.969)
log(PerCapitaIncome)	-0.126***	-0.120***	-0.058***	-0.131***	-0.127***	-0.083***
	(-2.893)	(-2.718)	(-4.213)	(-3.001)	(-2.854)	(-5.294)
log(CountyGDP)	0.019	0.020	0.011	0.028	0.029	0.028***
	(0.960)	(1.027)	(1.588)	(1.343)	(1.416)	(3.640)
log(TotalPop)	0.203***	0.223***	0.588***	0.162**	0.186**	0.613***
	(2.660)	(2.773)	(14.992)	(2.144)	(2.325)	(13.710)
log(#Banks)	-0.013	-0.012	-0.101***	-0.013	-0.013	-0.092***
	(-0.557)	(-0.491)	(-10.981)	(-0.552)	(-0.509)	(-9.277)
3G $\text{Coverage}_{2007}\times$ Quarter			0.026***			0.021***
			(9.992)			(7.249)
Bank FE	✓			✓		
County FE	✓			✓		
Bank×County FE		✓	✓		✓	✓
State×Quarter FE	✓	✓	✓	✓	✓	✓
Observations	332,614	332,054	312,417	304,801	304,248	286,625
Adjusted R^2	0.906	0.909		0.921	0.924	

Table 5 Impact of 3G on Loan Pricing

This table reports the impact of 3G coverage on loan pricing. The underlying sample includes loan-level observations of all bank originated loans recorded in HMDA in 2018. The outcome variable is the origination fee. $log(1+\#Branches)$ is the logarithm of one plus the number of branches a bank has for a given county. The Unreported *Controls* include the natural logarithm of loan size, loan type, loan purpose (home purchases, refinancing or others), loan-to-value ratios, debt-to-income ratios, and borrowers' income in natural logarithm, gender, age, and race. Standard errors are clustered at the county level. Numbers in parentheses are t-statistics. *, **, *** represent statistical significance at 10%, 5% and 1% level, respectively.

	Origination Fees (%)	
	(1)	(2)
3G Coverage × log(1+#Branches)	0.322***	0.315***
	(5.39)	(5.33)
3G Coverage	-0.797***	
	(-16.28)	
Controls	✓	✓
State FE	✓	
County FE		✓
Bank FE	✓	✓
Observations	1,815,347	1,815,322
Adjusted R^2	0.211	0.214

Table 6 The Impact of 3G Coverage on Lending Competition

This table reports the effect of 3G coverage on lending competition in the lending market. The dependent variable is HHI in the first two columns and the number of lenders in the last two columns. Both HHI and the number of lenders are constructed using all lenders. Standard errors are clustered at the county level. Numbers in parentheses are t-statistics. *, **, *** represent statistical significance at 10%, 5% and 1% level, respectively.

	HHI		log(#Lenders)	
	(1)	(2)	(3)	(4)
3G Coverage	-47.254***		0.032***	
	(-3.147)		(6.895)	
3G $\widehat{\text{Coverage}}$		-717.305**		0.194*
		(-2.013)		(1.650)
log(PerCapitaIncome)	86.013	-159.951	0.001	0.043
	(0.992)	(-1.573)	(0.053)	(1.271)
log(countyGDP)	13.134	38.031	0.020*	0.008
	(0.371)	(1.145)	(1.742)	(0.774)
log(TotalPop)	-671.912***	-992.039***	0.821***	0.778***
	(-5.375)	(-7.291)	(17.568)	(18.373)
3G $\text{Coverage}_{2007} \times$ Year		-62.768*		0.014
		(-1.929)		(1.335)
County FE	✓	✓	✓	✓
State×Year FE	✓	✓	✓	✓
Observations	30,493	30,493	30,493	30,493
Adjusted R^2	0.757		0.987	

Table 7 Distributional Effect of 3G on Loan Pricing across Age Groups

This table reports the interaction effect between 3G coverage and borrower age on loan pricing. The underlying sample includes all loans originated by banks in 2018 from HMDA. The analysis unit is at the loan level. The dependent variable is the loan origination fees in columns (1)-(2), and the loan interest rates in columns (3)-(4). The key independent variables of interest are the interaction term between 3G coverage and indicator variables for borrowers' age range. The Unreported *Controls* include the natural logarithm of loan size, loan type, loan purpose (home purchases, refinancing or others), loan-to-value ratios, debt-to-income ratios, and borrowers' income in natural logarithm, gender, age, and race. All columns include bank fixed effects. Odd columns include state fixed effects, and even columns include county fixed effects. Standard errors are clustered at the county level. Numbers in parentheses are t-statistics. *, **, *** represent statistical significance at 10%, 5% and 1% level, respectively.

	Origination Fees (%)		Interest Rate (%)	
	(1)	(2)	(3)	(4)
3G Coverage	-1.071***		-0.0615**	
	(-14.20)		(-2.12)	
34<Borrower Age<55 ×3G Coverage	0.213***	0.215***	0.0798***	0.0588**
	(3.57)	(3.67)	(2.94)	(2.28)
Borrower Age>54 ×3G Coverage	0.543***	0.502***	0.0636*	0.0287
	(6.80)	(6.16)	(1.80)	(0.83)
34<Borrower Age<55	-0.170***	-0.182***	-0.0691**	-0.0494*
	(-3.01)	(-3.26)	(-2.56)	(-1.92)
Borrower Age>54	-0.619***	-0.595***	-0.126***	-0.0924***
	(-8.30)	(-7.82)	(-3.87)	(-2.93)
Controls	✓	✓	✓	✓
State FE	✓		✓	
County FE		✓		✓
Bank FE	✓	✓	✓	✓
Observations	1,815,347	1,815,322	1,809,045	1,809,020
Adjusted R^2	0.210	0.214	0.677	0.677

Table 8 Distributional Effect of 3G on Financial Inclusion

The table presents results of the impact of 3G coverage on consumers' access to banking services, using FDIC Survey of Consumers Use of Banking and Financial Services. "Unbank" refers to consumers who do not have a bank account, and "Losing Banking Access" refers to consumers who once had a bank account but turn unbanked. "Young" refers to consumers under 45 years old; "High-income" refers to consumers with more than $50,000 annual income; "Education" refers to consumers with college education; "White" refers to white consumers. The observations are weighted to account for non-response and under-coverage. Numbers in parentheses are t-statistics. *, **, *** represent statistical significance at 10%, 5% and 1% level, respectively.

	Unbank		Losing Banking Access	
	(1)	(2)	(3)	(4)
3G Coverage	0.014		−0.073	
	(0.392)		(−1.501)	
3G Coverage×Young	−0.037***	−0.034***	−0.043***	−0.043***
	(−3.266)	(−3.042)	(−5.354)	(−5.424)
3G Coverage×High-income	−0.080**	−0.116***	0.028	0.013
	(−2.504)	(−2.993)	(1.409)	(0.393)
3G Coverage×Education	−0.001	−0.002	0.022	0.022
	(−0.130)	(−0.141)	(1.389)	(1.460)
3G Coverage×White	0.097***	0.111***	0.056	0.067
	(6.706)	(7.248)	(1.017)	(1.188)
Young	0.074***	0.072***	0.060***	0.060***
	(6.852)	(6.564)	(8.017)	(7.972)
High-income	−0.008	0.027	−0.074***	−0.060**
	(−0.252)	(0.712)	(−5.140)	(−2.286)
Education	−0.079***	−0.079***	−0.053***	−0.052***
	(−7.361)	(−7.203)	(−2.618)	(−2.689)
White	−0.185***	−0.199***	−0.091*	−0.101*
	(−13.165)	(−13.289)	(−1.730)	(−1.882)
MSA	✓		✓	
Year FE	✓		✓	
MSA×Year FE		✓		✓
Observations	144,794	144,794	139,910	139,910
Adjusted R^2	0.118	0.120	0.048	0.051

Table 9 Reasons of Becoming Unbank

This table studies the drop-out reason of previous banked individuals after 3G expansion. The dependent variables are indicators of a certain reason is chosen by interviewees. "Losing Banking Access" refers to consumers who once had a bank account but turn unbanked. The observations are weighted to account for non-response and under-coverage. Numbers in parentheses are t-statistics. *, **, *** represent statistical significance at 10%, 5% and 1% level, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	NoMoney	NoTrust	HighFee	NoPrivacy	Inconvenience	AccountProblem	NoProducts
3G Coverage	0.046	0.016	-0.121^{***}	0.016	-0.016	0.007	-0.005
	(0.630)	(0.336)	(-2.600)	(0.695)	(-0.560)	(0.157)	(-0.325)
3G Coverage$\times\mathbb{1}$(Losing Banking Access)	0.002	-0.006	0.135^{**}	-0.023	0.009	-0.030	0.001
	(0.025)	(-0.101)	(2.486)	(-0.830)	(0.260)	(-0.624)	(0.035)
MSA FE	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓
Observations	9,472	9,472	9,472	9,472	9,472	9,472	9,472
Adjusted R^2	0.109	0.042	0.049	0.018	0.019	0.045	0.016

Table 10 Estimated Parameters for the Structural Model

This table presents the estimated parameter values. Panel A shows the demand parameters. Panel B shows the supply parameters.

Panel A: Demand Parameters

α_y	α_o	β_y^{pre}	β_y^{post}	β_o	$\Delta\gamma_y$	λ_T	λ_F
0.91	0.70	$9.25*10^{-3}$	$1.20*10^{-3}$	$1.89*10^{-3}$	0.49	0.14	0.14

Panel B: Supply Parameters

c_T	c_F	κ_T	κ_F	CF_T	CF_F
-0.60	-0.90	$3.79*10^{-5}$	$4.27*10^{-5}$	0.02	0.05

Bank Competition amid Digital Disruption: Implications for Financial Inclusion

Internet Appendix

A Survey Data

The FDIC Survey of Household Use of Banking and Financial Services has been conducted by the FDIC biennially since 2009. Each survey collects responses from around 33,000 consumers, including their bank account ownership, like whether they are bank or unbanked, the primary methods they access their bank accounts if they are banked, why they are unbanked if they don't have a bank account, and saturated set of demographic information. Specifically, respondents' answer the question "unbank" with choices between "Unbanked" and "Has bank account;" the question "Previously banked" with choices between "Once had bank account" and "Never had bank account;" the question "most common way to access account" with the following six choices: "Bank teller," "ATM/Kiosk," "Telephone banking," "Online banking," "Mobile banking," and "Other;" the question "Main reason unbanked" with the following choices: "Inconvenient hours," "Inconvenient locations," "Account fees too high," "Account fees unpredictable," "Banks do not offer needed products or services," "Do not trust banks," "Do not have enough money to keep in account," "Avoiding bank gives more privacy," "ID, credit, or former bank account problems."

B Empirical Evidence for Digital Divide

In this section, we provide statistical analysis linking the major ways of accessing banks to consumers' characteristics and 3G coverage in their residence areas.

The dependent variable in Panel A of Table IA.2 is the difference in the share of consumers who rely mainly on branches versus those who prefer online and mobile banking services in each characteristic category. The negative coefficients on the interaction term between 3G coverage and indicators suggest that younger, richer, and more educated consumers shift away from branches and towards online and mobile banking after 3G expands to their residential areas. The diverging preference is robust to different specifications of fixed effects. Moreover, to confirm that the results are not driven purely by the increasing share of mobile and online banking, we conduct the same analysis with the share of branches as the dependent variable in panel B. The results are largely consistent with panel A.

The evidence on the higher-income group is not statistically significant. To better understand the difference between the two panels, we run the analysis for the share of mobile banking and online banking separately, and present results in Table IA.3. Panel B shows that, after 3G expansion, more higher-income consumers shift to online banking, compared to poorer consumers.

Overall, the results show that heterogeneous changes to access banking services across digital and non-digital consumers following 3G expansion.

C Event Study for Bank Branch Closure

To validate the parallel trend assumption, we conduct an event study focusing on sharp increases in county 3G coverage. We define a treatment event as a county's 3G coverage increasing by more than 50% from the previous year. Given the monotonic increasing feature of 3G coverage, such event can happen at most once for one county.

For each treated county, we construct a control county if a county has the closest matching score based on county characteristics but did not experience a sharp increase in 3G coverage ever or reach 30% 3G coverage within three years upon the treatment event. We acknowledge that the matching outcome is not ideal because controlled counties are economically less developed than the treated group. However, we show graphically in Figure 6 that there is no pretrend between the two groups.

Focusing on the sample constructed, we estimate a DiD specification as below:

$$Y_{cohort,c,t} = \alpha_{cohort,t} + \alpha_{cohort,c} + \beta \text{Treat} * Post + \lambda X_{c,t} + \epsilon_{cohort,c,t}, \tag{IA.A}$$

where Treat refers to the treatment counties with sharp increase in 3G coverage. We consider a three-year window around the treatment year [-3, -2, -1, 0, 1, 2, 3], and assign Post to be 1 if the event year is ≥ 0 and zero otherwise. We use *cohort* to indicate the matched group for each treated county. The analysis compares county-level number of branches between the treated and control groups within a window around the event that counties experienced a sharp 3G coverage increase.

We report the estimation results in Table IA.4. Columns 1-2 consider the treatment group only, where we regress the number of branches on post dummy and year dummies relative to the year of the event and all the controls. The coefficient on the post and post-event dummies are significantly negative when explaining the number of local branches; however, those on pre-event dummies are not statistically significant. We observe the same pattern in columns 3-4 which consider both treatment and control groups: the coefficients on the interaction term Treat×Post and those involving post-event dummies are significantly negative; in contrast, interaction terms with pre-event dummies have much smaller magnitude and are statistically indistinguishably from zero.

We illustrate the dynamic differences between the two groups in Figure 6. The treated counties start to experience significant branch closures at the time of sharp 3G networks expansion and this effect gets stronger over the years. In contrast, the differences during the 3-year pre-event window are small in magnitude and statistically indistinguishable from zero. Hence, there exit no pretrends.

Overall, the results in this section support Prediction 1 and suggest a causal impact of 3G networks on banks' decision to shut down branches.

D Model Analysis

D.1 The General Case

Due to the nested structure, the likelihood $s_{i,}$ can be decomposed into two parts, 1) the likelihood that one-type is chosen, and 2) conditional on that, bank j is selected. Following formula based on the property of the generalized extreme value distribution, the conditional probability 2) is given by

$$Pr_i(j|j \in t) = \frac{A_{i,j}}{Z_{i,t}}, \quad t \in \{T, F\},$$

where

$$A_{i,j} = \exp(\frac{1}{\lambda_t}(-\alpha_i r_j + \beta_i b_j + \gamma_i d_j)), \qquad Z_{i,t} = \sum_{j=1}^{J_t} \exp(\frac{1}{\lambda_t}(-\alpha_i r_j + \beta_i b_j + \gamma_i d_j)).$$

The term $A_{i,j}$ captures the consumer type i's exponential utility from accessing the bank j's service, and the them $Z_{i,t}$ is the sum of her exponential utility assuming she have access to all $t-$type banks. Since we assume all banks in each type are the same, this conditional probability equals to $\frac{1}{J_t}$. The marginal probability that $t-$type bank is chosen is

$$Pr(j \in t) = \frac{Z_{i,t}^{\lambda_t}}{1 + \sum_{t \in \{T,F\}} Z_{i,t}^{\lambda_t}},$$

where we standardize the utility from the outside option to be 1. Intuitively, if $t-$type bank' service generates a higher utility, consumer i is more likely to choose that type of bank. These two terms pin down $s_{i,j}$ where bank j is one of type-t banks as

$$s_{i,j} = \frac{A_{i,j}}{Z_{i,t}} \frac{Z_{i,t}^{\lambda_t}}{1 + \sum_{t \in \{T,F\}} Z_{i,t}^{\lambda_t}}. \tag{IA.A}$$

The first-order condition for banks' optimization problem gives rise to the following equations:

$$FOC_{r_j} : \; r_j = c_j + \frac{\sum_{i \in y,o} \mu_i s_{i,j}}{\sum_{i \in y,o} \mu_i \frac{\alpha_i}{\lambda_t} s_{i,j} \Big(1 + (\lambda_t - 1)\frac{A_{i,j}}{Z_{i,t}} - \lambda_t s_{i,j}\Big)};$$

$$FOC_{b_j} : \; b_j = \frac{1}{\kappa_j}(r_j - c_j) \sum_{i \in y,o} \mu_i \frac{\beta_i}{\lambda_t} s_{i,j} \Big(1 + (\lambda_t - 1)\frac{A_{i,j}}{Z_{i,t}} - \lambda_t s_{i,j}\Big).$$

The difference of $r_j - c_j$ captures the markup of bank j.

Proof. We first derive this derivative $\frac{\partial s_{i,j}}{\partial r_J}$.

$$\begin{aligned}
\frac{\partial \ln s_{i,j}}{\partial r_j} = \frac{1}{s_{i,j}}\frac{\partial s_{i,j}}{\partial r_j} &= \frac{\partial \ln A_{i,j}}{\partial r_j} + (\lambda_t - 1)\frac{\partial \ln Z_{i,t}}{\partial r_j} - \frac{\partial \ln(1 + \sum_{t\in\{T,F\}} Z_{i,t}^{\lambda_t})}{\partial r_j} \\
&= \frac{1}{A_{i,j}}\Big(-\frac{\alpha_i}{\lambda_t}\Big)A_{i,j} + (\lambda_t - 1)\frac{1}{Z_{i,t}}\Big(-\frac{\alpha_i}{\lambda_t}\Big)A_{i,j} - \lambda_t\frac{Z_{i,t}^{\lambda_t-1}\Big(-\frac{\alpha_i}{\lambda_t}\Big)A_{i,j}}{1+\sum_{t\in\{T,F\}} Z_{i,t}^{\lambda_t}} \\
&= \Big(-\frac{\alpha_i}{\lambda_t}\Big)\Big(1 + (\lambda_t - 1)\frac{A_{i,j}}{Z_{i,t}} - \lambda_t s_{i,j}\Big) \\
\Longrightarrow \frac{\partial s_{i,j}}{\partial r_j} &= \Big(-\frac{\alpha_i}{\lambda_t}\Big)s_{i,j}\Big(1 + (\lambda_t - 1)\frac{A_{i,j}}{Z_{i,t}} - \lambda_t s_{i,j}\Big).
\end{aligned}$$

Similarly, we have

$$\frac{\partial s_{i,j}}{\partial b_j} = \Big(\frac{\beta_i}{\lambda_t}\Big)s_{i,j}\Big(1 + (\lambda_t - 1)\frac{A_{i,j}}{Z_{i,t}} - \lambda_t s_{i,j}\Big).$$

Then, it is straightforward to derive the the first-order conditions for banks:

$$\begin{aligned}
r_j &= c_j + D_j\Big(-\frac{\partial D_j}{\partial r_j}\Big)^{-1} = c_j + D_j\Big(-\sum_{i\in\{y,o\}}\mu_i\frac{\partial s_{i,j}}{\partial r_j}\Big)^{-1}, \\
b_j &= \frac{1}{\kappa_j}(r_j - c_j)\frac{\partial D_j}{\partial b_j} = \frac{1}{\kappa_j}(r_j - c_j)\sum_{i\in\{y,o\}}\mu_i\frac{\partial s_{i,j}}{\partial b_j}.
\end{aligned}$$

□

D.2 A Simplified Case

The model has a simple closed-form solution when both banks and consumers are homogeneous, which is when $\lambda_t = 1$, $\alpha_i = \alpha$, $\beta_i = \beta$, $\gamma_i = 0$, $c_j = c$, and $\kappa_j = \kappa$. In this case, the market share of each bank j (denoted as s_j) among the total J banks is the same, and it is easy to show that Equations (14) and (15) collapse to be

$$r_j = c + \frac{\sum_{i\in y,o}\mu_i s_{i,j}}{\sum_{i\in y,o}\mu_i\alpha_i s_{i,j}\Big(1 - s_{i,j}\Big)} = c + \frac{s_j}{\alpha s_j(1 - s_j)} = c + \frac{1}{\alpha(1 - s_j)}$$

$$b_j = \frac{1}{\kappa_j}\frac{\sum_{i\in y,o}\mu_i s_{i,j}\sum_{i\in y,o}\mu_i\beta_i s_{i,j}\Big(1-s_{i,j}\Big)}{\sum_{i\in y,o}\mu_i\alpha_i s_{i,j}\Big(1-s_{i,j}\Big)} = \frac{1}{\kappa}\frac{s_j\times\beta s_j(1-s_j)}{\alpha s_j(1-s_j)} = \frac{1}{\kappa}\frac{\beta}{\alpha}s_j,$$

where $s_j = \frac{exp(-\alpha r_j+\beta b_j)}{1+Jexp(-\alpha r_j+\beta b_j)}$.

Relationship between r_j and b_j We can rewrite the relationship between r_j and b_j as

$$b_j = \frac{\beta(\alpha(r_j-c)-1)}{\alpha^2\kappa(r_j-c)}$$

Then it is easy to show that $\frac{\partial b_j}{\partial r_j} = \frac{\beta}{(c-r)^2\alpha^2\kappa} > 0$. This is intuitive: to cover the cost to operate more branches, banks have to charge a higher service fee.

Derivative of r_j and b_j in respect with β We take implicit differentiation of β for both r_j and b_j, we get

$$\alpha\kappa\frac{\partial b_j}{\partial\beta} = s_j + \beta\frac{\partial s_j}{\partial\beta}$$
$$\frac{\partial r}{\partial\beta}(1-s_j) = (r_j-c)\frac{\partial s_j}{\partial\beta}$$
$$\frac{\partial s_j}{\partial\beta} = s_j\frac{-\alpha\frac{\partial r_j}{\partial\beta}+\beta\frac{\partial b_j}{\partial\beta}+b_j}{1+Jexp(-\alpha r_j+\beta b_j)}$$

Combine above equations, we get

$$\frac{\partial r}{\partial\beta}(1-s_j) = (r_j-c)s_j\frac{-\alpha\frac{\partial r_j}{\partial\beta}+\beta\frac{\partial b_j}{\partial\beta}+b_j}{1+Jexp(-\alpha r_j+\beta b_j)}$$
$$\implies \frac{\partial r}{\partial\beta}\frac{(1-s_j)}{(r_j-c)s_j} = \frac{-\alpha\frac{\partial r_j}{\partial\beta}+\beta\frac{1}{\alpha\kappa}\Big(\frac{\partial r}{\partial\beta}\frac{\beta(1-s_j)}{r_j-c}+s_j\Big)+b_j}{1+Jexp(-\alpha r_j+\beta b_j)}$$
$$\implies \Big(\frac{(1-s_j)}{(r_j-c)s_j} - \frac{-\alpha+\beta\frac{1}{\alpha\kappa}\frac{\beta(1-s_j)}{r_j-c}}{1+Jexp(-\alpha r_j+\beta b_j)}\Big)\frac{\partial r_j}{\partial\beta} = \frac{\beta\frac{1}{\alpha\kappa}s_j+b_j}{1+Jexp(-\alpha r_j+\beta b_j)}$$
$$\implies \Bigg(\alpha+\frac{(1-s_j)}{(r_j-c)s_j}Jexp(-\alpha r_j+\beta b_j)+\frac{(1-s_j)}{(r_j-c)}\Big(\frac{1}{s_j}-\frac{\beta^2}{\alpha\kappa}\Big)\Bigg)\frac{\partial r_j}{\partial\beta} = \beta\frac{1}{\alpha\kappa}s_j+b_j$$

If $\alpha + \frac{(1-s_j)}{(r_j-c)}\left(\frac{1}{s_j} - \frac{\beta^2}{\alpha\kappa}\right) > 0$, then we have $\frac{\partial r_j}{\partial \beta} > 0$.

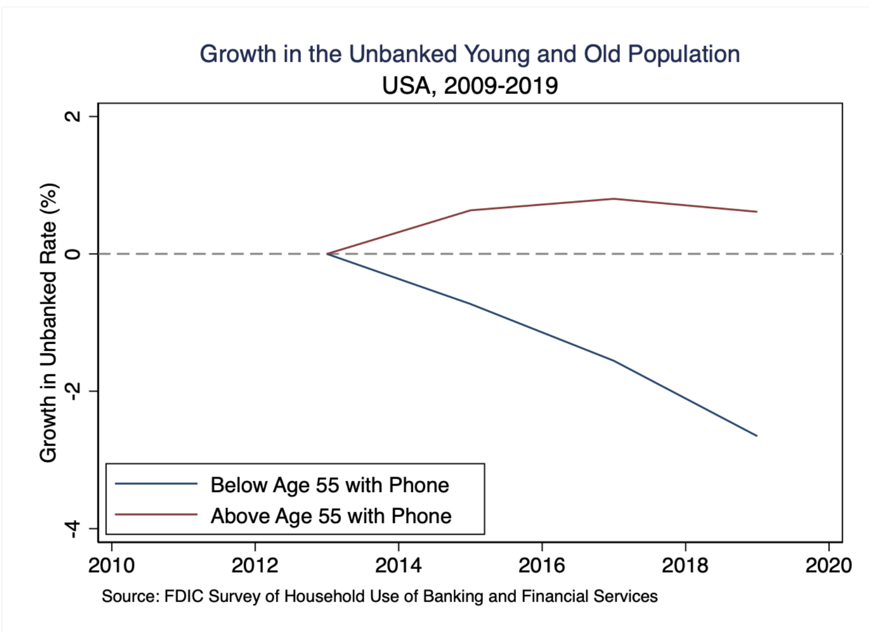
$$\alpha + \frac{(1-s_j)}{(r_j-c)}\left(\frac{1}{s_j} - \frac{\beta^2}{\alpha\kappa}\right) > 0 \Longrightarrow \frac{\alpha(r_j-c)}{1-s_j} + \frac{1}{s_j} - \frac{\beta^2}{\alpha\kappa} > 0 \Longrightarrow \frac{1}{(1-s_j)^2} + \frac{1}{s_j} - \frac{\beta^2}{\alpha\kappa} > 0$$

As s_j is bounded by $[0, \frac{1}{J}]$, and $\frac{1}{(1-s_j)^2} + \frac{1}{s_j}$ is monotonically decreasing in s_j. Therefore, when $J + \frac{J^2}{(J-1)^2} \geq \frac{\beta^2}{\alpha\kappa}$, we have $\frac{\partial r_j}{\partial \beta} > 0$, $\frac{\partial s_j}{\partial \beta} > 0$, and $\frac{\partial b_j}{\partial \beta} > 0$.

E Figures

Figure IA.1. Growth in the Unbanked Young and Old Population with Phone

This figure plots the growth rate of unbanked consumers under 55 versus above 55 over years with phones. Source: FDIC Survey of Household Use of Banking and Financial Services.



F Tables

Table IA.1 Impact of Lightning Strikes on Local Economic Conditions

The table presents results on the impact of lightening strikes on local economic conditions. High lightening strikes represent counties whose average population-weighted frequency of lightning strikes across 2007 to 2018 is higher than the state median. Standard errors are clustered at the county level. Numbers in parentheses are t-statistics. *, **, *** represent statistical significance at 10%, 5% and 1% level, respectively.

	log(CountyGDP)	log(TotalPop)	Unemployment Rate	Share of Pop Under 40
	(1)	(2)	(3)	(4)
$\mathbb{1}$(High Lightning) × $Year$	−0.001	0.0004	0.013	−0.00002
	(−1.012)	(1.533)	(0.986)	(−0.305)
County FE	✓	✓	✓	✓
State×Year FE	✓	✓	✓	✓
Observations	33,586	34,070	34,081	34,070
Adjusted R^2	0.993	1.000	0.915	0.984

Table IA.2 Impact of 3G Coverage on Consumers' Main Access to Banking Services

The table presents results of the impact of 3G coverage on consumers' access to banking services using branches versus mobile banking, using FDIC Survey of Consumers Use of Banking and Financial Services. In panel A, the dependent variable is the difference in the proportion of consumers using bank teller v.s. both mobile and online baking as main access to banking services in each MSA region. In panel B, the dependent variable is the proportion of consumers using bank teller as main access to banking services in each MSA region. "Young" refers to consumers under 45 years old; "High-income" refers to consumers with more than $50,000 annual income; "Education" refers to consumers with college education. Numbers in parentheses are t-statistics. *, **, *** represent statistical significance at 10%, 5% and 1% level, respectively.

	Panel A: Share difference between using bank teller v.s. online + mobile banking					
	(1)	(2)	(3)	(4)	(5)	(6)
3G Coverage	0.236***		0.396***		0.245***	
	(7.075)		(3.302)		(12.05)0	
Young	0.065*	0.064***				
	(1.883)	(4.476)				
3G Coverage×Young	−0.322***	−0.321***				
	(−8.774)	(−18.894)				
High-income			0.046	0.106		
			(0.407)	(0.811)		
3G Coverage×High-income			−0.327***	−0.391***		
			(−2.831)	(−2.892)		
Education					0.065***	0.066***
					(4.395)	(4.461)
3G Coverage×Education					−0.316***	−0.318***
					(−17.324)	(−17.471)

	Panel B: Share using bank teller					
	(1)	(2)	(3)	(4)	(5)	(6)
3G Coverage	0.121***		0.163**		0.102***	
	(6.994)		(2.007)		(8.446)	
Young	0.031*	0.030***				
	(1.718)	(3.677)				
3G Coverage×Young	−0.154***	−0.153***				
	(−8.107)	(−16.053)				
High-income			0.009	0.029		
			(0.113)	(0.348)		
3G Coverage×High-income			−0.108	−0.130		
			(−1.378)	(−1.510)		
Education					0.015**	0.016**
					(2.431)	(2.541)
3G Coverage×Education					−0.091***	−0.092***
					(−10.488)	(−10.610)
MSA FE	✓		✓		✓	
Year FE	✓		✓		✓	
MSA×Year FE		✓		✓		✓

Table IA.3 Impact of 3G Coverage on Consumers' Access to Banking Services

The table presents results of the impact of 3G coverage on consumers' access to banking services using branches versus mobile banking, using FDIC Survey of Consumers Use of Banking and Financial Services. The dependent variable is the proportion of consumers in each group using mobile banking (panel A) and online banking (panel B) in each MSA region. "Young" refers to consumers under 45 years old;"High-income" refers to consumers with more than \$50,000 annual income; "Education" refers to consumers with college education. Numbers in parentheses are t-statistics. *, **, *** represent statistical significance at 10%, 5% and 1% level, respectively.

	Panel A: Share using mobile banking as main way to access bank services					
	(1)	(2)	(3)	(4)	(5)	(6)
3G Coverage	-0.079^{***}		-0.070^{*}		-0.037^{***}	
	(-5.403)		(-1.778)		(-6.421)	
Young	-0.032^{**}	-0.032^{***}				
	(-2.136)	(-5.721)				
3G Coverage×Young	0.131^{***}	0.132^{***}				
	(8.176)	(17.270)				
High-income			-0.037	-0.046		
			(-0.992)	(-1.134)		
3G Coverage×High-income			0.061	0.070^{*}		
			(1.597)	(1.686)		
Education					-0.011^{***}	-0.011^{***}
					(-4.073)	(-4.102)
3G Coverage×Education					0.044^{***}	0.045^{***}
					(10.698)	(10.720)

	Panel B: Share using online banking as main way to access bank services					
	(1)	(2)	(3)	(4)	(5)	(6)
3G Coverage	-0.036^{*}		-0.233^{***}		-0.106^{***}	
	(-1.907)		(-3.400)		(-8.655)	
Young	-0.002	-0.002				
	(-0.100)	(-0.717)				
3G Coverage×Young	0.036^{*}	0.037^{***}				
	(1.746)	(5.484)				
High-income			-0.037	-0.077		
			(-0.579)	(-0.979)		
3G Coverage×High-income			0.219^{***}	0.261^{***}		
			(3.321)	(3.208)		
Education					-0.040^{***}	-0.040^{***}
					(-5.251)	(-5.283)
3G Coverage×Education					0.181^{***}	0.181^{***}
					(19.104)	(19.214)
MSA FE	✓		✓		✓	
Year FE	✓		✓		✓	
MSA×Year FE		✓		✓		✓

Table IA.4 Event Study for Branch Closure

The table reports DiD analysis results for county-level branch closures. The dependent variables are the same as those in Table 2. The treatment group includes counties that had a sharp increase in 3G coverage, more than 50% in a single year. For each treated county, we construct a control county if a county has the closest matching score based on county characteristics but did not experience a sharp increase in 3G coverage ever nor reach 30% 3G coverage three years after the treatment year. The sample covers a three-year window around the shock year, [-3, -2, -1, 0, 1, 2, 3]. Post equals to 1 if the window is above 0 and zero otherwise. Standard errors are clustered at the county level. Numbers in parentheses are t-statistics. *, **, *** represent statistical significance at 10%, 5% and 1% level, respectively.

	log($\#Branches_{c,t}$)				
	Treatment only			Treatment + Control	
	(1)	(2)		(3)	(4)
Post	-0.004^{**}		Treat×Post	-0.007^{**}	
	(-1.980)			(-2.078)	
Window $t-3$ or earlier		0.003	Treat×Window $t-3$		-0.005
		(0.818)			(-1.446)
Window $t-2$		0.003	Treat×Window $t-2$		0.001
		(1.324)			(0.410)
Window t		-0.003^{*}	Treat×Window t		-0.004
		(-1.681)			(-1.483)
Window $t+1$		-0.006^{**}	Treat×Window $t+1$		-0.007**
		(-2.386)			(-2.158)
Window $t+2$		-0.004	Treat×Window $t+2$		-0.012***
		(-1.282)			(-2.930)
Window $t+3$ or later		-0.006	Treat×Window $t+3$		-0.012***
		(-1.459)			(-2.683)
log(PerCapitaIncome)	0.047***	0.047***	log(PerCapitaIncome)	0.046**	0.044**
	(3.283)	(3.240)		(2.469)	(2.362)
log(CountyGDP)	-0.012^{*}	-0.012^{*}	log(countyGDP)	0.006	0.007
	(-1.728)	(-1.700)		(0.494)	0.504
log(TotalPop)	0.115***	0.108***	log(TotalPop)	0.143***	0.147***
	(3.423)	(3.144)		(2.783)	(2.841)
log(# Bank)	0.631***	0.632***	log(# Bank)	0.349***	0.348***
	(36.099)	(36.085)		(17.928)	(17.880)
County FE	✓	✓		✓	✓
Year FE	✓	✓		✓	✓
Observations	27,853	27,853		20,655	20,655
Adjusted R^2	0.997	0.997		0.997	0.997

Table IA.5 Impact of 3G on Loan Pricing (All lenders)

This table reports the impact of 3G coverage on loan pricing and is analogous to Table 5. The underlying sample includes loan-level observations of all originated loans from all kinds of lenders recorded in HMDA in 2018. *Branch* equals 100 if the lender has a branch in the county and 0 otherwise. $\log(1 + \#Branches)$ is the logarithm of one plus the number of branches a bank has for a given county. The Unreported *Controls* include the natural logarithm of loan size, loan type, loan purpose (home purchases, refinancing or others), loan-to-value ratios, debt-to-income ratios, and borrowers' income in natural logarithm, gender, age, and race. All columns include lender fixed effects. Odd columns include state fixed effects, and even columns include county fixed effects. Standard errors are clustered at the county level. Numbers in parentheses are t-statistics. *, **, *** represent statistical significance at 10%, 5% and 1% level, respectively.

	Origination Fees (%)			
	(1)	(2)	(3)	(4)
Branch×3G Coverage	0.523***	0.495***		
	(8.89)	(8.86)		
log(1+#Branches)×3G Coverage			0.657***	0.678***
			(11.57)	(10.53)
log(1+#Branches)			-0.732***	-0.746***
			(-13.03)	(-11.68)
Branch	-0.811***	-0.757***		
	(-14.01)	(-13.75)		
3G Coverage	-0.880***		-0.905***	
	(-20.72)		(-21.17)	
Controls	✓	✓	✓	✓
State FE	✓		✓	
County FE		✓		✓
Bank FE	✓	✓	✓	✓
Observations	5,095,074	5,095,062	5,095,074	5,095,062
Adjusted R^2	0.239	0.240	0.239	0.240

Table IA.6 Impact of 3G on Market Share of Banks with Branches

This tables tabulates the effect of 3G on the market share of bank lenders for all loans. The analysis unit is at county-year level. The dependent variable is the loan market share of lenders with at least one branch for a given county-year pair. Columns (1)-(2) include all lenders; columns (3)-(4) include entry lenders; and columns (5)-(6) include incumbent banks. Even columns include the interaction term between 3G and young county which is a dummy variable indicating that a county's median age is below 40. All columns include year, state-year, and county fixed effects. Standard errors are clustered at the county level. Numbers in parentheses are t-statistics. *, **, *** represent statistical significance at 10%, 5% and 1% level, respectively.

	Market Share of Banks with Branches	
	(1)	(2)
3G Coverage	-1.080**	-0.582
	(-2.218)	(-1.013)
Young County × 3G Coverage		-1.174**
		(-2.004)
Young County		0.477
		(0.746)
log(PerCapitaIncome)	-4.035*	-4.388**
	(-1.843)	(-2.005)
log(countyGDP)	-0.887	-0.878
	(-0.815)	(-0.809)
log(TotalPop)	-17.900***	-15.898***
	(-3.793)	(-3.311)
log(TotalLoan)	3.210***	3.186***
	(5.181)	(5.142)
County FE	✓	✓
Year FE	✓	✓
Observations	30,501	30,479
Adjusted R^2	0.791	0.793

Table IA.7 Digital Inequality

The table presents results of the impact of 3G coverage on consumers' access to banking services, using FDIC Survey of Consumers Use of Banking and Financial Services. "Unbank" refers to consumers who do not have a bank account, and "Losing Banking Access" refers to consumers who once had a bank account but turn unbanked. "Young" refers to consumers under 45 years old; "High-income" refers to consumers with more than $50,000 annual income; "Education" refers to consumers with college education; "Phone" refers to consumers with a mobile phone; "White" refers to white consumers. The observations are weighted to account for non-response and under-coverage. Numbers in parentheses are t-statistics. *, **, *** represent statistical significance at 10%, 5% and 1% level, respectively.

	Unbank									
	Young		High-income		Educated		Phone		White	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
3G Coverage	0.019**		0.089***		0.039***		0.513***		−0.043***	
	(2.104)		(2.608)		(3.889)		(3.817)		(−3.015)	
Young	0.075***	0.073***								
	(6.639)	(6.391)								
3G Coverage×Young	−0.029**	−0.027**								
	(−2.467)	(−2.290)								
High-income			−0.034	0.014						
			(−1.064)	(0.361)						
3G Coverage×High-income			−0.094***	−0.144***						
			(−2.876)	(−3.622)						
Education					−0.063***	−0.063***				
					(−5.754)	(−5.641)				
3G Coverage×Education					−0.054***	−0.054***				
					(−4.750)	(−4.675)				
Phone							0.231**	0.217**		
							(2.278)	(2.124)		
3G Coverage×Phone							−0.340***	−0.326***		
							(−3.298)	(−3.137)		
White									−0.185***	−0.198***
									(−13.040)	(−13.047)
3G Coverage×White									0.061***	0.074***
									(4.142)	(4.729)
MSA	✓		✓		✓		✓		✓	
Year FE	✓		✓		✓		✓		✓	
MSA×Year FE		✓		✓		✓		✓		✓
Observations	144,794	144,794	144,794	144,794	144,794	144,794	75,337	75,337	144,794	144,794
Adjusted R^2	0.022	0.024	0.061	0.064	0.057	0.060	0.031	0.033	0.062	0.064

Table IA.8 Preference for Branches by Borrowers' Age

This table reports the effect of 3G coverage on borrowers' choices of lenders with branches by borrowers' age using loan-level data from HMDA in 2018. The outcome variable equals 100 if the lender has a branch in the county and 0 otherwise. The independent variables of interest are the indicator variables for the borrowers' age range. For example, *34<Borrower Age<45* equals one if the borrower is between 34 and 45 years old. The Unreported *Controls* include the natural logarithm of loan size, loan type, loan purpose (home purchases, refinancing or others), loan-to-value ratios, debt-to-income ratios, and borrowers' income in natural logarithm, gender, and race. The underlying sample includes loan-level observations of all originated loans recorded in HMDA in 2018. Standard errors are clustered at the county level. Numbers in parentheses are t-statistics. *, **, *** represent statistical significance at 10%, 5% and 1% level, respectively.

	Likelihood of choosing a lender with a branch			
	(1)	(2)	(3)	(4)
34<Borrower Age<45	-0.105	-0.213***	-6.963***	-2.717***
	(-1.24)	(-6.52)	(-8.99)	(-5.49)
44<Borrower Age<55	1.741***	-0.032	-9.125***	-4.214***
	(15.08)	(-0.75)	(-9.58)	(-7.18)
54<Borrower Age<65	4.916***	0.345***	-7.223***	-5.436***
	(30.08)	(6.01)	(-7.35)	(-9.06)
Borrower Age>64	10.957***	1.172***	-5.361***	-5.049***
	(46.18)	(14.13)	(-4.33)	(-7.24)
34<Borrower Age<45×3G Coverage			6.968***	2.545***
			(8.80)	(5.08)
44<Borrower Age<55×3G Coverage			11.036***	4.248***
			(11.24)	(7.14)
54<Borrower Age<65×3G Coverage			12.338***	5.877***
			(11.70)	(9.62)
64<Borrower Age×3G Coverage			16.600***	6.328***
			(12.43)	(8.84)
Controls	✓	✓	✓	✓
County FE	✓	✓	✓	✓
Lender FE		✓		✓
Observations	6,125,807	6,125,767	6,125,807	6,125,767
Adjusted R^2	0.149	0.780	0.149	0.780

Table IA.9 Distributional Effect of 3G on Loan Pricing across Age Groups (All Lenders)

This table reports the interaction effect between 3G coverage and borrower age on loan pricing using loans originated by all lenders. The table is analogous to Table 7. The analysis unit is at the loan level. The dependent variable is the loan origination fees in columns (1)-(2), and the loan interest rates in columns (3)-(4). The key independent variables of interest are the interaction term between 3G coverage and indicator variables for borrowers' age range. The Unreported *Controls* include the natural logarithm of loan size, loan type, loan purpose (home purchases, refinancing or others), loan-to-value ratios, debt-to-income ratios, and borrowers' income in natural logarithm, gender, age, and race. All columns include lender fixed effects. Odd columns include state fixed effects, and even columns include county fixed effects. Standard errors are clustered at the county level. Numbers in parentheses are t-statistics. *, **, *** represent statistical significance at 10%, 5% and 1% level, respectively.

	Origination Fees		Interest Rate	
	(1)	(2)	(3)	(4)
3G Coverage	-0.978***		-0.137*	
	(-17.47)		(-1.91)	
34<Borrower Age<55×3G Coverage	0.142***	0.147***	0.205**	0.209**
	(3.43)	(3.65)	(2.34)	(2.33)
Borrower Age>54×3G Coverage	0.318***	0.300***	-0.0168	0.00927
	(5.05)	(4.84)	(-0.08)	(0.05)
34<Borrower Age<55	-0.0687*	-0.0789**	-0.141**	-0.142**
	(-1.73)	(-2.05)	(-2.39)	(-2.31)
Borrower Age>54	-0.242***	-0.234***	0.117	0.0915
	(-4.10)	(-4.04)	(0.37)	(0.31)
Controls	✓	✓	✓	✓
State FE	✓		✓	
County FE		✓		✓
Lender FE	✓	✓	✓	✓
Observations	5,095,074	5,095,062	5,035,302	5,035,290
Adjusted R^2	0.239	0.240	0.0459	0.0456