Grand Theft Identity: The Privacy Costs of Digitalization*

Kenny Phua Chishen Wei Gloria Yang Yu

Abstract

We study whether greater digital engagement increases the risk of identity theft by exploiting bank branch closures as a shock that shifts economic activity online. Using two quasi-natural experiments, we find causal evidence that branch closures increase identity theft, particularly in more vulnerable communities. Exposed consumers spend more time on mobile apps and shift their expenditures from offline to online channels. Adversarial activities associated with identity theft, such as unwanted calls and phishing attempts, increase after branch closures. Overall, our evidence suggests that digitalization offers consumer benefits, but also imposes hidden privacy costs.

*Yu (corresponding author): gloriayu@smu.edu.sg, Singapore Management University. Phua: kenny.phua@uts.edu.au, University of Technology Sydney and Digital Finance Cooperative Research Center. Wei: chishen.wei@polyu.edu.hk, Hong Kong Polytechnic University. We are grateful for comments from Bo Becker, William Cong, Marcin Kacperczyk, Robert Marquez, Christine Parlour, Manju Puri, Adriano Rampini, Alberto Rossi. We also thank seminar participants at the Hong Kong Polytechnic University and the University of Technology Sydney, as well as participants at the Monash Winter Finance Conference.

Grand Theft Identity: The Privacy Costs of Digitalization

Abstract

We study whether greater digital engagement increases the risk of identity theft by exploiting bank branch closures as a shock that shifts economic activity online. Using two quasi-natural experiments, we find causal evidence that branch closures increase identity theft, particularly in more vulnerable communities. Exposed consumers spend more time on mobile apps and shift their expenditures from offline to online channels. Adversarial activities associated with identity theft, such as unwanted calls and phishing attempts, increase after branch closures. Overall, our evidence suggests that digitalization offers consumer benefits, but also imposes hidden privacy costs.

1 Introduction

As more economic activities shift online, privacy concerns have become a significant consumer issue. Safeguards such as firm-level data security policies, platform-level protection measures, and government-level regulations are effective in protecting privacy, but these protections govern the legitimate collection of consumer data (Ramadorai, Uettwiller, Walther, 2025). Adversarial actors, however, can unlawfully exploit vulnerabilities to harvest personal identifying information (PII). Our paper focuses upstream of these institutional protections—we pinpoint consumers' increasing engagement with the digital economy as a structural source of exposure to privacy costs. Digital engagement is a modern necessity, but it imposes privacy risks that are hidden and difficult to quantify. For example, 16 billion login credentials, including those linked to Apple, Facebook, and Google accounts, have been leaked (Forbes Magazine, 2025). We hypothesize that digital engagement exposes consumers to greater privacy costs and assess whether vulnerable communities are disproportionally affected.

The digitalization of banking services provides an opportunity to test our hypothesis. Banks are uniquely positioned to shape consumer behavior in the digital economy because they facilitate many economic activities that involve payments and transfers. Traditionally, banks had extensive physical branch networks, but are now closing branches as they offer more digital services (Amberg and Becker, 2024; Jiang, Yu, Zhang, 2025; Narayanan, Ratnadiwakara, Strahan, 2025). Consumers affected by branch closures must learn to adopt digital banking and payment tools. These tools can reduce the marginal costs of using other digital services, sparking network effects that reinforce digital adoption and drive greater digital engagement.

We expect bank branch closures to push consumers toward greater digital engagement, hence exposing them to privacy risks through two pathways. First, branch closures nudge consumers to use online services and conduct more transactions digitally. Conducting digital activities typically requires the transfer of sensitive PII, which could leak and be compromised. Digital banking in itself is unlikely to incur privacy costs because financial institutions invest heavily in cybersecurity infrastructure. However, many digital transactions involve third parties that process digital transactions and may have weaker data security. Some third parties act as data brokers, who harvest digital footprints for sale and have been prosecuted for selling PII to scammers (Federal Trade Commission, 2015). Even sophisticated consumers can be affected because these data breaches often occur outside of one's control.

Second, branch closures eliminate a crucial alternative for consumers who prefer in-person services, exposing vulnerable communities. Physical branch services can limit PII exposure, allow face-to-face verification, and provide personalized security advice. For example, a financial advice columnist for New York Magazine recounts her experience with a bank teller, who warned that her $50,000 cash withdrawal was likely related to a scam (Cowles, 2024). Without these physical touchpoints, some consumers may become more susceptible to adversarial tactics such as phishing attacks using phone calls, text messaging, emails, SIM card swaps, and even generative artificial intelligence (AI) tools. These attacks can be especially harmful during the transition to digital services, where some consumers are less familiar with security practices or warning signs of fraud.

Privacy costs are multifaceted. We focus on identity theft because it is the most common data privacy concern (Armantier, Doerr, Frost, Fuster, et al., 2024) and arises from the abuse of user data. Under 18 U.S.C. §1028(a)(7), identity theft is any crime that misuses personal information for financial gain, including credit card fraud, fraudulent loan applications, and unauthorized access to bank accounts.[1] Identity theft has serious long-lasting consequences, and the recovery of stolen identities is difficult. Victims may suffer severe emotional distress (Harrell and Langton, 2013) and face significant financial repercussions including reduced credit access and greater bankruptcy risk (Hamdi, Kalda, Sovich, 2024). The problem is so severe that specialized insurance products now exist to protect against losses related to identity theft.

Before turning to our main analysis, we examine changes in consumer banking behavior after branch closures. First, we estimate the elasticity of substitution to neighboring bank branches after the closure of a focal branch. Granular footfall data from `pass_by` show that, somewhat surprisingly, only 18% of foot traffic redirects to nearby branches—82% of footfall disappears entirely. This pattern suggests that most consumers cease physical banking after branch closures. It is also consistent with evidence that physical banking is hyper-localized, as minor inconveniences cause customers to adopt online payments and transactions (Choi and Loh, 2024). Second, using microdata on mobile application (app) usage from Global Wireless Solutions, we find that consumers spend 24.2% more time on the

[1] The U.S. Congress passed the Identity Theft and Assumption Deterrence Act in 1998 to criminalize identity theft. In 2024, the U.S. House of Representatives passed another act to provide additional assistance to victims of identity theft.

mobile app of a bank that closes a branch in that area. These preliminary findings validate our premise that branch closures can change consumer behavior.

Using metropolitan statistical area (MSA) level data from the FTC Consumer Sentinel Network database, we test whether branch closures trigger more cases of identity theft in the local area. The main threat to identification is that branch closures may be correlated with unobserved, time-varying local factors that also affect identity theft. For example, banks may close branches in areas experiencing economic distress or increasing digital adoption. However, for reasons unobservable to us, these areas may also carry a higher risk of identity theft. In these scenarios, our estimated effect of branch closures on identity theft would be entangled with changes in local conditions. To address this concern, we use staggered exposures to postmerger consolidation between large regional/national banks as an instrument for branch closures (Nguyen, 2019). Intuitively, an MSA with *both* acquirer and target branches is more likely to experience postmerger branch closures due to duplication in branch service for the consolidated bank.

The key identifying assumption is that MSA-level exposures to mergers between large banks are as good as randomly assigned with respect to local factors. We focus on mergers where both acquirer and target banks have at least U.S. $1 billion in premerger assets because these mergers are typically driven by broader strategic and synergistic objectives. Further tests suggest that these bank mergers are unlikely to be motivated by local factors because (i) the acquirers and targets are large and geographically diversified, and (ii) exposed MSAs account for only a small share of their overall deposits. Moreover, Narayanan, Ratnadiwakara, and Strahan (2025) show that branch closure decisions are unrelated to local customer usage. Finally, we find no statistically significant pretrends in net branch closures between exposed and unexposed MSAs.

Our results from the Callaway and Sant'Anna (2021) difference-in-differences estimator indicate that branch closures lead to more identity theft. On average, MSAs exposed to large bank mergers have 2.79 more branch closures and 455.56 more reports of identity theft. The Wald estimate implies that each branch closure leads to an increase of +163.28 (= 455.56/2.79) identity theft reports. This estimate of privacy costs is economically meaningful—a one standard deviation shock to branch closures corresponds to an increase of 2,318 identity theft cases, representing 1.6 times of its unconditional sample mean. In 2022, our imputed financial losses from identity theft due to branch closures stand at U.S. $1.4 billion, or 21.7% of all reported losses to consumer fraud in that year.

For our results to inform policy or be representative of the average MSA, we must assess whether the local average treatment effect (LATE) generalizes to the population. We use the Marbach and Hangartner (2020) framework to profile the subpopulation of "compliers", which are MSAs that (do not) encounter branch closures due to the presence (absence) of merger exposures.[2] Because the LATE is an estimate of the treatment effect only for complier MSAs, its generalizability to other MSAs is ex ante unclear. Reassuringly, we find that complier MSAs resemble the average MSA in many measurable dimensions, including household income, education attainment, internet penetration, and ownership of computing devices. Thus, our diagnostics suggest that our LATE has external validity and is unlikely to be driven by pre-existing local differences in demographics and technological adoption rates.

To establish a more direct link between branch closures and the focal bank's customers, we turn to the Consumer Complaint Database (CCD) administered by the Consumer Financial Protection Bureau (CFPB). Although the CCD covers only the largest banks, a key advantage lies in its highly granular data, which record individual geotagged consumer complaints against specific banks. This granularity allows us to precisely trace how recent branch closures of a particular bank affect its own local customers. Following a branch closure, we find a significant increase in identity theft complaints from local customers of that particular bank, relative to other consumers and banks in that county.

We expect that the transition to digital platforms and services can affect consumers in unequal ways because technological shocks impact bank customers differently (Fuster, Goldsmith-Pinkham, Ramadorai, and Walther, 2022; Jiang, Yu, and Zhang, 2025). Although some sophisticated consumers adapt seamlessly, vulnerable communities may face serious challenges, particularly from the loss of physical touchpoints they have often relied on. First, we test whether consumers with limited digital capabilities are more vulnerable to the adverse effects of forced digitalization. We find that branch closures lead to more identity theft cases when (i) consumers are more reliant on bank branches, (ii) banks are less digitally focused to start with, and (iii) the local area has lower internet penetration rates. Next, we direct our attention to a community that is particularly vulnerable to identity theft —U.S. military veterans. The American Association of Retired Persons (AARP) finds that veterans are more likely than civilian Americans to fall

[2] Ultimately, the merged bank must select which branches to close, but this layer of selection does not invalidate our identification strategy, which only requires that MSA exposures to bank mergers are as good as randomly assigned. However, it could affect the interpretation of our LATE.

prey to identity theft and suffer greater financial losses when victimized. To generate plausibly exogenous variation in veteran populations, we exploit MSA-level exposures to the Vietnam War draft lotteries (Angrist and Chen, 2011). We find causal evidence that the effect of branch closures on identity theft is significantly stronger in areas with more military veterans.

If branch closures increase identity theft through our proposed pathway of digital engagement, we expect to observe shifts in online and offline consumption patterns. Consistent with this prediction, we find that an average consumer exposed to a branch closure spends 38.4 more hours per month on all mobile apps, excluding banking ones.[3] Furthermore, using data from Safegraph, we estimate that a branch closure tilts consumer expenditures and transaction volumes toward online channels by 54.1% and 60.5%, respectively. These findings indicate that branch closures lead consumers to increase overall digital engagement, which structurally exposes them to identity theft risks.

Identity theft poses a significant challenge to law enforcement because adversaries often operate outside of legal jurisdictions and employ an ever-evolving suite of tactics. To shed light on their otherwise opaque activities, we examine two common adversarial tactics. Phone calls are often the first attack vector used to target potential victims. Adversaries can impersonate trusted institutions on these calls to extract PII such as Social Security numbers and bank account details. To compound the problem, advances in telecommunications technology have enabled the rise of "robocalls", which are programmatically automated calls that can reach large numbers of consumers at low cost. Using the FTC Do-Not-Call (DNC) Reported Calls database, we estimate that a one-standard deviation increase in branch closures leads to a 6.75% increase over the unconditional probability of unwanted calls. These calls also impose additional social costs by wasting time and disrupting the provision of legitimate services.

Another adversarial tactic is phishing attacks. According to the FTC, a significant portion of identity theft cases stem from phishing attacks, where adversaries deceive consumers into providing personal information through deceptive emails or websites (Federal Trade Commission, 2024a). To identify phishing attacks, we exploit the idea that adversaries often clone the legitimate website as a template for malicious use. Using the search engine Shodan, we sweep the internet for suspicious websites that share features of their legitimate counterparts but lack

[3] This implies an increment of over 1.2 hours per consumer per day. As a benchmark, Americans spend an average of 4.65 hours per day on their mobile phones (Statista, 2025).

digital security certificates. Using a two-stage estimation approach, we find that branch closures in an MSA can increase the exposure of its residents to identity theft through phishing attacks.

A limitation of our study is that we do not evaluate the *overall* welfare effects of digitalization. Identity theft is only one of many facets of privacy costs, and we do not address any benefits of digitalization. Moreover, we have not considered the nondigital counterfactual. Cash transactions create security and logistic costs, and the maintenance of bank branches requires significant resources. A welfare analysis that balances these costs against the convenience, cost savings, and efficiency brought about by the digital economy is beyond the scope of our study. Nevertheless, our findings suggest that bank branches can serve as a vital social good in the digital age, offering security and consumer protection.

Our study contributes to the literature on the data economy (Farboodi and Veldkamp, 2023), particularly the economics of data privacy (Goldfarb and Tucker, 2012; Acquisti, Taylor, Wagman, 2016; Tirole, 2023; Bian, Ma, Tang, 2023). Users bear privacy costs from digital businesses that over-collect data and under-invest in consumer data protection (Cong, Xie, Zhang, 2021; Fainmesser, Galeotti, Momot, 2023; Chen, Huang, Ouyang, and Xiong, 2025). However, large-scale evidence on privacy costs remains elusive (Johnson, 2022).[4] Ramadorai, Uettwiller, and Walther (2025) examine the privacy policies of individual US firms. Bian, Pagel, Raval, and Tang (2024) assess privacy costs by analyzing restrictions on personal data collection from the Apple App Tracking Policy. Our contribution is upstream of data protection policies as we show that digital engagement is a structural source of exposure to identity theft risks.

We also add to the growing evidence that digitalization and technological disruption affect bank consumers unequally. Jiang, Yu, and Zhang (2025) show that bank digitalization tends to unbank elderly consumers. Koont (2024) finds that the surplus from the digital banking revolution is mostly captured by the wealthier consumers. Fuster, Goldsmith-Pinkham, Ramadorai, and Walther (2022) show that digital financial technologies may have different effects on racial minorities. We examine the privacy costs of digitalization and find that bank branch closures disproportionately harm U.S. military veterans and communities with poorer digital-savviness.

Finally, our findings suggest that bank branches provide social benefits to the

[4] In the context of online lending, Tang (2019) quantifies the monetary value required for borrowers to share their personal data.

local community. For banks, bank branches offer market presence, funding stability, and information sharing (Drechsler, Savov, Schnabl, 2017; Benmelech, Yang, Zator, 2023; Amberg and Becker, 2024; Keil and Ongena, 2024; Qi, De Haas, Ongena, Straetmans, et al., 2024). For local residents, physical branches increase income, employment, entrepreneurship, financial inclusion, and improve health outcomes.[5] Using a pre-2000 sample, Garmaise and Moskowitz (2006) find that branch closures are accompanied by more property crime. We study the later era of bank digitalization and show that bank branches can provide a social benefit by mitigating the risk of identity theft.

2 Identity theft in the digital economy

This section motivates our empirical analysis by developing the hypothesis that bank branch closures increase the risk of identity theft. We begin by defining identity theft and documenting its prevalence. Next, we describe how branch closures push economic activity toward digital platforms. Finally, we discuss how this shift exposes consumers to greater risks of identity theft.

2.1 What is identity theft?

Identity theft is any illegal activity involving the unlawful acquisition of an individual's personal information, typically for financial gain. The Identity Theft and Assumption Deterrence Act of 1998 established identity theft as a distinct federal crime. Under 18 U.S.C. §1028(a)(7), identity theft is defined as any act of

> "knowingly transfer[ring] or us[ing], without lawful authority, a means of identification of another person with the intent to commit, or to aid or abet, any unlawful activity [...]"

Identity theft is widespread in the United States. According to FTC data, there were more than 1.1 million reports of identity theft in 2023 alone. The top three types, credit card fraud (33%), loan or lease fraud (13%), and bank account fraud (11%), comprise over half of all cases. FTC data also highlights that this crime affects a broad demographic. Individuals aged 30 to 49 are the most frequently targeted group, accounting for 470,663 reports or 43% of all identity theft

[5] See, for example, Jayaratne and Strahan (1996), Célerier and Matray (2019), Nguyen (2019), Martín-Oliver, Toldrà-Simats, Vicente (2020), Bonfim, Nogueira, Ongena (2020), Ji, Teng, Townsend (2023), Sakong and Zentefis (2024), Cramer (2024), and Fonseca and Matray (2024).

cases in 2023. Identity theft cases steadily decline beyond these prime years of workforce participation. We tabulate the breakdowns of identity theft reports by types and age groups in the Internet Appendix.

Identity thieves use both online and offline strategies to acquire personal data. Common online tactics include impersonation scams, social media data mining, cyberattacks exploiting platform vulnerabilities, and intercepting data through unsecured public Wi-Fi networks. Offline methods include physical theft such as retrieving sensitive information from discarded documents and installing skimming devices at ATMs and point-of-sale terminals.

2.2 Branch closures push consumers to increase digital engagement

Consumers affected by branch closures must learn to adopt digital banking and payment tools. Once adopted, these tools can reduce the marginal costs of using other digital services. A 2021 report by fintech firm Plaid suggests that such adoption boosts users' financial confidence with technology, encouraging broader use of online services. Digital payments are also increasingly embedded in everyday platforms, such as e-commerce, food delivery, and subscription services. Thus, branch closures can expand the use of digital tools, serving as a gateway to deeper, habitual engagement with the digital economy.

The two-sided nature of digital platforms accelerates the shift of economic activity from the physical realm to the digital economy. As more consumers engage with digital tools for payments, shopping, and services, businesses have stronger incentives to expand their digital presence.[6] In turn, the wider availability of digital services reinforces consumer reliance on these platforms. Consumers and businesses reflect this mutually reinforcing dynamic in their behavior and strategy. The Federal Reserve's Diary of Consumer Payment Choice shows that the share of remote purchases more than doubled between 2016 and 2023 (Cubides and O'Brien, 2023). Major retailers such as Nike and restaurants such as Domino's Pizza have also expanded their digital platforms to meet consumer preferences.

[6] Branch closures may also force local businesses to travel further for cash management services, increasing logistical complexity and security risks. Debt covenants and insurance contracts often prohibit businesses from storing cash in stores. These added burdens make accepting and handling physical cash less viable, particularly for small businesses. In response, businesses affected by branch closures may have greater incentives to adopt digital payment systems and even shift their operations online, where cash handling is no longer required.

2.3 Branch closures, digital engagement, and identity theft

Bank branch closures push consumers toward greater digital engagement, increasing their exposure to privacy risks through two pathways. The first pathway concerns the background risk of exposure inherent in digitalization. As consumers engage in more digital activities, their personal identifying information (PII) is transmitted and stored across a fragmented and less secure network of retailers, mobile apps, and service providers. Thus, their expanding digital footprints structurally increase the odds that their PII is leaked, stolen, or abused. Third parties that process digital transactions may have inadequate data infrastructure and poor security practices. For example, the Federal Trade Commission has prosecuted data brokers for selling PII to scammers. Even sophisticated consumers can be affected because these breaches or data leaks often occur outside of one's control. A 2019 Pew Research survey reflects these concerns, finding that over 80% of Americans feel they have little or no control over their personal data collected online.

The second pathway reflects the transition risks that emerge as branch closures remove a crucial alternative for some vulnerable consumers who prefer in-person services. Bank branches provide face-to-face verification, secure document storage, controlled access, and continuous surveillance to deter and detect fraud in real time. Human interaction also helps prevent fraud, as bank staff are trained to spot suspicious behavior and advise customers on security best practices. With the loss of these physical touchpoints, some consumers may become more susceptible to adversarial tactics such as phishing attacks using phone calls, emails, SIM card swaps, and even generative artificial intelligence (AI) tools. These tactics can be especially harmful during the transition period, when some consumers are less familiar with security practices or warning signs of fraud. Despite the widespread rollout of digital financial services, a 2016 survey by the National Telecommunications and Information Administration finds that 45% of U.S. households avoid online financial transactions due to privacy and security concerns.

As branch closures drive a digital shift in economic activity and eliminate physical touchpoints, we hypothesize that they lead to an overall increase in identity theft. These risks are unlikely to be evenly distributed across the population. The transition may be relatively seamless for tech-savvy consumers, but it poses a significant behavioral adjustment for many others, particularly those with limited digital literacy or a long-standing reliance on offline transactions.

3 Data and validation

We describe the primary datasets used in our main empirical analysis and validate that branch closures change the banking patterns of consumers.

3.1 Data and descriptive statistics

We obtain identity theft reports (*ID theft reports*) from the Consumer Sentinel Network (CSN), which is maintained by the Federal Trade Commission (FTC).[7] The CSN database is a collection of consumer fraud reports, a subset of which includes identity theft. Identity theft is categorized into the following seven types: credit cards, loans or leases, bank accounts, government documents or benefits, employment or tax, phone or utilities, and others. We provide a detailed summary of CSN identity theft reports by types in the Internet Appendix. Our sample covers 2009 to 2022, providing a panel with variation at the Metropolitan Statistical Area (MSA) and year level.

We collect bank branch closures from the Federal Deposit Insurance Corporation's (FDIC) Summary of Deposits (SOD) files, which provides annual branch-level data for all U.S. depository institutions. We compute *net branch closures* at the MSA-year level by subtracting the current year's branch count from the previous year's. The average MSA in our sample experiences 2.9 net branch closures per year with a standard deviation of 14.2, pointing to significant heterogeneity in branch activity across regions.

Our demographic variables at the MSA-year level are sourced from the U.S. Census Bureau's American Community Survey (ACS) database. These variables include population size, proportion of people aged 60+, unemployment rate, median household income, gender and racial compositions, and educational attainment (over high school).

- Figure 1 here -

[7] The CSN compiles complaints from a wide range of data contributors including law enforcement agencies (e.g., F.B.I. and U.S. Postal Inspection Service), state attorneys general and consumer protection offices (e.g., New York State Attorney General and Ohio Attorney General), federal government agencies (e.g., U.S. Department of Justice and U.S. Social Security Administration), consumer advocacy organizations (e.g., AARP Fraud Watch Network and National Consumers League), financial institutions (e.g., Mastercard International and JPMorgan Chase), and technology and telecommunications companies (e.g., Apple and Verizon Wireless). The Internet Appendix contains a full list of these data contributors.

Figure 1 shows consumer fraud complaints from 2009 to 2022, with the subcategory of identity theft presented in blue. At the start of our sample in 2009, consumer fraud complaints total 1.33 million reports and increases steadily over the subsequent decade. By the end of our sample in 2022, the overall number of fraud reports exceeded 5.1 million, with identity theft accounting for 21.4% of all incidents. Notably, consumer fraud and identity theft reports surged after the COVID-19 pandemic.

In terms of financial losses, Figure 1 shows a steady upward trend in consumer fraud over the period. In 2009, total losses were approximately $1.38 billion, with relatively minor fluctuations in the following years. However, losses spike in recent years, reaching $6.47 billion in 2022. This sharp rise suggests an escalation in the financial impact of consumer fraud, which may be driven by more sophisticated fraud schemes and the increasing stakes of digital fraud, particularly identity theft.

- Figure 2 here -

Figure 2 shows the geographic distribution of identity theft reports across MSAs in 2022. Larger circles indicate a greater number of reports. MSAs that are hit hard by identity theft are typically found in urban and coastal areas, like New York City and Los Angeles. These affected regions usually have higher population densities, better access to digital services, and rely heavily on online transactions. They also tend to have higher income levels, making them attractive to adversaries. Given the substantial variation across regions, it is crucial to control for these demographic traits in our analysis to estimate the relation between bank branch closures and identity theft.

- Table 1 here -

Table 1 summarizes the variables and pairwise correlations. Panel A shows that MSAs average 1,420 identity theft reports annually with a standard deviation of 5,320. The 90^{th} percentile of MSAs reports 2,470 reports, highlighting concentration in specific areas. MSAs experience an average of 2.9 net branch closures per year with a substantial standard deviation of 14.2. Decomposing this statistic, the average MSA has 3.9 branch closures and 1.0 branch opening. The average MSA has 18.4% of its population over 60, 50.1% male, 6.6% unemployed, 76.6% White, and 63.4% with at least a high school education. Average household income is $56,100 with significant dispersion (σ = $13,000). MSA population size varies widely, averaging 730,000 with a large standard deviation of 1,580,000.

Panel B of Table 1 shows that identity theft correlates positively with net branch closures (57%) and strongly with population size (76%), indicating a link between urbanization and identity theft. A 28% correlation with household income suggests higher-income areas may be more prone to fraud. In contrast, identity theft shows a negative association with the proportions of white residents and individuals who finished high school, suggesting possible distributional effects.

3.2 Validation: Consumer response to branch closures

Before presenting our main results, we first validate the premise that consumers affected by branch closures change their banking behavior through two tests. First, we investigate whether these consumers substitute visits to nearby branches. Second, we examine whether consumers use more digital banking services after branch closures in the local area.

3.2.1 Do affected consumers switch to nearby branches?

In principle, consumers who are affected by branch closures could travel to another branch, hence avoiding the need to engage more with the digital economy. On the other hand, if physical banking is hyper-localized with high switching costs (Choi and Loh, 2024), consumers may not readily switch to other nearby branches.

To estimate the elasticity of branch visits between the closed branch and other nearby branches, we collect weekly branch-level footfall data from the `pass_by` database. Using the Google Geocoding API, we map the street address, zipcode, city, and state of every `pass_by` branch to its latitude and longitude coordinates to precisely compute the distances between `pass_by` branches and closed branches.[8] Then, for every `pass_by` branch, we merge in the total *net branch closures* that occur (i) within the past 180 days and (ii) within a 1-mile distance band. Finally, for branch i in week t and for branch closures/openings i' within the $(\omega, \omega+1)$ mile distance band, we estimate equation (1):

$$\begin{gathered} \textit{bank branch visits}_{i,t} = \alpha + \beta \cdot \sum_{i'} \sum_{d=t-180}^{t-1} \textit{net branch closures}_{i',d} + \mathbf{X}_{z,t}^{\top} \boldsymbol{\lambda} + \varepsilon_{z,t}, \\ \text{s.t.} \quad \omega \text{ mi} \le \text{distance}(i, i') \le (\omega+1) \text{ mi}. \end{gathered} \tag{1}$$

[8] We focus on savings banks by filtering for "stores" tagged with the NAICS code for savings institutions (522120). We obtain the latitude and longitude coordinates of the closed branches from the FDIC Summary of Deposits dataset.

$\mathbf{X}_{i,t}$ is a vector of control variables at the zipcode-year level, and $\boldsymbol{\lambda}$ is a vector representing their corresponding coefficients.

- Figure 3 here -

Figure 3 shows that most consumers affected by branch closures do not readily switch to other nearby branches. A branch closure is associated with just 14.77 (t = 2.61) more visits per week to another branch within 0–1 miles. To put this estimate into context, the average branch closure/opening is surrounded by 8.14 other branches within 1 mile, and the average bank branch receives 665.89 visits per week. Somewhat surprisingly, only 18% (= 14.77 × 8.14/665.89) of the foot traffic from a closed branch redirects to nearby branches within a distance band of 0–1 mile. We find a weaker and statistically insignificant foot traffic redirection from closed branches to more distant locations. For example, the implied elasticity of branch visits drops sharply from 18% to 5.1% between the 0–1 mile and 1–2 mile distance bands.

Overall, our findings support the view that physical banking patterns are hyper-localized. Consumers inconvenienced by branch closures largely choose not to switch to other branches even if the alternatives are in close proximity. This pattern could reflect the high costs of switching banks or the shift toward digital banking methods away from branch-based services.

3.2.2 Consumers spend more time on banking mobile apps

Next, we test whether customers spend more time on the mobile app of a bank after it closes branches in the local area. We obtain consumer-level mobile app consumption from the Global Wireless Solutions (GWS) Magnify database. GWS curates a panel of opted-in Android smartphone users that is demographically representative of the United States population. Within the panel, GWS continuously tracks every consumer's mobile activity 24 hours a day, seven days a week. When a consumer uses her smartphone, we can observe among many statistics the (i) name of app used, (ii) time and duration of app usage, and (iii) her current latitude and longitude coordinates. Crucially, we can infer the zipcode of a consumer's primary residence by her most frequented location between 9.00 PM and 6.00 AM in her local time zone. The GWS Magnify database covers 194,530 consumers between 2019 and 2022.[9] We aggregate the bank mobile app usage data

[9] We hand-match banks to their mobile apps on the Google Play Store using the following procedure. First, we collect a bank's website URL from the FDIC BankFind Suite. Next, we carefully search the website for the URL to the bank's mobile app. This URL contains a unique identifier of

to the bank-county-month level.

The granularity of the GWS Magnify database allows us to examine the relation between bank mobile app usage patterns and branch closures. For all branch closures of bank b within the past 180 days in county c, we estimate in equation (2) whether there is a greater usage of bank b's mobile app among consumers in that county in a given month m:

$$\begin{aligned}\log(\textit{bank mobile app usage}_{c,b,t}) = \alpha + \beta \cdot \sum_{t=m-180\text{ days}}^{m-1\text{ day}} \textit{net branch closures}_{c,b,t} \\ + \mathbf{X}_{c,m}^{\top}\boldsymbol{\lambda} + \varepsilon_{c,m}.\end{aligned} \tag{2}$$

The vectors $\mathbf{X}_{c,m}$ and $\boldsymbol{\lambda}$ represent vectors of county-level control variables and their corresponding coefficients, respectively. We saturate the model with state × month and bank × month fixed effects, and cluster standard errors at the state, bank, and month levels.

- Figure 4 here -

Figure 4 presents the binned scatterplot corresponding to equation (2).[10] On average, consumers spend 24.2% ($t = 8.37$) more time on the mobile app of a bank that closes a branch in the local area. Our estimates suggest that branch closures push consumers to use significantly more digital banking. Although our results are grounded in granular mobile activity data, two caveats remain. First, our analysis necessarily focuses only on surviving banks. Second, we observe mobile app usage only among Android users, so our results may not generalize to behavioral shifts among iOS users. Nevertheless, we find evidence that local branch closures spur greater usage of digital banking, which can facilitate the shift of economic activities to the digital economy.

the mobile app on the Google Play Store. We manage to match 2,493 (out of 3,823) banks to their respective mobile apps. Most of the shortfall arises from banks that have ceased operations as of June 2025. Such banks typically delist their apps from the Google Play Store, but we can only observe apps that are currently hosted on it. A small number of banks do not offer mobile banking services.

[10] To construct the binned scatterplot, we first regress separately log(*bank app usage*) and *net branch closures* on the control variables, as well as state × month and bank × month fixed effects. Next, we sort the residualized *net branch closures* into bins and compute the averages of residualized log(*bank app usage*) within these bins. Finally, we plot these residuals and the OLS best-fit line through them. By the Frisch-Waugh-Lovell theorem, the slope of this best-fit line equals β in equation (2).

4 Main analysis

In this section, we estimate the effect of branch closures on identity theft.

4.1 Branch closures and identity theft

We examine whether MSAs with bank branch closures report more cases of identity theft. To test this hypothesis, we estimate specification (3) for MSA i in year t:

$$\textit{ID theft reports}_{i,t} = \alpha + \beta \cdot \textit{net branch closures}_{i,t} + \mathbf{X}_{i,t}^{\top} \boldsymbol{\lambda} + \varepsilon_{i,t}. \tag{3}$$

The vectors $\mathbf{X}_{i,t}$ and $\boldsymbol{\lambda}$ represent vectors of control variables at the MSA-year level and their corresponding coefficients, respectively.

- Table 2 here -

Our estimates in Table 2 show that MSAs with more bank branch closures experience more identity theft. In column 1, we find that an additional bank branch closure is associated with 213.42 ($t = 5.91$) more *ID theft reports*. Our estimated effect is economically significant. A standard deviation shock to *net branch closures* leads to an increase of 3,030.6 (= 14.2 × 213.42) identity theft cases, representing 2.13 times of its unconditional sample mean. In column 2, our findings remain unchanged as we control for various MSA demographic characteristics and saturate our models with MSA fixed effects and year fixed effects.

To sharpen our findings, we verify that our results are driven by MSAs with diminishing branch presence in our sample. To do so, we define Δ *branch* $(-)$ and Δ *branch* $(+)$ as the absolute values of *net branch closures* when it is positive and negative, respectively. Consistent with our priors, column 3 shows that a unit increase in Δ *branch* $(-)$ is associated with 166.92 ($t = 6.33$) more *ID theft reports*. In contrast, the effect of Δ *branch* $(+)$ on identity theft is much smaller and statistically insignificant.

Overall, we find that identity theft is more prevalent in MSAs with bank branch closures. MSA and year fixed effects ensure that our findings cannot be explained by persistent, unobserved local factors that affect both bank branching decisions and identity theft.

4.2 Identification strategy

The main threat to identification is that branch closures may coincide with unobserved, time-varying local factors that also affect identity theft.

For example, banks may close more branches in areas experiencing economic distress, which reduces the profitability of maintaining a local branch presence. At the same time, such distress may be correlated with lower financial sophistication among consumers, which could drive both economic vulnerability and susceptibility to identity theft. Alternatively, branch closures may be more prevalent in areas where digital adoption is on the rise. Consumers in these areas may be wealthier or use more digital services, making them more attractive targets for cybercrime in the first place. In this case, even absent branch closures, we would expect to find more identity theft in these areas.

To make causal inferences, we need variation in the incidence of branch closures that is plausibly exogenous to time-varying local factors. Our main identification strategy in this section is an instrumental variable (IV) approach within a staggered difference-in-differences framework.

4.2.1 Instrument for bank branch closures

Our instrument for branch closures is staggered exposures to postmerger consolidation of large banks following the approach used in Nguyen (2019). Bank mergers create operational redundancy such that a merged institution often closes branches in areas where the two previously separate banking networks overlap. Therefore, MSAs with both acquirer and target branches in situ are more exposed to postmerger branch closures.

The key identifying assumption is that MSA-level exposures to bank mergers are as good as randomly assigned with respect to local factors. To ensure that this assumption holds, we focus on mergers where both acquirer and target banks have at least U.S. $1 billion in premerger assets. These mergers are often driven by broader goals to enter new markets and achieve synergy across business lines. By construction, these are large institutions with extensive branch networks. The median acquirer (target) in our sample holds $8.6 billion ($2.2 billion) in assets and controls 81 (24) branches across 10 (4) MSAs. The median bank in the U.S. in comparison holds only $0.19 billion in assets and controls only 3 branches in a single MSA. Moreover, the median percentage of the acquirer (target) banks' deposits held in exposed MSAs before the mergers is only 1.03% (2.12%). Thus, these mergers are unlikely to be motivated by local factors.

Overall, we identify 227 mergers between large banks, led by 136 unique acquirers between 2009 and 2022. These mergers are geographically diverse, covering 200 unique MSAs across 40 states.

4.2.2 Staggered difference-in-differences estimation

An MSA is first exposed in the year of a merger between large banks with branches in the area. Because the merger exposures are staggered across years and MSAs, we use the doubly robust Callaway and Sant'Anna (2021) difference-in-differences estimator. The average treatment effects on treated (ATTs) are the building blocks of this estimator and are defined as follows,

$$\mathrm{ATT}(g,t) = E[Y_t - Y_{g-1} \mid G_g = 1] - E[Y_t - Y_{g-1} \mid C = 1]. \tag{4}$$

Among MSAs that are first exposed to merger shocks at year g (i.e., $G_g = 1$), the first expectation takes the difference in the outcomes of Y at year t and at year $g-1$. The second expectation computes the same difference among control MSAs that are never exposed to the merger shocks (i.e., $C = 1$). These ATTs form the building blocks of the estimator.

- Figure 5 here -

We illustrate our results using event-time plots in Figure 5. At every event-time $\tau \in [-10\,..\,+10]$ in years, we aggregate the $\mathrm{ATT}(g,t)$'s by averaging over all exposed MSAs that have been observed at that event-time.

$$\begin{aligned} \mathrm{ATT}(\tau) &= \frac{1}{\sum_{g,t} 1(t-g=\tau)} \sum_{g,t} 1(t-g=\tau) \cdot \mathrm{ATT}(g,t) \\ &\text{for } \tau \in [-10\,..\,+10] \end{aligned} \tag{5}$$

The top subfigure shows no statistically significant differences in net branch closures in the (blue) preexposure period.[11] However, we find a mostly increasing trend of net branch closures in the (red) postexposure period. These patterns validate the "first-stage" of our empirical design by showing exposed MSAs have significantly more branch closures after the shocks. In the bottom subfigure, we find a statistically significant and steady rise in identity theft reports in the postexposure period. The lack of pretrend differences supports our identifying assumption

[11] For brevity, we truncate the plot at $\tau = -5$ by summing up the $\mathrm{ATT}(\tau)$, $\forall \tau \leq -5$.

that the assignment of merger exposures is orthogonal to local factors that affect identity theft.

- Panel A of Table 3 here -

We summarize the ATTs of merger exposures on branch closures and identity theft reports in Panel A of Table 3. The preexposure ATT averages the ATT(τ)'s for event-time τ between -10 and -1. The postexposure ATT does likewise but for τ between 0 and +10. Equation (6) defines the overall ATT, which is the simple average of all ATT(g,t)'s for all $t \geq g$.

$$\text{Overall ATT} = \frac{1}{\sum_{g,t} 1(t \geq g)} \sum_{g,t} 1(t \geq g) \cdot \text{ATT}(g,t) \tag{6}$$

Figure 5 shows that treatment effects obtain only after the merger exposures with large and statistically significant postexposure ATTs. Consistent with the assumption of parallel trends, the preexposure ATTs in branch closures and identity theft reports are small and statistically insignificant. The overall ATT on identity theft reports is positive and statistically significant ($+455.56, t = 2.72$). Given the overall ATT on branch closures of +2.79 ($t = 2.83$), a single branch closure corresponds to an increase of +163.28 (= 455.56/2.79) identity theft reports. Reassuringly, this estimate is smaller but relatively close in magnitude compared to its OLS counterpart (column 1 of Table 2) because we generally expect unobserved heterogeneity to induce an upward bias in the OLS estimate.

4.2.3 External validity of LATE

Up to this point, our diagnostic tools support the validity of our identifying assumption that merger exposures are as good as randomly assigned to MSAs. We observe no significant pre-trends and IV estimates that are within reasonable bounds. However, if we wish to inform policy or make predictions about the broader population, we need to evaluate the external validity of our estimates.

With heterogeneous treatment effects, the LATE we identify is the treatment effect on compliers. Drawing from the potential outcomes framework of Angrist, Imbens, Rubin (1996), compliers are MSAs that (do not) encounter branch closures due to the (absence) presence of merger exposures. However, the population also contains (i) "always-takers" where branches would have closed regardless of merger exposures, (ii) "never-takers" where no branches close even with merger exposures, and (iii) "defiers" which do the opposite of the assignment. The concern is that MSAs that subsequently close branches from exposure to mergers might systematically differ from those that do not, in ways that affect identity theft. To examine the generalizability of our LATE to the population beyond compliers, we use the framework of Marbach and Hangartner (2020) to estimate characteristic means within subpopulations. We defer details of this framework to the Internet Appendix. Similarity in characteristics between the compliers and the full sample would increase confidence in the external validity of our LATE.

- Panel B of Table 3 here -

Panel B of Table 3 presents the estimated characteristic means in 2010 for the full sample and subpopulations. Against the secular trend of branch closures in the U.S., always-taker MSAs make up a large proportion of our sample. Always-taker MSAs tend to have smaller, older populations and fewer bank branches, suggesting that they are relatively unattractive banking markets. In contrast, never-taker MSAs that have no branch closures irregardless of merger exposures have significantly larger populations, more branches, and higher household incomes. These MSAs are likely population centers where branch presence may be profitable or strategically important.

Complier MSAs have slightly larger populations but are otherwise representative of the average MSA in our sample. Drawing on Census Bureau data from 2015 (earliest year available), we further verify that complier MSAs resemble the average MSA in terms of technological adoption. They have comparable internet subscription rates and ownership of internet-enabled computing devices (i.e., desktops, laptops, and smartphones). These patterns suggest that our LATE is generalizable and is unlikely to be driven by pre-existing local differences in demographics and technological adoption rates.

4.2.4 Quantifying the effects of branch closures on identity theft

Given that the complier characteristics analysis supports the generalizability of our LATE, we attempt to quantify the effects of branch closures on the prevalence of and losses from identity theft. First, we compute the Callaway and Sant'Anna (2021) calendar ATTs, defined as the average treatment effect for MSAs that are or are already exposed to large bank mergers in that year. For MSAs that were first exposed to merger shocks at year g, the calendar ATT in year t is the average over all ATT(g,t) with $t \geq g$:

$$\begin{gathered}\text{Calendar ATT}(t) = \frac{1}{\sum_g 1(t \geq g)} \sum_g 1(t \geq g) \cdot \text{ATT}(g,t) \\ \text{for } t \in [2011\,..\,2022].\end{gathered} \tag{7}$$

To obtain a measure of the increase in identity theft per instrumented branch closure, we then compute the (calendar) yearly Wald estimate ω_t as the following ratio,

$$\omega(t) = \frac{\text{Calendar ATT}^{\text{ID theft reports}}(t)}{\text{Calendar ATT}^{\text{net branch closures}}(t)}. \tag{8}$$

Next, we impute the yearly total increase in identity theft due to branch closures by multiplying $\omega(t)$ by the actual *net branch closures* in the MSAs and summing them up. This imputation, in spirit, applies the LATE to all MSAs, even though it is defined as the treatment effect for only complier MSAs. We are comfortable with this application because Panel B of Table 3 shows that complier MSAs closely resemble the average MSA across most observable dimensions. Subscripting MSAs by i, we compute the number of identity theft cases attributable to branch closures as follows,

$$\textit{total num. ID theft reports}(t) = \sum_i \omega(t) \times \textit{net branch closures}_{i,t}. \tag{9}$$

Finally, we impute the total annual losses stemming from the increase in identity theft reports by factoring in the average per-report dollar loss (*avg. loss*) and the proportion of reports with reported losses (*% reports with loss*) in the state-year:

$$\begin{aligned}\textit{total losses}(t) = \sum_i & \,\omega(t) \\ & \times \textit{net branch closures}_{i,t} \\ & \times \textit{avg. loss}_{\text{state}(i),t} \\ & \times \textit{\% reports with loss}_{\text{state}(i),t}.\end{aligned} \tag{10}$$

The FTC Consumer Sentinel Network provides these statistics only at the state level, so we match them with their constituent MSAs. Not all fraud reports involve financial losses because some victims may ultimately be able to recover or reverse their damages (e.g., banks waiving fraudulent transactions). The proportions of fraud reports accompanied by dollar loss amounts vary significantly across states and years.

- Figure 6 here -

Figure 6 shows that the imputed financial losses to identity theft have risen steadily over the years, peaking at over U.S.$1.8 billion in 2021 during the COVID-19 pandemic. During the pandemic, consumers experienced an accelerated push onto the digital economy as many in-person economic activities were curtailed. Even as the pandemic began to ease in 2022, the losses remained high at almost U.S.$1.4 billion, or 21.7% of all reported losses to consumer fraud in that year. The steady rise in identity theft cases is not only due to the multi-year decline in physical branch banking—our Wald estimates show that the impact per branch closure is also stronger in later years. This pattern may reflect the increasing risk of identity theft faced by consumers as more economic activities digitalize over time. We tabulate the annual statistics from this analysis in the Internet Appendix.

4.2.5 Linking banks to customer harm

Despite the strength of our IV design, it remains possible that the rise in identity theft is driven by a subgroup of local consumers who are *not* actually customers of banks that close branches. Given the exogenous nature of our IV, such misattribution is unlikely to be systematic across MSAs. Nevertheless, this possibility challenges us to establish a more direct link between branch closures and the focal bank's customers.

To this end, we turn to the Consumer Complaint Database (CCD) administered by the Consumer Financial Protection Bureau (CFPB). Although the CCD covers only the largest banks, its key advantage lies in its highly granular data, which record individual geotagged consumer complaints against specific banks.[12] This granularity allows us to observe both the exposure (i.e., branch closures) and the outcome (i.e., complaints) on a single group of consumers—the bank's own customers. Thus, we can more directly identify the harm from branch closures that falls on the bank's own customers, rather than being misattributed to other local consumers.

[12] Only banks with at least U.S.$ 10 billion in assets are subject to the supervisory authority of the CFPB. Identity theft in financial services is also a particularly serious problem. According to the FTC, there were 614,711 reports of identity theft related to credit card and bank fraud in 2022. These reports account for 44.2% of all identity theft cases, marking a 16.8% year-over-year increase.

We process the CCD data by first hand-matching the names of banks on the CCD to their FDIC identifiers. These banks are inherently large because only depository institutions with over $10 billion in assets are subject to CFPB oversight. Altogether, we match 386,549 complaints to 136 banks that operate branches in 2,533 counties. Although we observe a consumer's zipcode, we cannot know whether she used a specific bank branch in her residential zipcode. Thus, we aggregate at the bank-county-month level, (i) consumer complaints and (ii) branch closures within the past 180 days.

We hypothesize that branch closures are followed by a higher incidence of complaints filed by consumers in the area. To test this hypothesis, we estimate equation (11) for county c, bank b, in month m:

$$1(complaint)_{c,b,m} = \alpha + \beta \cdot \sum_{t=m-180\text{ days}}^{m-1\text{ day}} net\ branch\ closures_{c,b,t} + \mathbf{X}_{c,m}^{\top}\boldsymbol{\lambda} + \varepsilon_{c,m}. \quad (11)$$

The vectors $\mathbf{X}_{c,m}$ and $\boldsymbol{\lambda}$ represent vectors of control variables at the county-year level and their corresponding coefficients, respectively.

- Table 4 here -

The results in Table 4 indicate that branch closures are followed by a higher incidence of consumer complaints. Column 1 shows that a branch closure in the county is associated with a 326 bps (t = 9.78) higher probability of complaints lodged by customers of the bank in the area. As we have both county and month fixed effects, this finding is not explained by unobserved heterogeneity across counties or variation in complaint incidence across time. All complaints have the potential to be relevant to consumer fraud even when they are not explicitly labeled in the CCD as such (Bian, Ma, and Tang, 2023). For example, an "incorrect information on your report" label may indicate that an adversary has applied for a loan using a customer's stolen identity. Likewise, a complaint about a bank "closing your account" could stem from fraudulent activity if the closure was due to suspicious transactions or identity theft.

To sharpen our analysis, we distill consumer narratives from complaints to identify those most likely related to identity theft. We use a zero-shot-learning model, the `bart-large-mnli` model developed by `Facebook`, to classify the nature of every complaint. For every complaint narrative, we apply the hypothesis format "I am reporting a case of {label}" with the following labels: "fraud", "harassment", "inaccuracy", and "identity theft". The model produces a probabilistic score for each label, indicating the likelihood that the complaint matches each category. Then, we classify each complaint by assigning it the label with the highest score.

Having classified complaints that are related to identity theft, we reestimate equation (11). Column 2 shows that a bank branch closure in the county increases the probability of identity-theft-related complaints from local customers by 34 bps ($t = 3.71$). This estimate is economically meaningful—it nearly doubles the baseline probability of 18.8 bps per bank-county-month.

Finally, we saturate the model with bank × county, county × month, and bank × state × month fixed effects. This stringent specification help us address endogeneity concerns related to (i) banks matching to specific geographies and (ii) unobserved heterogeneity in consumers across counties in a given month. With this stringent specification in column 3, we continue to find a significantly positive relation (+17 bps, $t = 2.00$) between branch closures and identity theft complaints.

5 Vulnerable communities

The transition to digital platforms and services can affect consumers in unequal ways. Although sophisticated consumers may be more able to navigate the digital economy safely, some vulnerable communities can face significant challenges. In this section, we examine whether branch closures disproportionately affect consumers who are less digitally savvy. Second, we focus on military veterans, a community that is particularly vulnerable to identity theft.

5.1 Less digitally savvy consumers

To test the idea that branch closures disproportionately affect vulnerable segments of the population, we condition our merger-exposure IV on the ability of bank customers to engage safely with the digital economy. We expect consumers who are less able or ready to make the digital transition to experience a stronger effect of branch closures on identity theft.

We first classify each large bank merger by the acquirer-bank customers' reliance on bank branches. Following Jiang, Yu, and Zhang (2025), we measure *branch reliance* of a bank as the ratio of the number of branches to its total deposits. A high *branch reliance* measure of a bank implies that its customers are likely more accustomed to physical banking and hence less able or prepared to go online. Among the set of large bank mergers, we classify a merger as having high (low) *branch reliance* if its value is above (below) the yearly median. We predict that branch closures has a stronger effect on identity theft when consumers rely more on bank branches in the first place.

- Table 5 here -

The results in Panel A of Table 5 support our prediction. The IV in columns 1 and 2 is an MSA's exposure to large bank mergers in which the acquirer has high *branch reliance*. On average, this exposure leads to 11.67 (t = 2.70) net branch closures and 2,362.36 (t = 2.92) ID theft reports. Thus, the Wald estimate reveals that a branch closure causes 202.43 (= 2,362.36/11.67) more cases of identity theft. We find weaker results in columns 3 and 4 where the IV is an MSA's exposure to mergers characterized by low *branch reliance*. The second-stage estimate is less than one-fifth the size of the estimate in column 2. Moreover, the Wald estimate is noisy and uninformative because the first-stage estimate is statistically insignificant (Jiang, 2017).

In Panel B, we classify bank mergers by the *digital focus* of the acquirers. We measure the *digital focus* of a bank as the ratio of the all-time download volume of its mobile app on the `Google Play Store` to its number of branches. A digitally focused bank likely has fewer number of branches per mobile app user, compared to another bank that has a digital presence but also caters to the physical banking preferences of some customers. So, customers of a bank with lower (higher) *digital focus* are more (less) likely to be pushed online after branch closures. Columns 1 and 2 indicate that a branch closure leads to 145.96 (= 281.70/1.93) more cases of identity theft in MSAs exposed to mergers with low *digital focus*.[13] We find a much weaker effect (86.11) when we switch to bank mergers marked by high *digital focus*.

In Panel C, we use Census Bureau data to separate exposed MSAs by the proportion of people who have internet subscriptions (*consumer tech-savviness*). Columns 1 and 2 show that a branch closure leads to 208.54 (= 398.31/1.91) more cases of identity theft in MSAs with low *consumer tech-savviness*. In contrast, this effect is 15.4% smaller among MSAs with high *consumer tech-savviness*.

Overall, we find that the treatment effect is stronger when (i) consumers are more reliant on bank branches, (ii) banks are less digitally focused to start with, and (iii) consumers are less tech-savvy. Therefore, our findings support the view

[13] We classify mergers by the *digital focus* of their acquirers. An acquirer is labelled as low (high) *digital focus* if its value is below (above) the yearly median.

that branch closures may disproportionately harm some consumers who are ill-equipped or unable to safely engage with the digital economy.

5.2 Military veterans

Identity theft is a serious issue for U.S. military veterans. According to the FTC, veterans reported a record 45,882 bank-related identity theft cases in 2022, marking a 28.8% surge over three years. They are more than twice as likely to be victims of identity theft than civilian Americans. Moreover, the American Association of Retired Persons (AARP) estimates that victimized veterans are around 40% more likely to lose money and often lose more money than civilians.

Veterans are particularly vulnerable to identity theft for four reasons. First, veterans are attractive targets for adversaries due to the significant financial value of veteran benefits such as disability compensation, pensions, and healthcare. Second, veterans are accustomed to sharing their PII. For example, to claim a military discount, veterans must present their DD214, a document that includes a veteran's full name, SSN, and date of birth, and address. Third, the transition to civilian life often involves financial adjustments and increased interactions with financial institutions. Fourth, many veterans are elderly and may be unfamiliar with digital security practices, making them more susceptible to online fraud.

We hypothesize that the effect of branch closures on identity theft is stronger in areas with more veterans. The structural relation of interest is β_1, which is the coefficient on the interaction between the veteran population and branch closures in MSA i.

$$\begin{aligned} \textit{ID theft reports}_{i,t} = \alpha &+ \beta_1 \cdot (\textit{veterans}_{i,t} \times \textit{net branch closures}_{i,t}) \\ &+ \beta_2 \cdot \textit{veterans} + \beta_3 \cdot \textit{net branch closures} \\ &+ \mathbf{X}_{i,t}^{\top} \boldsymbol{\lambda} + \varepsilon_{i,t}. \end{aligned} \tag{12}$$

The vectors $\mathbf{X}_{i,t}$ and $\boldsymbol{\lambda}$ represent vectors of control variables at the MSA-year level and their corresponding coefficients, respectively.

However, there may be a spurious link between branch closures and the geographic concentration of veterans. Veterans facing financial hardship may concentrate in economically vulnerable areas. These areas might face more branch closures and inherently have higher rates of identity theft due to unobserved risky financial behaviors, such as using unsecured credit or seeking aid through insecure channels.

To generate plausibly exogenous variation in veteran populations, we construct MSA-level exposures to the Vietnam War draft lotteries (e.g., Angrist and Chen, 2011). The draft lotteries held in 1970 and 1971 involved males aged 19 years old in those years. Thus, males born in 1951 and 1952 would have been at risk of conscription in the respective lotteries. Our idea is that these lotteries "seeded" more veterans in MSAs that happen to have more eligible-age males in those years. To operationalize the instrument, we define *draft share* as the MSA share of males aged eight and nine in the 1960 U.S. Census. The MSA shares of veterans are highly persistent across time. On average, a percentage point of *draft share* (i.e., constructed from the 1960 Census data) is associated with 0.74 percentage points of MSA shares of veterans in our sample period. Overall, absent any systematic displacement patterns, an MSA with higher *draft share* is likely to have more veterans in our sample period.

The central identifying condition is that *draft share* must be plausibly exogenous to local factors that affect identity theft. This condition is likely satisfied for two reasons. First, the distribution of males aged eight and nine in 1960 across MSAs predates the emergence of modern identity theft and its determinants by more than six decades. Specifically, childbearing decisions in the 1950s—and the largely random gender composition of those births—are unlikely to be connected to recent identity theft trends. Second, the draft lotteries were implemented uniformly in all MSAs. So, their effects on veteran population changes are driven by national policy rather than local economic or social conditions. In a regression of *draft share* on MSA characteristics, no characteristic has a partial R^2 of more than 0.01% except for *population* with a partial R^2 of 35.9%. This exception is unsurprising because *population* is mechanically correlated with *draft share* by construction, but it behooves us to control for it in our tests.

- Table 6 here -

We estimate a two-stage regression in Table 6, using *draft share* to instrument for *veterans* (abbreviated by v).[14] The first-stage regression in column 1 shows that *draft share* positively predicts *veterans* (481.49, $t = 88.24$), with a high R^2 of 84.5%. The F-statistic of 1,123.32 exceeds the Stock and Yogo (2005) critical value at the 10% maximal IV size threshold (16.38). Following Jiang (2017), we also perform a partial R^2 decomposition on the first-stage regression and find that *draft share*

[14] The variable *veterans* is the number of male veterans ($\times 10^5$) aged 55–74 in the MSA. We focus on this age range because males subjected to the draft lotteries would have been between 58 and 70 years old in our sample period (2010–2022). Our Census Bureau data do not have a more detailed breakdown of MSA-level veteran populations by age.

alone explains 65.4% of the variance in *veterans*. Therefore, our diagnostics indicate that *draft share* is a strong instrument for *veterans*, fulfilling the relevance condition.

Next, we interact the instrumented *veterans* (i.e., $\widehat{v}$) with *net branch closures* in the second-stage regression. The estimated coefficient on the interaction term in column 2 ($+3.65, t = 5.49$) is positive and statistically significant. Thus, conditional on a branch closure, MSAs with more veterans experience more identity theft. This estimate is also economically significant. Evaluated at the mean of $\widehat{v}$ ($= 1.8 \times 10^5$), a unit standard deviation increase in *net branch closures* leads to 336.87 more *ID theft reports*. This incremental increase represents 23.7% of the unconditional mean of *ID theft reports*.

In summary, we find that branch closures lead to a greater increase in identity theft in areas with more military veterans. Our evidence highlights the consumer costs of reduced access to bank branches, especially in vulnerable communities.

6 Pathways

We perform additional tests to investigate two pathways that underpin our main findings. First, we examine whether branch closures lead consumers to engage more deeply with the digital economy, increasing the margin over which identity theft can occur. Second, we demonstrate that consumers' exposure to adversarial activities makes the transition to the digital economy perilous.

6.1 Evidence of digital engagement

If branch closures increase identity theft through our digital engagement pathway, we expect consumer's consumption patterns to change as a result. To measure consumption, we track the mobile app usage of consumers and analyze their online and offline expenditures.

6.1.1 Mobile app usage

Mobile app usage is a behavioral measure of how consumers engage with the digital economy. Many economic activities, such as finance, commerce, and entertainment, are now conducted on mobile platforms. Thus, changes in time spent on mobile apps are likely to reflect shifts in consumption from offline to online channels. We turn to the GWS Magnify database introduced in Section 3.2.2. Our structural relation of interest is β—the effect of local branch closures in county c on a consumer's (i) time spent on all mobile apps (excluding banking apps) in month m:

$$\textit{mobile app usage}_{i,c,m} = \alpha + \beta \cdot \textit{net branch closures}_{c,b,t} + \mathbf{X}_{i,c,m}^{\top}\boldsymbol{\lambda} + \varepsilon_{i,c,m}. \tag{13}$$

The vectors $\mathbf{X}_{i,c,m}$ and $\boldsymbol{\lambda}$ represent vectors of county and consumer characteristics and their corresponding coefficients, respectively.

To address endogeneity concerns, we instrument branch closures with consumers' staggered exposures to large bank mergers (Section 4.2). Although our panel of mobile app usage has a relatively short sample period, its high frequency (monthly) and granularity (consumer-level) ensure that we have sufficient statistical power to perform this test. We present the Callaway and Sant'Anna (2021) ATTs of merger exposures in Table 7.

- Table 7 here -

We find that branch closures in the local area lead consumers to increase mobile app usage. The first stage in column 1 shows that consumers exposed to large bank mergers encounter 0.072 ($t = 3.96$) more branch closures per month. This is an economically sizable effect, representing a 16.7% increase over the number of branch closures (0.432) encountered by the average consumer in a month. In the second stage contained in column 2, we find that an exposed consumer increases her mobile app usage by 2.76 ($t = 3.21$) hours per month. Taken together, the Wald estimate implies that one branch closure causes the average consumer to spend an additional 38.42 (= 2.760/0.072) hours on mobile apps per month, or over 1.2 hours more per day. As a benchmark, an industry survey finds that Americans spend an average of 4.65 hours per day on their mobile phones in 2022 (Statista, 2025).

Overall, using microdata on mobile app usage, we find a causal link between branch closures and engagement with the digital economy. Our findings also alleviate the concern that local economic conditions alone drive the observed changes in consumer behavior. Although banks may strategically close branches in areas with greater digital adoption, our IV strategy is unlikely to suffer from this reverse causality bias.

6.1.2 Consumer expenditures

To complement our analysis of mobile app usage, we examine whether branch closures are accompanied by measurable changes in consumers' spending and transaction activity between online and offline channels. Changes in consumption patterns reflect how economic activity reallocates across transaction modes. If branch closures push consumers to engage with the digital economy, we expect to see a greater reliance on online transactions at the expense of offline ones.

To measure shifts in consumption patterns, we use SafeGraph data to construct two aggregate measures at the MSA-month level. The *online spending gap* is the difference in dollar values between online transactions and offline transactions, and the *online transaction gap* is the corresponding difference in transaction volume. Both measures capture the relative tilt of consumer activity toward online channels and abstract away from changes in aggregate local economic activity.

- Table 8 here -

As before, we instrument for branch closures with consumers' staggered exposures to large bank mergers and estimate the effects using the Callaway and Sant'Anna (2021) doubly-robust estimator. Table 8 shows that an MSA exposed to these mergers have 0.311 ($t = 3.79$) additional branch closures per month. These exposures increase the *online spending gap* and *online transaction gap* by \$2.774 million ($t = 2.65$) per month and 2.622 million ($t = 2.58$) transactions per month, respectively. The Wald estimates indicate that a branch closure widens the online spending gap by \$8.92 (= 2.774/0.311) million per month and the online transaction gap by 0.254 (= 0.079/0.311) million transactions per month. These shifts to online channels are economically meaningful, representing increases of 54.1% and 60.5% over the respective means of total expenditures and total transaction volume.[15]

In summary, we find causal evidence that branch closures shift consumption activity toward online channels, reflecting a deeper engagement with the digital economy. Although this change can improve convenience and access, it also expands the surface over which adversaries can target consumers for identity theft.

[15] In a month, the average MSA has total expenditures of \$16.5 million and 0.42 million total transactions.

6.2 Adversarial tactics

Law enforcement faces significant challenges in combating identity theft because adversaries continually evolve their tactics to exploit security vulnerabilities. To shed light on their otherwise opaque operations, we focus on two common adversarial tactics used in identity theft—unwanted calls and phishing attacks.

6.2.1 Unwanted calls

In many identity theft schemes, phone calls serve as the initial attack vector used to target consumers. Adversaries are known to impersonate trusted institutions (e.g., banks and government agencies) on these calls to extract PII such as Social Security numbers and bank account details. Greater digital engagement after branch closures make consumers more vulnerable to these attacks for two reasons. First, consumers' phone contact details are often shared in the delivery of digital services. Security breaches on the provider's end may then expose these details to adversaries. Second, as digital services often involve phone-based interactions, such as SMS verification or customer support, adversaries can readily exploit consumers' trust placed on these systems.

To compound the problem, advances in telecommunications technology have enabled the rise of "robocalls", which are programmatically automated calls that can reach a large number of consumers at a low cost. This trend prompted the FTC to alert consumers to the use of robocalls in phone scams, especially in relation to identity theft (Federal Trade Commission, 2024b). In a notable enforcement action, the FTC fined VoIP (Voice-over-Internet-Protocol) provider XCast Labs U.S. $10 million in January 2024 for enabling billions of illegal robocalls, many of which impersonated government agencies to commit fraud.[16]

We collect data to test whether unwanted calls become more prevalent after branch closures in the local area. From the FTC Do-Not-Call (DNC) Reported Calls database, we first collect 1,015,744 complaints of unwanted calls and robocalls lodged between 2014 and 2022. For brevity, we refer to both categories collectively as "unwanted calls". Next, we aggregate these complaints at the MSA-day level and merge in net branch closures in the MSA over the past 180 days. Because unwanted calls are not always related to identity theft, we also collect MSA-day-level telemarketing call volumes to control for baseline telemarketing activity. To

[16] Under the FTC's Telemarketing Sales Rule, an illegal robocall is one that, among other violations, fails to obtain prior consent from the recipient, disregards the National Do-Not-Call (DNC) registry, or uses false or misleading information to induce sales or payments.

test our hypothesis, we estimate the following model for MSA i on day d:

$$unwanted\ calls_{i,d} = \alpha + \beta \cdot \sum_{t=d-180}^{d-1} net\ branch\ closures_{i,t} + \mathbf{X}_{c,d}^{\top}\boldsymbol{\lambda} + \varepsilon_{c,d}. \tag{14}$$

The vectors $\mathbf{X}_{i,d}$ and $\boldsymbol{\lambda}$ represent vectors of control variables at the MSA-year level and their corresponding coefficients, respectively.

- Table 9 here -

Our results in Table 9 indicate that consumers receive more unwanted calls following branch closures in the local area. The dependent variable is *unwanted calls*, the number of unwanted call or robocall reported by consumers in an MSA on a given day. Column 1 shows that a branch closure in the local area is associated with a 1.509 ($t = 2.28$) more unwanted calls in an MSA-day. This estimate is economically significant: A one standard deviation increase in *net branch closures* is associated with a 5.5 ($= (1.51 \times 6.88)/1.88$) standard deviation increase in *unwanted calls*. To sharpen our findings, we then distinguish between unwanted calls received on wireless devices and wired devices. We posit that consumers share more PII through mobile apps and websites, which are typically accessed on wireless rather than wired devices. Consistent with this prediction, column 2 shows a significantly positive relation between branch closures and unwanted calls to wireless devices. In contrast, the effect based on unwanted calls to wired devices, in column 3, is 56% smaller and statistically insignificant.

Overall, we find that consumers are targeted more after local branch closures, particularly on wireless devices. Our evidence reflects how the transition to the digital economy expands the exposure of PII, acting as a pathway to more identity theft.

6.2.2 Phishing attacks

According to the FTC, a significant portion of identity theft cases stem from phishing attacks, where adversaries deceive consumers into providing PII through deceptive emails or websites (Federal Trade Commission, 2024a). The adversaries then use the information to open bogus accounts or access existing ones. As branch closures push consumers toward digital banking platforms, they may become more exposed to such phishing attacks.

In phishing attacks, adversaries often buy unlicensed developer kits from black markets to clone legitimate websites for malicious use. Thus, many phishing websites have the same exact "favicons" (i.e., abbreviation for favorite icons) as their legitimate counterparts. Favicons are small logos shown on internet browser tabs to provide a recognizable brand identity to consumers.

By tracking the appearances of suspicious websites using the favicon of a bank's website, we can measure the intensity of phishing attacks on its customers. We begin by hand-collecting "favicons" from the websites of the top 100 U.S. banks by total assets in 2022. Next, we represent every favicon as an integer hash using the MMH3 (MurmurHash3) hashing algorithm. For example, the favicon of the Wells Fargo Bank website maps to the hash value "`893468414`".[17] We then query this hash value on the Shodan search engine to obtain the monthly number of web servers that host websites with the same favicon. Shodan indexes devices connected to the internet, including their open ports, running services, and potential vulnerabilities. To exclude legitimate websites from our query, we filter out those with secure-socket-layers (SSL) certificates, which authenticate website identities and enable secure data transmission. Some of the remaining websites might not be related to phishing attacks but are still suspicious from a cybersecurity standpoint. If anything, this measurement noise biases our estimates toward zero.

We create a novel measure of consumers' exposure to phishing attacks at the MSA-bank-month level. For bank b servicing the set S_b of MSAs indexed by i in month m, we construct a MSA-bank-month measure of *phishing exposure* as the product of (i) the month-over-month change in the number of phishing sites impersonating a bank, and (ii) the population weight of the MSA relative to all MSAs serviced by the bank,

$$\textit{phishing exposure}_{i,b,m} := \Delta_{m-1,m}(\textit{num. phish sites})_b \times \frac{\textit{population}_{i\in S_b,m}}{\sum_{i\in S_b}\textit{population}_{i,m}}. \quad (15)$$

Intuitively, phishing attacks on a bank have a greater effect on an MSA with a larger population weight. With more potential victims, consumers in these MSAs are exposed more to any given phishing attack. Thus, *phishing exposure* could vary over time due to the emergence of phishing attacks and changes in the bank's operational exposure to different MSAs.

- Table 10 here -

To investigate the relation between *phishing exposure* and branch closures, we

[17] This favicon can be found at `https://www.wellsfargo.com/favicon.ico`.

estimate the equation (16) at the MSA-bank-month level:

$$\textit{phishing exposure}_{i,b,m} = \alpha + \beta \cdot \textit{net branch closures}_{i,b,m} + \varepsilon_{i,b,m}. \tag{16}$$

Column 1 of Table 10 shows a positive and statistically significant relation between *net branch closures* and *phishing exposure*. Thus, when a bank closes branches in an MSA, its customers in that area are potentially exposed to more phishing attacks.

Next, we examine whether greater exposure to phishing attacks translates to more identity theft. We aggregate *phishing exposure* to the MSA-year level by computing the average of fitted values from equation (16), weighted by the number of branches each bank operates in the MSA:

$$\textit{MSA phishing}_{i,t} := \sum_{\substack{i, \\ \text{year}(m)=t}} \textit{phishing exposure}^{(\text{pred})}_{i,b,m} \times \frac{(\#\ \textit{branches})_{i,b,m}}{(\#\ \textit{branches})_{i,m}}. \tag{17}$$

Column 2 shows that *MSA phishing* has a positive and statistically significant relation with identity theft. A standard deviation increase in *MSA phishing* is associated with 375.40 (= 0.0245 × 15,300) more reports of identity theft, representing 26.4% of its unconditional sample mean. In columns 3 and 4, we repeat our analysis using an out-of-sample variant of *MSA phishing*. Using predictive loadings from equation (16) estimated between 2018–2019, we continue to find that *MSA phishing* is strongly associated with identity theft in 2020–2022.

Overall, our results indicate that branch closures in an MSA could expose local consumers to more phishing attacks. In turn, higher exposures to these attacks are associated with more identity theft reports.

7 Conclusion

The digitalization of financial services has reshaped the conduct of economic activities. We examine the privacy costs borne by consumers in the shift from brick-and-mortar banking to online financial services. Following branch closures, we show that consumers increase their overall engagement with the digital economy and face greater exposure to adversarial attacks. These changes expand the margin over which consumers' PII is exposed, increasing the risks of identity theft. Digitalization often affects communities in unequal ways. We find that vulnerable communities, such as military veterans and the less digitally savvy, are disproportionately affected by the transition to the digital economy.

Our findings underscore the need for caution as societies transition towards a digital economy. This includes promoting financial literacy, providing support to consumers experiencing financial difficulties, and investing in cybersecurity measures to prevent cybercrime. In addition, our results highlight the complexity of navigating the digital landscape and the need for proactive policy responses. Policymakers now spend substantial resources to educate the public and have created cybercrime networks to quickly alert and shut down fraudulent transactions. More research is needed to understand the long-term impacts of digitalization and to develop strategies to mitigate the risks associated with it.

Finally, we caution that we do not evaluate the *overall* welfare effects of digitalization. We have not considered the additional logistical, security, and operational costs in a non-digital counterfactual (e.g., a cash-based economy). These costs must also be balanced against the convenience, cost savings, and efficiency brought about by the digital economy. This exercise is outside the scope of our study. However, our evidence indicates that consumers who are pushed into the digital economy may suffer significant privacy costs in the form of identity theft. Thus, bank branches can remain a vital social good in the digital age, providing security and consumer protection.

References

Acquisti, A., Taylor, C., and Wagman, L. (2016). “The economics of privacy”. Journal of economic Literature 54, 442–492.

Amberg, N. and Becker, B. (2024). “Banking without branches”. Working paper.

Angrist, J. and Chen, S. (2011). “Schooling and the Vietnam-era GI Bill: Evidence from the draft lottery”. American Economic Journal: Applied Economics 3, 96–118.

Angrist, J., Imbens, G., and Rubin, D. (1996). “Identification of causal effects using instrumental variables”. Journal of the American Statistical Association 91, 444–455.

Armantier, O., Doerr, S., Frost, J., Fuster, A., and Shue, K. (2024). “Nothing to hide? Gender and age differences in willingness to share data”. Working paper.

Benmelech, E., Yang, J., and Zator, M. (2023). “Bank branch density and bank runs”. Working paper.

Bian, B., Ma, X., and Tang, H. (2023). “The supply and demand for data privacy: Evidence from mobile apps”. Working paper.

Bian, B., Pagel, M., Raval, D., and Tang, H. (2024). “Consumer surveillance and financial fraud”. Working paper.

Bonfim, D., Nogueira, G., and Ongena, S. (Nov. 2020). ““Sorry, we’re closed” Bank branch closures, loan pricing, and information asymmetries”. Review of Finance 25, 1211–1259.

Callaway, B. and Sant’Anna, P. (2021). “Difference-in-differences with multiple time periods”. Journal of Econometrics 225, 200–230.

Célerier, C. and Matray, A. (2019). “Bank-branch supply, financial inclusion, and wealth accumulation”. The Review of Financial Studies 32, 4767–4809.

Chen, L., Huang, Y., Ouyang, S., and Xiong, W. (2025). “The data privacy paradox and digital demand”. Working paper.

Choi, H.-S. and Loh, R. K. (2024). “Physical frictions and digital banking adoption”. Management Science 70, 6597–6621.

Cong, L. W., Xie, D., and Zhang, L. (2021). “Knowledge accumulation, privacy, and growth in a data economy”. Management Science 67, 6480–6492.

Cowles, C. (2024). “How I fell for an Amazon scam call and handed over $50,000”. New York Magazine. URL: https://www.thecut.com/article/amazon-scam-call-ftc-arrest-warrants.html.

Cramer, K. F. (2024). “Bank presence and health”. Working paper.

Cubides, E. and O’Brien, S. (2023). “2023 Findings from the Diary of Consumer Payment Choice”. Working paper.

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

Fainmesser, I., Galeotti, A., and Momot, R. (2023). “Digital privacy”. Management Science 69, 3157–3173.

Farboodi, M. and Veldkamp, L. (2023). “Data and markets”. Annual Review of Economics 15, 23–40. ISSN: 1941-1391.

Federal Trade Commission (2015). *FTC Charges Data Brokers with Helping Scammer Take More Than* 7*MillionfromConsumers′Accounts*. URL: https://www.ftc.gov/news-events/news/press-releases/2015/08/ftc-charges-data-brokers-helping-scammer-take-more-7-million-consumers-accounts (visited on 08/12/2015).

— (2024a). *Phishing Scams*. Accessed: 2024-07-11. URL: https://www.ftc.gov/news-events/topics/identity-theft/phishing-scams.

— (2024b). *Unwanted Calls*. Accessed: 2024-10-18. URL: https://consumer.ftc.gov/unwanted-calls-emails-and-texts/unwanted-calls.

Fonseca, J. and Matray, A. (2024). “Financial inclusion, economic development, and inequality: Evidence from Brazil”. Journal of Financial Economics 156, 103854.

Forbes Magazine (2025). *16 Billion Apple, Facebook, Google And Other Passwords Leaked*. URL: https://www.forbes.com/sites/daveywinder/2025/06/20/16-billion-apple-facebook-google-passwords-leaked---change-yours-now/ (visited on 06/20/2025).

Fuster, A., Goldsmith-Pinkham, P., Ramadorai, T., and Walther, A. (2022). “Predictably unequal? The effects of machine learning on credit markets”. Journal of Finance 77, 5–47.

Garmaise, M. J. and Moskowitz, T. J. (2006). “Bank mergers and crime: The real and social effects of credit market competition”. Journal of Finance 61, 495–538.

Goldfarb, A. and Tucker, C. (2012). “Shifts in privacy concerns”. American Economic Review 102, 349–353.

Hamdi, N., Kalda, A., and Sovich, D. (2024). “The costs of financial fraud victimization”. Working paper.

Harrell, E. and Langton, L. (2013). *Victims of identity theft, 2012*. US Department of Justice, Office of Justice Programs, Bureau of Justice.

Jayaratne, J. and Strahan, P. (1996). “The finance-growth nexus: Evidence from bank branch deregulation”. Quarterly Journal of Economics, 639–670.

Ji, Y., Teng, S., and Townsend, R. M. (2023). “Dynamic bank expansion: spatial growth, financial access, and inequality”. Journal of Political Economy 131, 2209–2275.

Jiang, E. X., Yu, G. Y., and Zhang, J. (2025). “Bank competition amid digital disruption: Implications for financial inclusion”. Journal of Finance, Forthcoming.

Jiang, W. (2017). “Have instrumental variables brought us closer to the truth”. Review of Corporate Finance Studies 6, 127–140.

Johnson, G. (2022). “Economic research on privacy regulation: Lessons from the GDPR and beyond”. Working Paper.

Keil, J. and Ongena, S. (2024). “The demise of branch banking - Technology, consolidation, bank fragility”.

Koont, N. (2024). “The Digital Banking Revolution: Effects on Competition and Stability”. Working paper.

Marbach, M. and Hangartner, D. (2020). “Profiling compliers and noncompliers for instrumental-variable analysis”. Political Analysis 28, 435–444.

Martín-Oliver, A., Toldrà-Simats, A., and Vicente, S. (2020). “The real effects of bank branch closings and restructurings”. Working paper.

Narayanan, R., Ratnadiwakara, D., and Strahan, P. (2025). “The Decline of Branch Banking”. Working paper.

Nguyen, H.-L. (2019). “Are credit markets still local? Evidence from bank branch closings”. American Economic Journal: Applied Economics 11, 1–32.

Qi, S., De Haas, R., Ongena, S., Straetmans, S., and Vadasz, T. (July 2024). “Move a little closer? Information sharing and the spatial clustering of bank branches”. Review of Finance 28, 1881–1918.

Ramadorai, T., Uettwiller, A., and Walther, A. (2025). “The market for data privacy”. Working paper.

Sakong, J. and Zentefis, A. K. (2024). “Bank branch access: Evidence from geolocation data”. Working paper.

Statista (2025). *Time spent with nonvoice activities on mobile phones every day in the United States from 2019 to 2024*. URL: https://www.statista.com/statistics/1045353/mobile-device-daily-usage-time-in-the-us/ (visited on 06/26/2015).

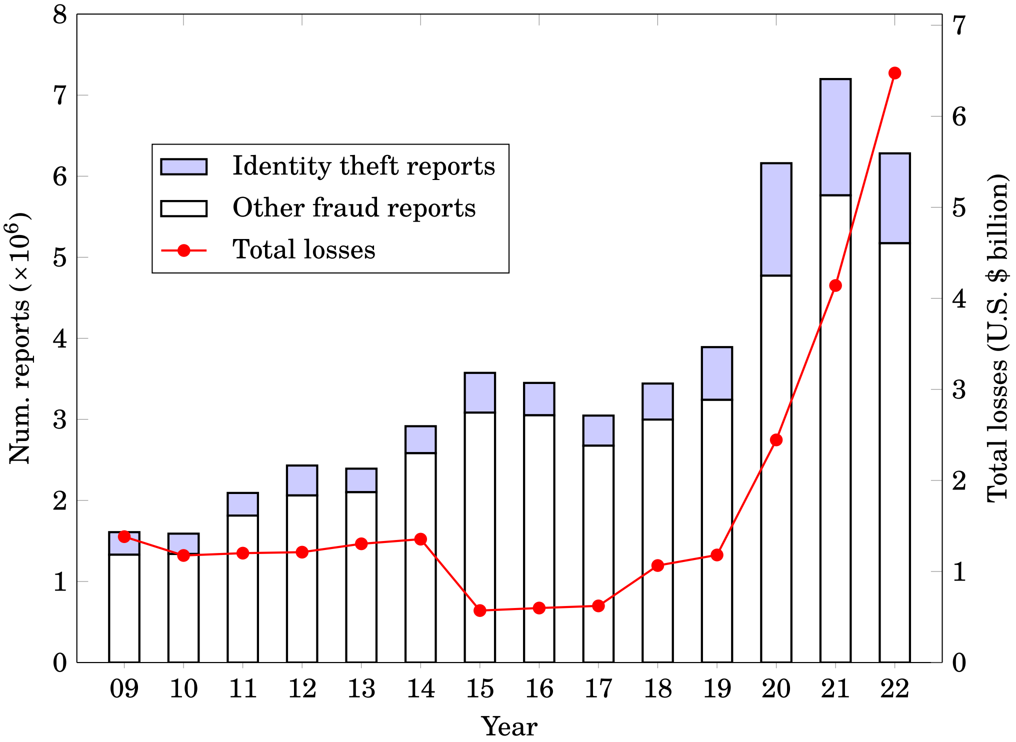
Stock, J. and Yogo, M. (2005). “Testing for Weak Instruments in Linear IV Regression”. *Identification and Inference for Econometric Models*. Ed. by D. Andrews. New York: Cambridge University Press, 80–108.

Tang, H. (2019). “The value of privacy: Evidence from online borrowers”. Available at SSRN 3880119.

Tirole, J. (2023). “Competition and the industrial challenge for the digital age”. Annual Review of Economics 15, 573–605.

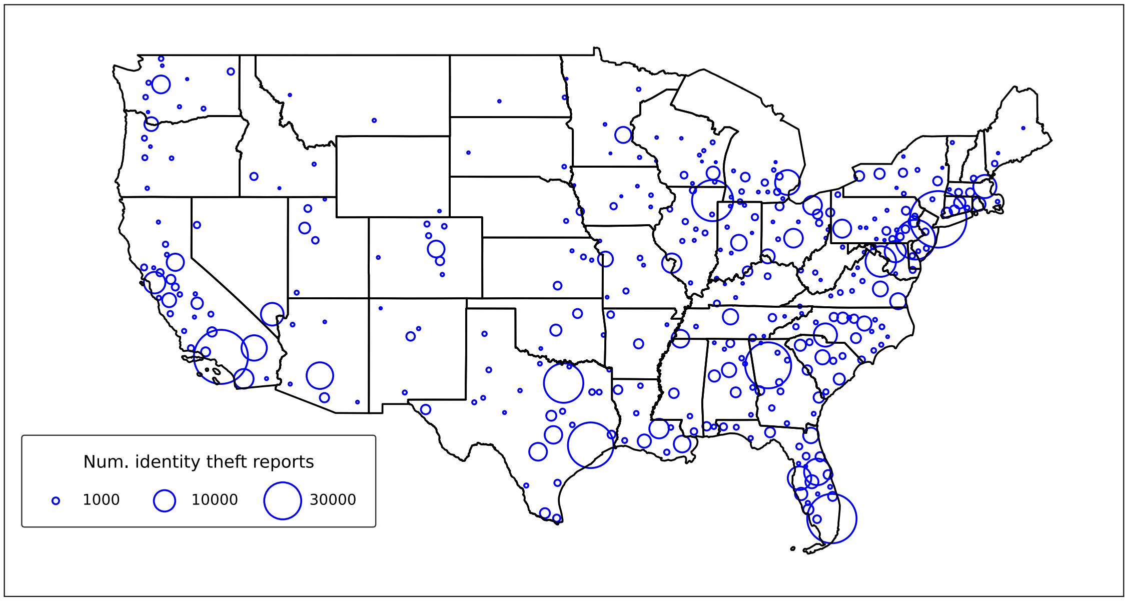
Yin, W., Hay, J., and Roth, D. (2019). “Benchmarking zero-shot text classification: Datasets, evaluation and entailment approach”. arXiv preprint arXiv:1909.00161.

Figure 1. Annual trends of consumer fraud in the U.S.



This figure plot the national annual trends of consumer fraud statistics. The left axis present the number of reports (in millions) of identity theft and all other consumer fraud plotted using stacked barcharts. The right axis reported total losses in U.S. billion dollars plotted using the red line. Data on consumer fraud are sourced from the Consumer Sentinel Network, administered by the U.S. Federal Trade Commission.

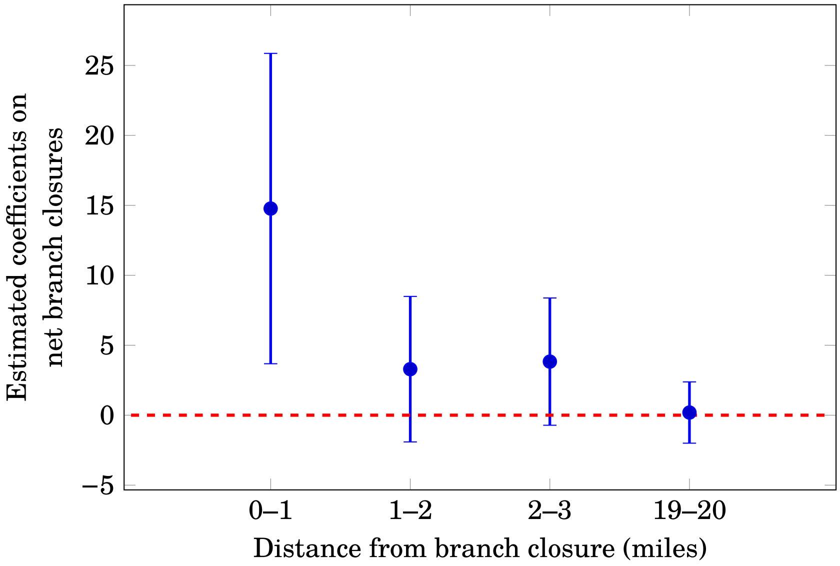
Figure 2. Geography of identity theft reports in the U.S.



This figure plots the numbers of identity theft reports at the MSA level in the year 2022. Larger circles reflect higher numbers of identity theft reports. Data on identity theft volume are sourced from the Consumer Sentinel Network database, administered by the U.S. Federal Trade Commission.

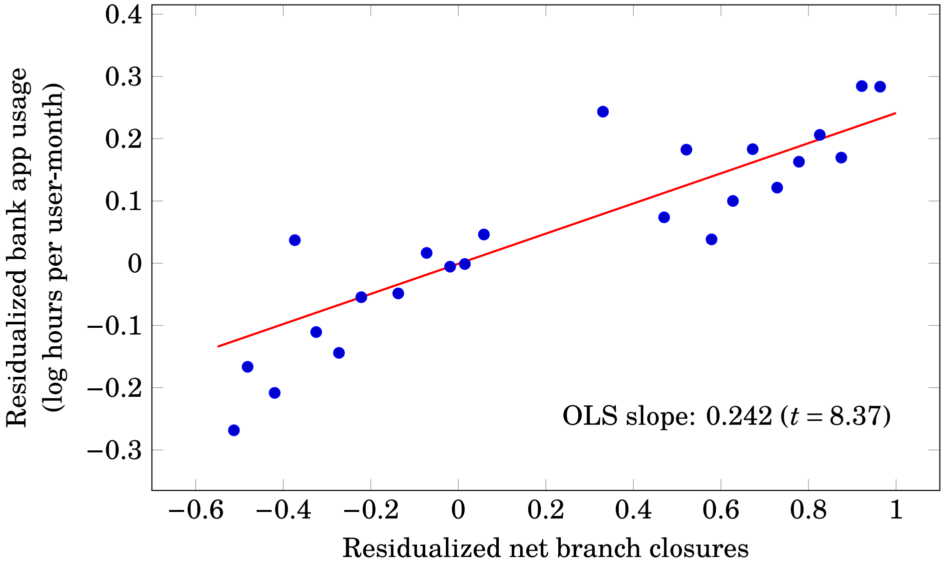
Figure 3. Elasticity of bank branch visits

Dep. variable: Bank branch visits



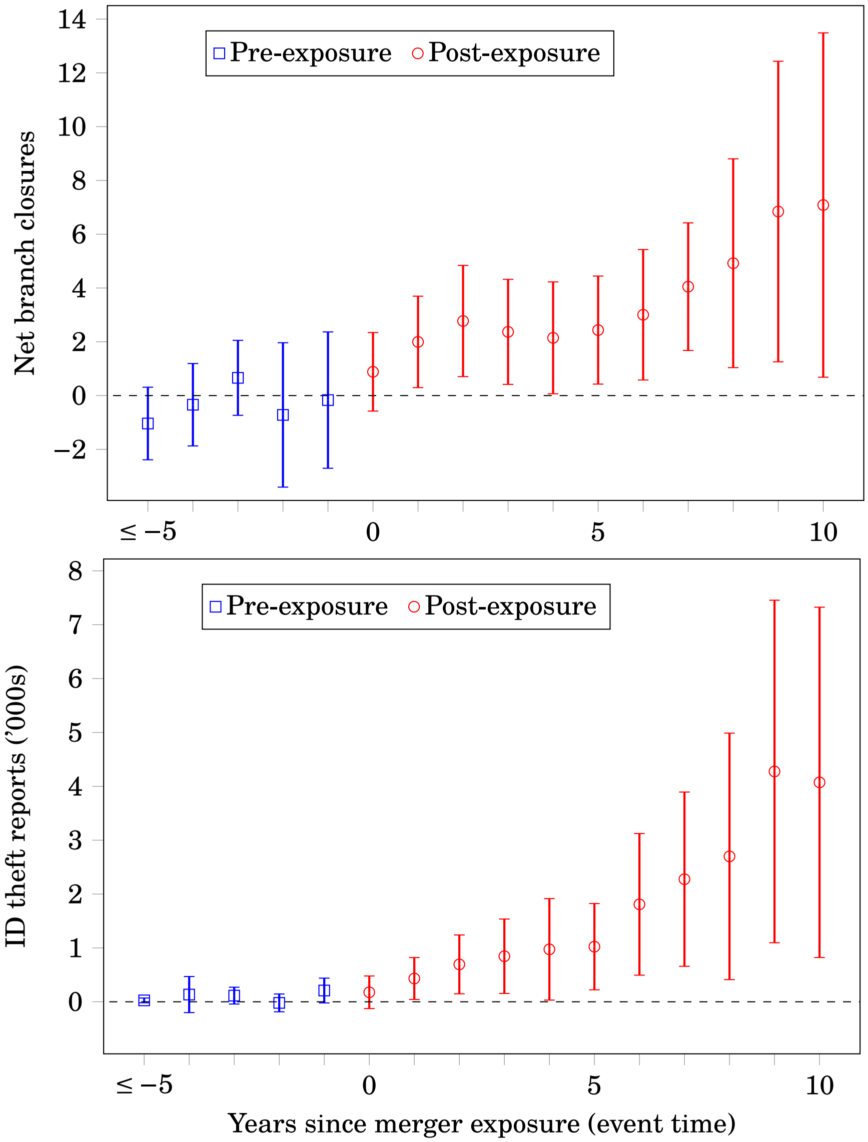
This figure presents coefficient estimates on *net branch closures* from OLS regressions of the form in equation (1). The point estimates are presented in circles and the 95% confidence intervals are represented by end points of the line. The dependent variable is *bank branch visits*, which is defined as the weekly number of visits received by a bank branch, compiled from the `pass_by` database. For every bank branch, we merge in the *net branch closures* of neighboring bank branches that are within a particular distance band. The variable *net branch closures* is the negative change in the number of bank branches over the past 180 days. We calculate distances between bank branches using latitude and longitude coordinates obtained from the FDIC SOD dataset and by geocoding `pass_by` branch addresses with the Google Maps API.

Figure 4. Banking mobile app usage after branch closures



This figure presents a binned scatterplot of *bank app usage* against *net branch closures*. The variable *bank app usage* is the estimated average duration (in hours) spent by consumers in county on the mobile app of a specific bank in a month. The variable *net branch closures* is the negative change in the number of bank branches in the month, matched to consumers' residences at the county level. To construct the binned scatterplot, we first regress separately *bank app usage* and *net branch closures* on county-year control variables, as well as fixed effects at the state × month and bank × month dimensions. Next, we sort the residualized *net branch closures* into bins and compute the averages of residualized *bank app usage* within these bins. Finally, we plot these residuals and the OLS best-fit line. Mobile app usage data of individual Android smartphone users is sourced from an opted-in panel from the Global Wireless Solutions Magnify database.

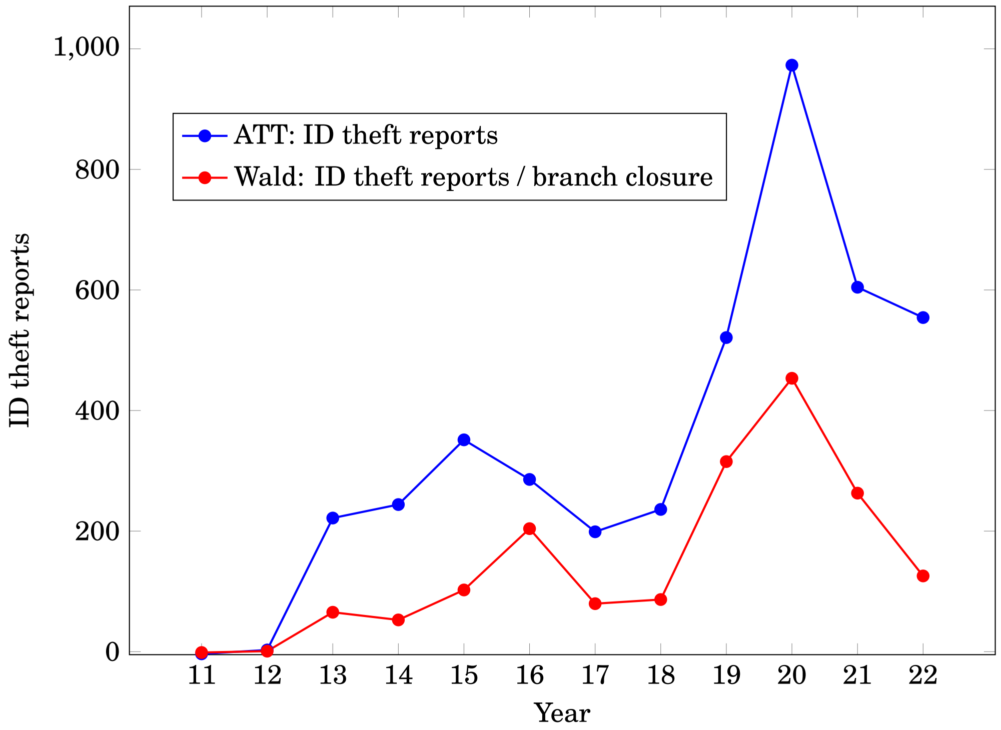
Figure 5. Dynamic effects of merger exposures on identity theft



This figure plots the average treatment effects on treated (ATTs) on *net branch closures* and *ID theft reports*, and their respective 95% confidence intervals by event time (years). The variable *net branch closures* is the negative year-on-year change in bank branches within a MSA. ATTs are estimated from the Callaway and Sant'Anna (2021) doubly-robust estimator. An MSA is first exposed in the year when a merger occurs between large banks that both have branches in the area. The MSA-year number of *ID theft reports* is sourced from the Consumer Sentinel Network database, administered by the U.S. Federal Trade Commission.

Figure 6. Quantifying the effects of branch closures on identity theft

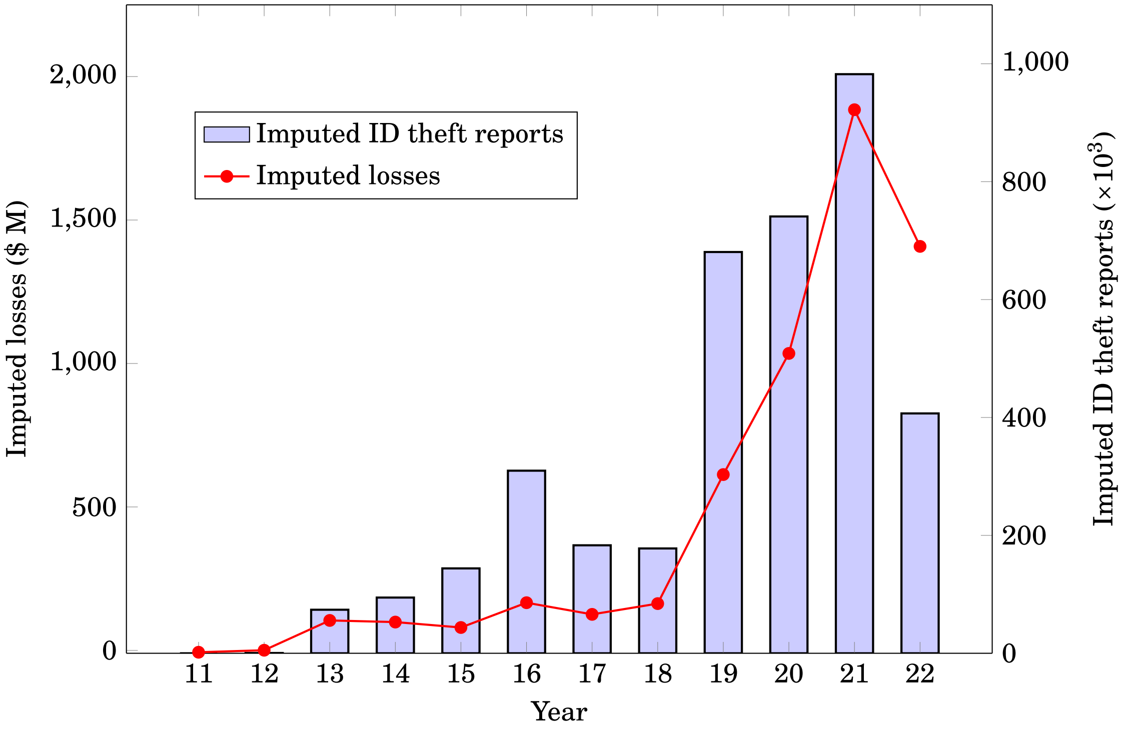
(a) Calendar ATTs and Wald estimates (MSA level)



This subfigure presents the Callaway and Sant'Anna (2021) calendar ATTs for *net branch closures* (i.e., the first stage) and the number of *ID theft reports* (i.e., the second stage). The calendar ATT in a year is the average treatment effect for a MSA that is or is already exposed to large bank mergers in that year. The subfigure also presents the annual Wald estimates, which are the ratios of the second-stage estimates to the first-stage estimates. The Wald estimate represents the causal effect of one instrumented *net branch closure* on the number of *ID theft reports*.

Figure 6. (Continued)

(b) Imputed outcomes by calendar year (whole of U.S.)



This subfigure imputes the number of *ID theft reports* by multiplying the Wald estimates in Figure 6(a) by the actual *net branch closures* in MSA-years and aggregating them to the year level. We also impute the total losses by multiplying the imputed number of ID theft reports in the MSA-year by the average loss per report in the state-year and aggregating them to the year level.

Table 1. Descriptive statistics

Panel A: Summary statistics ($N = 4{,}266$)

	Mean	S.D.	P10	P50	P90
ID theft reports ($\times 10^3$)	1.42	5.32	0.08	0.28	2.47
Net branch closures	2.9	14.2	–1	1	8
Δ Branch (−)	3.9	12.0	0	1	8
Δ Branch (+)	1.0	7.0	0	0	1
Over 60	18.4	7.1	10.5	17.2	26.8
Male	50.1	1.3	48.6	50.1	51.6
Unemployed	6.6	2.9	3.6	6.0	10.5
White	76.6	13.5	60.0	80.0	90.0
High school	63.4	35.4	9.3	85.7	92.0
Household income ($\times 10^3$)	56.1	13.0	41.8	53.8	72.8
Population ($\times 10^5$)	7.3	15.8	1.2	2.7	15.6

Panel B: Pairwise correlations (% pts.)

		A	B	C	D	E	F	G	H
A	ID theft reports								
B	Net branch closures	57							
C	Over 60	4	10						
D	Male	(9)	(11)	(40)					
E	Unemployed	(5)	(10)	(39)	28				
F	White	(28)	(16)	5	7	(16)			
G	High school	(16)	(11)	(54)	32	33	25		
H	Household income	28	22	29	(21)	(50)	(23)	(40)	
I	Population	76	57	(6)	(2)	(3)	(27)	(1)	31

Panel A reports the summary statistics of the variables used in our main analysis. The MSA-year number of *ID theft reports* is compiled by the Consumer Sentinel Network of the U.S. Federal Trade Commission. Branch closures are computed from the Summary of Deposits files of the Federal Deposit Insurance Corporation. MSA-year demographic data are collected from the U.S. Census Bureau. The statistics for the variables *over 60*, *male*, *unemployed*, *white*, and *high school* are reported in percentage points. Panel B reports the Pearson pairwise correlation coefficients between the variables used in the main test. Correlation coefficients are rounded to their nearest integers and expressed in percentage points. Negative values are contained in parentheses.

Table 2. Bank branch closures and identity theft

Dependent variable: ID theft reports

	(1)	(2)	(3)
Net branch closures	213.42	87.01	
	(5.91)	(3.23)	
Δ Branch (−)			166.92
			(6.33)
Δ Branch (+)			4.97
			(0.91)
Over 60		–53.35	–64.33
		(1.00)	(1.28)
Male		–136.07	–120.92
		(3.46)	(3.26)
Unemployed		83.78	74.92
		(1.98)	(1.84)
White		–45.01	–42.55
		(1.81)	(1.81)
High school		90.19	83.27
		(2.88)	(2.73)
log(Household income)		168.47	68.41
		(0.15)	(0.06)
log(Population)		1,150.77	1,328.55
		(0.86)	(1.08)
# Obs.	4,266	4,266	4,266
R^2	0.324	0.690	0.709
MSA FE		✓	✓
Year FE		✓	✓

This table presents estimates from OLS regressions. The dependent variable is the number of *ID theft reports* at the MSA-year level, collected from the Consumer Sentinel Network database of the U.S. Federal Trade Commission. The key independent variable is *net branch closures*, which is the negative year-on-year change in the number of bank branches within an MSA. The variables Δ *branch* (−)/(+) are the absolute values of *net branch closures* when it is positive and negative, respectively. Standard errors are clustered at the MSA level. Absolute values of t-statistics are reported in parentheses.

Table 3. Merger exposures and identity theft

Panel A. Average treatment effects on treated (ATTs)

	(1)	(2)
Outcome:	Net branch closures	ID theft reports
Pre-exposure	–0.76	34.85
	(1.83)	(1.51)
Post-exposure	3.51	608.18
	(3.57)	(2.81)
Overall	2.79	455.56
	(2.83)	(2.72)
Wald estimate	163.28	

This table presents ATTs on *net branch closures* and *ID theft reports* from the Callaway and Sant'Anna (2021) difference-in-differences estimator. The variable *net branch closures* is the negative year-on-year change in the number of bank branches within an MSA. The MSA-year number of *ID theft reports* is collected from the Consumer Sentinel Network database of the U.S. Federal Trade Commission. The *pre* (*post*)-*exposure averages* are the aggregated ATTs before (after) merger exposures. An MSA is first exposed in the year when a merger occurs between large banks that both have branches in the area. Standard errors are clustered at the MSA level. Absolute values of t-statistics are reported in parentheses.

Table 3. (continued)

Panel B. Complier characteristics analysis

Characteristics	Full sample	Compliers	Never-takers	Always-takers
Demographics				
Over 60	11.7	11.2	11.0	12.1
Male	50.8	50.5	50.7	50.9
Unemployed	10.2	11.6	9.7	10.1
White	80.2	83.7	78.0	80.3
High school	86.1	84.0	87.0	86.2
Household income ($\times 10^3$)	45.9	46.6	46.9	45.2
Population ($\times 10^5$)	4.15	4.94	6.05	3.04
Num. branches	120.9	119.8	174.0	95.8
Digital adoption (having:)				
Internet subscription	74.7	76.2	76.0	73.6
Any computing devices	85.7	86.1	86.4	85.3
Desktop/laptop	76.6	77.0	77.4	76.1
Smartphone	72.8	72.6	73.8	72.3
Proportion of sample (% pts.)	100	19	26	55

This table compares the characteristic means of complier MSAs with those of the full sample and non-compliers. Characteristic means of compliers and non-compliers are computed using the methodology outlined in Marbach and Hangartner (2020). The demographic and identity theft characteristics are taken from the year 2010. The digital adoption characteristics are taken from the year 2015 (earliest year available). We report in percentage points the proportions of the sample made up of compliers, never-takers, and always-takers.

Table 4. Bank branch closures and consumer complaints

Dependent variable ($\times 10^2$): 1(Complaint)

	(1)	(2)	(3)
Type of complaint	Any	ID theft	ID theft
Net branch closures	3.26	0.34	0.17
	(9.78)	(3.71)	(2.00)
Over 60	0.19	0.00	
	(5.99)	(0.30)	
Male	–0.06	–0.03	
	(2.54)	(4.13)	
Unemployed	0.26	0.00	
	(4.60)	(0.48)	
White	–0.16	–0.01	
	(9.75)	(8.09)	
High school	0.07	0.01	
	(2.80)	(3.83)	
log(Household income)	7.36	–0.18	
	(8.12)	(2.33)	
log(Population)	4.90	0.12	
	(28.07)	(6.47)	
# Obs.	1,374,000	1,374,000	1,374,000
R^2	0.115	0.011	0.251
County FE	✓	✓	
Month FE	✓	✓	
Bank-County FE			✓
County-Month FE			✓
Bank-State-Month FE			✓

This table presents estimates from OLS regressions. The dependent variable is 1(*complaint*), an indicator that switches on if a bank receives a CFPB complaint in a county-month. In columns 2 and 3, we use the `bart-large-mnli` model (Yin, Hay, Roth, 2019) to identify CFPB complaints that are related to identity theft. The key independent variable is *net branch closures*, which is the negative change in the number of bank branches over the past 180 days in the county. All standard errors are clustered at the county level. Standard errors in column 3 are additionally clustered at the bank-month level. Absolute values of t-statistics are reported in parentheses.

Table 5. Heterogeneous effects of bank branch closures on identity theft

Average treatment effects on treated (ATTs)

	(1)	(2)	(3)	(4)
	Net branch closures	ID theft reports	Net branch closures	ID theft reports
Panel A. Bank mergers by branch reliance:				
	High		Low	
Pre-exposure	0.47	64.28	0.42	53.93
	(1.14)	(2.25)	(2.26)	(1.81)
Post-exposure	19.52	4,444.48	1.70	656.88
	(2.76)	(2.58)	(2.05)	(3.03)
Overall	11.67	2,362.36	1.03	440.27
	(2.70)	(2.92)	(1.48)	(3.09)
Wald estimate	202.43		—	
Panel B. Bank mergers by digital focus:				
	Low		High	
Pre-exposure	0.15	12.15	0.27	5.35
	(0.49)	(1.38)	(1.29)	(0.72)
Post-exposure	3.27	555.73	2.66	119.29
	(2.79)	(2.78)	(1.59)	(1.01)
Overall	1.93	281.70	2.66	229.06
	(2.20)	(2.30)	(1.72)	(1.94)
Wald estimate	145.96		86.11	

(Continued next page)

Table 5. (continued)

Average treatment effects on treated (ATTs)

	(1)	(2)	(3)	(4)
	Net branch closures	ID theft reports	Net branch closures	ID theft reports
Panel C. Bank mergers by consumer tech-savviness:				
	Low		High	
Pre-exposure	0.15	−4.34	0.43	45.67
	(0.37)	(0.97)	(1.55)	(1.50)
Post-exposure	2.47	534.21	5.21	1,120.46
	(2.20)	(2.17)	(3.19)	(2.74)
Overall	1.91	398.31	3.79	668.13
	(2.02)	(2.31)	(2.79)	(2.62)
Wald estimate	208.54		176.29	

This table presents ATTs on *net branch closures* and *ID theft reports* from the Callaway and Sant'Anna (2021) difference-in-differences estimator. The variable *net branch closures* is the negative year-on-year change in the number of bank branches within an MSA. The MSA-year number of *ID theft reports* is collected from the Consumer Sentinel Network database of the U.S. Federal Trade Commission. The *pre* (*post*)*-exposure averages* are the aggregated ATTs before (after) merger exposures. An MSA is first exposed in the year when a merger occurs between large banks that both have branches in the area. We classify bank mergers based on the (i) *branch reliance* of the acquirer (Panel A), (ii) *digital focus* of the acquirer (Panel B), and (iii) *consumer tech-savviness* in the exposed MSA (Panel C). The *branch reliance* of a bank is the ratio of its number of branches to its total deposits. The *digital focus* of a bank is the ratio of the download volume of its mobile app on the `Google Play Store` to its number of branches. The *consumer tech-savviness* of an MSA is the proportion of consumers who have internet subscriptions. Standard errors are clustered at the MSA level. Absolute values of t-statistics are reported in parentheses.

Table 6. Military veterans, branch closures, and identity theft

	(1)	(2)
	1st stage	2nd stage
Dependent variable:	Veterans (v)	ID theft reports
Draft share	481.49 (88.24)	
$\widehat{v}$ × Net branch closures		3.65 (5.49)
$\widehat{v}$		1,531.35 (1.12)
Net branch closures		17.17 (0.92)
# Obs.	4,138	4,138
R^2	0.845	0.703
F	1,123.32	11.29
Controls	✓	✓
MSA FE		✓
Year FE	✓	✓

This table presents estimates from a two-stage IV regressions. The variable *veterans* (abbreviated as v) is the number of male veterans aged 55–74 in an MSA. Both *veterans* and its instrumented variant are scaled by a factor of 10^5. The variable *draft share* is the MSA's population share of males who were eight or nine years old during the 1960 census. The variable *net branch closures* is the negative year-on-year change in the number of bank branches within an MSA. The number of *ID theft reports* is collected from the Consumer Sentinel Network database of the U.S. Federal Trade Commission. For brevity, we include but suppress the estimates of control variables (used in Table 2) in our models. Standard errors are clustered at the MSA level. Absolute values of t-statistics are reported in parentheses.

Table 7. Merger exposures and consumer-level mobile app usage

Average treatment effects on treated (ATTs)

	(1)	(2)
Outcome:	Net branch closures	Mobile app usage
Pre-exposure	0.005	0.042
	(1.76)	(0.07)
Post-exposure	0.052	5.344
	(0.99)	(1.71)
Overall	0.072	2.766
	(3.96)	(3.21)
Wald estimate	38.42	

This table presents ATTs on *net branch closures* and *mobile app usage* from the Callaway and Sant'Anna (2021) difference-in-differences estimator. The variable *mobile app usage* is the number of hours spent on all mobile applications, except mobile banking apps, in a month by each Android smartphone user in an opted-in panel from the Global Wireless Solutions Magnify database. The variable *net branch closures* is the negative change in the number of bank branches in the month, matched to the consumers' residences at the county level. The *pre* (*post*)-*exposure averages* are the aggregated ATTs before (after) merger exposures. An MSA is first exposed in the year when a merger occurs between large banks with branches in the area. Standard errors are clustered at the consumer level. Absolute values of t-statistics are reported in parentheses.

Table 8. Merger exposures and consumption patterns

Average treatment effects on treated (ATTs)

	(1)	(2)	(3)
Outcome:	Net branch closures	Online spending gap ($\$ \times 10^6$)	Online transaction gap ($\times 10^6$)
Pre-exposure	0.003	–0.211	–0.004
	(0.16)	(2.20)	(1.91)
Post-exposure	0.415	4.499	0.124
	(4.42)	(3.10)	(3.23)
Overall	0.311	2.774	0.079
	(3.79)	(2.65)	(2.97)
Wald estimate	—	8.920	0.254

This table presents ATTs on *net branch closures*, *online spending gap*, and *online transaction gap* from the Callaway and Sant'Anna (2021) difference-in-differences estimator. The variable *online spend gap* (*online transaction gap*) is the monthly dollar value (number) of online transactions less that of offline transactions in an MSA, compiled by the `SafeGraph` database. The variable *net branch closures* is the negative change in the number of bank branches in the month, matched to the consumers' residences at the county level. The *pre* (*post*)*-exposure averages* are the aggregated ATTs before (after) merger exposures. An MSA is first exposed in the year when a merger occurs between large banks that both have branches in the area. Standard errors are clustered at the MSA level. Absolute values of t-statistics are reported in parentheses.

Table 9. Bank branch closures and unwanted calls

Dependent variable: Unwanted calls

	(1)	(2)	(3)
Unwanted call type	All	Wireless	Wired
Net branch closures	1.509	0.792	0.362
	(2.28)	(1.95)	(1.65)
Marketing calls	–32.798	–17.152	–4.357
	(10.26)	(11.27)	(4.62)
Over 60	–6.584	–4.592	–1.038
	(3.28)	(4.02)	(2.09)
Male	–5.369	–3.118	–1.395
	(3.31)	(3.50)	(3.38)
Unemployed	1.974	0.750	0.473
	(1.75)	(1.12)	(1.67)
White	–0.241	–0.095	–0.049
	(1.29)	(0.94)	(0.94)
High school	0.412	–0.045	0.067
	(1.68)	(0.43)	(0.63)
log(Household income)	88.100	42.194	25.868
	(2.85)	(2.42)	(3.39)
log(Population)	62.158	37.937	7.831
	(1.70)	(2.06)	(0.82)
# Obs.	1,113,936	1,113,936	1,113,936
R^2	0.553	0.489	0.360
MSA FE	✓	✓	✓
Year FE	✓	✓	✓

This table present estimates from OLS regressions. The dependent variable is *unwanted calls*, the number of unwanted calls or robocalls reported by consumers in an MSA to the U.S. Federal Trade Commission on a given day. The key independent variable in all columns is *net branch closures*, which is the negative year-on-year change in the number of bank branches in an MSA over the past 180 days. Standard errors are clustered at the MSA level. Absolute values of t-statistics are reported in parentheses.

Table 10. Bank branch closures and phishing sites

	(1)	(2)	(3)	(4)
	In-sample		Out-of-sample	
Dependent variable:	Phishing exposure	ID theft reports ($\times 10^3$)	Phishing exposure	ID theft reports ($\times 10^3$)
Net branch closures	0.024		0.058	
	(6.09)		(9.08)	
MSA phishing		15.30		4.11
		(2.07)		(3.00)
Over 60		0.21		0.03
		(2.13)		(0.41)
Male		0.05		–0.17
		(0.51)		(1.74)
Unemployed		0.37		0.22
		(2.36)		(1.88)
White		–0.03		0.02
		(1.61)		(1.04)
High school		0.06		0.01
		(2.63)		(0.54)
log(Household income)		–0.95		1.94
		(0.71)		(1.19)
log(Population)		0.02		13.38
		(0.01)		(1.70)
Unit of obs.	Bank-MSA-Month	MSA-Year	Bank-MSA-Month	MSA-Year
Sample period	2018–22	2018–22	2018–19	2020–22
# Obs.	49,888	1,725	12,723	1,063
R^2	0.001	0.834	0.006	0.946
MSA FE		✓		✓
Year FE		✓		✓

This table presents estimates from OLS regressions. The dependent variable *phishing exposure* in column 1 is the monthly number of websites that use the favicon of the legitimate bank website but lack secure-sockets-layer (SSL) certification, weighted by the population of a MSA where the bank operates. The independent variable *net branch closures* is the negative month-on-month change in the number of bank branches within an MSA. The dependent variable in column 2 is the number of *ID theft reports*, collected from the Consumer Sentinel Network database of the U.S. Federal Trade Commission. The key independent variable *MSA phishing* is the weighted sum of predicted *phishing exposure*, aggregated to the MSA-year level. The sum is weighted by the number of branches operated by banks in a given MSA-year. Columns 3 and 4 are the out-of-sample analogs of columns 1 and 2. Standard errors are clustered at the MSA level. Absolute values of t-statistics are reported in parentheses.

Internet Appendix to:

Grand Theft Identity: The Privacy Costs of Digitalization

Kenny Phua Chishen Wei Gloria Yang Yu

Abstract

The Internet Appendix contains supplementary information and additional tests for the paper "Grand Theft Identity: The Privacy Costs of Digitalization". The contents of the Internet Appendix are organized as follows. Section IA.1 reports detailed statistics of identity theft reports. Section IA.2 tabulates the list of entities that contribute data to the FTC Consumer Sentinel Network database. Section IA.3 provides additional details on the consumer response to branch closures. Section IA.4 presents additional details on the complier characteristics analysis. Section IA.5 details annual imputed financial losses from identity theft due to branch closures.

Internet Appendix to:

Grand Theft Identity: The Privacy Costs of Digitalization

Abstract

The Internet Appendix contains supplementary information and additional tests for the paper “Grand Theft Identity: The Privacy Costs of Digitalization”. The contents of the Internet Appendix are organized as follows. Section IA.1 reports detailed statistics of identity theft reports. Section IA.2 tabulates the list of entities that contribute data to the FTC Consumer Sentinel Network database. Section IA.3 provides additional details on the consumer response to branch closures. Section IA.4 presents additional details on the complier characteristics analysis. Section IA.5 details annual imputed financial losses from identity theft due to branch closures.

IA.1 Detailed statistics on identity theft reports

We present detailed statistics on identity theft reports from the Federal Trade Commission (FTC) Consumer Sentinel Network database in 2023. Panel A provides a breakdown of these reports by the types and subtypes of identity theft. Panel B provides a breakdown of these reports by identity theft types and age groups.

- Table IA.1 here -

IA.2 Data contributors of the FTC Consumer Sentinel Network database

We provide the list of organizations that provide data to the FTC Consumer Sentinel Network database.

- Table IA.2 here -

IA.3 Details on consumer response to branch closures.

Table IA.3 contains the full regressions results that underpin Figure 3 in the main text.

- Table IA.3 here -

IA.4 Details on the complier characteristics analysis

In this section, we provide details on our complier characteristics analysis. To fix ideas, every MSA has two unobserved potential treatment indicators $D(0)$ and $D(1)$ that manifest an observed treatment $D \in \{0,1\}$ representing the presence of branch closures. The indicator Z switches on if an MSA is exposed to mergers. The matrix below maps the treatment responses to Z of the subpopulations.

	$D(Z=0)$	$D(Z=1)$
Compliers	0	1
Always-Takers	1	1
Never-Takers	0	0
Defiers	1	0

Because we only observe the realized treatment D but not $D(0)$ and $D(1)$, we cannot classify individual MSAs into subpopulations. Specifically, compliers and always-takers assigned to the exposure group ($Z = 1$) are observably identical. The same applies to compliers and never-takers assigned to the control group ($Z = 0$). Marbach and Hangartner (2020) proposes a framework to estimate the mean of a characteristic X within subpopulations by imposing four assumptions.

Assumption 1 (Monotonicity). $D(1) \geq D(0)$.

Assumption 2 (Independence of instrument). $D(0), D(1), X \perp\!\!\!\perp Z$.

Assumption 3 (Relevance condition). $E[D \mid Z=1] \neq E[D \mid Z=0]$.

Assumption 4 (Probability bounds on assignment). $0 < \Pr(Z=1) < 1$.

Assumption 1 is standard in IV analysis and posits that the instrument cannot have an opposite effect on any subpopulation, thereby ruling out defiers (Angrist, Imbens, and Rubin, 1996). Assumption 2 implies the independence of Z with both X and $D(Z)$, which holds if merger exposures are randomly assigned across MSAs. Assumption 3 is the relevance condition, which guarantees that the fraction of compliers is nonzero. The assumption 4 strictly bounds the probability of assignment between 0 and 1 to ensure there is variation in Z across MSAs.

Under assumptions 1 and 2, observable and unobservable always-takers draw from the same distribution of X. So, we can profile the characteristic mean for always-takers by focusing on the observable subset of nonencouraged ($Z = 0$) MSAs that experience branch closures:

$$E[X \mid D(0) = D(1) = 1] = E[X \mid D = 1, Z = 0] \tag{IA.1}$$

By the same logic, we can profile the characteristic mean for never-takers by focusing on encouraged ($Z = 1$) MSAs that do not experience branch closures:

$$E[X \mid D(0) = D(1) = 0] = E[X \mid D = 0, Z = 1] \tag{IA.2}$$

We cannot immediately estimate the characteristic means for compliers because they are observably identical to always-takers and never-takers when $Z = 1$ and $Z = 0$, respectively. Instead, we first decompose the population mean into a linear combination of subpopulation means by the Law of Iterated Expectations:

$$\begin{aligned} E[X] = \;& E[X \mid D(1) > D(0)] \cdot \Pr[D(1) > D(0)] \\ + \;& E[X \mid D(1) = D(0) = 1] \cdot \Pr[D(1) = D(0) = 1] \\ + \;& E[X \mid D(1) = D(0) = 0] \cdot \Pr[D(1) = D(0) = 0] \end{aligned} \tag{IA.3}$$

Substituting in equations (IA.1) and (IA.2) and expanding all conditionals, we can express the characteristic mean for compliers as a function of observables.

$$\begin{aligned} E[X \mid D(1) > D(0)] = & \left(E[X] - \frac{E[X 1_{\{D=0,Z=1\}}]}{\Pr[Z=1]} - \frac{E[X 1_{\{D=1,Z=0\}}]}{1 - \Pr[Z=1]} \right) \\ & \left(1 - \frac{\Pr[D=0,Z=1]}{\Pr[Z=1]} - \frac{\Pr[D=1,Z=0]}{1-\Pr[Z=1]} \right)^{-1} \end{aligned} \tag{IA.4}$$

IA.5 Details on quantifying the effects of branch closures on identity theft

Table IA.4 tabulates the annual statistics used to impute financial losses from identity theft due to branch closures.

- Table IA.4 here -

Table IA.1. Detailed statistics on identity theft reports

Panel A. Breakdown of identity theft reports by types

Identity theft type	Subtype	Num. reports
Credit card	New accounts	381,122
	Existing accounts	44,855
Loan or lease	Apartment or house rented	13,201
	Auto loan/lease	52,070
	Business/personal loan	81,342
	Federal student loan	6,815
	Non-federal student loan	10,921
	Real estate loan	7,551
Bank account	Debit cards, electronic funds transfer, or ACH	42,148
	Existing accounts	18,723
	New accounts	84,335
Govt. documents or benefits	Driver's license issued/forged	8,977
	Govt. benefits applied for/received	82,419
	Other govt. documents issued/forged	9,096
	Passport issued/forged	1,623
Employment or tax-related	Employment or wage-related	31,207
	Tax	60,970
Phone or utilities	Landline telephone – existing accounts	1,125
	Landline telephone (new accounts)	4,578
	Mobile telephone (existing accounts)	7,853
	Mobile telephone (new accounts)	43,225
	Utilities (existing accounts)	1,896
	Utilities (new accounts)	28,725
Other identity theft	Email or social media	19,534
	Evading the law	5,526
	Insurance	11,402
	Medical services	13,683
	Online shopping or payment account	18,058
	Other	205,505
	Securities accounts	5,513

(Continued next page)

Table IA.1. (continued)

Panel B. Breakdown of identity theft reports by age groups

Age group	< 19	20-29	30-39	40-49	50-59	60-69	70-79	> 80
Identity theft type								
Bank account	1,833	20,558	36,859	28,804	20,589	14,339	7,076	2,132
Credit card	2,501	71,900	122,246	84,604	53,438	27,974	10,899	2,852
Employment or tax-related	13,774	16,826	17,827	13,765	10,869	7,877	3,899	1,261
Govt. documents or benefits	1,989	11,373	21,791	19,095	16,061	9,692	3,274	987
Loan or lease	874	26,152	44,611	30,730	20,437	11,264	4,117	1,069
Other identity theft	2,415	42,673	66,702	45,545	27,768	13,665	5,174	1,417
Phone or utilities	622	13,653	23,066	16,770	11,801	6,974	2,582	650

This table presents detailed statistics on identity theft reports from the Federal Trade Commission (FTC) Consumer Sentinel Network database in 2023. Panel A provides a breakdown of these reports by the types and subtypes of identity theft. Panel B provides a breakdown of these reports by identity theft types and age groups.

Table IA.2. Data contributors to the FTC Consumer Sentinel Network database

AARP Fraud Watch Network
Alaska Attorney General
Apple Inc.
Arvest Bank
AT&T Corporation
Australian Competition and Consumer Commission
Best Buy Co. Inc.
Canada Competition Bureau
Capital One Bank
Colorado Attorney General
Comcast Corporation
Connecticut Department of Consumer Protection
Consumer Financial Protection Bureau
Corporation for National and Community Service
Costco Wholesale Corporation
Craigslist
Cybercrime Support Network
Discover Bank
Dominion Energy
eBay
FedEx
First National Bank of Omaha
Florida Attorney General, Office of Citizen Services
Florida Department of Agriculture and Consumer Services
Grants.gov
Handshake
Hawaii Office of Consumer Protection
Hewlett-Packard
Idaho Attorney General
Indeed
Indiana Attorney General
International Association of Better Business Bureaus
Internet Crime Complaint Center
Iowa Attorney General
Iowa, Clinton County Sheriff's Office
JPMorgan Chase & Co.
LinkedIn
Los Angeles County Department of Consumer Affairs
Louisiana Attorney General
Maine Attorney General
Massachusetts Attorney General
MasterCard International
Michigan Attorney General
Microsoft Corporation Cyber Crime Center
Mississippi Attorney General
MoneyGram International

(Continued next page)

Table IA.2. (Continued)

National Consumers League
National Council on Aging
National Grid
Nebraska Attorney General
Nevada Attorney General
Nevada Department of Business and Industry
New Mexico, Albuquerque
New York State Attorney General
North Carolina Department of Justice
Ohio Attorney General
Ohio, Cuyahoga County Department of Consumer Affairs
Oregon Department of Justice
Pennsylvania Attorney General
PeopleClaim
PepsiCo, Inc.
Petscams.com
PrivacyStar
Prosperity Bank
Publishers Clearing House
Rent Group, Inc.
Sages Theater, Inc.
Scam Advisor
Scam Detector
Society of Citizens Against Relationship Scams
South Carolina Department of Consumer Affairs
Tennessee Division of Consumer Affairs
U.S. Bureau of Prisons
U.S. Citizenship and Immigration Services
U.S. Customs and Border Protection
U.S. Department of Defense
U.S. Department of Education
U.S. Department of Health and Human Services, Office of Inspector General
U.S. Department of Justice, Consumer Protection Branch
U.S. Department of Justice, Disaster Fraud Task Force
U.S. Department of Justice, Elder Fraud Hotline
U.S. Department of Justice, Executive Office for Immigration Review
U.S. Department of Justice, Task Force on Market Integrity and Consumer Fraud
U.S. Department of the Treasury, Internal Revenue Service
U.S. Department of Veterans Affairs
U.S. Drug Enforcement Administration
U.S. Federal Bureau of Investigation
U.S. Federal Communications Commission
U.S. Marshals Service
U.S. Parole Commission
U.S. Patent and Trademark Office
U.S. Postal Inspection Service
U.S. Social Security Administration

(Continued next page

Table IA.2. (Continued)

United Parcel Service
USA.gov
Utilities United Against Scams
Valve Corporation
Verizon Wireless
Walmart Corporation
Washington State Attorney General
Western Union Company
Wisconsin Department of Agriculture, Trade and Consumer Protection
Xerox Corporation
Zelle
Zillow Group

This table contains the list of data contributors to the Federal Trade Commission (FTC) Consumer Sentinel Network database.

Table IA.3. Elasticity of bank branch visits

Dependent variable: Bank branch visits

	(1)	(2)	(3)	(4)
Distance band	0–1 miles	1–2 miles	2–3 miles	19–20 miles
Net branch closures	14.77	3.29	3.83	0.19
	(2.61)	(1.24)	(1.65)	(0.17)
Over 60	–2.37	–2.40	–2.39	–2.42
	(4.66)	(4.63)	(4.63)	(4.71)
Male	0.13	0.19	0.20	0.20
	(0.17)	(0.24)	(0.26)	(0.26)
Unemployed	0.00	0.00	0.00	0.00
	(6.35)	(6.33)	(6.32)	(6.33)
White	–0.01	–0.01	–0.01	–0.01
	(5.71)	(5.73)	(5.73)	(5.71)
High school	–1.26	–1.20	–1.21	–1.16
	(1.97)	(1.93)	(1.93)	(1.82)
log(Household income)	–97.66	–97.99	–97.91	–98.73
	(6.41)	(6.44)	(6.48)	(6.45)
log(Population)	49.47	48.66	48.60	48.64
	(8.49)	(8.05)	(8.06)	(8.06)
# Obs.	11,676,567	11,676,567	11,676,567	11,676,567
R^2	0.595	0.595	0.595	0.595
County-Week FE	✓	✓	✓	✓
Bank-Week FE	✓	✓	✓	✓
County cluster	✓	✓	✓	✓
Bank cluster	✓	✓	✓	✓
Week cluster	✓	✓	✓	✓
Implied elasticity	18.1%	5.1%	7.8%	0.7%

This table presents estimates from OLS regressions. The dependent variable is *bank branch visits*—the weekly number of visits received by a bank branch, compiled by the pass_by database. For every bank branch, we merge in the *net branch closures* of neighboring bank branches that are within the distance band stated in each column. The variable *net branch closures* is the negative change in the number of bank branches over the past 180 days. We calculate the distances separating bank branches using their latitude and longitude coordinates. These coordinates are sourced from the FDIC SOD dataset and through geocoding the pass_by branch addresses via the Google Maps geocoding API. t-statistics are reported in parentheses.

Table IA.4. Quantifying the effects of branch closures on identity theft

Imputed effects from calendar ATTs

	(1)	(2)	(3)	(4)	(5)
Year	ATT (Net branch closures)	ATT (ID theft reports)	Wald estimate = (2)/(1)	Total num. reports	Total losses (U.S.$ million)
2011	2.60	−3.60	−1.38	4,973	6.40
2012	3.40	2.79	0.82	476	0.60
2013	3.39	221.65	65.37	73,868	104.46
2014	4.63	244.01	52.66	94,525	98.76
2015	3.43	351.36	102.35	144,005	79.83
2016	1.40	285.66	204.04	309,740	166.27
2017	2.50	198.85	79.61	183,186	125.60
2018	2.73	235.93	86.46	177,838	163.15
2019	1.65	521.01	315.28	680,686	612.84
2020	2.15	972.88	453.48	740,994	1,035.07
2021	2.30	604.49	262.94	982,351	1,884.40
2022	4.41	554.20	125.62	406,888	1,408.09

This table presents effects of branch closures on identity theft imputed from the Callaway and Sant'Anna (2021) calendar ATTs. The calendar ATT in a year is the average treatment effect for a MSA that is or is already exposed to large bank mergers in that year. Columns 1 and 2 present the calendar ATTs for *net branch closures* (i.e., the first stage) and the number of *ID theft reports* (i.e., the second stage), respectively. Column 3 presents the Wald estimates, which are the ratios of the second-stage estimates to the first-stage estimates. The Wald estimate represents the causal effect of one instrumented *net branch closure* on the number of *ID theft reports*. Column 4 imputes the number of *ID theft reports* by multiplying the Column 3 estimates by the actual *net branch closures* in MSA-years and aggregating them to the year level. Column 5 imputes the total losses by multiplying the imputed number of ID theft reports in the MSA-year by the average loss per report in the state-year and aggregating them to the year level.