Investing in Lending Technology: An Anatomy of Bank IT Spending*

Zhiguo He, Sheila Jiang, Douglas Xu, Xiao Yin

September 7, 2021

Abstract

This paper investigates the lending technology of the traditional banking sector after the arrival of information age, by examining the investment in information technologies (IT) by U.S. commercial banks. Given the distinctive natures of banks' dealing with information throughout the process of their lending activities, we link banks' IT spending in different categories to different aspects of their lending technologies. Investment in communication IT is shown to be more associated with improving banks' ability of soft information production and transmission, while investment in software IT helps prompt banks' hard information processing capacity. By exploiting polices that affect geographic regions differentially, we show that banks respond to an increased demand for small business credit (mortgage refinance) by increasing their spending on communication (software) IT spending. We also find an asymmetric impact of technological development on the labor employment in the banking sector.

Keywords: Information Technology, Small Business Lending, Mortgage Refinance, Communication Equipment, Software
JEL codes: G21, G51, J24, O32

---

*Preliminary and references incomplete; please do not quote. He: Booth School of Business, University of Chicago and NBER, zhiguo.he@chicagobooth.edu. Zhiguo He acknowledges financial support by the John E. Jeuck Endowment at the University of Chicago Booth School of Business. Jiang: Warrington College of Business, University of Florida, sheila.jiang@warrington.ufl.edu. Xu: Warrington College of Business, University of Florida, douglas.xu@warrigton.ufl.edu. Yin: PhD program, Haas School of Business, University of California, Berkeley.

1 Introduction

Commercial banks have long relied on cutting-edge technology to deliver innovative products such as ATMs and online banking, streamline the loan making processes, and improve back-office efficiency. According to a 2012 Mckinsey Report, across the globe commercial banks are spending about 4.7% to 9.4% of their operating income on information technology (later called IT); for comparison, insurance companies and airlines only spend 3.3 percent and 2.6 percent of income, respectively. This trend has been accelerated in an unprecedented pace in recent years, especially after the COVID-19 pandemic, as industry professionals often attribute top commercial banks to be more like "technology companies" than actual technology companies by virtue of the enormous amounts on IT spending.[1] Recently, the impact of information technology on banking sector and financial stability has been a hot topic in the policy discussion (Banna and Alam (2021), Pierri and Timmer (2020)).

Although the financial services industry–especially the banking industry–is increasingly becoming a tech-like business, the academic literature is lagging behind in understanding the economics of IT spending in the banking industry. Which banks, large or small, have been investing more? How IT spending affects banks' lending technology and loan making? Our study is a first step toward understanding the basic pattern of these banking IT expenditures, and explore the connections and underlying mechanisms between these expenditures and the core functioning of the banking system.

To place our research in the established banking literature, think about the information transmission between the loan officer and the borrower, or across layers of loan officers within a bank organization. As highlighted by Stein (2002), soft information tends to be more effectively transmitted and thus lending decisions are more easily determined when a bank has

[1] For instance, an interesting article shows that IT spending in most of top banks, including JP Morgan, Bank of America, Citi, Goldman Sachs and international leaders like Deutsche Bank and Barclays, exceeds 17% of their total operating costs, while Amazon and Alphabet devote 12% and 20% of their operating costs respectively to IT. This article stresses that banks are still lagging quite a bit in their IT spending in absolute terms, and cautions further that the above-mentioned IT spending numbers do not include compensation for IT staff members.

less hierarchical structure. The fast developing technologies in recent decades provide more options for the banking sector to cope with such problems. Indeed, in recent years banks are observed to pay more attention to strengthening their internal communication through the wide use of private branch exchanges and intranet installment.[2] And technology advances are far beyond internal communication; for instance, the explosive big data analytics market—which combines "hard" information like credit scores and other alternative data—calls for automatic information processing as a pressing need for banks today.

Our study relies on a comprehensive dataset, Harte Hanks Market Intelligence Computer Intelligence Technology database. This data set has been used in academic research in economics (e.g., Bloom et al. (2014) and Forman et al. (2012)). Focusing on commercial banks, we are the first—to the best of our knowledge—to analyze this dataset with detailed branch-level information on detailed spending categories.

We explain in Section 2.2 the four major categories of IT expenditure in the Harte Hanks dataset—hardware, software, communication, services—but in the context of banking industry. As become clear later, our study focuses on two of these four categories. First, *Software* includes the desktop applications (e.g., Microsoft Office), information management software, and risk and payment management software. By greatly improving the efficiency in document assembly, digitization, information classification, these software products automate information processing through specialized programming and AI technologies and thus improve accuracy and shortening processing speed.

On the other hand, *Communication* which includes radio and TV transmitters, private branch exchanges, and video conferencing, is defined as the network equipment that banks operate to support their communication needs. For instance, a set of advanced communication equipment allow bankers to more effectively interact with their borrowers; and private branch exchanges facilitates smoother exchanges of information and opinions within the bank branching network.

[2] For instance, a news article on Finextra reports the installment of intranet by First Citizens National Banks.

In Section 3, we start our analysis by first documenting that over the past 10 years, IT expenditure of the US banking sector has been growing rapidly and has reached to an economically significant level– the average IT spending as a share of total expenses has grown from nearly nil in 2010 to around 5% since 2015. A further examination across bank size groups reveals a somewhat different pattern in terms of the time trend dynamics as well as profile structure of IT investment made by smaller versus bigger banks. Larger banks' IT spending has been steadily growing, while this upward trend in smaller banks' IT spending is only observed during 2010-2015, which has significantly slowed down since 2015. Interestingly, medium-sized banks saw the fastest growth in their IT spending during 2010-2014, but dramatically slowed down during 2010-2015. In terms of the IT spending profile, a noticeable distinction is that smaller banks consistently allocate a higher share of their IT budget towards communication technology than bigger banks do.

To study the connection between IT spending and lending technology, Section 3.3 examines how banks' loan categories on their balance sheets vary with their IT spending profiles. Among the three large loan types as categorized in Call Report, we find that the shares in banks' C& I loans and agriculture loans are positively and significantly correlated with their communication spending intensity, but are uncorrelated with software spending; whereas the share of personal loans appears to be positively correlated with banks' software spending intensity, but not with communication spending. Furthermore, within C&I loans, it is small business lending that stands out to drive the positive correlation with communication IT spending; whereas within the personal loans, mortgage refinance stands out to drive the positive correlation between personal loans and software spending. These findings are robust at both bank and bank-county-year level.

Based on the above findings, we develop three main hypotheses in Section 4 linking the nature of information in banks' lending activities to their IT adoption behavior. Conceptually, we differentiate two fundamental types of bank lending technologies. The first heavily relies on the gathering and augmentation of the soft information from the borrower; in the context of Berger and Udell (2002), "relationship-lending" is a concrete example of the first type. The second fundamental type of lending technology, on the other hand, relies primarily on the processing and quantification of hard information. "Transactions lending" in Berger and Udell (2006), i.e., loans that are based on specific credit scoring system and quantified financial statement metrics are standard examples of the second type.

As the first testable hypothesis, we posit that a demand shock of small business credit will lead banks to invest more in communication technologies. This is because communication technologies—say video conferencing—enable banks to more effectively interact with small business borrowers (whose information structure is often rather opaque) to gather more soft information, as well as allow for a smoother transmission of these otherwise hard-to-verify soft information within a bank organization. Taking advantage of an arguably exogenous demand shifter, we find that an increase in banks' small business credit demand– due to a higher ex-ante exposure of the local counties to the policy shock exploited in our analysis– leads to a positive and significant growth in banks' communication spending ex-post, without much impact on the bank's software spending.[3]

The second testable hypothesis is that a positive demand shock for mortgage refinancing should push banks to engage more IT spending on software. This is because software is particularly useful for dealing with existing or readily accessible information that is relatively "hard," and especially the credit scoring software utilized by banks when making refinancing decisions. To identify the causal relationship between a demand for "hard" information processing and banks' software spending, we construct a shifter to the mortgage refinance demand faced by banks across different regions. Specifically, we utilize the cross-county variation in the interest payment gap of outstanding mortgage due to the presence of local mortgage rate gap– a higher interest payment gap naturally implies a higher mortgage refinance demand by local households. Thanks to this instrument variable which has been used

[3]The policy shock we utilize to construct the demand shifter for small business credit is "Small Business Health care Tax Credit." As a part of the Affordable Care Act, this program was enacted between 2014 and 2015 and provided beneficial tax treatment specifically targeted at small business establishments in the US economy.

in the literature (e.g., Eichenbaum et al. (2018)), we show that a one standard deviation higher mortgage refinance lent out by a bank (due to its local exposure to high refinance savings) leads to around 8% higher on average software spending intensity measured as software spending scaled by total revenue, while no significant impact on local banks' communication spending is observed.

Our last testable hypothesis concerns how the IT spending affects the labor in the banking industry, as development in information technologies offers more options for banks to perform "tasks" that are essential in their lending technology, though traditionally undertaken by employing human labor (Acemoglu and Restrepo (2018)). Specifically, utilizing the county-level availability of land-grant colleges as an instrument variable for local supply of IT-related employees, we find that counties with more abundant supply of IT-related employees have significantly slower software spending growth compare to those located in regions where the supply of IT-related employees is relatively scarce. In contrast, we do not see such a difference in banks' communication spending growth. Quantitatively, a one standard deviation higher growth in IT-employee hiring results in an average of 5.13% slower growth in software spending as a share of branch-level gross revenue. These results indicate an important yet asymmetric impact of the development in information technologies on the labor employment outcomes in the traditional banking sector.

Related Literature

Bank lending technology and the nature of information This paper is closely related to the literature of relationship and transactions lending. Berger and Udell (2006) provide a comprehensive framework of the two fundamental types of bank lending technology in the SME lending market.[4] A fundamental difference between these two types of lending is related to the role played by information in the process of lending as highlighted by Stein

[4]Bolton et al. (2016) study the joint determination of relationship lending and transactions lending. They find that firms relying more on relationship banking are better able to weather a crisis than firms relying on transactions banking, suggesting a higher capital requirement for relationship banks.

(2002), who provides an explanation to why soft information production favors organizational structure of fewer hierarchical layers and why banks cut their lending to small businesses after consolidation. This notion that credible communication is essential for the production of soft information is in line with our finding that spending on communication technology facilitates the generation and transmission of soft information and meanwhile smaller banks tend to allocate a higher proportion of their IT budget to communication compared with larger banks.[5]

We contribute to this strand of literature by linking information technology to bank lending technology, especially on the distinction between soft information production/transmission and hard information processing. Methodologically, we demonstrate causal linkages between the different informational components in credit demand to banks' endogenous decision on lending technology adoption.[6]

Information technology and banking industry While the exact mechanism through which information technology affects banks' lending technology remains somewhat a black box, the interactions between the development of information technology and the evolution of banking industry have been well explored in the literature. For instance, Hannan and McDowell (1984) document that larger banks and banks in more concentrated banking industry are more likely to adopt automated teller machines. Stepping into the twentieth century, Berger (2003) shows that technological progress in both information technology and financial technology led to significant improvement in banking services productivity and quality. Petersen and Rajan (2002) document that development in communication technology greatly increased the lending distance of small business loans, reflected in that firms receiving credit

[5]Along these lines, Liberti and Mian (2009) find empirically that greater hierarchical distance leads to less reliance on subjective information but more on objective information and that more frequent communication between information collecting agents and loan approving officers can mitigate the effects of hierarchical distance on information use. Paravisini and Schoar (2016) document that credit scores, which serve as "hard information," improve the productivity of credit committees, reduce managerial involvement in the loan approval process, and increase the profitability of lending.

[6]Hard and soft information components are never black and white. Berger and Black (2011) emphasize that hard lending technologies often have both hard and soft information components, thus large banks' comparative advantage in hard lending technology is dependent on relative importance of hard versus soft information components involved in the technology.

in more distant areas no longer need to be the ones with the highest credit score.[7]

The emergence of fintech is a signature result of recent development in information technologies. Our paper can also be related to a fast growing literature studying how the emergence of fintech industry is affecting (or has affected) the traditional banking sector. Related research in this strand of literature includes Lorente et al. (2018), Hornuf et al. (2018), Calebe de Roure and Thakor (2019), Tang (2019), Erel and Liebersohn (2020), Aiello et al. (2020), Schnabl and Gopal (2020), and He et al. (2021). While a common theme and angle of these research has mostly been focused on examining how the uprising fintech industry is affecting the bank-fintech competition, in which process traditional banks are largely viewed as a *passive* player, little attention has been paid on how banks are *actively* responding to the emergence of this group of new challengers. To the best of our knowledge, our paper is the first to study whether and how the traditional banking sector is catching up with the penetrating fintech companies, through examining the IT investment behavior in the U.S. banking sector.

More broadly, our paper is also related to the strand of literature studying financial technology and the transformation of financial sector.[8] Our paper contribute to this literature by providing the first detailed categorization of banks' IT spending in the information age, as well as identifying the bank's endogenous adoption of different technologies in order to cater to different natures of information involved in various lending activities.

Endogenous technology adoption and its real impact Recently, there has been an large literature studying the endogenous adoption of IT across non-financial firms or geographical units and its impact on real economic outcomes, such as firm productivity, em-

[7]There is also a vast theoretical literature on the topic of the interactions among information technology, banking market competition and bank lending; see Freixas and Rochet (2008) for a review. Hauswald and Marquez (2003) analyzes how two dimensions of technological progress affect competition in financial services. They show that although technological progress will lower the cost of information processing, it also implies lower entry cost and higher competition. So the overall effect of technological progress on interest rate is mixed. More recently, Vives and Ye (2021) study how diffusion of information technology affect competition in bank lending market and banking sector stability.

[8]The related works include but are not limited to Di Maggio and Yao (2020), Frost et al. (2019), Hughes et al. (2019), Stulz (2019), Fuster et al. (2019), Buchak et al. (2018), Jagtiani and Lemieux (2017), and Philippon (2020).

ployment and local wages. This literature include McElheran and Forman (2019), Bloom and Pierri (2018), Brynjolfsson and Hitt (2018), Akerman et al. (2015), Bloom et al. (2012), Beaudry et al. (2010), Autor et al. (2003), Brynjolfsson and Hitt (2003).

In particular, to the best of our knowledge, our paper is the first to document that as information technology develops, banks with more availability of IT-related employees will accordingly invest less in technology that aims at processing hard information processing; but more higher availability of IT-related employees does not lead to lower degree of adoption of communication technology, which primarily deals with soft information generation. These findings suggest that the development in the information technology could have an important yet *asymmetric* impact on the labor employment in the banking sector.

2 Data and Background

We explain our main data sources in this section, together with detailed explanations of various categories of IT spending.

2.1 Data Source and Sample

The data on banks' IT spending comes from Harte Hanks Market Intelligence Computer Intelligence Technology database, which includes over three million establishment-level IT spending information from 2010 to 2019 to conduct information technology consulting for firms. Harte Hanks collects and sells this information to technology companies, who use this information for marketing purpose or to better serve their clients. Firms with IT spending information have incentives to report truthfully to Hart Hanks because they also want to receive the most appropriate IT products and services. This data set has been used in academic research in economics; leading examples include Bloom et al. (2014) who study the impact of information communication technology on firms' internal control, and Forman et al. (2012) who study firms' IT adoption and regional wage inequality.

We focus on commercial banks in the database. To the best of our knowledge, we are the first to analyze this data set with detailed branch-level information as well as detailed spending categories in the context of banking industry. Specifically, the data not only includes establishment-level information on banks' total IT spending budget, but also detailed information on different categories of IT budget. The categories include hardware, software, communication, services and others.

Our sample consists of 1806 commercial banks in the U.S.. The sum of assets of the banks in our sample covers more than 80% of the total asset size of the overall banking sector across 2010-2019, which allows us to draw a full picture of the IT investment behavior of the entire US banking sector. Figure 1 displays the comparison between the total asset size of banks in the sample and the total asset size of the overall banking industry in U.S. from 2010 to 2019. Table 1 reports the coverage of our sample by bank asset size group. For banks with asset size of over $250 billion, $10 billion-$250 billion, and $1 billion-$10 billion, the coverage in frequency and asset are both over 80%, for banks with asset size of $100 million-$ 1 billion , our sample overs 22.64% of total number of commercial banks in this size group, and covers 29.44% of total asset of banks in this size group; for asset size of less than $100 million, our sample covers 14.45% of total number and 14.23% of total assets.

The average total IT spending as a percentage of net income ranges from 1.11% (25-th percentile) to 10.60% (75-th percentile), which suggests that investment in IT is an economically significant portion of spending for US banks and it features large variation across different banks. Table 2 displays the summary statistics of banks' IT spending. This is consistent with the report by Mckinsey in 2012, which documented that according to their survey, banks' IT spending as a share of operating income ranges from 4.7% to 9.4%.[9]

[9]A screenshot of the report is in the Appendix Figure A1.

2.2 IT Investment Categorization

A major strength of out data set is that it gives us information on the detailed decomposition of banks IT investment into different types, with four major categories: hardware, software, communication, services. We now provide detailed explanations for these categories; the formal definition is also provided in 3(a) to 3(d) of Figure A2.

Software is defined as software purchased from third parties. This could be packaged or semi-packaged software delivered on CD and installed with the company, or offered on an SaaS from a multi-tenant shared-license server accessible by a browser, or custom-created for a company by third-party contractors or consultants. More specifically, the category of software includes the desktop applications, information management software, processing software, ePurchase, risk and payment management software.

One representative example of desktop applications is the Microsoft Office software package, and especially Excel, Powerpoint, Python. These software products are easy to grasp by bank employees and allows employees to conduct basic calculations and visualization of data that's associated with lending business.[10]

Examples of processing software includes Trapeze Mortgage Analytics, Treeno Software, Kofax, eFileCabinet etc. The specialty of these software products lies at automated information processing from loan applicants' paper document packets through specialized programming and AI technologies, which greatly improve the efficiency in document assembly, digitization, information classification, which are otherwise done manually by loan officers, and thus improve accuracy and shortening processing speed.

ePurchase software products allow banks' customers to make fund transfers more easily online and through mobile app. Examples of ePurchase include Zelle and Stripe. For instance, large banks (say BOA, CitiBank, and Wells Fargo) as well as smaller regional banks (say First Tennessee Bank and SunTrust Bank) use Zelle.

[10]For example, on Mendeley.com, the job postings of loan officers or project managers by many banks, basic skills requirements always include the applicants to be proficient with Microsoft Office.

Risk management software provides on-going risk assessment after loans have been issued, through augmenting borrowers' repayment information as well as industrial and economic condition information. These software products allow banks to better monitor loans in progress. Examples of risk management software products that banks use include Actico, ZenGRC, Equifax, Oracle ERP etc. Other software products include security trading system and operating systems that are typically bundled with the specific software products.

Communication is defined as the network equipment that banks operate to support their communication needs. It includes routers, switches, private branch exchanges, radio and TV transmitters, Wi-Fi transmitters, desktop telephone sets, wide-area network, local-area network equipment, video conferencing, and mobile phone devices.

When there is a need for bankers to contact or interact directly with borrowers, a set of advanced communication equipment allows bankers to more effectively talk to and see the borrowers. In addition, communication equipment such as private branch exchanges allow the exchange of information, opinion, and decision more effective within the bank branching network.

Hardware as a form of IT investment includes classic computer hardware such as PCs, monitors, printers, keyboard, USB devices, storage devices, servers, mainframes, etc. In terms of lending services, hardware is a fundamental type of tech investment that complements and facilitate both the gathering of borrower information and the processing of the information. This is because hardware devices, such as PC and servers, help provide storage and transmission of data, and meanwhile, they are also the carrier of software and toolboxes that need to be installed on them in order to process, calculate and analyze the data. In the supply of deposit-based services, use of computer is also important in the sense that with computers, bank tellers can more quickly locate depositors' data more quickly, instead of having to manually find the folder of files storing the depositors' information, they can simply type in the name and bring up the relevant data instantly.

Services are defined as project-based consulting or systems integration services that vendors provide to banks. Specifically, these include consulting services for IT strategy, security assessments, system integration, project services, hardware support and maintenance services. The services are mainly provided by IT outsourcing companies on contractual basis. Similar as hardware, services work as complements to other categories of information technology investment to facilitate banks' lending, although these services do not directly correlate to banks' information gathering or processing. For example, Aquiety is an IT service company based in Chicago that provide cybersecurity services to banks and other firms. Iconic IT is an IT service company headquartered in New York that provides software and hardware procurement, installment and upgrade services.

2.3 Bank Balance Sheet

We obtain bank-level balance sheet information from Call Report from 2010-2019, and for bank-county-year analysis, we utilize information on banks' revenue at county level from Summary of Deposits. The bank-level control variables include revenue per employee, net income scaled by assets, equity scaled by assets, deposits scaled by assets and total salary scaled by assets.

2.4 Loans and Local Characteristics

To supplement our study on banks' lending technology and banks' IT investment, we combine loan level information from multiple sources. We obtain syndicated loan information on frequency of a bank acting as lead bank in syndicated loan packages from LPC Dealscan. Small business loan origination data are from the Community Reinvestment Act (CRA), which is at the bank-county-year level covering the sample period of 2010-2019. Mortgage refinance information is reported in Home Mortgage Disclosure Act (HMDA) from 2010-2019. When constructing an instrument variable which serves as a demand shifter, we download the average mortgage interest rate at the county level before 2010 from Freddie Mac.

3 Empirical Patterns of Banks' IT Spending

We start our analysis by reporting some basic statistics of banks' investment in IT over the last decade in the US economy, and then explore how banks' IT investment correlates with its size. Finally, we show that banks' IT investment is closely related to their lending activities, by examining how the profile structure of banks' IT investment varies i) when banks have different specialization in loan types, e.g., commercial loans versus personal loans; and ii) when banks engage in lending activities of different nature, e.g., loan refinancing versus loan origination.

3.1 Time Trend of Banks' IT Investment

As the starting point of our analysis, in this section we demonstrate the following two facts: 1) banks' investment in information technology (IT) has been growing dramatically over the last decade; 2) up to now, banks' spending on IT has grown to an economically significant level, compared to other variables such as total expense, income, etc.

Figure 2 displays the average IT spending as a share of total expenses from 2010 to 2019. Overall, banks have drastically increased their investment in information technologies after the financial crisis. As a share of total expenses, IT budget increased from nearly zero in 2010 to about 5% after 2015. There was a slight slowdown in 2015, which is primarily driven by the medium and small-size banks and will be shown in the next subsection. In 2016, there was a dramatic pick-up. This could be potentially driven by the release of a "white paper" by the Office of the Comptroller of the Currency's (the "OCC," the regulator of federally chartered national banks and savings associations), that sets forth OCC's perspective on supporting responsible innovation in the federal banking system on March 16, 2016. The "White Paper" encourages banks of all sizes to integrate responsible innovation into their strategic planning and encourages banks to collaborate with non-banks in developing responsible innovative products that satisfy regulator requirements. In response to this "White Paper," banks might have gained more freedom and have been more actively investing in information technology in order to better catch up with Fin-tech development since 2016.[11]

We also report summary statistics on the detailed structure of the banks' IT spending profile. In particular, we report how banks' investments on IT are distributed across different categories as we defined in Section 2.2. By size software and services are the highest among all categories of IT spending, each constituting 33% of total IT budget. Hardware constitutes about 17% of total IT budget, and communication is on average 9% of total IT budget. The average size of storage of all banks in the sample is 3.52PB (3520 TB), and average number of IT-related employees is 133.

As another important input in banks' lending technology investment, banks' IT-related labor hiring is also of interests to our study. Figure 7 shows the line plots of banks' IT-related employees and IT spending intensity. Similar to the IT investment, we also see an overall upward trend in the importance of IT-related employment plays in the operation of banking sector. We postpone more detailed elaboration and discussion on the interaction between bank IT spending and labor employment in the later sections, after we have established more concrete economic meanings of bank IT investment.

3.2 Bank IT Spending across Bank Size

We now examine how the IT investment vary across different bank groups, by conducting the same set of analysis as in the last section–but for different banks size groups. We also briefly discuss how bank size could affect the profile structure of their IT spending, as well as the time trend.

Table 2 reports the summary statistics of banks' IT spending and Table A2 reports the summary statistics by bank size. Overall, we find that larger banks tend to make more IT investment as a share of non-interest expense than smaller banks do. As can be seen in Table A2, banks with total asset less than $0.1 Billion has an average IT/Non-interest

[11] The article by McKinsey documented the Fin-tech investment boom by venture capitalists and Fin-tech IPO boom since 2016.

expense ratio of 1.9%, whereas banks with asset size in the range of $1 Billion-$10-Billion and banks with asset size within $10 Billion-$250 Billion have an average IT/Non-interest expense of 6.2% and 6.7% respectively.

Another noticeable feature revealed by cross-size summary statistics is that smaller banks tend to allocate a higher fraction of their IT budget towards communication technology than larger banks do, while there are no significant differences in the software spending as a share of total IT spending across asset groups. To be more concrete, the average communication/Total spending ratio is 15.9% for banks with asset less than $0.1 Billion; 9% for banks with asset in the bin of $ 0.1 Billion and $ 1 Billion; 6.4% for banks in the asset bin $1 Billion-$10-Billion; 5.5% for banks in asset bin $10 Billion-$250 Billion and 4.6% for banks with asset higher than $250 Billion. A full comparison of the spending on communication and software (as a share of total IT spending) across different bank size group is shown in Figure 4.

Figure 3 displays the over-time trend of banks' IT investment by each bank size group. Overall, an upward trend in IT spending as a share of total non-interest expense is observed in all bank size groups.[12] Despite this common upward trend over the past decade, there are also some noticeable differences in the detailed dynamics across different bank size groups. Among the five size groups, banks with asset size greater than $250 Billion and banks with asset size $10 Billion-$250 Billion experienced steadily increasing their IT budget, while banks in asset size bin $1 Billion-$10 billion, $0.1 Billion-$1 Billion and those with asset less than $100 million seemed to have gradually decreased their IT budget since 2015, compared with their apparent up-ward trend during 2010-2014.

3.3 Bank IT Investment and Bank Lending

A major objective of our research is to achieve a better understanding of of the formation of banks' lending technology during the information age, hence a natural yet important step

[12]The magnitude of IT budget as a share of non-interest expense in this figure is also in line with Hitt et al. (1999), in their survey banks' IT spending could be as high as 15% of non-interest expense.

of our study is to examine the relationship between banks' investment in IT and their lending activities.

As an important dimension of banks' asset side activities, it is worthwhile to explore how the profile of banks' IT investment varies when banks are faced with different types of borrowers. Due to the differences in the identity and characteristics of borrowers, conducting lending to different types of borrowers often involves different treatment of information. As a consequence, if banks specialize in different types of loan making, we should naturally expect that they will also demonstrate differences in their IT investment profiles. To this end, in this section we ask and provide answer to the following empirical question: if a bank has a higher proportion of its loan portfolio allocated to certain specific type of loan, does the bank also demonstrate a higher share of IT budget allocated to certain specific type of information technology?

In what follows in this section, we examine how banks' IT investment is related to their (relative) specialization in three major types of loans: commercial and industrial (C&I) loans, personal loans, and agriculture loans. In studying banks' IT investment patterns, we conduct our analysis at both bank level and bank-county level. We report bank-level regression in this section and postpone the discussion of bank-county evidences to the next section after we have a more clear picture of the economic meaning of banks' investment in IT. Specifically, we run the following bank-level regression:

$$\frac{\text{Type S IT Spending}}{\text{Revenue}}_{i,10-19} = \alpha_i + \beta \frac{\text{Type L loan}}{\text{Total loan}}_{i,10-19} + \gamma X_i + \epsilon_i \tag{1}$$

where i refers to bank, the outcome variable of interests is $\frac{\text{Type S IT spending}}{\text{Total IT Spending}}_{i,10-19}$ is the average investment intensity in a specific type of IT spending as a share of bank i's revenue between 2010 and 2019. The main explanatory variable $\frac{\text{Type L loan}}{\text{Total loan}}_{i,10-19}$, which captures bank i's loan specialization, is measured by the average share of a specific type of loan of bank i of total loan size. Control variables include net-income scaled by total assets, total deposits scaled by total assets, revenue per employee, total equity scaled by total assets and total salaries scaled by total assets. All of the control variables are the ten-year average between 2010 and 2019 at the bank level. We report the estimation results of of the above regression for commercial and industrial loan, personal loan and agricultural loan in Table 3, and Table 4. Table 3 provides a complete demonstration of control variables and fixed effects, whereas Table 4 combines together the results of banks' IT spending on all categories of loan specialization using the same methodology.

A. Commercial and Industrial (C&I) Loans

Table 3 demonstrates the correlation between commercial and industrial loan and different types of banks' IT budget. Overall, our findings suggest that specialization in commercial and industrial loan is most positively and significantly correlated with banks' spending in communication technology. A one standard deviation higher loan portfolio share allocated to commercial and industrial loans predicts a 0.053 standard deviation higher communication budget as a share of total gross revenue, which is 0.33% higher as a share of revenue. In dollar terms, a 25% increase in commercial and industrial loan as a share of total loan means an average of $0.1 million higher budget per year to be spent on communication. Higher degree of specialization in commercial and industrial loan also predicts more spending on hardware, although the magnitude is slightly smaller than the prediction on communication budget. On the other hand, specialization in commercial and industrial loan does not predict a higher share of IT budget allocated to software spending. A one standard deviation higher specialization in commercial and industrial lending only predicts 0.015 standard deviation higher of software as a share of revenue, and is statistically insignificant.

Within C&I Loans In the next step, we decompose commercial and industrial loans into "Small Business Loans," as measured by a bank's small business lending reported in CRA, and "Other C& I loans". Row 2 and row 3 of Table 4 illustrate how small business loans compared to other C&I loans in terms of their correlations with banks' IT spending– in which we examine whether a higher portfolio share in small business loans or other C&I loans predicts any difference in banks' IT spending profiles. By comparing column (1) and column (2) of the two rows respectively, we see a sharp comparison: while higher small business loans in a bank' loan portfolio is positively and significantly correlated to its' communication spending intensity, it is negatively correlated to a bank's software spending; on the other hand, C&I lending to big firms is positively correlated to software spending, but shows no significant correlation with communication spending.

We also investigate a special type of loan (or lending form) – the syndicated loans, in which being a frequent lead bank versus playing more of the role as a participant bank implies different economic nature of the lending activities that banks engage. The results are reported in row 4 of Table 4. We postpone more detailed discussion about this comparison to Section 4.1, where we lay out a more clear conceptual framework to think about the distinctive economic meanings of different types of banks' IT investment.

B. Personal Loans

The second major category of loan type we examine are personal loans and mortgages. Row 5 of Table 4 reports the correlation between specialization in personal loan and mortgage and banks' IT spending. Contrary to commercial and industrial loans, a higher share of loan portfolio allocated to personal loans and mortgages appears to be predicting more investment on software rather than on communication technologies. A one standard deviation higher in the loan portfolio share on personal loan and mortgage predicts 0.0485 standard deviation higher software budget as a share of total gross revenue. That is to say, 7% increase in personal loan and mortgage in banks' loan portfolio is associated with 0.22% higher software as a share of total gross revenue, which means an average of \$0.65 million more software budget per year. On the other hand, higher personal and mortgage loan share does not have qualitatively significant predicting power on communication, hardware or services budget.

Within Personal Loans Paralleling our analysis within C& I loans, we also decompose personal loans and mortgage into two sub-categories. The first is mortgage refinancing and the other is everything else inside personal loans and mortgage. As shown in row 6 and row 7 (and compared with row 5) of Table 4 , it can be seen clearly that among the broad category of personal loans (including mortgage), mortgage refinancing as a sub-category is particularly positively correlated with banks' software spending. Motivated by these findings, in the next section where we study the economic meanings of banks' IT investment, we pay a particular attention to mortgage refinancing as a specific type of lending activity in which the processing of hard information plays a critical role.

Additionally, the abundance of mortgage data allows us to gain insights about the following important aspect in lending activities– originating a new loan versus refinancing an existing loan. In the context of mortgage lending, we investigate from banks' perspective whether being a refinancing lender versus an origination lender will be associated with different investment profile in information technologies. The results are reported in row 8 of Table 4. Again, we postpone more detailed discussion about this comparison to Section 4.1, under the context of understanding the economic meanings of bank IT investment.

C. Agriculture loans

Finally, we examine the correlation between agriculture loan specialization and banks' IT spending profile, the results of which is reported in row 9 of Table 4. Overall, we find that higher proportion of agriculture loan in banks' loan portfolio is positive correlated with spending on communication. A one standard deviation higher allocation of loan towards agriculture loan is associated with 0.056 standard deviation increase in communication budget. This means a 4.8% higher agriculture loan as a share of total loan is associated with 0.35% higher communication as a share of total gross revenue, which amounts to an average of $0.11 million more spending on communication per year.

4 Economics of Banks' IT Investment

Having elaborated the detailed definition and categorization of bank information technology investment, as well as having demonstrated basic patterns of the IT investment in the U.S. banking sector and the determinants of different types of IT investment, we now move on to a perhaps more important question. That is, what are the economics behind these bank IT spending, and in particular, how can they be related–and contribute–to the development of banks' lending technology? In this section, we seek answers to this question, which is important for a better understanding of how banks develop their lending technology in the information age.

We start with Section 4.1 by explaining how the bank lending technology during the information age could be related to banks' investment in information technology. We then provide some motivating evidences that give us some clues about the economic meaning of different types of IT investment made by banks, based on which we formalize several testable hypotheses. Finally, by exploiting policies that affect geographic regions differentially, we identify the causal relationships highlighted in these hypotheses based on the relevant cross-sectional analysis.

4.1 Bank Lending Technology, Information and IT Investment

What is a bank's lending technology? While the answer likely depends on the exact definition of "technology," the lending activities that banks engage in has always been associated with dealing with information about borrowers in relevant ways. In this sense, the evolution and development of a bank's lending technology could be largely viewed as how the bank's ability of dealing with information regarding their borrowers has evolved.

The linkage between the role played by information in banks' lending activity and banks' lending technology development in the information age is important yet unexplored. Broadly speaking, in conducting their lending business, the activities lenders need to engage in regarding the information on borrowers often can be categorized into two types/stages: information *production/transmission* and information *processing*. More specifically, information *production/transmission* refers to the stage in which information needs to be created or gathered on borrowers and then relayed to the hands of who later makes decisions based on these information; while information *processing* is more about the stage in which existed (or readily available) information on borrowers needs to be properly utilized by the lender for more efficient decision making.

Besides building the mapping between banks' lending technology development (in terms of information treatment capability) and banks' IT investment, identifying a causal relationship between the two is meaningful. Previous literature has shown how credit supply will positively affect non-financial firms' technology adoption or innovation (Amore et al. (2013), Chava et al. (2013), Bircan and De Haas (2019)), yet how credit demand associated with different information nature will induce the banking sector to upgrade their lending technology is still in the black box. The establishment of such causal relationship could help explain why credit cycles tend to be sustained, and can lead us to think more about whether unbalanced development across different types of information technology will tilt banks' lending towards or away from certain types of lending in the long run.

4.1.1 Communication Technology and Soft Information Production/Transmission

Let us start with information *production/transmission*. When faced with borrowers they have never dealt with or borrowers whose information structure is relatively opaque, the first thing lenders need to do is often to communicate with borrowers, so that they can at least draw a picture of the borrower that they can later on examine in detail. Such communications are often essential as the first step that allows the lender to gather information about their borrowers, through talking to them face-to-face or seeing borrowers' project themselves.

Once these first-hand information about borrower have been gathered, which often can be somewhat subjective and thus difficult to be conveyed to others, effective transmission of the gathered information within the lending organization can be another crucial factor affecting the lending efficiency. This problem associated with the credible transmission of hard-to-verify information within an organization has been recognized in previous studies (e.g. Stein (2002)). Efficient internal communication and exchange is particularly important when the relevant information is relatively soft, as soft information by definition is often less objective and often hard to be verified by someone who is not the first-hand collector of the information.

One concrete example of how communication technology can help in the two aforementioned dimensions is video conferencing, which has become an important way banks communicate with borrowers and customers during the past decade. New account openings, loan origination, and problem resolutions were done through in-person visit to the physical branches in the past, but are now being replaced by video conferencing technology as it makes the direct contact between loan officers and borrowers more timely and cost-saving.[13] Moreover, video conferencing among the employees within banks has also been welcomed by banking sector for its advantage of facilitating effective internal collaboration.[14]

Small Business Lending and Mortgage Initiation The lending to small business borrowers is one concrete example of such a situation. Sahar and Anis (2016) document that in the context of lending to small and medium size enterprises, direct contact with borrowers and frequent visit by loan officers to the borrower allow loan officers to produce soft information. Agarwal et al. (2011) highlight that soft information, such as what the borrower plans to do with the loan proceeds, plays an important role in the loan origination and underwriting process, and soft information like this is always the product of multiple rounds of interactions between the lender and the borrower.[15]

[13] See a real-world example of the communication tool "Liveoak" designed for banking services.

[14] See an article on Bankingdive for a detailed description of how video conferencing helps within-bank communication.

[15] The mapping between communication technology and soft information or information production can also be seen in our findings on how being a frequent lead bank in syndicate loan market will imply about a bank's technology adoption as shown in Table 4. The finding is in line with discussion on how the nature of interactions between lenders and borrowers differ if the lender is a lead bank versus if the bank is a participant bank: being a frequent lead bank means frequent communication, reporting, coordination.

The production and transmission of soft information is not only important from the angle of lenders, but also crucial from the perspective of borrowers. Recent industry reports show that 90% of home buyers, especially the *first-time* home buyers, cite that they want to directly speak with a loan officer to understand the product details and pricing options. A survey done by Wells Fargo reveals that even though they spend $500 millions on digitizing mortgage experience, 73% of home buyers wanted assurance that they could speak with a real person through the process.[16] Mobile phone devices, desktop telephone sets, Wi-Fi transmitters, video conferencing devices are all important communication infrastructure that ensures the communication between borrowers and lenders for the generation of soft information.[17]

Lead versus Participant (in Syndicated Loans) Syndicated loan market provides a special environment to explore the relationship between communication technology and soft information production. In syndicated loan lending, the nature of interactions between lenders and borrowers differs drastically if the lender is a lead bank versus if the lender is a participant bank (Sufi (2007), Ivashina (2009)). Lead bank are mandated by the borrower to acquire other lending participants, conducting compliance reports, and negotiate loan terms. After the loan is issued, they also have the responsibility to conduct monitoring, distribute repayments among and provide overall reporting among all lenders within the deal. In this regard, lead banks' job performing involves significantly heavier effort in information generation and sharing as well as coordination negotiations than that of a participant bank. In short, effective communication plays a more central role in the functioning of lead banks

[16] The report is summarized in this article.

[17] Not only does communication infrastructure help soft information generation at the loan origination stage, but also greatly enable lenders to closely monitor borrowers' business conditions after the loan is issued, especially for information that is not directly observable or verifiable. A concrete example is the agriculture loan. Effective monitoring of agriculture loan involves a good estimate of farmers' income and riskiness. But these are greatly dependent on the the observation of what's going on in the farm, is there pests, flood, is the weather normal and is the growth of products at its expected output level. Traditionally this is done by frequent in-person visit to the farmland by the loan officers, with the advances of technology, more and more of this monitoring is relying on video conferencing and even satellite imaging generation and transmission, to better achieve the monitoring. See this article describing the difficulty of monitoring agricultural borrowers faced by traditional lenders.

than that for participant banks.

Given this conceptual difference between the roles performed by lead bank versus participant bank in conduct syndicated lending, one empirical examination that naturally follows is whether the frequency that a bank participates in syndicated loans as a lead arranger can predict the IT investment behavior of the bank.

We described our empirical specification for answering this question in Section 3.3. Row 4 of Table 4 shows our finding. We find that banks' investment in communication technology exhibits a strong positive correlation with frequency they perform as lead bank in syndicate loans, whereas a negative correlation is seen between banks' software investment and their frequency in performing the role as a lead bank. To be more precise, a 0.5 standard deviation increase in the lead bank frequency is associated with 0.88 standard deviation increase in communication spending as a share of total revenues, and a 0.31 standard deviation decrease in software spending as a share of total revenues. This is equivalent to an increase of 0.56% of communication as a share of revenue, which amounts to $1.67 million higher in communication budget per year and $4.03 million lower software budget per year.

These findings motivate us to propose the following hypothesis on how banks' investment in IT could be mapped to their lending technology development, which will be causally identified in our later analysis in this section.

Hypothesis I

Banks' investment in communication IT is more associated with improving their ability of soft information production/transmission. Therefore, investing in communication IT is more effective when banks are faced with soft-information intensive borrowers. Specifically, in the context of small business lending, banks will respond by increasing their spending on communication technology when facing an increased credit demand from small business borrowers.

4.1.2 Software Technology and Hard Information Processing

We now move to information *processing*. Once information has been produced or is readily accessible, the next concerns for the lender is then how to most efficiently utilize these information and make wise decisions base on it. In the context of credit allocation, banks need to be able to properly evaluate the creditworthiness of their borrowers so that efficient decisions on the amount and pricing of credit can be made. More specifically, when banks are facing borrowers whose information structure is relatively transparent or those they already have some knowledge about from previous interactions, efficient credit making simply boils down to efficient utilization and processing of the available information.

Accurate evaluation of borrowers' credit risk often requires complicated calculations based comprehensive modeling and simulations, which are often impossible to be done without the support of sophisticated software tools. In terms of facilitating lending activities, the application of software is particularly useful for dealing with existed or readily accessible information that is relatively "hard." Banks have been actively adapting new software-based technologies to store, organize and analyze large chunk of loan applicants' data. Furthermore, many software products nowadays enable banks to conduct risk assessment and help banks make lending decisions based on the information augmented. In this regard, investing in software can greatly help bank to more promptly and accurately process the available information so that they can make their lending decision in a more efficient and reliable way.[18] As a sharp comparison, several decades ago all credit application information had to be stored on paper application form, reviewed and processed manually by loan officers.

From banks' perspective, a defining difference of software compared with communication technology is that while communication devices can facilitate the gathering and dissemination of information, the application of software is more targeted at utilizing the information

[18]For example "nCino" is operating system software that offers financial institutions to replace manual collection of loan/account applications with automated and AI based solutions. "Finaxtra" and "Turnkey" are both comprehensive loan origination systems that offer solutions for the whole lending process.

that is readily available at hand.[19] Because of this difference, software as a category of banks' IT budget should be more relevant in dealing with loans that are primarily about processing and assessment of existing information that is relatively "hard." In practice, a specific form of software technology product is the credit scoring software utilized by banks when making refinancing decisions. Some concrete examples of these credit scoring software include SAS Credit Scoring, GinieMachine, RNDPoint, etc.[20]

Following the logic discussed above, a natural examination of the distinctive roles played by software IT would be in the context of comparing the origination of new loans with the refinancing of existing loans, which we explore next.

Refinancing versus Origination (of Mortgage Loans) The comparison between refinancing and origination in the mortgage loan market provides a special environment to study how the processing of hard information could be related to banks' IT adoption.

In our analysis of the relationship between IT spending and banks' loan specialization in Section 3.3, we examined mortgage refinance as a particular type of loans and compare it to other personal loans. There we find mortgage refinance stands out as a special type of loan that appears to be particularly strongly correlated with banks' spending on software, as shown in rows 6 and 7 of Table 4. To sharpen our understanding of the role play by software in lending technology, we move one step further and conduct a similar analysis within banks' mortgage lending businesses, which can be broadly categorized into mortgage origination and mortgage refinancing. We find that if a bank is a frequent "refinance" lender in the mortgage market versus serving as an originator, which is measured by the number of refinance loan issued within the mortgage lending in the HMDA, this bank also tends to spend more on software. The detailed results are reported in row 8 of Table 4.

A one standard deviation increase in refinance business frequency is associated with

[19] This difference also explains why Fin-tech companies, which are equivalently the suppliers of new banking software products, have been expanding dramatically in mortgage refinance market Fuster et al. (2019), Seru (2019).

[20] To use these software, Banks usually just need to import the available demographic and historical information of borrowers, the software calculates credit scores and conduct statistical testings for robustness using AI and machine learning methodologies, which save banks from manual work and fasten the processing.

0.0805 standard deviation increase in software spending intensity. In real terms, this is about 0.36% of total gross revenue per year and amounts to an average of $1.07 million higher expense on software per year. Contrary to the positive significant correlation between software spending and mortgage refinance, communication spending shows no correlation with intensity in mortgage refinance business: a one standard deviation increase in refinance intensity is associated with a weakly of 0.0022 standard deviation of decrease in communication. Hardware and services, which are both complements to software, also shows positive correlation with mortgage refinance. A one standard deviation higher intensity in mortgage refinance lending is associated with an average of $0.22 million higher spending on hardware and $0.63 million higher spending on services.

Motivated by these findings, we propose and test the following hypothesis:[21]

Hypothesis II

Banks' investment in software IT is more associated with improving their ability of hard information processing. Therefore, investing in software IT is more effective when banks are faced with hard-information intensive borrowers. Specifically, in the context of mortgage refinance, banks are likely to increase their spending on software IT when facing an increased credit demand for mortgage refinancing purposes.

4.1.3 Labor Employment and Bank IT Investment

In addition to these two aspects of banks' lending technology and their relation to banks' investment in IT, we also pay attention to a third dimension in understanding the economics of banks' lending technology formation– the labor employment in the banking sector. In tandem with the wide application of automation and artificial intelligence, there has been an ongoing debate on how these technological progress would impact the demand for human workforce. In the context of banks lending technology investment, if we treat banks'

[21] Verification of this hypothesis will also be consistent with recent literature showing that the fintech expansion is particularly pronounced in the refinancing segment of mortgage, auto loans, and student loans market. (Drechsler et al. (2016)).

different types of employees as traditional input that deals with information production and information processing, one question naturally arises. That is, how would the adoption of new information technology that deals with information production/processing impact the job security of these bank employees? In particular, will banks' investment in IT substitute out or complement different types of bank employees that perform information related tasks?

Co-movement between IT-related Labor and IT Investment One natural conjecture is that the superior computing power and algorithmic calculation would enable IT to effectively take over the task of processing relatively hard information, while the generation of new and soft information would still have to rely (at least to some degree) on human labor. Thus, in the long run, high software spending tends to induce banks to decrease hiring of IT-related employees while high communication spending won't predict lower hiring of IT-related employees. Figure 7 displays the co-movement between the dynamics of banks' IT spending and IT-related employee hiring. Banks in the sample are divided into those with high communication/revenue, and high software/revenue during the ten-year from 2010-2019. It can be seen clearly that banks with high communication spending intensity have been steadily increasing their hiring of IT-related employees, especially after 2014; while in contrast, banks with high software spending intensity have started to decrease their IT-related employees since 2015. We formalize this conjecture as the following hypothesis.

Hypothesis III

Banks' adoption of information technologies has an asymmetric relationship with banks' IT related employment. Specifically, i) investment in software IT is a substitute to banks' IT-related employment, and therefore, banks tend to reduce their spending on software IT when the supply of IT-related labor is abundant; while ii) no such substitution effect exists between banks' investment in communication IT and their IT-related employment.

4.2 Soft Information Production and Bank IT Investment

In this section, we identify the causal linkage between demand of credit that relies heavily on the production/transmission of soft information and banks' investment in communication technology.

Small Business Credit Demand and Bank IT Investment Conceptually, credit demand of different economic nature is likely to entail different lending technology to be developed from the perspective of the bank sector. As a result, one should expect banks to react differently in their IT investment when faced with different types of credit demand, if different types of IT investment indeed map into different dimensions in banks' lending technology development.

One situation where the production of soft information plays a critical role is the lending to small business borrowers (Berger and Udell (2002), Liberti and Petersen (2018)). Small business loans refer to a special type of bank loans where banks mainly deal with firms that are unlisted, young and opaque, who often do not have publicly available credit history. The positive correlation between specialization in small business lending and banks' spending on communication technology has been demonstrated in Section 3.3, where we find that within the commercial and industrial loans, banks with higher shares of loan portfolio allocated to small business lending also exhibit a more intensive investment in communication technology. Following this logic, we regard an increased credit demand from small business borrowers as a demand for more intensive production of soft information. The goal of our identification analysis in this section is then to establish a causal relationship between a higher demand for credit from small business borrowers and an increase in banks' spending on communication technology.

To motivate our identification analysis, which will be based on exploiting certain exogenous variations across regions, we start off by demonstrating a set of scatter plots displaying banks' IT investment patterns and lending technology adoption when they are operating in different economic environments. In Figure 6 we display correlation between banks' IT investment and characteristics of local economy that banks operate in along two dimension along two dimension: 1) share of small business in local counties; and 2) growth rate of refinance mortgage refinance (will be discussed in the next section).[22] As can be seen clearly from upper panel of the figure, higher presence of small businesses (as measured by establishment with less or equal to 5 employees) in a local county is strongly positively correlated with communication spending intensity of banks operating in the county, but shows no significant correlation with software spending. These patterns suggest that in local areas featuring more small business operation and hence a higher demand for credit from small business borrowers, banks tend to tilt their investment in information technology more towards communication technology.

While these scatter plots provide a useful picture for us to think about the economic meanings of banks' IT investment, such a correlation is obviously subject to endogeneity issues as the credit demand itself is an equilibrium outcome that often depends on some latent factors which may also affect banks' IT investment. To overcome such endogeneity issues in our identification analysis, in the remainder of this section we exploit a policy event to construct cross-section variations in the small business credit demand.

Constructing Shifter in Small Business Credit Demand Our identification strategy relies on an event-driven credit demand shock that heterogeneously hit the US banking sector across different regions in terms of the demand for soft information generation. This feature allows us to extract and utilize variations that come from banks' exogenous exposure to such a credit demand shocks.

To achieve the identification goal, we utilize a special program in the Affordable Care Act enacted between 2014 and 2015, which is the "Small Business Health care Tax Credit." The Small Business Health Care Tax Credit Program was aimed at encouraging small businesses to provide health care coverage to their employees. The program offers tax credit to small business employers who pay health insurance premiums on behalf of employees. To qualify

[22] We put the regression results between banks' IT spending and local characteristics as well as additional results about local banking market structure in the Appendix Table A3, Table A6, and Table A4.

for the tax credit, the employer needs to satisfy the following requirements: (1) the employer has 25 or fewer employees; (2) the employer pays average wages of less than $50,000 a year per full-time equivalent; (3) pays at least 50% of its full-time employees' premium costs; (4) the health plan provided to employees is qualified under the SHOP program coverage requirements.[23]

Previous literature in economics has documented increases in labor demand, firm expansion, productivity growth, etc after the implementation of corporate tax cut or launch of subsidy (Cerqua and Pellegrini (2014), Rotemberg (2019)).[24] Economically, the launch of this program could boost local small businesses' credit demand because: (1) small business owners who have not provided their employees health coverage before the availability of the program might now have higher debt capacity to borrow from banks and fund their working capital; (2) with the tax credit subsidized to them, small businesses could potentially initiate the expansion and development projects that they wanted to do but were constrained by the limited working capital and debt capacity before the program was launched.

While the implementation of this program tends to shift up the credit demand from small business borrowers universally across the entire US economy, the exposure to this credit demand shock may vary substantially across different regions, depending on their pre-event characteristics. In particular, the number of qualified establishments at (or right before) the event date, which features substantial variation across different counties, plays an important role in determining the magnitude of credit demand from small business borrowers faced by banks in the local area.[25] Such a variation in the pre-event number of qualified

[23]See the link here directing to the introduction of the policy. According to IRS, small business employers should apply for the tax credit by filling in Form 8941. The tax credit can be carried backward or forward to other tax years. Also, since the amount of the health insurance premium payments is more than the total credit, eligible small businesses can still claim a business expense deduction for the premiums in excess of the credit, which means both a credit and a deduction for employee premium payments. This tax credit is also refundable, for small businesses that's tax-exempt and have no taxable income, they may be eligible to receive the credit as a refund so long as it does not exceed their income tax withholding and Medicare tax liability. According to IRS, the amount of the credit small businesses receive works on a sliding scale. The smaller the employer, the bigger the credit. The tax credit is highest in particularly for small companies with fewer than 10 employees, with an average annual salary of $25,000 or less.

[24]Ivanov et al. (2021) documented the increase in small businesses' credit demand after tax cut.

[25]Qualified establishments are those small businesses with less than 25 employees, whom the program is

establishments can thus help us isolate variations in small business credit demand across regions, from that of other types of credit demand which may otherwise affect our causal identification. Finally, since the policy is only explicitly targeting at small businesses, its impact on other types of credit demand would be relatively limited or at least less direct.

Empirical Design: 2SLS Regression Following our discussion above on the construction of county-level shifters that help us to extract variation in the small business credit demand across regions, we run the following diff-in-diff regression. In this diff-in-diff specification, we use the number of small businesses with less than 20 employees in 2013 as an instrumental variable of local county's exposure to the program, and importantly we control for the total number of establishments in 2013.

$$\begin{aligned} \Delta \ln(\text{CRA})_{i,c,post} &= \mu_i + \mu_1 \ln(1 + \text{Qualified Small Buzs})_{c,pre} + \mu_2 X_{i,c} + \epsilon_{i,c} \\ \Delta \ln \text{IT}_{i,c,post} &= \alpha_i + \beta \Delta \widehat{\ln(\text{CRA})}_{i,c,post} + \gamma X_{i,c} + \epsilon_{i,c} \end{aligned} \tag{2}$$

The first equation in Equation 2 is the first- stage regression. The outcome variable $\Delta\text{Ln}(\text{CRA}_{i,c,\text{post}})$ is the change in the logarithm of bank i's small business loans in county c in the three-year time window after the program implementation compared to the its small business loans in the same county in the three-year time window before the program implementation. $ln(1 + \text{Qualified Small Buzs})_{c,pre}$ is the logarithmic of the total number of small businesses with less than 20 employees. X is a vector of control variables which include the bank-level control variable, and the county-level control variable.

The second equation in 2 is the second-stage regression equation. The left-hand side variable $\Delta\text{Ln IT}_{i,c,t}$ is the change in logarithm of a specific type of IT spending of bank i in county c after the program implementation compared with before the program implementation. The explanatory variable of the second-stage regression is the fitted value of changes in small business loans made by bank i in county c. Control variables include "Net income/Assets," "Equity/Assets," Deposits/Assets," and "Revenue per Employee," the first

specifically targeted at.

three control variables are from Call Report and is at bank-year level, because there's no equity, net income, salary, and total asset information of a bank at county level. Revenue per employee is at bank-county-year level, the total revenue is from Summary of Deposits, and total number of employees is from Aberdeen. County level control variables include the labor force, population growth rate, and total number of establishments.

Estimation Results The estimation results for the above Diff-in-Diff identification analysis are reported in the first three columns of Table 5. Column (1) shows the regression in the first-stage regression, and column (2) and (3) show results of the second stage. In column (1) we find that a one standard deviation increase in the number of qualified small businesses in a county where a bank is located explains 0.171 standard deviation higher increase in that banks' small business lending in that county compared with the change in small business lending of the same bank in other counties. That is, a one-standard deviation higher exposure to the number of qualifies small businesses means 32% higher increase in small business loans of the same bank compared with the spending of its branches in other county. The F-statistics of the first stage regression is 15.18, which is well above the conventional threshold for weak instruments (Stock and Yogo (2005)). The magnitude of this regression parallels the magnitude in similar research findings such as Adelino et al. (2017), where in their Bartik-style instrumental variable approach, they find that a 1 percentage point increase in exposure to manufacturing employment growth is associated with 1.146 percentage point increase in local income growth.

As reported in column (2) and column (3), we find that there is no significant response in banks' software spending post the program launch, but there is a positive and statistically significant response in banks' communication investment across counties. In particular, banks saw one standard deviation higher growth in their small business loan, due to the higher shock exposure of the counties they operate in, experienced a 0.0745 standard deviation higher growth in their communication spending. That is to say, a one standard deviation higher growth in banks' small business loans post the event (due to higher exposure of the county towards the policy event ex ante), translated into 26% higher communication spending growth compared with banks located in other lower exposure counties. Furthermore, thanks to the granularity of our data which allows us we to include bank fixed effects, the above interpretation can be viewed as the within-bank and cross-county estimation results.

Comparison: OLS Estimates To gain more insights about our 2SLS estimates, we also compare them to the estimates from the OLS regression, which is specified as follows:

$$\Delta \ln(\mathrm{IT})_{i,c,post} = \alpha_i + \beta \times \Delta \ln \mathrm{CRA}_{i,c,post} + \mu_c + \gamma X_{i,c} + \epsilon_{i,c} \quad (3)$$

The results of this paralleling regression are reported in column (4) and column (5) of Table 5. Bank and county level control variables are the same as in the 2SLS. Overall, the results of OLS specification are consistent with that in 2SLS in terms of the sign and statistical significance. Within a bank, its branches in counties seeing higher growth rate in small business loans is also seeing higher growth rate in communication spending compared with its branches in other counties with lower growth rate in small business loans. But there is no significant correlation between small business loan growth rate and software spending. Specifically, a one standard deviation higher growth rate in small business loans predicts 0.0146 standard deviation higher growth rate in communication spending, but only 0.00307 standard deviation higher growth rate in software spending.

However, there exists a noticeable difference between the two estimates. The magnitudes of the coefficients are significantly smaller in OLS compared with that in 2SLS. One potential explanation for such a downward bias in OLS estimator could be the following "omitted variable" problem, in which counties with faster growth in the overall amount of small business loans are those with even faster growth in some unobservable economic variables—say mortgage demand—that drives local banks to spend less on communication. More specifically, if mortgage demand is correlated with that of small business loans, and if banks have a fixed amount of IT budget each year, then they will allocate more IT spending, say software, to cater mortgage demand. The omitted-variable problem then leads the OLS estimator to underestimate the responsiveness of communication spending towards small business loan demand. The cross-region shifter constructed above based on the tax policy targeting only at small business thus allows us to distill the variation in small business credit demand from that of other types of credit.

4.3 Hard Information and Bank IT Investment

In this section, we identify the causal linkage between credit demand that requires intensive processing of hard information and banks' investment in software technology.

Mortgage Refinance Demand and Bank IT Investment By definition, "hard" information refers to the kind of information that is existent, easy to quantify and does not require further verification. In terms of lending, loan origination and refinance gives an ideal comparing environment to investigate the relationship between "hard" information and banks' lending technology investment.

Mortgage refinance is a special type of loan that primarily involves processing of existing information, as refinance applicants by definition have already been granted loans before. Typically refinance application requires applicants to provide their W-2, Pay stubs, financial account statements, most recent monthly statement on the mortgage, and sometimes the appraisal documents of the house. Many banks are offering complete online refinance applications so that borrowers do not even need to visit bank offices in person. These types of procedures often can be done solely through digitization software and risk assessment software,etc. This is in drastic contrast to the origination of mortgages, as discussed before, in which case both borrowers and lenders have a non-trivial incentive to communicate with each other. Results on the correlations between the weight of mortgage refinance in banks' loan portfolio and banks' IT spending as well as the frequency of refinance in all mortgage lending and banks' IT spending were in Section 3.3 and Section 4.1 respectively. These findings indicate that software spending is especially useful for processing hard information involved in mortgage refinance.

Similar to our analysis of the small business credit demand in the previous section, here we also motivate our cross-region identification with a set of scatter plots displaying how banks' IT investment correlates with the mortgage refinance demand in the local area they operate in. The lower panel of Figure 6 displays these correlations– in which we can clearly see that the average refinance growth rate in a local area exhibits quite different correlation with the software and communication IT spending made by local banks. In the remainder of this section, to causally identify this relationship, we construct shifter that allows us to extract exogenous variation in the refinancing demand at across county level.

Constructing Shifter in Mortgage Refinance Demand Based on the analysis above discussion, we treat mortgage refinance as a special type of credit which particularly involves banks' dealing with "hard" information and investigate the causal linkage between hard information processing and banks' response in their software spending. To achieve the goal of identification, we need to find exogenous sources of variation in mortgage refinance demand across different regions during a specific period. For this purpose, we utilize the post-crisis period as the period during which the mortgage interest rate is systemically low and aggregate demand for mortgage refinance is high.

Figure 5 displays the evolution of nation-wide average interest rate of mortgage before and after the financial crisis. The blue line shows the aggregate 30-year mortgage rate of U.S. from 2003 to 2019. There has been a dramatic decrease in mortgage rate since 2009– the interest rate on mortgages plummeted from the pre-crisis level of 6.5% to 3.75%. The big nation-wide decrease in mortgage rate therefore greatly pushes existing home owners with mortgage balances to refinance their mortgage. We utilize the 3-year episode of 2010-2012 as the time window and utilize the heterogeneity of each county's refinance propensity to identify whether a given bank's branches located in counties with high mortgage refinance propensity also increased their spending on software dramatically in those three years, compared to branches in low-propensity counties, in order to cater to the increased refinancing demand.

Beraja et al. (2018) also documented the strong refinancing activities during the low interest rate episode and used a similar empirical specification as follows in our paper.

Given the systemically and nation-wide low interest rate in the post-crisis period, an important determinant of homeowners' refinance propensity would be the pre-crisis mortgage characteristics that are existent before the low-interest episode kicks in while not determined by concurrent economic factors. Based on this logic, we construct the following county-level refinance propensity as an instrumental variable for the post-crisis local refinance activities:

$$\text{Payments gap}_c = Ave.(\text{Payments}_{\text{Old interest rate}} - \text{Payments}_{\text{New interest rate}})_c$$

which is the average total remaining mortgage payment gap under old versus new interest rates at county level. In constructing this measure, we use information of all the outstanding mortgage loans of local household and their mortgage rate at issuance at a county since 2000. Loans that were defaulted and prepaid were removed to ensure the measurement captures only refinance propensity from local household with outstanding loans.

Importantly for our identification purpose, the county-level payment gap measure as constructed above features significant variation across regions. Eichenbaum et al. (2018) also showed that the long-term low interest rate had a substantial but heterogeneous impact on the regional refinancing by households located in different states, depending on how much savings households can make under the different interest rate scenario. This variation in local homeowners' payment gap from interest saving allows us to construct exogenous shifter on the mortgage refinance demand faced by banks operating in the local economy.

Identification Design: 2SLS Regression The regression specification using Payments gap_c as the instrumental variable is:

$$\begin{aligned} \ln \text{Refinance}_{i,c} &= \mu_i + \mu_1 \text{Payments gap}_c + \mu_2 X_{i,c,t} + \epsilon_{i,c,t} \\ \left(\frac{\text{Type S IT Spending}}{\text{Revenue}}\right)_{i,c,t} &= \alpha_i + \beta \ln \widehat{\text{Refinance}}_{i,c} + \gamma X_{i,c,t} + \epsilon_{i,c,t} \end{aligned} \tag{4}$$

The first equation in equation (4) is the first-stage regression using Payments gap_c as the instrumental variable. The left-hand side variable in the first stage is the average of the logarithmic refinance loan volume of a bank i in county c during 2010-2013, compared with 2007-2009. The second equation is the second-stage regression regressing the IT spending intensity of bank i in county c on the fitted value of $\Delta \ln \text{Refinance}_{i,c}$, which identifies whether IT investment that specializes at processing existing information increases when there's higher local refinance demand.

To construct the left-hand side variable of the second-stage regressions, we aggregate all branches spending on a specific category of bank i in county c in year t, and scale it by the aggregated revenue of all branches of that bank in that county. Total revenue of a branch is from FDIC Summary of Deposit Database. To guarantee the accuracy of this measure, we require that a banks' total revenue is not missing in that county and that the total employee of a banks' branches in this county is not missing in the Aberdeen database when it was surveyed. Control variables include "Net income/Assets," "Equity/Assets," "Deposits/Assets," and "Revenue per Employee," the first three control variables are from Call Report and is at bank-year level, because there's no equity, net income, salary, and total asset information of a bank at county level. Revenue per employee is at bank-county-year level, the total revenue is from Summary of Deposits, and total number of employees is from Aberdeen. County level control variables include the unemployment rate, population growth and local house price index.

Estimation Results The first three columns of Table 6 reports the result of the above set of equations. Column (1) shows the result of first-stage regression, a one standard deviation higher interest payment saving of county c's mortgage borrowers is associated with 0.863 standard deviation higher logarithm of mortgage refinance during the low interest rate period of a bank's local branches compared with banks located in counties with lower interest saving. The F-statistics is 2408.16.

Column (2) and (3) show the results of second-stage regressions. A one standard deviation higher mortgage refinance lent out by a bank (due to its local exposure to high refinance savings) leads to 0.0672 standard deviation higher (around 8% higher on average) software spending intensity measured as software spending scaled by total revenue compared with other banks located in potentially low exposure counties. This result can also be interpreted as the within-bank cross-county impact of mortgage refinance demand on banks' software spending. On the other hand, communication spending intensity as a share of total revenue does not demonstrate significant changes in response to the refinancing demand captured by the pre-determined refinance propensity during the low mortgage rate episode.

Comparison: OLS Estimates As a comparison with the 2SLS, we also run the paralleling OLS specification as follows:

$$\frac{\text{Type S IT Spending}}{\text{Revenue}}_{i,c,t} = \alpha_i + \beta \times \ln \text{Refinance}_{i,c,t} + \mu_t + \gamma X_{i,t} + \epsilon_{i,c,t} \quad (5)$$

The regression results are reported in the last two columns of Table 6. We find that in response to a one standard deviation increase in mortgage refinance in a county c in year t, banks in that county in that year increase their software spending intensity by about 0.02 standard deviation on average. That is roughly equivalent to 2% higher software as a share of total revenue and bank local branch level. But there is no significant response by local banks in that county in that year in terms of their spending on communication. On the other hand, if local small business credit demand goes up in a county c in year t by one standard deviation, local banks communication spending intensity goes up by 0.00739 standard deviation. Both the magnitude and the significance are lower compared with that correlation with software/revenue. The results are robust when bank level control variables and county control variables are added.

Again, we find that the estimate of OLS regression is smaller in magnitude compared with those of 2SLS.[26] The reason is similar to that discussed in the previous section. Coun-

[26]Table 4 show the results of the same OLS specification with bank, year and county fixed effects and bank×year and county fixed effects.

ties seeing more mortgage refinance issued by local banks could also have other loan demand which recovered more significantly during the post-crisis period, say commercial and industrial loans, agriculture loans or even mortgage origination, which might tilt banks' IT budget towards other types of IT spending. Likewise, this omitted-variable problem will make OLS underestimate local banks' software spending response towards mortgage refinance propensity. In the 2SLS estimation, fitted value constructed from the first-stage regression resolves the issue, by extracting the cross-region variation in mortgage refinance issuance that is driven only by the pre-period local mortgage market characteristics determining ex-post refinance propensity.

4.4 Bank IT Investment and Labor

How firms' technology adoption and automation of tasks will affect their labor employment has been one of the focal points of recent economic studies. The labor employment in the banking sector is also subject to, if not more, the impact of the fast developing information technologies. In this section, we study the relationship between banks' IT investment and their labor employment outcomes. Specifically, we investigate whether IT investment is a substitute or complement to banks' IT employees and Non-IT employees.

As a starting point, table A9 shows the panel regression of the correlation between banks' labor hiring outcomes and banks' IT spending decisions. Consistent with our examination of the time trends of banks' IT investment and labor employment in section 3.1, we find that banks' IT-related labor employment exhibits a clear negative correlation with banks' software spending but a somewhat positive correlation with the spending on communication technology. For the main part of this section, we investigate whether there is a causal linkage between banks' IT spending and their labor employment structure. A key challenge in this causal identification analysis is that the observed IT employee growth growth in a county's local banking sector could be related to certain demand side factors in the local area which may also have an impact on local banks' IT investment, making the cause or the direction of causal effect unclear.

Our identification strategy is based on exploiting variation in the supply of skilled labor across difference regions. To be more concrete, we examine whether banks located in counties with more abundant supply of high-skilled labor, which are the the main source of IT-related employees of the local banking sector, will display different IT investment patterns when compared with banks in operating in other areas.

In this regard, we utilize an instrumental variable approach that has been widely utilized in previous literature of labor studies such as Moretti (2004) and Liu (2015). The main idea is to instrument for IT-related employees growth using the presence of land-grant colleges in the local area. Land-grant colleges were the universities supported by the first federal program to support higher education.[27] The local availability of a land-grant college is a strong predictor of skilled labor supply in local area Moretti (2004) and is shown to be unrelated to other local economic condition changes Kantor and Whalley (2019). We instrument for the growth of IT-related employees by a measure of the local availability of land-grant college: this measure is either a dummy variable which equals 1 if there is a land-grant college located in the county; or this measure is the distance between this county and the nearest land-grant college. The 2SLS regression specification is specified as follows:

$$\begin{aligned} \Delta\left(\text{IT Emp/Total Emp}\right)_{i,c} &= \alpha_i + \gamma \cdot \text{Land-grant college availability}_c + \mu X + \epsilon_{i,c} \\ \Delta\text{Software}_{i,c} \text{ or } \Delta\text{Communication/Revenue}_{i,c} &= \alpha_i + \beta \Delta \left(\widehat{\text{IT Emp/Total}} \text{ emp}\right)_{i,c} + \phi X + \epsilon_{i,c} \end{aligned}$$

$\Delta\ (\text{IT Emp/Total Emp})_{i,c}$ is the difference of the 2015-2019 IT employee share in total employee of bank i in county c and the 2010-2014 IT employee share. Land-grant college availability$_c$ is either $\mathbb{1}[\text{Land-grant college}]$, an indicator variable which equals to 1 if there's a land-grant

[27] In 1862, the U.S. Congress passed the Morrill Act and granted federally controlled land to the states establish colleges that focuses the teaching of agriculture, science and engineering. A second follow-up act was launched in 1890. In 1994, the act was further amended and supplemented. Altogether, there are 106 land-grant colleges in the United States. Many of the land-grant colleges have grown into large public universities have educated almost one-fifth of all students seeking degrees in the United States.

college, or (Ln min distance), the logarithm of the distance between the county and its nearest land-grant college. Table 7 shows the results of the above regression specification. Column (1) is the result of the first-stage regression and column (2) and (3) are the second-stage regression. Panel A of the table shows the results using the indicator variable whether there's a land-grant college in the county as an instrumental variable and Panel B shows the results using the logarithm of distance between the county where the bank is located and its nearest land-grant college.

The first-stage results show that availability of land-grant college strongly predict local banks' IT employees growth from 2015-2019 relative to 2010-2014. The second-stage regression results show that banks located in counties with faster growth in IT-employees (due to the fact that the county has higher land-grant college availability) saw significantly slower growth in software spending, while there's no statistically significant difference in communication spending growth compared with counties with lower degree of land-grant college availability. Quantitatively, our results show that a one standard deviation higher growth in IT employee share due to the availability of Land-grant colleges results in 0.0463 standard deviation lower growth in software spending as a share of revenue, which is about 5.13% of the branch-level revenue and about 55.1% of local branches' non-interest expenses. result implies that when there's more demand on generation of new information, employee hiring related to information processing will not necessarily change dramatically compared with employees that directly deal with information generation.

5 Conclusion

Development of information technologies over the past several decades has dramatically revolutionized the way lending is conducted by the traditional banking sector. In this paper, we provide the first comprehensive study of banks' spending on information technologies – as an investment to develop and improve their lending technology in the information age.

The detailed information on the IT spending profile available in our unique data set enables us to uncover several novel findings. First, at the aggregate level, we document an overall fast growing trend in banks' spending on modern information technologies, and how that varies across banks of different sizes. Second, we show that the different types of information technology is closely related to the information nature embedded in different types of loan making: the production and transmission of "soft" information, which plays a crucial role in conducting small business lending, is strongly correlated with banks' communication spending; whereas "hard" information processing, which is the most relevant for conducting mortgage refinancing, strongly correlates with banks' software spending. Finally, we provide the causal identification between the credit demand associated with different information nature and banks' IT spending response using granular bank-county level data, as well as causally identify the substitute/complement relationship between the labor employment and IT investment in the US banking sector.

Our findings open up several possible directions for future research; and we list a few here. How does the endogenous technology adoption in the banking sector transform the banking/credit market structure? Does the presence of Fintech impede or spur banking sector's technology adoption? How does banking sector's technology upgrading affect banks' deposit taking activities, loan outcomes, properties of credit cycle, monetary policy transmission?

References

Acemoglu, D. and Restrepo, P. (2018). The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment. *American Economic Review*, 108(6):1488–1542.

Adelino, M., Ma, S., and Robinson, D. (2017). Firm age, investment opportunities, and job creation. *The Journal of Finance*, 72(3):999–1038.

Agarwal, S., Ambrose, B., Chomsisengphet, S., and Liu, C. (2011). The role of soft information in a dynamic contract setting: Evidence from the home equity credit market. *Journal of Money, Credit and Banking*, 43:633–655.

Aiello, D., Garmaise, M. J., and Natividad, G. (2020). Competing for deal flow in local mortgage markets. *Working Paper*.

Akerman, A., Gaarder, I., and Mogstad, M. (2015). The skill complementarity of broadband internet. *The Quarterly Journal of Economics*, 130(4):1781–1824.

Amore, M. D., Schneider, C., and Žaldokas, A. (2013). Credit supply and corporate innovation. *Journal of Financial Economics*, 109(3):835–855.

Autor, D. H., Levy, F., and Murnane, R. J. (2003). The Skill Content of Recent Technological Change: An Empirical Exploration. *The Quarterly Journal of Economics*, 118(4):1279–1333.

Banna, H. and Alam, M. R. (2021). Is digital financial inclusion good for bank stability and sustainable economic development? evidence from emerging asia. *Working Paper*.

Beaudry, P., Doms, M., and Lewis, E. (2010). Should the personal computer be considered a technological revolution? evidence from u.s. metropolitan areas. *Journal of Political Economy*, 118(5):988–1036.

Beraja, M., Fuster, A., Hurst, E., and Vavra, J. (2018). Regional Heterogeneity and the Refinancing Channel of Monetary Policy. *The Quarterly Journal of Economics*, 134(1):109–183.

Berger, A. N. (2003). The economic effects of technological progress: Evidence from the banking industry. *Journal of Money, Credit and Banking*, 35(2):141–176.

Berger, A. N. and Black, L. K. (2011). Bank size, lending technologies, and small business finance. *Journal of Banking and Finance*, 35(3):724–735.

Berger, A. N. and Udell, G. F. (2002). Small business credit availability and relationship lending: The importance of bank organisational structure. *The Economic Journal*, 112(477):F32–F53.

Berger, A. N. and Udell, G. F. (2006). A more complete conceptual framework for sme finance. *Journal of Banking and Finance*, 30(11):2945 – 2966.

Bircan, C. and De Haas, R. (2019). The Limits of Lending? Banks and Technology Adoption across Russia. *The Review of Financial Studies*, 33(2):536–609.

Bloom, N., Garicano, L., Sadun, R., and Reenen, J. V. (2014). The distinct effects of information technology and communication technology on firm organization. *Management Science*, 60(12).

Bloom, N. and Pierri, N. (2018). Cloud Computing Is Helping Smaller, Newer Firms Compete. *Working Paper*.

Bloom, N., Sadun, R., and Van Reenen, J. (2012). Americans Do IT Better: US Multinationals and the Productivity Miracle. *American Economic Review*, 102(1):167–201.

Bolton, P., Freixas, X., Gambacorta, L., and Mistrulli, P. E. (2016). Relationship and Transaction Lending in a Crisis. *The Review of Financial Studies*, 29(10):2643–2676.

Brynjolfsson, E. and Hitt, L. (2003). Computing Productivity: Firm-Level Evidence. *The Review of Economics and Statistics*, 85(4):793–808.

Brynjolfsson, E. and Hitt, L. (2018). Investment in ICT, Productivity, and Labor Demand : The Case of Argentina. *Policy Research Working Paper No. 8325.*

Buchak, G., Matvos, G., Piskorski, T., and Seru, A. (2018). Fintech, regulatory arbitrage, and the rise of shadow banks. *Journal of Financial Economics*, 130(3):453–483.

Calebe de Roure, L. P. and Thakor, A. V. (2019). P2P Lenders versus Banks: Cream Skimming or Bottom Fishing? *SAFE Working Paper No. 206.*

Cerqua, A. and Pellegrini, G. (2014). Do subsidies to private capital boost firms' growth? A multiple regression discontinuity design approach. *Journal of Public Economics*, 109:114–126.

Chava, S., Oettl, A., Subramanian, A., and Subramanian, K. V. (2013). Banking deregulation and innovation. *Journal of Financial Economics*, 109(3):759–774.

Di Maggio, M. and Yao, V. (2020). Fintech borrowers: lax Screening or cream-skimming? *The Review of Financial Studies.*

Drechsler, I., Savov, A., and Schnabl, P. (2016). "The Deposits Channel of Monetary Policy". *The Quarterly Journal of Economics, vol 132(4), pages 1819-1876.*

Eichenbaum, M., Rebelo, S., and Wong, A. (2018). State dependent effects of monetary policy: the refinancing channel. Working Paper 25152, National Bureau of Economic Research.

Erel, I. and Liebersohn, J. (2020). Does finTech substitute for banks? Evidence from the paycheck protection program. *Working Paper.*

Forman, C., Goldfarb, A., and Greenstein, S. (2012). The internet and local wages: A puzzle. *American Economic Review*, 102(1):556–75.

Freixas, X. and Rochet, J.-C. (2008). *Microeconomics of Banking, 2nd Edition.* MIT Press Books. The MIT Press.

Frost, J., Gambacorta, L., Huang, Y., Shin, H. S., and Zbinden, P. (2019). Investment in ict, productivity, and labor demand : The case of argentina. *BIS Working Papers.*

Fuster, A., Plosser, M., Schnabl, P., and Vickery, J. (2019). The role of technology in mortgage lending. *The Review of Financial Studies*, 32(5).

Hannan, T. H. and McDowell, J. M. (1984). The determinants of technology adoption: The case of the banking firm. *The RAND Journal of Economics*, 15(3):328–335.

Hauswald, R. and Marquez, R. (2003). Information technology and financial services competition. *The Review of Financial Studies*, 16(3):921–948.

He, Z., Huang, J., and Zhou, J. (2021). Open banking: Credit market competition when borrowers own the data. *Working Paper*.

Hitt, L., Frei, F., and Harker, P. (1999). How financial firms decide on technology. *Brookings Wharton Papers on Financial Services*.

Hornuf, L., Klus, M. F., Lohwasser, T. S., and Schwienbacher, A. (2018). How do banks interact with fintechs? forms of alliances and their impact on bank value. *CESifo Working Paper*.

Hughes, J., Jagtiani, J., and Moon, C.-G. (2019). Consumer lending efficiency:commercial banks versus a fintech lender. *FRB of Philadelphia Working Paper No. 19-22*.

Ivanov, I., Pettit, M. L., and Whited, T. (2021). Taxes depress corporate borrowing: Evidence from private firms. *Working Paper*.

Ivashina, V. (2009). Asymmetric information effects on loan spreads. *Journal of Financial Economics*, 92(2):300–319.

Jagtiani, J. and Lemieux, C. (2017). Fintech lending: Financial inclusion, risk pricing, and alternative information. *FRB of Philadelphia Working Paper No. 17-17*.

Kantor, S. and Whalley, A. (2019). Research proximity and productivity: Long-term evidence from agriculture. *Journal of Political Economy*, 127(2):819–854.

Liberti, J. M. and Mian, A. R. (2009). Estimating the effect of hierarchies on information use. *The Review of Financial Studies*, 22(10):4057–4090.

Liberti, J. M. and Petersen, M. A. (2018). Information: hard and soft. *The Review of Corporate Finance Studies*, 8(1):1–41.

Liu, S. (2015). Spillovers from universities: Evidence from the land-grant program. *Journal of Urban Economics*, 87:25–41.

Lorente, C., Jose, J., and Schmukler, S. L. (2018). The fintech revolution: A threat to global banking? Research and Policy Briefs 125038, The World Bank.

McElheran, K. and Forman, C. (2019). Firm organization in the digital age: It use and vertical transactions in u.s. manufacturing. *Working Paper*.

Moretti, E. (2004). Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data. *Journal of Econometrics*, 121(1):175–212. Higher education (Annals issue).

Paravisini, D. and Schoar, A. (2016). The incentive effect of scores: Randomized evidence from credit committees,. *NBER Working Paper, 19303*.

Petersen, M. A. and Rajan, R. G. (2002). Does distance still matter? the information revolution in small business lending. *The Journal of Finance*, 57(6):2533–2570.

Philippon, T. (2020). On fintech and financial inclusion. *Working Paper*.

Pierri, N. and Timmer, Y. (2020). Tech in fin before fintech: Blessing or curse for financial stability? *IMF Working Paper, No. 20/14*.

Rotemberg, M. (2019). Equilibrium effects of firm subsidies. *American Economic Review*, 109(10):3475–3513.

Sahar, L. and Anis, J. (2016). Loan officers and soft information production. *Cogent Business & Management*, 3(1):1199521.

Schnabl, P. and Gopal, M. (2020). The rise of finance companies and finTech lenders in small business lending. *Working Paper*.

Seru, A. (2019). Regulating banks in the era of fintech shadow banks. *BIS Speach 2019*.

Stein, J. C. (2002). Information production and capital allocation: Decentralized versus hierarchical firms. *The Journal of Finance*, 57(5):1891–1921.

Stock, J. and Yogo, M. (2005). *Testing for Weak Instruments in Linear IV Regression*, pages 80–108. Cambridge University Press, New York.

Stulz, R. M. (2019). FinTech, BigTech, and the future of banks. *Journal of Applied Corporate Finance*, 31(4):86–97.

Sufi, A. (2007). Information asymmetry and financing arrangements: Evidence from syndicated loans. *The Journal of Finance*, 62(2):629–668.

Tang, H. (2019). Peer-to-Peer lenders versus banks: Substitutes or complements? *The Review of Financial Studies*, 32(5):1900–1938.

Vives, X. and Ye, Z. (2021). Information technology and bank competition. *Working Paper*.

Figure 1. Total Asset of Banks in Sample

This figure shows the sum of total asset size of all banks in our sample from 2010 to 2019 U.S. The red dashed line is the sum of call commercial banks' asset size in U.S., data source is Board of Governors of the Federal Reserve System (US), Total Assets, All Commercial Banks [TLAACBW027SBOG]. The red solid line is the sum of total asset sizes of banks in our sample. The black solid line is the sample bank size out as a share of total nation-wide banks' total asset size.

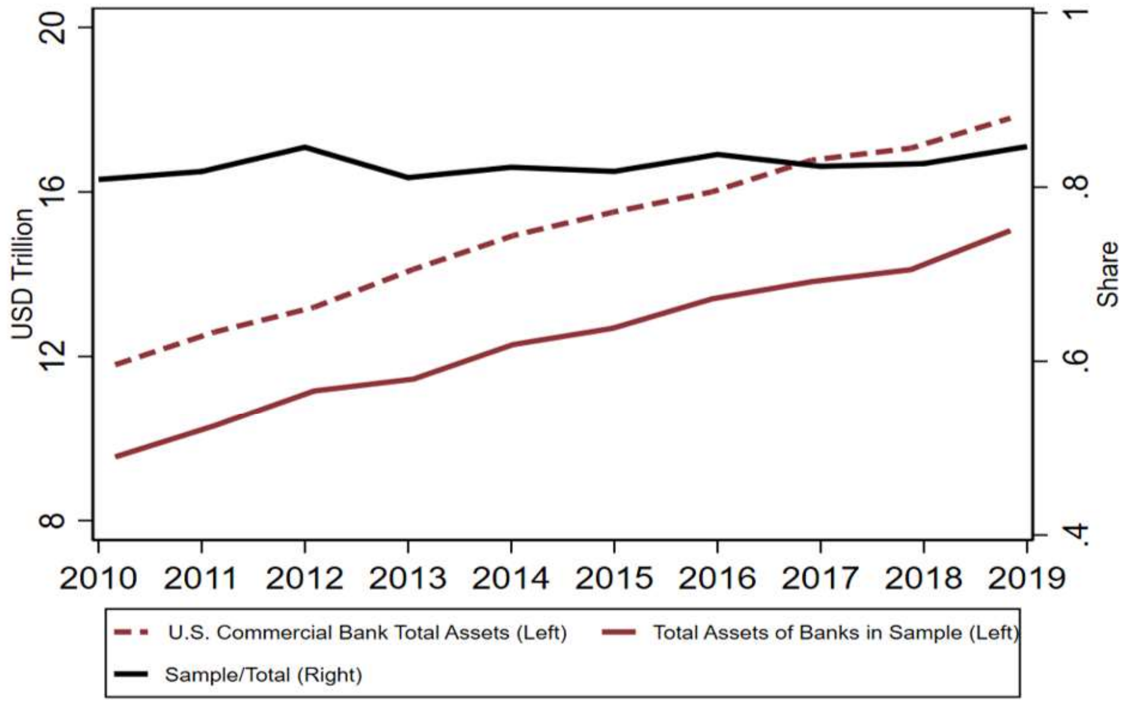

Figure 2. IT Spending: Time Trend

This figure shows the time trend of banks' IT spending from 2010 to 2019. The solid line shows the weighted average of banks' IT Spending as a share of banks' total income, weighted by banks' asset size; the dashed line shows the weighted average of banks' IT spending as a share of banks' total income, weighted by banks' asset sizes in each given year. "Total income" is constructed using the item RIAD4000 in Call Report, and "Total Expenses" is the non-interest expenses reported by item RIAD4093 in Call Report.

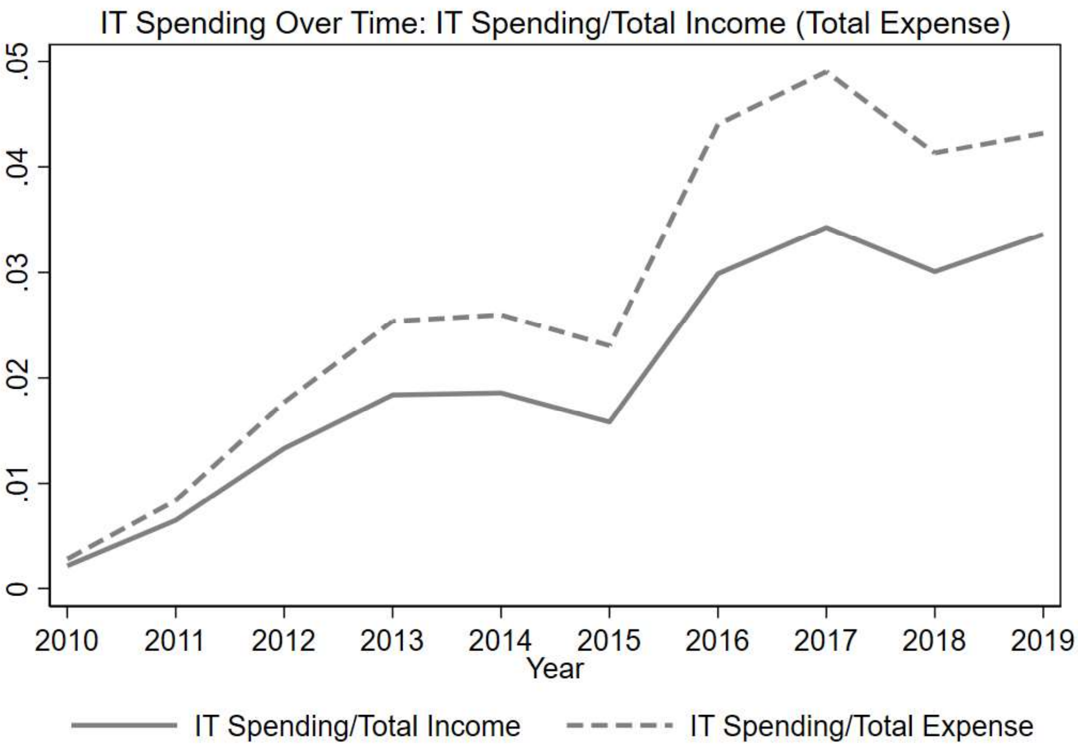

Figure 3. IT Spending Time Trend: By Bank Size

This figure shows the time trend of banks' IT spending from 2010 to 2019 by the five categories of bank asset size groups. The vertical axis is banks' total IT spending scaled by non-interest expenses. The asset size groups are categorized based on a bank's average asset size during 2010 and 2019. Non-interest expenses are calculated using banks' balance sheet item "RIAD4093" in Call Report.

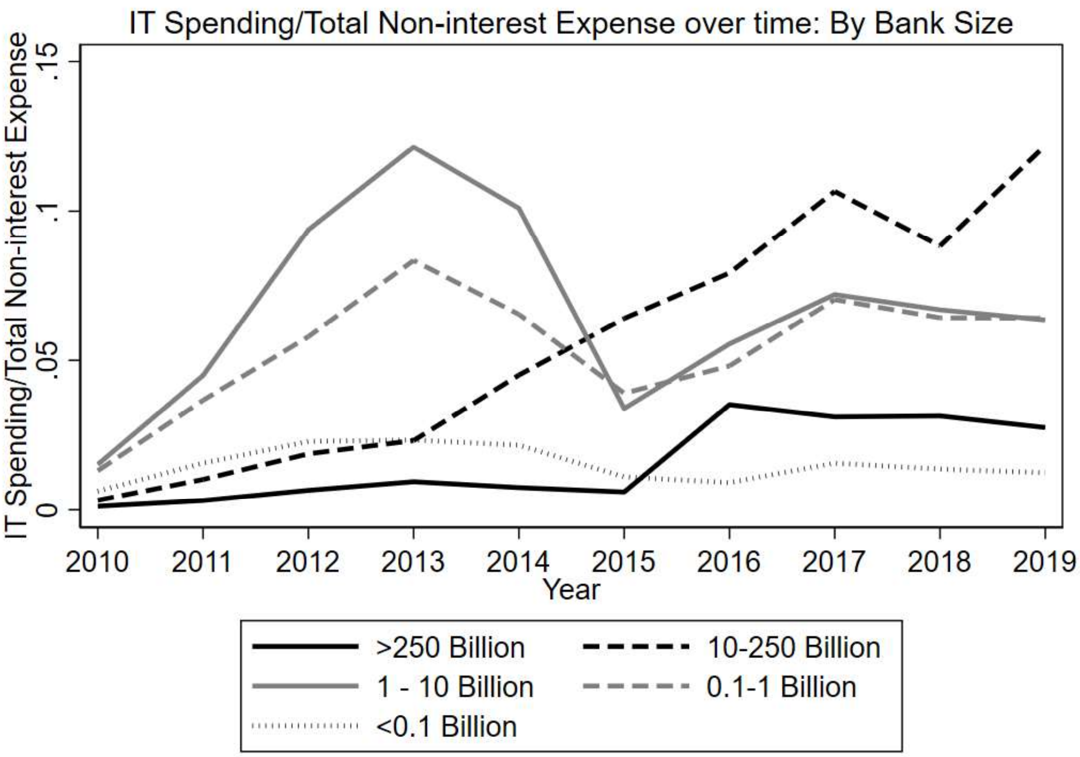

Figure 4. Communication and Software Spending by Bank Size

This figure shows the comparison of communication technology and software technology as a share of total IT spending across different bank size groups. The asset size groups are categorized based on a bank's average asset size during 2010 and 2019.

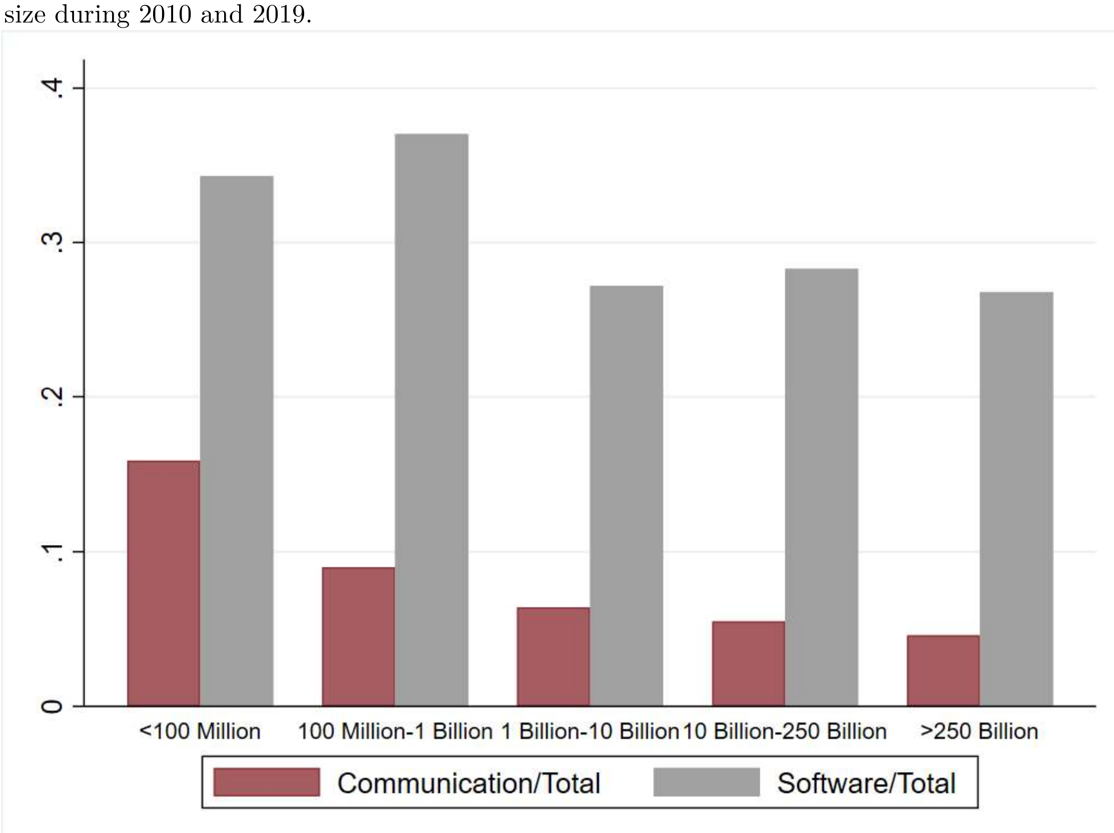

Figure 5. “Low Mortgage Rate Episode”

This figure shows the time-series of aggregate mortgage interest rate and the Federal Funds Rates. ”MORTGAGE30US” is the 30-Year Fixed Rate Mortgage Average in the United States from Freddie Mac. ”FEDFUNDS” is the effective Federal Funds Rate by Board of Governors of the Federal Reserve System.

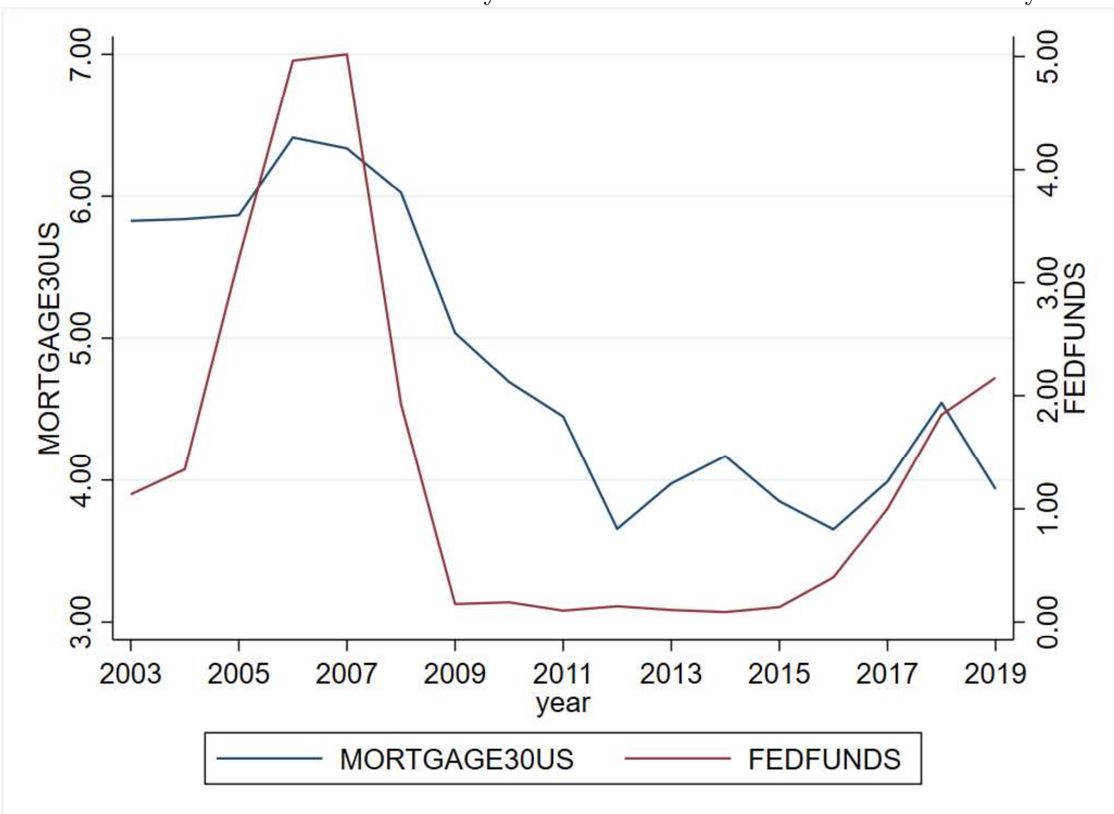

Figure 6. IT Spending Profile and Local Profiles

The figures below show the correlation between banks' IT spending pattern and local business characteristics. The two figures in the first row shows the bin-scatter plot of the correlation between the proportion of small businesses with less than or equal to 5 employees and the local banks' communication spending/software spending as a share of their total revenue. The IT spendings are weighted by banks' revenue in the county. The two figures in the second row show the scatter plot of correlation between counties' mortgage refinance growth rate and local banks' communication/software spending as a share of their total revenue between 2010-2019. All variables in the figures are standardized. Variables on both panels of figures are winsorized at 5% and 95%.

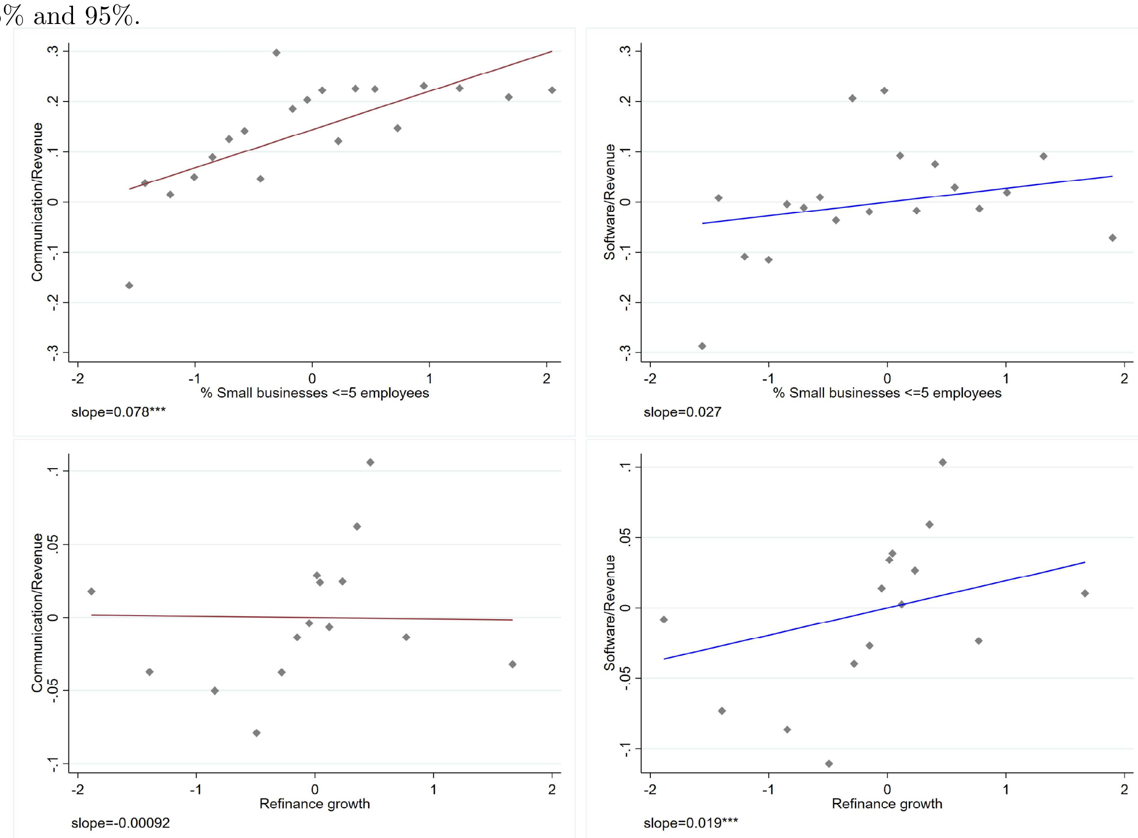

Figure 7. Bank Employment Trend and Banks' IT Spending

The figures below show the over-time trend of different types of banks' employees and IT spending trend. "Ln IT Employees" shows the weighted average of logarithm of IT related employees, weighted by banks' asset size. The green dashed line is the overall sample, the solid green line is the IT-employee dynamics of the sub-sample of banks within the top 25^{th} percentile of communication spending as a share their total revenue during 2010 and 2019, the solid black line is IT-employee dynamics of the sub-sample of banks within top 25^{th} percentile of software spending as a share of total revenue during 2010 and 2019.

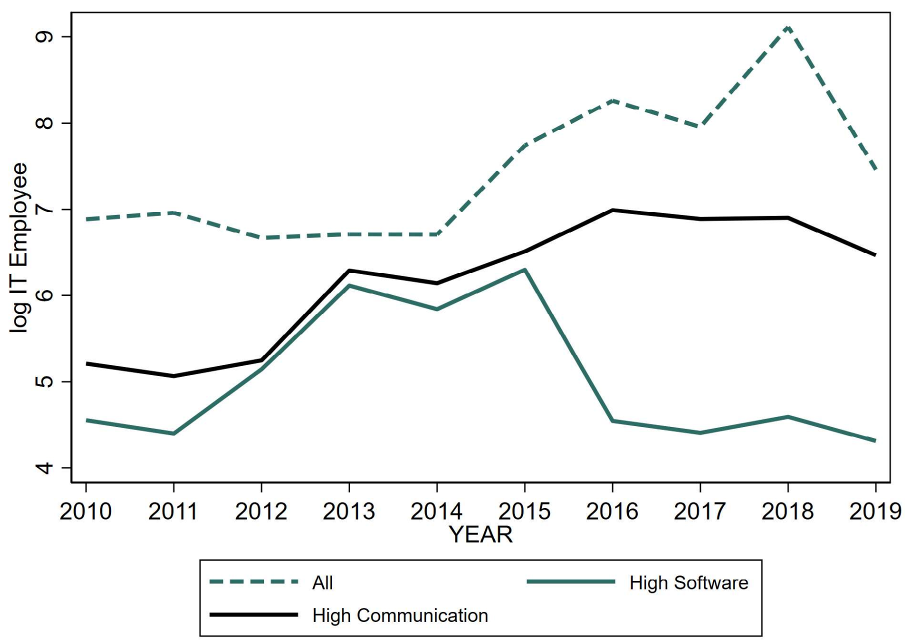

Table 1. Sample Coverage

This table demonstrates the sample coverage of banks across five categories of banks' size groups. All banks in the sample are commercial banks. The Call Report bank population includes only commercial banks ("Charter Type" being 200) following FFIEC definition. The first two columns shows the number of banks and the average asset sizes of banks in our sample, across five size groups. Column 3 and column 4 show the total number of banks and average asset sizes of all banks in the Call Report. Column 5 shows the percentage of sample coverage in terms of frequency compared with the population in Call Report, and column 6 shows the percentage of sample coverage in terms of total asset size compared with the population in Call Report.

Coverage of data	Sample		Call Report		Freq %	Asset %
Average Assets 2010-2019 (Billion)	Num banks	Ave Assets	Num banks	Ave Assets		
>$250 Billion	6	1184.24	7	787.34	85.71%	96.66%
$10 Billion-$250 Billion	88	42.30	106	43.69	83.02%	73.22%
$1 Billion-$10 Billion	474	2.90	590	2.78	80.34%	89.43%
$100 Million-$1 Billion	942	0.40	4161	0.32	22.64%	29.44%
<$100 Million	296	0.06	2048	0.05	14.45%	14.23%

Table 2. IT Spending Summary Statistics

This table shows the summary statistics of banks' IT Spending. Total IT Spending is the sum of all types of IT spending in millions of dollars. No. of IT employees is the total amount of IT-related employees. IT Spending/Income is total IT Spending scaled by banks' total income, IT Spending/Non-interest expense is total IT spending scaled by non-interest expenses; IT spending/Net income is total IT spending scaled by total income minus the gross total expenses. The different categories of IT spending are the four categories of IT spending scaled by total IT spending.

	Mean	S.d.	p(25)	Median	p(75)
Total IT Spending (Million)	7.31	111.35	0.03	0.21	1.06
No. of IT Employees	133.43	872.10	7.00	21.58	56.40
Storage Amount(PB)	3.52	25.52	0.11	0.48	1.78
IT Spending/Income	0.03	0.15	0.00	0.01	0.02
IT Spending/Non-interest Expenses	0.04	0.21	0.01	0.01	0.04
IT Spending/Net Income	0.16	3.65	0.01	0.04	0.11
Hardware/Total	0.17	0.12	0.06	0.16	0.24
Communication/Total	0.09	0.12	0.03	0.05	0.11
Software/Total	0.33	0.16	0.22	0.32	0.47
Services/Total	0.33	0.14	0.24	0.32	0.42
Other/Total	0.06	0.10	0.01	0.01	0.06

Table 3. C& I Loans and Banks' IT Spending

This table presents the results of regression of banks' C&I loan on the four major categories of banks' IT spending and a vector of control variables at bank-year level. The sample period is 2010 to 2019.

$$\frac{\text{C\&I Loan}}{\text{Total loan}}_{i,10-19} = \alpha + \beta \frac{\text{Type S IT spending}}{\text{Revenue}}_{i,10-19} + \gamma X + \epsilon_i$$

C&I Loan/Total Loan is commercial and industrial loan of bank i scaled by total loan between 2010-2019, Software/Rev is software spending scaled by total revenue, Communication/Rev is communication spending scaled by total revenue, Hardware/Rev is Hardware spending scaled by total Revenue, Services/Rev. Control variables include net income scaled by total assets, deposits scaled by total assets, revenue per employee, salaries scaled by total assets and equity scaled by total assets. Both the left-hand side and the right-hand side variables are taken the average across 2010-2019 within bank i. Fixed effects include Bank size, and banks' headquarter state fixed effects. $*p < 0.1, **p < 0.05, ***p < 0.01$.

	Panel B			
	Software/Revenue	Communication/Revenue	Hardware/Revenue	Services/Revenue
	(1)	(2)	(3)	(4)
C& I loans/total loans	0.0312	0.0497**	0.0492**	0.0432*
	(0.0251)	(0.0242)	(0.0244)	(0.0249)
Net income/Total Assets	-0.142***	-0.217***	-0.242***	-0.107***
	(0.0289)	(0.0279)	(0.0281)	(0.0286)
Deposits/Assets	-0.0126	0.0284	0.0320	-0.00793
	(0.0312)	(0.0301)	(0.0304)	(0.0310)
Revenue per Employee	-0.267***	-0.347***	-0.301***	-0.298***
	(0.0352)	(0.0340)	(0.0342)	(0.0349)
Salaries/Assets	-0.0175	-0.132***	-0.0993***	-0.0344
	(0.0258)	(0.0249)	(0.0251)	(0.0256)
Equity/Assets	0.0702**	0.0510*	0.0457*	0.0712***
	(0.0278)	(0.0269)	(0.0270)	(0.0276)
Size FE	Y	Y	Y	Y
State FE	Y	Y	Y	Y
AdR-squared	0.0981	0.162	0.150	0.110
N	1798	1798	1798	1798

Table 4. Loan Specialization and Banks' IT Spending

This table presents the results of regression of banks' C&I loan on the four major categories of banks' IT spending and a vector of control variables at bank-year level. The sample period is 2010 to 2019.

$$\frac{\text{Type S IT spending}}{\text{Revenue}}_{i,10-19} = \alpha + \beta \frac{\text{Type L Loan}}{\text{Total loan}}_{i,10-19} + \gamma X + \epsilon_i$$

Type L Loan/Total Loan is the average of a specific type of loan scaled by total loan, among them, Personal loan/Total Loan is the sum of personal loan and real estate loan to 1-4 family units scaled by total loan, Agriculture/Total loan is the agriculture loan scaled by total loan. CRA/Total loan is the sum of small business loans reported in CRA scaled by total loan, "Other C&I/Total loan" is the total C&I loan minus small business loans reported in CRA, scaled by total loan. "Mortgage refinance" is the total amount of mortgage refinance reported in HMDA scaled by the bank's total loan, "Other personal loans" is the deduction of "Mortgage refinance" from "Personal and mortgage loans". All of the loan profile variables are calculated as the average of the loan profile of a bank between 2010 and 2019. %Lead bank is the frequency of a banks' showing up as a lead bank in the syndicated loan market as a share of total number of syndicated loans lent out. %Refinance is the frequency of refinance as a percent of total number of mortgage issuance that are reported in HMDA. Software/Rev is software spending scaled by total revenue, Communication/Rev is communication spending scaled by total revenue, Hardware/Rev is Hardware spending scaled by total Revenue, Services/Rev. Control variables include net income scaled by total assets, deposits scaled by total assets, revenue per employee, salaries scaled by total assets and equity scaled by total assets. Fixed effects include Bank size, and banks' headquarter state fixed effects. $*p < 0.1, **p < 0.05, ***p < 0.01$.

	Software/Revenue	Communication/Revenue	Hardware/Revenue	Services/Revenue
	(1)	(2)	(3)	(4)
C&I loans/Total loan	0.0312	0.0497**	0.0492**	0.0432*
	(0.0251)	(0.0242)	(0.0244)	(0.0249)
CRA/Total loan	-0.150***	0.113***	0.0625**	0.0387
	(0.0291)	(0.0281)	(0.0285)	(0.0291)
Other C&I/Total loan	0.0495**	0.0355	0.0413*	0.0387
	(0.0249)	(0.0240)	(0.0242)	(0.0247)
% Lead bank/Total Syndicate	-0.612*	1.759***	1.393***	1.083***
	(0.333)	(0.295)	(0.278)	(0.262)
Personal loan/Total loan	0.0452*	0.0325	-0.0135	0.0247
	(0.0271)	(0.0261)	(0.0230)	(0.0234)
Mortgage refinance/Total loan	0.0933***	-0.0324	0.0333	0.0267
	(0.0329)	(0.0340)	(0.0350)	(0.0298)
Other personal loans/Total loan	-0.0276	0.0398	0.0186	0.0105
	(0.0260)	(0.0281)	(0.0345)	(0.0261)
% Refinance /Total Mortgage	0.0805***	-0.00220	0.0414*	0.0550**
Agricultural loans/Total loans	0.0263	0.0735**	0.0486	0.0438
	(0.0312)	(0.0300)	(0.0302)	(0.0310)

Table 5. Soft Information and Banks' IT Spending

This table presents the results of 2SLS and OLS discussed in Section 4.2.
The first three columns show the results for the following specification:

$$\Delta \ln(\text{CRA})_{i,c,post} = \mu_i + \mu_1 \ln(1 + \text{Qualified Small Buzs})_{c,pre} + \mu_2 X_{i,c} + \epsilon_{i,c}$$

$$\Delta \ln \text{IT}_{i,c,post} = \alpha_i + \beta \Delta \ln(\widehat{\text{CRA}})_{i,c,post} + \gamma X_{i,c} + \epsilon_{i,c}$$

The last two columns show the following OLS specification:

$$\Delta \ln(\text{IT})_{i,c,post} = \quad \alpha_i + \beta \times \Delta \ln \text{CRA}_{i,c,post} + \mu_c + \gamma X_{i,c} + \epsilon_{i,c}$$

ΔLn CRA$_{i,c,post}$ is the change in average natural log of small business loans reported in CRA of bank i at county c in year during 2014-2017 compared with the 2010-2013, Software/Revenue is software spending scaled by total revenue, Communication/Revenue is communication spending scaled by total expenses, Hardware/Revenue is Hardware spending scaled by total expenses, Services/Revenue. The four IT spending scaled by revenue is at bank-county-year level. Control variables include net income scaled by total assets, deposits scaled by total assets, revenue per employee, and equity scaled by total assets, control variables are at bank level. County level control variables include the labor force, population growth rate, and total number of establishments. Fixed effects include bank fixed effects. $*p < 0.1, **p < 0.05, ***p < 0.01$.

	2SLS			OLS	
	ΔLn CRA	ΔLn Software	ΔLn Communication	ΔLn Software	ΔLn Communication
	(1)	(2)	(3)	(4)	(5)
Ln(Qualified Small Busi)	0.171***				
	(0.0237)				
ΔLn(CRA)				0.00307	0.0146*
				(0.00863)	(0.00866)
$\Delta\widehat{\text{Ln}(\text{CRA})}$		0.0204	0.0745***		
		(0.0173)	(0.0173)		
Revenue per Employee	0.00223	0.186***	0.188***	0.186***	0.189***
	(0.00776)	(0.00894)	(0.00897)	(0.00894)	(0.00897)
Deposits/Assets	0.0149	-0.0592	-0.145	-0.0570	-0.137
	(0.0948)	(0.109)	(0.109)	(0.109)	(0.109)
Equity/Assets	-0.00263	0.134	0.0398	0.133	0.0378
	(0.0720)	(0.0829)	(0.0832)	(0.0829)	(0.0832)
Net income/Assets	0.0164	-0.101	-0.301***	-0.0975	-0.289***
	(0.0543)	(0.0625)	(0.0631)	(0.0625)	(0.0631)
Bank FE	Y	Y	Y	Y	Y
R-squared	0.431	0.157	0.148	0.157	0.147
F Statistics	15.18				
N	21202	21200	21195	21200	21195

Table 6. Hard Information and Banks' IT Spending

This table presents the results of regressions discussed in Section 4.3.
The first three columns show the results for the 2SLS specification below:

$$\ln \text{Refinance}_{i,c} = \mu_i + \mu_1 \text{Payments gap}_c + \mu_2 X_{i,c,t} + \epsilon_{i,c,t}$$

$$\left(\frac{\text{Type S IT Spending}}{\text{Revenue}}\right)_{i,c,t} = \alpha_i + \beta \ln \widehat{\text{Refinance}}_{i,c} + \gamma X_{i,c,t} + \epsilon_{i,c,t}$$

The last two columns show the results of OLS specification below:

$$\frac{\text{Type S IT spending}}{\text{Revenue}}_{i,c,t} = \alpha_i + \mu_c + \text{LnMortgage Refinance}_{i,c,t} + \pi_t + \beta + \gamma X_{i,c,t} + \epsilon_{i,c,t}$$

Ln Mortgage Refinance$_{i,c,t}$ is the amount of mortgage refinance loan issued by bank i in county c in year t. Payments gap is the hypothetical amount of interest payments that could be saved due to the interest rate gap, if local households choose to refinance their mortgages during the year of 2011 and 2015. Software/Revenue is software spending scaled by total revenue, Communication/Revenue is communication spending scaled by total expenses. Control variables include net income scaled by total assets, deposits scaled by total assets, revenue per employee, salaries scaled by total assets and equity scaled by total assets, control variables are at bank-year level. Fixed effects include bank fixed effects, year fixed effects and county fixed effects. Standard errors are clustered at county and bank level. $*p < 0.1, **p < 0.05, ***p < 0.01$.

	2SLS			OLS	
	Ln Refinance	Software/Revenue	Communication/Revenue	Software/Revenue	Communication/Revenue
	(1)	(2)	(3)	(4)	(5)
Payment gap	0.863***				
	(0.0370)				
Ln(Refinance)				0.0166***	0.00739*
				(0.00335)	(0.00277)
$\text{Ln}(\widehat{\text{Refinance}})$		0.0672**	0.0373		
		(0.0267)	(0.0251)		
Revenue per Employee	-0.329***	-0.334***	-0.329***	-0.332***	-0.328**
	(0.00466)	(0.00495)	(0.00466)	(0.0462)	(0.0786)
Net income/Assets	0.00896	0.00162	0.00896	-0.000293	0.00783
	(0.00908)	(0.00962)	(0.00908)	(0.0209)	(0.0196)
Equity/Assets	-0.0277**	-0.0139	-0.0277**	-0.0136	-0.0276
	(0.0115)	(0.0122)	(0.0115)	(0.0431)	(0.0215)
Deposits/Assets	0.0631***	0.153***	0.0631***	0.141*	0.0562
	(0.0162)	(0.0171)	(0.0162)	(0.0649)	(0.0344)
Bank FE	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y
F Statistics	2408.16				
AdR-squared	0.0176	0.00833	0.0176	0.438	0.450
N	63189	63159	63189	63159	63189

Table 7. IT Spending and IT Employment and Availability of Potential IT Employees

This table presents the evidence of identification of correlation between banks' IT spending and IT employment during the mortgage refinancing boom period. The regression equation is as follows

$$\Delta \text{IT Emp/Total Emp}_{i,c,15-19\rightarrow 10-14} = \alpha_i + \gamma \text{Land-grant college availability} + \mu X + \epsilon_{i,c}$$
$$\Delta \text{Software}/\Delta \text{Communication/Revenue} = \alpha_i + \beta \Delta \text{ IT } \widehat{\text{Emp/Total}} \text{ emp}_{i,c} + \phi X + \epsilon_{i,c}$$

Column (1) shows the results of the first stage and column (2) and (3) shows the results of the second stage. Control variables include net income scaled by total assets, deposits scaled by total assets, revenue per employee, salaries scaled by total assets and equity scaled by total assets, control variables are at bank level. Bank fixed effects are included. $*p < 0.1, **p < 0.05, ***p < 0.01$.

	Panel A		
	Δ IT Emp/Total Emp	Δ Software/Revenue	Δ Communication/Revenue
	(1)	(2)	(3)
𝟙[Land Grant College]	0.0615**		
	(0.0256)		
Δ IT emp/Total emp		-0.0463***	-0.0240
		(0.0178)	(0.0156)
Revenue per Employee	0.0129	0.173***	0.285***
	(0.00801)	(0.0190)	(0.0162)
Net income/Assets	0.0299	-0.0980	-0.0734
	(0.0559)	(0.0861)	(0.0796)
Equity/Assets	-0.0268	0.152	0.0428
	(0.0599)	(0.101)	(0.0975)
Deposits/Assets	0.0335	-0.0619	-0.237*
	(0.0816)	(0.148)	(0.140)
Bank FE	Y	Y	Y
F-statistics	296.0	9.627	36.04
AdR-squared	0.280	-0.219	-0.193
N	19588	19588	19588
	Panel B		
	Δ IT Emp/Total Emp	Δ Software/Revenue	Δ Communication/Revenue
	(1)	(2)	(3)
Ln(min distance)	-0.0209**		
	(0.00918)		
Δ IT emp/Total emp		-0.0442**	-0.0237
		(0.0179)	(0.0154)
Revenue per Employee	0.0153*	0.175***	0.284***
	(0.00798)	(0.0191)	(0.0162)
Net income/Assets	0.0335	-0.106	-0.0785
	(0.0553)	(0.0863)	(0.0799)
Equity/Assets	-0.0334	0.155	0.0518
	(0.0595)	(0.102)	(0.0978)
Deposits/Assets	0.0289	-0.0738	-0.238*
	(0.0809)	(0.148)	(0.140)
Bank FE	Y	Y	Y
F-statistics	301.4	9.649	35.35
R-squared	0.296	-0.223	-0.197
N	19588	19588	19588

Appendix

Figure A1. McKinsey (2012) "Breakthrough IT Banking"

BUSINESS TECHNOLOGY OFFICE

Breakthrough IT banking

Some Asian banks achieve superior returns despite relatively low IT expenditures. What's their secret?

Sai Gopalan, Gaurav Jain, Gaurav Kalani, and Jessica Tan

Banks have long relied on technology to introduce products such as online banking, ATMs, and mobile payments, and to improve back-office efficiency. But that reliance comes with a price. Globally, the banking sector spends an average of 4.7 percent to 9.4 percent of operating income on IT, while other sectors spend less: insurance companies and airlines, for example, spend 3.3 percent and 2.6 percent of income, respectively.

Our Asian Banking IT Benchmarking Study[1] finds, however, that a bank's high IT expenditures do not always correlate with superior performance. Some banks with large IT budgets often have trouble leveraging investments to generate commensurately high revenue growth and operational efficiency. Survey data show that 66 percent of banks with higher-than-average IT spending relative to income generated lackluster results, with revenue growth 0.4 percentage points lower than the industry standard and a cost-income (C/I) ratio 2.5 percentage points higher.

By contrast, 23 percent of the 44 banks surveyed outperformed the market on both revenue growth (up 10.9 percentage points) and C/I ratio (down 4.6 percentage points) while spending 29 percent less on IT than other banks in our study. These outperforming banks are more likely to view IT as a strategic enabler, and their investments mirror this outlook. Outperformers direct a higher share of spending toward technologies designed to create new business value and a lower share of spending on support operations, such as finance and human resources. These banks are also more likely than the lower performers to promote efficiency through a consolidated IT footprint as well as formal vendor- and demand-management practices.

The common denominator linking high-performing Asian banks is a commitment to strong governance and spending alignment with the needs of the business. This finding supports our experience with bank clients in Europe and the Americas, and prompted us

[1] The 2010 biennial McKinsey Asian Banking IT Benchmarking survey comprised 44 banks across 11 Asia-Pacific countries, with the results tracked against prior year benchmarks from 2006 onward.

Figure A2. Definition of Different Types of IT Spending

- ***COMM_BUDGET***

The modeled IT budget for communication services at this site.
It is defined as the network equipment that companies operate to support their communications needs.
It includes:

  - routers
  - carrier line equipment
  - fiber optic equipment
  - switches
  - private branch exchanges (PBXes)
  - radio and TV transmitters
  - Wi-Fi transmitters
  - desktop telephone sets; wide-area network (WAN) and local-area network (LAN) equipment
  - videoconferencing and telepresence equipment
  - cable boxes
  - other network equipment.
  - end-client mobile devices like cell phones/iPhones that are bought by individuals

(a) Figure A

- ***SOFTWARE_BUDGET***

The modeled IT budget for software at this site.
It is defined as software from third parties, whether that software is packaged or semipackaged software delivered on CD and installed within the company, hosted by a third party, offered on a SaaS basis from a multitenant shared-instance server accessible by a browser, or custom-created for a company by third-party contractors or consultants.
It includes:

  - license, maintenance, subscriptions and software vendor-provided services revenues for all categories of middleware software (including storage management systems, database management systems, IT management systems, security software, application servers and application development software)
  - application software such as :
    - desktop applications
    - information management software (like business intelligence and enterprise content management)
    - process applications (like ERP, CRM, SCM or PLM)
    - ePurchasing software
    - risk and payment management software

    . We also include vertical industry applications (like banking management systems, security trading systems, insurance underwriting or claims management software, retail management software, or hospital information systems). Finally, we include computer operating systems software, even though that cost is often bundled
  - vertical industry applications (like banking management systems, security trading systems, insurance underwriting or claims management software, retail management software, hospital information systems)
  - computer operating systems software (even though that cost is often bundled)

(b) Figure B

- ***SERVICES_BUDGET***

The modeled IT budget for IT-related services at this site.
It is defined as project-based consulting or systems integration services that vendors provide to businesses and Governments, whether on or off-site.
It includes:

  - contractors, consulting services for IT strategy, security assessments and process change
  - systems integration
  - project services
  - mainframe outsourcing, desktop support outsourcing, distributed systems outsourcing, network outsourcing, application hosting, application management outsourcing and application testing. These applications are single-instance software deployments, generally owned rather than subscribed to, and thus are different from SaaS.
  - computer hardware support and maintenance services.

(c) Figure C

- ***HARDWARE_BUDGET***

The modeled IT budget for hardware at this site.
It includes the classic computer hardware that IT departments buy and support, regardless of whether the IT department itself operates that equipment (such as servers) or oversees the use of this equipment by employees (such as PCs):

  - PCs: personal computers (laptops, desktops, and tablets)
  - Servers/Mainframes
  - Peripherals: monitors, terminals, printers, keyboards, mice, USB devices, etc...
  - Storage: storage devices (NAS, DAS, tape)
  - Other hardware: hardware specific to the industry (point-of-sales equipment based on PCs, smart cards, embedded computer chips, etc...)

(d) Figure D

Table A1. Bank Characteristics and IT Spending

This table presents the results of regression of banks' IT spending structure between 2010 and 2019 on banks loan portfolio on balance sheet before the financial crisis. The regression specification is as follows:

$$\frac{\text{Type S IT spending}}{\text{Total}}_{i,2010-2019} = \alpha + \beta \frac{\text{Type l Loan}}{\text{Total loan}}_{i,2005-2009} + \gamma X + \epsilon_i$$

C&I loan/Total loan is commercial and industrial loan scaled by total loan; Pers & Mort/Total loan is personal loan and the real estate loan to 1-4 families scaled by total loan; agriculture loan/Total loan is agriculture loan scaled by total loan. All the three types of loans as a share of total loans are the bank-level average loan proportions from 2005 to 2009. IT spending profiles are defined as each type of IT spending scaled by total IT spending, and taking an average at the bank level between 2010 and 2019. Control variables include net income scaled by total assets, deposits scaled by total assets, revenue per employee, salaries scaled by total assets and equity scaled by total assets, control variables are at bank-year level. Fixed effects include bank fixed effects, year fixed effects and county fixed effects. Standard errors are clustered at county and bank level. $*p < 0.1, **p < 0.05, ***p < 0.01$.

Panel A

	$\frac{\text{Communication}}{\text{Total}}_{10-19}$	$\frac{\text{Software}}{\text{Total}}_{10-19}$	$\frac{\text{Hardware}}{\text{Total}}_{10-19}$	$\frac{\text{Services}}{\text{Total}}_{10-19}$
	(1)	(2)	(3)	(4)
C& I loan/Total Loan(05-09)	0.0471***	-0.00138	0.0412***	-0.0113
	(0.0154)	(0.0222)	(0.0149)	(0.0156)
Net income/Total Assets	-0.166	0.217	-1.878***	2.017***
	(0.200)	(0.351)	(0.210)	(0.289)
Revenue per Employee	0.0413**	0.0507**	0.0198	-0.0262
	(0.0168)	(0.0249)	(0.0138)	(0.0282)
Equity/Assets	-0.00808	0.103	-0.122*	0.0874
	(0.0708)	(0.100)	(0.0717)	(0.0855)
Salaries/Assets	-0.334	1.304***	0.215	0.580**
	(0.331)	(0.238)	(0.616)	(0.268)
Deposits/Assets	0.0177	-0.0444	0.0264	-0.0370
	(0.0297)	(0.0530)	(0.0297)	(0.0362)
R-squared	0.161	0.199	0.110	0.220
N	1649	1649	1649	1649

Panel B

	(1)	(2)	(3)	(4)
Pers & Mort/Total Loan(05-09)	-0.0584***	0.0524***	-0.0382***	-0.0103
	(0.0114)	(0.0141)	(0.00932)	(0.00990)
Net income/Total Assets	-0.335*	0.382	-1.985***	1.981***
	(0.201)	(0.346)	(0.213)	(0.294)
Revenue per Employee	0.0405***	0.0579**	0.0209	-0.0295
	(0.0152)	(0.0272)	(0.0139)	(0.0272)
Equity/Assets	0.00603	0.0858	-0.114	0.0921
	(0.0711)	(0.101)	(0.0719)	(0.0856)
Salaries/Assets	-0.269	1.244***	0.257	0.592**
	(0.306)	(0.226)	(0.599)	(0.266)
Deposits/Assets	0.0233	-0.0434	0.0315	-0.0389
	(0.0294)	(0.0535)	(0.0298)	(0.0366)
R-squared	0.172	0.205	0.114	0.220
N	1649	1649	1649	1649

Panel C

	(1)	(2)	(3)	(4)
Agriculture loan/Total Loan(05-09)	0.0577***	-0.0826***	0.0286**	0.0253**
	(0.0155)	(0.0152)	(0.0130)	(0.0118)
Net income/Total Assets	-0.307	0.441	-1.942***	1.945***
	(0.203)	(0.353)	(0.217)	(0.290)
Revenue per Employee	0.0489**	0.0503**	0.0264*	-0.0279
	(0.0192)	(0.0253)	(0.0155)	(0.0281)
Equity/Assets	0.00350	0.0790	-0.118*	0.0961
	(0.0709)	(0.100)	(0.0711)	(0.0851)
Salaries/Assets	-0.271	1.211***	0.245	0.609**
	(0.298)	(0.213)	(0.601)	(0.259)
Deposits/Assets	0.0318	-0.0550	0.0360	-0.0355
	(0.0297)	(0.0529)	(0.0296)	(0.0365)
R-squared	0.165	0.208	0.109	0.221
N	1649	1649	1649	1649

Table A2. Summary Statistics of Banks' IT Spending by Bank Size Group

This table presents the summary statistics of banks' IT spending by banks' size groups. Banks in the sample are split into five groups. Total IT Spending is the sum of all types of IT spending in millions of dollars. No. of IT employees is the total amount of IT-related employees. IT Spending/Income is total IT Spending scaled by banks' total income, IT Spending/Non-interest expense is total IT spending scaled by non-interest expenses; IT spending/Net income is total IT spending scaled by total income minus the gross total expenses. The different categories of IT spending are the four categories of IT spending scaled by total IT spending.

	Mean	S.d.	Median
< $100 Million			
IT Spending/Total Assets	0.001	0.003	0.000
IT Spending/Income	0.015	0.048	0.005
IT Spending/Expenses	0.019	0.075	0.007
Hardware/Total	0.206	0.133	0.203
PC/Total	0.102	0.132	0.075
Server/Total	0.100	0.127	0.066
Terminal/Total	0.023	0.081	0.004
Printer/Total	0.022	0.081	0.003
Other/Total	0.098	0.137	0.037
Storage/Total	0.092	0.135	0.040
Communication/Total	0.159	0.176	0.086
Software/Total	0.343	0.132	0.341
Services/Total	0.330	0.138	0.340

	Mean	S.d.	Median
$100 Million-$1 Billion			
IT Spending/Total Assets	0.001	0.007	0.000
IT Spending/Income	0.028	0.140	0.009
IT Spending/Expenses	0.040	0.200	0.014
Hardware/Total	0.167	0.115	0.158
PC/Total	0.066	0.090	0.053
Server/Total	0.070	0.087	0.052
Terminal/Total	0.012	0.047	0.004
Printer/Total	0.011	0.047	0.003
Other/Total	0.056	0.097	0.014
Storage/Total	0.051	0.091	0.022
Communication/Total	0.090	0.109	0.052
Software/Total	0.370	0.155	0.348
Services/Total	0.308	0.124	0.316

	Mean	S.d.	Median
$1 Billion-$10 Billion			
IT Spending/Total Assets	0.002	0.007	0.001
IT Spending/Income	0.043	0.193	0.014
IT Spending/Expenses	0.062	0.262	0.021
Hardware/Total	0.165	0.117	0.140
PC/Total	0.056	0.064	0.043
Server/Total	0.065	0.063	0.050
Terminal/Total	0.007	0.021	0.004
Printer/Total	0.007	0.022	0.003
Other/Total	0.036	0.072	0.012
Storage/Total	0.033	0.061	0.017
Communication/Total	0.064	0.078	0.042
Software/Total	0.272	0.166	0.231
Services/Total	0.361	0.152	0.336

	Mean	S.d.	Median
$10 Billion-$250 Billion			
IT Spending/Total Assets	0.001	0.006	0.000
IT Spending/Income	0.045	0.249	0.012
IT Spending/Expenses	0.067	0.310	0.019
Hardware/Total	0.147	0.105	0.108
PC/Total	0.047	0.041	0.029
Server/Total	0.057	0.041	0.038
Terminal/Total	0.006	0.009	0.004
Printer/Total	0.005	0.010	0.003
Other/Total	0.032	0.057	0.010
Storage/Total	0.027	0.036	0.011
Communication/Total	0.055	0.052	0.042
Software/Total	0.283	0.161	0.233
Services/Total	0.335	0.137	0.293

	Mean	S.d.	Median
> $250 Billion			
IT Spending/Total Assets	0.001	0.002	0.000
IT Spending/Income	0.019	0.049	0.005
IT Spending/Expenses	0.031	0.075	0.008
Hardware/Total	0.158	0.103	0.138
PC/Total	0.051	0.043	0.039
Server/Total	0.062	0.044	0.050
Terminal/Total	0.007	0.011	0.004
Printer/Total	0.006	0.012	0.003
Other/Total	0.036	0.061	0.012
Storage/Total	0.031	0.039	0.018
Communication/Total	0.046	0.041	0.031
Software/Total	0.268	0.137	0.228
Services/Total	0.357	0.149	0.328

Table A3. County Characteristics and Local Banks' IT Spending: I

This table presents the results of regression of banks' IT spending structure on local business characteristics. The regression specification is as follows:

$$\left(\frac{\text{Type S IT spending}}{\text{Total}}\right)_c = \alpha + \beta \text{Small business characteristics}_c + \gamma X + \epsilon_c$$

The left-hand side variable is the local banks' IT spending profile at county level weighted by banks' asset size between 2010 to 2019. In column (1) and (2) the main right-hand side variable is the proportion of local small businesses with less than 10 employees among all local small businesses, the average size of small businesses is also controlled for. Column (3) and (4) regress local banks' IT spending profile on the number of establishments under three years old, controlling for the total number of small businesses. County level control variables include real GDP per capita, population growth, change in unemployment rate, population size, deposit HHI and small business loan HHI. All county-level control variables are the average of the control variables between 2010 and 2019. $*p < 0.1, **p < 0.05, ***p < 0.01$.

	Software/Total	Communication/Total	Software/Total	Communication/Total
	(1)	(2)	(3)	(4)
Share of Small businesses(<10 emp)	0.669	5.855**		
	(3.431)	(2.397)		
Average size of small bus	-0.0849	-0.0278		
	(0.0529)	(0.0266)		
#Establishments (< 3 years)			0.212	0.670***
			(0.157)	(0.193)
#Establishments			-0.0591	-0.240***
			(0.0602)	(0.0682)
Real GDP/pc	0.0233	-0.0622	0.0162	-0.0635
	(0.0432)	(0.0462)	(0.0389)	(0.0451)
Pop growth	-41.15***	0.862	-41.95***	-3.286
	(13.03)	(11.33)	(13.39)	(11.82)
Δ Unemployment rate	5.254*	2.116	6.984**	3.591
	(2.969)	(2.215)	(3.164)	(2.227)
Population size	0.538**	-0.775***	-0.0401	-1.186***
	(0.217)	(0.234)	(0.415)	(0.383)
Deposit HHI	0.882**	0.855***	0.909**	0.833***
	(0.378)	(0.315)	(0.381)	(0.320)
CRA HHI	-3.104***	3.213***	-2.418**	3.793***
	(0.915)	(0.669)	(1.017)	(0.688)
AdR-squared	0.00817	0.0600	0.00475	0.0501
N	3070	3070	3070	3070

Table A4. Local Banking Market Structure and Local Banks' IT Spending: II

This table presents the results of regressing banks' software and communication spending as a share of total revenue on local banking market structure. The regression is at bank-county-year level and the regression specification is as follows:

$$\frac{\text{Type S IT spending}}{\text{Revenue}}_{i,c,t} = \alpha_i + \mu_c + \pi_t + \beta(\text{CRA mkt share}_{i,c,t-2\to t})\text{Or }(\text{Mortgage mkt share}_{i,c,t-2\to t})$$
$$\text{Or }(\text{Deposit mkt share}_{i,c,t-2\to t}) + \gamma X + \epsilon_{i,c,t}$$

The left-hand side variables are software spending as a share of total revenue at bank-county-year level and communication spending as a share of total revenue at bank-county-year level. CRA market share is bank i's market share in the county c's small business lending market during the past two years; Mortgage market share is bank i's market share in county c's mortgage market as reflected in HMDA during the past two years; and deposit market share is bank i's deposit market share in county c during the past two years. Control variables include net income scaled by total assets, deposits scaled by total assets, revenue per employee, salaries scaled by total assets and equity scaled by total assets, control variables are at bank-year level. Fixed effects include bank fixed effects, year fixed effects and county fixed effects. Standard errors are clustered at county and bank level. $*p < 0.1, **p < 0.05, ***p < 0.01$.

	Software/Revenue			Communication/Revenue		
	(1)	(2)	(3)	(4)	(5)	(6)
CRA mkt Share	0.0130			0.0119*		
	(0.0295)			(0.00619)		
Mortgage mkt share		0.143***			-0.0120*	
		(0.0342)			(0.00724)	
Deposit mkt Share			0.100***			0.0510***
			(0.0225)			(0.00469)
Revenue/Employee	-1.496***	-1.496***	-1.534***	-0.372***	-0.371***	-0.389***
	(0.0483)	(0.0485)	(0.0517)	(0.00982)	(0.00984)	(0.0106)
Net income/Assets	-2.716***	-3.025***	-2.659***	0.00675	-0.00625	0.0425
	(0.968)	(0.983)	(1.009)	(0.185)	(0.187)	(0.192)
Equity/Assets	-1.897***	-1.673***	-2.118***	-0.218**	-0.214**	-0.274***
	(0.448)	(0.455)	(0.475)	(0.0884)	(0.0903)	(0.0940)
Deposit/Assets	-0.829***	-0.722***	-0.828***	0.00228	0.00353	-0.0203
	(0.196)	(0.201)	(0.211)	(0.0375)	(0.0384)	(0.0405)
Salary/Assets	-3.497*	-4.078**	-3.439*	-0.205	-0.138	-0.203
	(1.819)	(1.953)	(1.896)	(0.372)	(0.396)	(0.387)
Bank FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y	Y
R-squared	0.464	0.456	0.465	0.489	0.480	0.491
N	163662	159540	148894	163689	159567	148920

Table A5. Bank Size, Loan Information Nature and Banks' IT Spending

This table presents the results of regression of Banks' IT spending response when small business loans and mortgage refinance volume increases, interacting with bank size. This table focuses on software spending and communication spending. The left-hand side variables are software spending as a share of total revenue and communication spending as a share of total revenue respectively. The right-hand side variables are banks' small business loans and mortgage refinance respectively, interacting with banks' size groups. "Small" is defined as banks with average asset size less than $100 million between 2010-2019, "Medium" is defined as banks with average asset size $100 million-$10 billion between 2010-2019, "Large" is defined as banks with average asset size greater than defined Control variables include net income scaled by total assets, deposits scaled by total assets, revenue per employee, salaries scaled by total assets and equity scaled by total assets, control variables are at bank-year level. Fixed effects include bank fixed effects, year fixed effects and county fixed effects. Standard errors are clustered at county and bank level. $*p < 0.1, **p < 0.05, ***p < 0.01$.

	Software/Revenue		Communication/Revenue	
	(1)	(2)	(3)	(4)
(Small) × Ln(CRA)	-0.0151		0.0236	
	(0.0450)		(0.0469)	
(Medium) × Ln(CRA)	-0.00897		0.0299***	
	(0.00709)		(0.00688)	
(Large) × Ln(CRA)	0.0159		0.0319***	
	(0.0121)		(0.0110)	
(Small) × Ln(Mortgage refinance)		-0.00259		-0.00699
		(0.00626)		(0.00608)
(Medium) × Ln(Mortgage refinance)		0.0162**		0.00751
		(0.00673)		(0.00674)
(Large) × Ln(Mortgage refinance)		0.0344***		-0.0465***
		(0.0106)		(0.0101)
Fixed effects	Bank, County, Year			
Bank Controls	Y	Y	Y	Y
AdR-squared	0.533	0.525	0.540	0.532
N	160214	156718	160259	156762

Table A6. Local House Prices Growth and Banks' IT Spending

This table presents the results of regression of local house price growth on the four major categories of banks' IT spending and a vector of control variables at bank-county-year level. The sample period is 2010 to 2019.

$$\frac{\text{Type S IT spending}}{\text{Revenue}}_{i,c,t} = \alpha_i + \mu_c + \pi_t + \beta \Delta HPI_{c,t} + \gamma X + \epsilon_{i,c,t}$$

Refinance/Origination$_{i,c,t}$ is the number of refinance of mortgage relative to the number of new mortgage origination of bank i at county c in year t in reported in HMDA, Software/Revenue is software spending scaled by total revenue, Communication/Revenue is communication spending scaled by total expenses, Hardware/Revenue is Hardware spending scaled by total expenses, Services/Revenue. The four IT spending scaled by revenue is at bank-county-year level. Control variables include net income scaled by total assets, deposits scaled by total assets, revenue per employee, salaries scaled by total assets and equity scaled by total assets, control variables are at bank-year level. Fixed effects include bank fixed effects, year fixed effects and county fixed effects. Standard errors are clustered at county and bank level. $*p < 0.1, **p < 0.05, ***p < 0.01$.

	Software/Revenue	Communication/Revenue	Hardware/Revenue	Services/Revenue
	(1)	(2)	(3)	(4)
ΔHPI	0.293*	-0.0622***	-0.0968*	-0.110
	(0.165)	(0.0235)	(0.0561)	(0.163)
Revenue per Employee	-1.573***	-0.347***	-0.777***	-1.806***
	(0.111)	(0.0183)	(0.0484)	(0.127)
Net income/Assets	0.292	0.749**	1.903**	2.805*
	(1.576)	(0.324)	(0.829)	(1.680)
Equity/Assets	0.509	-0.111	-0.647*	-0.245
	(1.324)	(0.129)	(0.368)	(1.040)
Deposits/Assets	1.337	0.0505	0.167	0.686
	(1.028)	(0.0886)	(0.206)	(0.841)
Salaries/Assets	-6.299**	-0.649	-1.980*	-5.762**
	(2.624)	(0.487)	(1.097)	(2.636)
Bank FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
County FE	Y	Y	Y	Y
AdR-squared	0.489	0.507	0.492	0.496
N	177496	177542	177500	177495

Table A7. Small Business Loan, Mortgage Refinance and Bank IT Spending: I

This table presents the results of regression of banks' new mortgage issuance on the four major categories of banks' IT spending and a vector of control variables at bank-county-year level. The sample period is 2010 to 2019.

$$\frac{\text{Type S IT spending}}{\text{Revenue}}_{i,c,t} = \alpha_i + \mu_c + \pi_t + \beta \text{lnMortgage refinance}_{i,c,t} \text{ or } \text{lnCRA}_{i,c,t} + \gamma X + \epsilon_{i,c,t}$$

ln Mortgage refinance$_{i,c,t}$ is the natural logarithm of newly mortgage refinance of bank i at county c in year t in reported in HMDA, ln CRA$_{i,c,t}$ is the natural logarith of small business loans issued by bank i in county c and in year t. Software/Revenue is software spending scaled by total revenue, Communication/Revenue is communication spending scaled by total expenses, Hardware/Revenue is Hardware spending scaled by total expenses, Services/Revenue. The four IT spending scaled by revenue is at bank-county-year level. Control variables include net income scaled by total assets, deposits scaled by total assets, revenue per employee, salaries scaled by total assets and equity scaled by total assets, control variables are at bank-year level. Fixed effects include bank fixed effects, year fixed effects and county fixed effects. Standard errors are clustered at county and bank level. $*p < 0.1, **p < 0.05, ***p < 0.01$.

Panel A

	Software/Revenue			Communication/Revenue		
	(1)	(2)	(3)	(4)	(5)	(6)
Ln(Mortgage refinance)	0.0196***	0.0276***	0.0271***	-0.0151*	-0.00954	-0.0106
	(0.00565)	(0.00716)	(0.00962)	(0.00841)	(0.00775)	(0.00704)
Revenue per Employee		-0.209***	-0.221***		-0.224***	-0.232***
		(0.0147)	(0.0154)		(0.0118)	(0.0125)
Net income/Assets		0.00198	-0.000529		0.0256**	0.0280*
		(0.0114)	(0.0133)		(0.0112)	(0.0145)
Equity/Assets		0.00926	0.000128		-0.0135	-0.0184
		(0.0284)	(0.0270)		(0.0138)	(0.0150)
Deposits/Assets		0.0962	0.107*		0.0147	0.0220
		(0.0728)	(0.0629)		(0.0293)	(0.0246)
Fixed effects	Bank, Year, County					
Bank Controls	N	Y	Y	N	Y	Y
County Controls	N	N	Y	N	N	Y
R-squared	0.460	0.487	0.503	0.477	0.504	0.511
N	183454	180273	163129	183501	180320	163176

Panel B

	Software/Revenue			Communication/Revenue		
	(1)	(2)	(3)	(4)	(5)	(6)
Ln(CRA)	0.00386	0.00501	0.00441	0.0124***	0.0131***	0.0129***
	(0.00406)	(0.00400)	(0.00468)	(0.00322)	(0.00296)	(0.00293)
Revenue per Employee		-0.208***	-0.221***		-0.224***	-0.232***
		(0.0146)	(0.0154)		(0.0117)	(0.0124)
Net income/Assets		0.00195	-0.00184		0.0239**	0.0243*
		(0.0108)	(0.0125)		(0.0108)	(0.0132)
Equity/Assets		0.00616	-0.00115		-0.0131	-0.0179
		(0.0283)	(0.0268)		(0.0133)	(0.0144)
Deposits/Assets		0.0900	0.104*		0.0164	0.0247
		(0.0717)	(0.0622)		(0.0296)	(0.0249)
Salaries/Assets		-0.805	-0.635		-0.242	-0.164
		(0.807)	(0.853)		(0.958)	(1.019)
Fixed effects	Bank, Year, County					
Bank Controls	N	Y	Y	N	Y	Y
County Controls	N	N	Y	N	N	Y
AdR-squared	0.468	0.494	0.511	0.487	0.513	0.520
N	187871	184668	167018	187919	184716	167066

Table A8. Small Business Loan, Mortgage Refinance and Bank IT Spending: II

This table presents the results of regression of banks' new mortgage issuance on the four major categories of banks' IT spending and a vector of control variables at bank-county-year level. The sample period is 2010 to 2019.

$$\frac{\text{Type S IT spending}}{\text{Revenue}}_{i,c,t} = \alpha_i + \mu_{c,t} + \beta \text{lnMortgage refinance}_{i,c,t} \text{ or } \text{lnCRA}_{i,c,t} + \gamma X + \epsilon_{i,c,t}$$

ln Mortgage refinance$_{i,c,t}$ is the natural logarithm of newly mortgage refinance of bank i at county c in year t in reported in HMDA, ln CRA$_{i,c,t}$ is the natural logarith of small business loans issued by bank i in county c and in year t. Software/Revenue is software spending scaled by total revenue, Communication/Revenue is communication spending scaled by total expenses, Hardware/Revenue is Hardware spending scaled by total expenses, Services/Revenue. The four IT spending scaled by revenue is at bank-county-year level. Control variables include net income scaled by total assets, deposits scaled by total assets, revenue per employee, salaries scaled by total assets and equity scaled by total assets, control variables are at bank-year level. Fixed effects include bank fixed effects, county × year. Standard errors are clustered at county and bank level. $*p < 0.1, **p < 0.05, ***p < 0.01$.

Panel A

	Software/Revenue			Communication/Revenue		
	(1)	(2)	(3)	(4)	(5)	(6)
Ln(Mortgage refinance)	0.0270***	0.0354***	0.0340***	-0.00292	0.00332	0.00246
	(0.00636)	(0.00784)	(0.0105)	(0.00850)	(0.00798)	(0.00731)
Revenue per Employee		-0.213***	-0.225***		-0.228***	-0.236***
		(0.0146)	(0.0154)		(0.0119)	(0.0126)
Net income/Assets		0.00459	0.00411		0.0289***	0.0333**
		(0.0110)	(0.0127)		(0.0109)	(0.0147)
Equity/Assets		0.0137	0.00321		-0.00787	-0.0124
		(0.0283)	(0.0274)		(0.0119)	(0.0124)
Deposits/Assets		0.104	0.106*		0.0222	0.0241
		(0.0700)	(0.0606)		(0.0293)	(0.0231)
Salaries/Assets		-1.560	-1.382		-0.899	-0.800
		(1.064)	(1.089)		(1.281)	(1.349)
Fixed effects	Bank, **County**×**Year**					
Bank Controls	Y	Y	Y	Y	Y	Y
County Controls	Y	Y	Y	Y	Y	Y
AdR-squared	0.449	0.477	0.487	0.468	0.496	0.501
N	179713	176486	159732	179759	176532	159778

Panel B

	Software/Revenue			Communication/Revenue		
	(1)	(2)	(3)	(4)	(5)	(6)
Ln(CRA)	0.00577	0.00689*	0.00616	0.0141***	0.0149***	0.0145***
	(0.00425)	(0.00418)	(0.00480)	(0.00349)	(0.00323)	(0.00313)
Revenue per Employee		-0.212***	-0.224***		-0.228***	-0.235***
		(0.0145)	(0.0153)		(0.0118)	(0.0125)
Net income/Assets		0.00398	0.00240		0.0261**	0.0282**
		(0.0104)	(0.0119)		(0.0107)	(0.0136)
Equity/Assets		0.0119	0.00259		-0.00601	-0.0101
		(0.0282)	(0.0271)		(0.0117)	(0.0122)
Deposits/Assets		0.0972	0.101*		0.0224	0.0251
		(0.0686)	(0.0597)		(0.0292)	(0.0232)
Salaries/Assets		-0.702	-0.548		-0.394	-0.304
		(0.825)	(0.880)		(0.988)	(1.049)
Fixed effects	Bank, **County**×**Year**					
Bank Controls	Y	Y	Y	Y	Y	Y
County Controls	Y	Y	Y	Y	Y	Y
AdR-squared	0.459	0.485	0.497	0.479	0.506	0.511
N	184314	181056	163775	184363	181105	163824

Table A9. IT Spending and IT Employment

This table presents the results of correlation between banks' IT investment and banks' employment, with a distinction between IT employees and non-IT employees. The regression equations are as follows:

$$\Delta \text{ IT Emp/Non-IT employees} = \alpha_{i,t} + \mu_{c,t} + \beta \text{Ln software spending}_{i,c,t} / \text{Ln Communication spending}_{i,c,t} + \gamma X + \epsilon_{i,c,t}$$

Δlog IT Emp and Δlog Non-IT Emp are the logarithm of IT related employees and Non-IT related employees respectively. Software/Revenue is software spending scaled by total revenue, Communication/Revenue is communication spending scaled by total revenue. The IT spending scaled by revenue is at bank-county-year level. Control variables include net income scaled by total assets, deposits scaled by total assets, revenue per employee, salaries scaled by total assets and equity scaled by total assets, control variables are at bank-year level. Fixed effects include bank-year fixed effects and county-year fixed effects. Standard errors are clustered at county and bank level. $*p < 0.1, **p < 0.05, ***p < 0.01$.

	Δ log IT Emp	Δ log Non-IT Emp	Δ log IT Emp	Δ log Non-IT Emp
	(1)	(2)	(3)	(4)
Δ log Software	-0.148***	-0.117***		
	(0.0510)	(0.0431)		
Δ log Communication			0.0432**	0.0865***
			(0.0221)	(0.0223)
County×Year FE	Y	Y	Y	Y
Bank×Year FE	Y	Y	Y	Y
Control	Y	Y	Y	Y
R-squared	0.237	0.231	0.236	0.232
N	148617	148617	148617	148617