U.S. Banks' Artificial Intelligence and Small Business Lending: Evidence from the Census Bureau's Technology Survey

Jeffery Piao
University of Missouri-Columbia

K. Philip Wang
University of Florida

Diana L. Weng
Baruch College-City University of New York

September 23, 2025

Acknowledgments: We thank Jesse Chan, Kevin Chen, Keith Czerney, Yonca Ertimur, Elia Ferracuti, Tom Hagenberg, Seil Kim, Kalin Kolev, Hoyoun Kyung, Edward Li, Sara Malik, Trung Nyugen, Min Shen, and Elizabeth Tori for their thoughtful comments and suggestions. We thank conference participants at the 2024 Federal Statistical Research Data Center Conference at the University of Utah and workshop participants at the University of Missouri-Columbia, Emerging Scholars Brownbag Series, Baruch College-CUNY, Georgia State University, the University of Florida, and Florida State University for their helpful suggestions and comments.

U.S. Banks' Artificial Intelligence and Small Business Lending: Evidence from the Census Bureau's Technology Survey

Abstract

Utilizing confidential microdata from the Census Bureau's new technology survey, we shed light on U.S. banks' use of artificial intelligence (AI) and its effect on their small business lending. We find that the percentage of banks using AI increases from 14% in 2017 to 43% in 2019. Linking banks' AI use to their small business lending, we find that banks with greater AI usage lend significantly more to distant borrowers, about whom they have less soft information. Using an instrumental variable based on banks' proximity to AI vendors, we show that AI's effect is likely causal. In contrast, we do not find similar effects for cloud systems, other types of software, or hardware surveyed by Census, highlighting AI's uniqueness. Moreover, AI's effect on distant lending is more pronounced in poorer areas and areas with less bank presence. Last, we find that banks with greater AI usage experience lower default rates among distant borrowers and charge these borrowers lower interest rates, suggesting that AI helps banks identify creditworthy borrowers at loan origination. Overall, our evidence suggests that AI helps banks reduce information asymmetry with borrowers, thereby enabling them to extend credit over greater distances.

1. Introduction

Artificial intelligence (AI) is transforming the U.S. banking industry. Since 2017, J.P. Morgan has used AI software to review commercial loans, saving 360,000 hours of lawyer and loan officer time annually. Bank of America leveraged machine learning and natural language processing tools to assess borrowers' default risk during the Pandemic. In 2024, M&T Bank partnered with Rich Data Corp., an AI-powered decisioning platform, to enhance its small business and commercial lending. Seattle Bank streamlined its lending process in 2023 by adopting an AI-driven solution developed by JUDI.AI. According to McKinsey's 2021 estimates, AI has the potential to generate $1 trillion in annual value for the banking industry (McKinsey 2021). Despite these examples highlighting AI's importance, there is little empirical evidence on the extent to which U.S. banks use AI and its impact on their lending practices. This study aims to fill this gap by utilizing confidential microdata from the Census Bureau's new technology survey.

The Census Bureau's annual technology survey, implemented in 2018, is "one of the largest and most up-to-date data set available on advanced technology adoption in the world" (Zolas et al. 2020, p.5). It offers several advantages for measuring bank AI use.[1] First, compared with studies that focus on a firm's overall investment in technologies (e.g., Charoenwong et al. 2024), this survey specifically asks about the extent to which a firm uses certain technologies, such as AI. Second, in contrast with firms' voluntary disclosure of AI (e.g., Jia et al. 2024), responding to Census is mandated by law, mitigating selection bias. The survey's confidential nature and penalties for misreporting also help elicit honest and accurate responses from firms.

[1] Census technology surveys define AI as "a branch of computer science and engineering devoted to making machines intelligent. Intelligence is that quality that enables an entity to perceive, analyze, determine response and act appropriately in its environment" (Census Bureau 2019, p.23). It primarily includes machine learning, natural language processing, machine vision, and voice recognition (Zolas et al. 2020; McElheran et al. 2024). These technologies largely overlap with those defined by bank regulators as AI (Bowman 2024).

Third, a few recent studies measure AI based on a firm's in-house AI experts (e.g., Babina et al. 2024). However, these measures may underestimate AI adoption because firms often outsource AI development via partnerships with FinTech companies (Puri et al. 2024). Using firms' direct response to Census overcomes this problem. Last, the survey covers both private and public companies, so our sample offers a more comprehensive picture than those who only focus on public firms' AI adoption (e.g., Chen and Srinivasan 2024).

We find that about 22% of U.S. banks use AI to some extent from 2017 to 2019. The percentage of banks using AI increases from 14% in 2017 to 43% in 2019. Moderate or high use of AI by banks rises from 8% in 2017 to 22% in 2019. The high percentage of banks using AI is likely because many banks can access AI without developing it on their own (Mason 2023). For example, Kendall Bank CEO Tim Barron states, "[w]e don't have the capabilities to develop that AI for ourselves because it would cost millions. But we can use vendors who are pushing this type of AI product into the market for small banks" (Dornbrook 2024). Merging Census AI-use microdata with banks' annual Call Reports, we show that larger banks are more likely to use AI, but other bank characteristics are not associated with AI use.

We examine how AI affects U.S. banks' lending practices in the small business markets. Small businesses account for 44% of U.S. gross domestic product (GDP) and 63% of new job creation (U.S. Chamber of Commerce 2023). However, according to a Federal Reserve survey, 85% of small businesses face financial challenges, yet only 42% of them have their financing needs met (Federal Reserve Banks 2022). An important reason for the credit supply gap is that small businesses have little public information, and their private information is costly to collect and analyze, making their risk particularly difficult to assess. Banks traditionally rely on "soft" information gathered through local interaction with borrowers, which constrains their lending to those who are geographically nearby (Petersen and Rajan 2002; Granja et al. 2022).

AI may affect banks' small business lending in three aspects. First, it reduces banks' cost of collecting hard information, thereby enabling them to acquire more hard information about borrowers. Anecdotal evidence suggests that banks use image recognition tools and generative AI assistants to collect borrowers' raw financial records (Biz2X 2022; Crosman 2024). Second, it helps banks standardize and centralize unstructured data, reducing their cost of incorporating hard information into decision-making (Bowman 2024; Gargano 2024). Third, it makes banks use available hard information more effectively. For example, machine learning enables banks to identify nonlinear patterns in data overlooked by traditional statistical models, enhancing their ability to predict borrowers' default (Dryer 2018; Son 2017). As a result of improvement in banks' hard information, AI may reduce banks' reliance on soft information in lending.

To empirically examine the effect of AI on small business lending, we merge Census AI-use confidential microdata to publicly available Community Reinvestment Act (CRA) data, which contains the annual county-level volume of small business loans originated by banks with total assets above $1 billion. Following prior literature (e.g., DeYoung et al. 2008; Agarwal and Hauswald 2010), we measure banks' reliance on soft versus hard information based on how their lending varies with distance to borrowers. We break down the analysis by loan size because prior studies suggest that larger loans may not benefit from technology improvements due to banks' on-site monitoring (Adams et al. 2023).

We find that banks with higher AI use on average do not provide more overall credit to small businesses. However, they lend significantly more to distant borrowers, about whom banks have less soft information. A one-standard-deviation increase in AI use attenuates the negative relation between lending distance and bank loan growth by 46% for small loans (i.e., loan amounts below $100,000) and by 64% for large loans (i.e., loan amounts between $100,000 and $1 million).[2] These results cannot be explained by differential local credit demand because we control for borrower county × year fixed effects to compare lending within the same county-year across banks. When we discretize lending distance using Granja et al. (2022) cutoffs, we find that AI's effect on lending increases monotonically with distance and becomes statistically significant when distance is over 1,000 miles. These results provide empirical evidence supporting the prediction that AI reduces banks' reliance on soft information in lending.

A potential concern is that banks' AI use is endogenous and thus our results may be driven by unobserved shocks that drive both AI adoption and distant lending. To address this concern, we exploit a unique question from Census' 2019 technology survey that identifies firms that *sell* AI solutions (i.e., AI vendors). Recent studies indicate a strong pattern of AI's geographic diffusion, suggesting that firms located closer to AI developers are more likely to use AI (Hunt et al. 2024; McElheran et al. 2024; Muro and Liu 2021). Based on this idea, we construct an instrumental variable using the proximity of AI vendors to a bank's *headquarters*. Conceptually, this instrument is unlikely to correlate with a bank's distant borrowers, who are defined by their distance from the bank's nearest *branch*. We first validate this instrument by showing that the presence of AI vendors near the bank significantly increases the bank's AI use (i.e., the first stage). In the second stage, we show that instrumented AI use is significantly and positively associated with banks' lending to distant borrowers, enhancing our findings' causal inference.

[2] Untabulated analysis shows that banks with higher levels of AI are not more likely to close branches, so the increase in distant lending is unlikely related to banks' branch closure.

Another potential concern is that any technological improvement would increase banks' distant lending, so AI would be no different from other types of technologies. To mitigate this concern, we conduct a falsification test by exploiting banks' reported use of *other* advanced technologies from Census technology use microdata, which includes cloud systems, other types of software, and hardware.[3] Similar to AI, many of these technologies are gaining popularity among banks. For example, banks increasingly use cloud systems for data storage to reduce costs (Nguyen 2022). We do not find that banks with more usage of cloud systems, other software, or hardware lend more to distant borrowers. These results highlight the unique impact of AI compared with other recent technological improvements and alleviate the concern that all technologies would increase banks' distant lending.

In addition, we propose a mechanism through which AI affects distant lending: AI helps banks identify small businesses that are creditworthy but have been predicted to be risky by traditional lending models. For example, a machine learning model developed by Mirador Inc. allows banks to relax the thresholds for traditional risk metrics for small businesses while maintaining the same level of risk (Dryer 2018); an AI solution by Abrigo also enables banks to "approve businesses that may have been overlooked by conventional systems" (Business Wire 2024). To explore this mechanism, we partition the sample based on the county's percentage of population living in poverty and based on the number of banks serving small businesses in that county. We find that our results are significantly stronger for poorer counties and counties with less bank presence, suggesting that AI helps banks to extend credit to borrowers in underserved areas. In addition, we find that counties served by distant banks with higher AI usage experience

[3] Cloud systems is the only other technology that is consistently surveyed by Census from 2018 to 2020. Other software technologies (e.g., automation, enterprise resource planning) are mostly only surveyed in one year, so we group them into other software. Examples of hardware surveyed include server technologies, internet-connected devices, and robotics.

significantly greater loan growth compared to other counties, suggesting that AI enables banks to expand credit access rather than merely taking market share from other banks.

Although we show that AI helps banks extend credit to distant borrowers, it does not necessarily mean that AI helps banks identify borrowers with high credit quality. Early technological advances in the 1990s, such as credit scoring, have been shown to lead to worse loan performance (Berger et al. 2005; DeYoung et al. 2008). Therefore, an alternative explanation for the high growth in distant lending is that banks with more AI usage take on more risk. To examine the effect of AI on loan quality and banks' risk assessment at origination, we follow Granja et al. (2022) in using Small Business Administration (SBA) loan-level data.[4]

We find that banks using more AI experience significantly lower charge-offs from distant borrowers, suggesting that AI helps banks improve these borrowers' loan performance. When we discretize distance using Granja et al. (2022) cutoffs, we find that AI's effect on loan performance increases monotonically with distance. Different from the credit supply analyses, the improvements in loan performance are statistically significant for all groups of borrowers who are over 50 miles away, suggesting that AI's effect on loan performance may apply to a broader borrower base. A one-standard-deviation increase in AI use reduces these distant borrowers' charge-offs by 0.96 to 1.20%, corresponding to 58 to 72% of the unconditional mean of charge-off.

To determine whether banks anticipate these distant borrowers to be low risk at loan origination, we examine the association between AI and interest rates. We find that banks with more AI usage charge significantly lower interest spreads to distant borrowers, suggesting that they correctly assessed these borrowers as less risky at loan origination. Consistent with the loan performance results, the lower interest spread becomes statistically significant beginning at

[4] The CRA dataset does not contain loan performance or interest rates.

distances over 50 miles. A one-standard-deviation increase in AI use reduces these distant borrowers' interest spread by 17 to 25 bps, which corresponds to 3.4 to 5.0% of unconditional mean of interest spread. The lower risk premium charged at loan origination, along with these loans' better *ex-post* performance, suggests that AI helps banks identify distant, creditworthy borrowers.

Our study contributes to the literature on the effects of technologies on lending in three aspects. First, Demerjian (2024, p.5) highlights that "the bulk of the academic literature (e.g., Berg et al. 2022; Di Maggio et al. 2022) examines FinTech lending in the context of FinTech-only lenders, leaving the role of traditional banks using FinTech relatively underexplored." We answer his call by providing the first empirical effort to shed light on AI adoption among U.S. banks and AI's effect on banks' lending practices. To our knowledge, the only other concurrent study that touches on banks' fintech and small business lending is Chen et al. (2023). They find that local banks gain a larger share of small business loans following local newspaper closure, but this finding disappears either after 2010 or among counties that receive loans from banks with data-driven technologies. Our study differs from Chen et al. (2023) in that we document a *direct* effect of a bank's AI on its own small business lending, whereas they document that general data-driven technologies *moderate* the effect of newspaper closures on competition between local and nonlocal lenders, irrespective of who owns the technology.

Second, most of the recent FinTech research focuses on the improvement in efficiency of the loan origination process. In the context of the Paycheck Protection Program during the COVID-19 Pandemic, studies show that FinTech lenders distribute funds more equally across areas and at a faster pace (e.g., Erel and Liebersohn 2022; Howell et al. 2024; Core and De Marco 2024). Because these programs are fully government-guaranteed and forgivable, lenders do not screen borrowers based on their credit risk. In contrast, our study shows that AI helps banks identify distant creditworthy borrowers, underscoring its role in enhancing borrower screening effectiveness.

Third, earlier studies find that technological improvements (e.g., credit scoring) increased lender-borrower distance but at the cost of higher default rates between the 1970s and 1990s (Petersen and Rajan 2002; Berger et al. 2005; DeYoung et al. 2008). In contrast, we show that AI enables banks to increase lending to distant borrowers while reducing default rates. Additionally, we find that other advanced technologies do not affect distant lending, underscoring the heterogeneous effects of different technologies on lending.

Last, our findings may be of interest to bank regulators, who are concerned that banks could use technological innovations to take on excessive risk. For example, Michelle Bowman, the Federal Reserve governor, recently questioned, "how can regulators best balance the risks AI may pose to bank safety and soundness and financial stability with the need to allow for continued innovation?" (Bowman 2024). We offer evidence that AI enables banks to expand credit access without increasing risks, alleviating regulators' concern about banks' AI use in small business lending.

2. How could AI affect banks' lending?

In making lending decisions to small business borrowers, banks collect two types of information about borrowers when evaluating their risk: hard and soft information (Liberti and Petersen 2019; Vashishtha 2019). Hard information is defined as numerical and standardized data that can be transmitted in impersonal ways, while soft information is defined as qualitative and unverifiable data that is intricately linked to its context. Banks have limited hard information about small business borrowers, which restricts their ability to evaluate borrowers' risk and thus increases their demand for soft information (Minnis 2011).

However, banks' limited hard information is often not due to a lack of availability, but because of the high cost of collecting and analyzing it. For example, a small restaurant may have a large volume of financial records, but they are less organized, some scanned or even hand-written, and the data are in heterogeneous formats across restaurants. Since collecting this data can be costly while the profit from small business loan origination is limited, many banks in practice screen small businesses only using a narrow set of hard information such as the debt service coverage ratio and the business owner's credit score (Dryer 2018).[5] Consistent with this anecdotal evidence, Liu (2022) uses loan-level data from a small business lender and shows empirically that loan officers process only a small portion of hard information that is useful for predicting borrower defaults.

AI can affect small business lending in three aspects. First, it reduces banks' cost of collecting data. For example, an AI module developed by Biz2X allows banks to read and automatically extract borrowers' raw financial records using image recognition and natural language processing (Biz2X 2022). Cascading AI, another FinTech startup, developed an AI assistant to help Bankwell Bank collect borrowers' information. By reducing information acquisition costs, AI allows banks to obtain more data about borrowers with the same budget constraints (Blankespoor et al. 2020; Even-Tov et al. 2024).

Second, AI reduces banks' cost of integrating data into decision-making. For example, Citigroup partners with Numerated, a FinTech company, to centralize borrower data "from

[5] For example, Dryer (2018) states that "bankers, being human, have limited time and brainpower. Faced with an edge-case scenario, they'll look at additional variables that complete the bigger picture. But if those additional variables require tracking down more data manually, bankers reach a limit to how much energy they can devote to one loan application."

disparate sources in diverse formats" into a singular dashboard (Gargano 2024). Similarly, Michelle Bowman, the Federal Reserve governor, points out that AI is frequently used by banks for "reviewing and summarizing unstructured data" (Bowman 2024). David Donovan, head of a digital consulting firm, observes that many banks adopt AI to integrate alternative data into their loan approval processes (Gargano 2024).

Third, AI helps banks use hard information more effectively (Crosman 2024). For example, a Moody's report states that "[a] machine learning model, unconstrained by some of the assumptions of classic statistical models, can yield much better insights that a human analyst could not infer from the data" (Moody's 2017). Similarly, Gord Baizley, CEO of JUDI.AI–a FinTech company that partnered with Seattle Bank–explains, "our small business-specific credit model, powered by machine learning, uses real-time, customer-permissioned bank transaction data to offer a more accurate snapshot of the financial health of the business" (Lau 2024).

AI's effect on banks' hard information has implications for banks' credit supply (i.e., extensive margin) as well as loan performance and pricing (i.e., intensive margin). Conceptually, the improvement in hard information allows banks to rely less on soft information while still maintaining the accuracy of their risk assessments. As a result, banks are now able to extend credit to borrowers about whom they have less soft information. However, whether AI improves banks' hard information in practice is not obvious. For example, a premature image recognition tool may assign numbers extracted to incorrect data fields, compromising input data quality. Machine learning may produce hallucinations due to insufficient training data or unrealistic model assumptions. Given these potential problems, whether banks are willing to trust AI and place significant weight on it (and thus less weight on soft information) is an empirical question.

Regarding loan performance and pricing, AI's effect is also not obvious. On one hand, if AI reduces the bank's demand for soft information, the loss of soft information can result in worse loan performance (Agarwal and Ben-David 2018; Gallo et al. 2023). On the other hand, hard and soft information can be correlated, so the newly collected hard information could substitute for some soft information (He et al. 2024). Also, more effective utilization of hard information can sometimes triangulate soft information in predicting borrower defaults (Fuster et al. 2022). Last, recent studies show that using soft information does not always improve banks' risk assessment, depending on loan officers' cognitive constraints and behavioral biases (Campbell et al. 2019).

Empirically, the most common way to disentangle a bank's reliance on hard versus soft information is to observe how a bank's lending varies with the distance between the bank and its borrowers (Liberti and Petersen 2019). As banks can generate soft information through frequent in-person contact with local borrowers but cannot do so with distant borrowers, they lend disproportionally less to distant borrowers (Petersen and Rajan 2002; Granja et al. 2022).

3. Data Sources

3.1 Census Bureau's Technology Survey

Lack of data has been a key impediment to researching AI's impacts (Seamans and Raj 2018). Partnering with the National Center for Science and Engineering Statistics, the Census Bureau introduced the Annual Business Survey (ABS) in 2018. The technology module in this new survey represents Census's first effort to identify adoption rates of advanced technologies and the extent to which each technology is used. Notably, Zolas et al. (2020, p.33) states that "one of the primary goals of the ABS technology module is to provide the first comprehensive look into the adoption rates of Artificial Intelligence (AI) by US firms."

ABS is conducted on a company basis rather than on an establishment (i.e., branch) basis. All surveyed firms are required by law to respond (*Census Act, 13 U.S.C. § 224* 1954), significantly alleviating selection and nonresponse bias. The first ABS was sent out in June 2018, with a reference period of calendar year 2017. It surveys a representative sample using stratified systematic sampling of approximately 850,000 firms every 5 years and approximately 300,000 firms annually (Zolas et al. 2020). As a result, the sample in the initial year of 2018 covers more firms than those in 2019 and 2020.

The list of technologies included in the module varies across years. For example, in 2019, the module asked respondents about their use of AI, cloud systems, robotics, specialized software, and specialized equipment. However, in 2020, it instead asked about AI, cloud systems, computer infrastructure, automation, internet-connected devices, mobile communication technologies, digital technologies for collaboration and communication, digital technologies for planning and management, and blockchain.

The Census's technology survey offers three unique advantages to our study. First, it allows us to directly measure a bank's AI use. Second, it provides data on banks' use of other advanced technologies, so we could run a falsification test and contrast AI's impact with those of other advanced technologies. Third, it also provides a list of AI vendors, which allows us to construct an instrumental variable and provide causal inferences.

3.2 The Community Reinvestment Act (CRA) dataset

Following prior literature (e.g., Granja et al. 2022; Cortés et al. 2020), we use data on small business loan originations collected by the Federal Financial Institutions Examination Council (FFIEC) pursuant to the Community Reinvestment Act (CRA). Small business loans are defined as non-farm commercial or industrial loans whose principal amount does not exceed $1 million. The CRA data is further disaggregated based on loan amount: less than $100,000, between $100,000 and $250,000, and between $250,000 and $1 million.

All commercial banks regulated by the Office of the Comptroller of the Currency (OCC), Federal Reserve System, and Federal Deposit Insurance Corporation (FDIC) that exceed certain asset thresholds are required to report the number and aggregated dollar amount of loans originated in each county on an annual basis. The asset threshold was raised from $250 million to $1 billion in 2005 and is adjusted annually for inflation. During our sample period, the asset threshold ranges from $1.226 billion (2017) to $1.305 billion (2020).

3.3 The Small Business Administration (SBA) dataset

Because the CRA dataset aggregates loans at the bank-county-year level, it does not contain loan-level information such as default or interest rates. Therefore, following Granja et al. (2022), we use the SBA dataset of loans originated under the 7(a) program, the SBA's primary lending program, to conduct loan-level tests. This dataset includes information about borrowers and lenders (e.g., identity, address, and industry), loan terms (e.g., approval date, loan amount, guarantee portion, maturity, and interest spread plus base rate), and *ex-post* loan performance (e.g., the loan balance that has been charged off).

SBA guarantees a portion of the loan principal balance against losses upon defaults. Specifically, SBA guarantees up to 50% for express loans, the largest category of the 7(a) program. For other regular programs, SBA guarantees up to 85% for loans under $150,000 and up to 75% for loans over $150,000. The guaranteed rate does not always reach the allowable maximum because borrowers often try to lower guarantee fees, which increase with the guaranteed rate (Huang 2024). Note that SBA does not take a first-loss position upon default; instead, it covers the guaranteed portion of the *remaining* outstanding balance and delinquent interest (Glennon and Nigro 2005). As a result, lenders have incentives to screen borrowers at loan origination to minimize their default risk.

3.4 Other datasets

Banks' financial data: We obtain commercial banks' financial information from their Call Reports. We merge banks' financial data to Census' ABS survey using the nine-digit Employer Identification Number (EIN) assigned by the IRS.

Banks' branch data: FDIC's Summary of Deposits (SOD) data provides branch-level data on geographic locations of banks' headquarters and branches on an annual basis. We use this data to calculate the distance between lenders and borrowers in the CRA and SBA datasets.

Economy data: We obtain GDP data from the Federal Reserve Economic Data (FRED) website of the Federal Reserve of St. Louis. We obtain the percentage of each county's population living in poverty from the Census' Small Area Income and Poverty Estimates (SAIPE) program.

4. Empirical Results

All results that use confidential Census microdata need to go through a disclosure avoidance review process by Census personnel prior to being disclosed to ensure that no individual bank responses can be inferred. Because each reported number is vetted to ensure confidentiality, we do not tabulate coefficients and *t*-statistics for control variables in regression analyses.

4.1 The AI adoption rate

To shed light on AI adoption in the U.S. banking industry, we rely on banks' responses to AI-related question(s) in each year's ABS survey. The ABS in both 2019 and 2020 directly asked about AI, but the 2018 ABS survey did not and instead asked about more specific technologies.

Therefore, following Zolas et al. (2020) and McElheran et al. (2024), we define "machine learning," "machine vision," "natural language processing," and "voice recognition" as AI.[6] Because banks' AI use can change across years, we do not extrapolate AI levels for unsurveyed years using values from surveyed years to avoid introducing measurement errors. Appendix B provides details of the construction of the AI variable.

Merging ABS with banks' Call Reports results in a sample of 1,500 (rounded) bank-year observations from 1,100 (rounded) unique banks. We find that, on average, 22% of banks use AI during our three-year sample period. When broken down by AI level, 11% of banks use AI to a low extent, 8% to a moderate extent, and 3% to a high extent.[7] The percentages of banks using AI in 2017, 2018, and 2019 are 14%, 17%, and 43%, showing increasing use of AI in the banking industry. Moderate or high use of AI by banks increases from 8% in 2017 to 22% in 2019.

4.2 What bank characteristics are associated with AI use?

Zolas et al. (2020, p.3) state that "our understanding of how and why firms adopt new technologies is still rather imprecise," suggesting limited research on what affects firms' use of AI. To explore this question in the banking industry, we run the following model using the bank-year sample:

$$\begin{aligned} \textit{AI} = {} & \textit{Size} + \textit{Public} + \textit{ROA} + \textit{Leverage} + \textit{Liquid Assets} + \textit{Securities} + \textit{Loans} \\ & + \textit{Residential Loans} + \textit{Commercial Loans} + \textit{Nonperforming Loans} + \textit{Derivatives} \\ & + \textit{Deposits} + \text{Year FEs} + \text{State FEs} + \varepsilon \end{aligned} \quad (1)$$

AI, the extent to which a bank uses AI, equals 0 if no use, 1 if testing, 2 if low use, 3 if moderate use, and 4 if high use. We include 12 bank characteristics as possible determinants of

[6] Zolas et al. (2020) also identify "automated guided vehicles" as AI, but we do not include it as it is not relevant to banks.
[7] Because there are multiple technologies in 2018, we define AI use based on the highest level of usage among them.

AI. These characteristics include basic financials (i.e., *Size, ROA, Leverage*), asset composition (i.e., *Liquidity, Securities, Loans*), loan composition (*Residential Loans, Commercial Loans, Nonperforming Loans*), and banks' other major business components (*Derivatives, Deposit*). Appendix A provides variable definitions, and Table 1, Panel A presents the descriptive statistics for the variables included in the determinants test. In addition, we include year and state fixed effects and cluster standard errors by bank.

Table 2 presents the results of this determinants analysis. We find that only *Size* is positively associated with *AI*, suggesting that larger banks are more likely to use AI. However, all other bank characteristics are not significantly associated with AI. These results suggest that banks' use of AI may not be related to technological demand for certain business lines or types of assets. The 11% R^2 suggests that nearly 90% of variation in AI use among banks cannot be explained by bank characteristics or fixed effects. Relatedly, He et al. (2022) examine the factors that affect banks' spending on software and communication IT products. While they show that real estate loans are associated with more software IT products, we find that it is not associated with AI use.

4.3 The effect of AI on the quantity of small-business loans

To examine the effect of AI on banks' supply of small business loans, we rely on the CRA dataset, which is at the bank-county-year level. Following Adams et al. (2023), we exclude bank-county observations in Alaska, Hawaii, and U.S. territories (e.g., U.S. Virgin Islands, Puerto Rico, etc.), because observations in these locations introduce discrete jumps in the distance between the borrowers' county and the bank. Merging the bank-year sample with AI and bank controls to the CRA dataset yields 172,000 (rounded) bank-county-year observations. Panel B of Table 1 reports the descriptive statistics of this sample. We find that the average distance between a bank and its borrower's county is 459.6 miles. The bank-county-years with distances between 50 and 250 miles, between 250 and 1,000 miles, and over 1,000 miles account for 22%, 36%, and 14% of the sample.

We run the following model using this sample:

$$\begin{aligned}\Delta Ln(SBL)_{i,c,t+1} = {} & \beta_1\, AI_{i,t} + \beta_2\, Ln(Distance_{i,c}) + \beta_3\, AI_{i,t} \times Ln(Distance_{i,c}) \\ & + \beta_4\, Size_{i,t} \times Ln(Distance_{i,c}) + \beta_5\, \Delta GDP_{t+1} \times Ln(Distance_{i,c}) \\ & + \text{Bank Controls}_{i,t} + \text{County} \times \text{Year FEs} + \varepsilon \end{aligned} \quad (2)$$

The model resembles that of Granja et al. (2022) except that we include *AI*, its interaction with *Ln*(*Distance*), and more control variables but exclude bank fixed effects. Consistent with Granja et al. (2022), we define Δ*Ln*(*SBL*) as the change in natural logarithm of one plus the total dollar amount (in thousands) of small business loans by bank *i* in county *c* from year *t* to *t*+1. A larger Δ*Ln*(*SBL*) implies a greater growth in credit supply. Because only borrowers' county is available in the CRA dataset, we calculate *Distance* as the geodetic distance between the centroid of the borrower's county and the bank's closest branch. If AI helps banks increase credit supply, we expect β_1 to be positive when *AI* × *Ln*(*Distance*) is not included in the model. If AI helps banks extend credit to distant borrowers, we expect β_3 to be positive.

We control for Δ*GDP* × *Ln*(*Distance*) because Granja et al. (2022) show that banks lend more to distant borrowers when the economy is booming. We do not include Δ*GDP* itself in the model because it would be absorbed by county × year fixed effects. We also control for *Size* × *Ln*(*Distance*) because *Size* is the only bank characteristic associated with *AI* in the determinants analysis. We cannot control for bank fixed effects because 68% of banks in our sample are only surveyed once. Because the absence of bank fixed effects may raise concerns about endogeneity, we conduct an instrumental variable analysis to address this issue (see Section 4.4). Last, we include all bank characteristics used in Equation (1) and borrower county × year fixed effects. We cluster standard errors by bank. Following Levine et al. (2020), Adams et al. (2023), and Bord et al. (2021), we estimate the regression separately for large versus small loans. Large (small) loans are defined as those with a principal amount greater than (less than) \$100,000.

Table 3 presents the regression results for Equation (2). Columns (1) and (2) show that AI is not associated with Δ*Ln*(*SBL*), suggesting that AI does not impact a bank's overall quantity of small business lending. In Columns (3) and (4) where we include *AI* × *Ln*(*Distance*), we find that banks with more AI lend significantly more to borrowers as their distance to the bank increases. This result holds for both small and large loans, despite generally greater monitoring of large loans. Because larger banks are more likely to use AI, we additionally control for *Size* × *Ln*(*Distance*) in Columns (5) and (6). We find that the results hold, suggesting that our results are not driven by bank size. Regarding economic magnitude, a one-standard-deviation increase in AI use attenuates the negative relation between lending distance and bank loan growth by 46% for small loans and by 64% for large loans.[8] Last, consistent with the notion that banks tend to lend more to local borrowers, we find a negative coefficient on *Ln*(*Distance*) in most columns. Note that the coefficients on *AI* in Columns (3) to (6), where we include *AI* × *Ln*(*Distance*), are not interpretable, because they reflect the effect of *AI* on loan growth only when *Ln*(*Distance*) equals zero, which barely exists in the sample period (Burks et al. 2019; Aiken and West 1991, p.38).

To investigate the distances at which borrowers are affected by AI, we discretize distance into four bins based on the three cutoffs (i.e., 50 miles, 250 miles, and 1,000 miles) used in

[8] The calculation details are as follows: 46% = 1.305 (standard deviation of AI from Panel B of Table 1) × 0.055 (coefficient on *AI* × *Ln*(*Distance*) in Column (3)) / 0.157 (coefficient on *Ln*(*Distance*) in Column (3)); 64% = 1.305 (standard deviation of AI) × 0.037 (coefficient on *AI* × *Ln*(*Distance*) in Column (4)) / 0.075 (coefficient on *Ln*(*Distance*) in Column (4)). Note that we rely on Column (3) and (4) to interpret the magnitude because adding *Size* × *Ln*(*Distance*) in Columns (5) and (6) means that the coefficient on *Ln*(*Distance*) reflects the slope of *Ln*(*Distance*) when *Size* equals 0 (i.e., bank total assets equals \$10 million).

Granja et al. (2022). We replace *Ln*(*Distance*) with the three binary variables and interact AI with these three variables. Table 4 reports the results. Consistent with the notion that banks on average lend more to local borrowers, we find that the coefficients on the distance bin indicators become more negative as lending distance increases. Interestingly, we observe that coefficients on *AI* × distance bin indicators are all positive and monotonically increase as lending distance increases. However, we observe that only the coefficient on *AI × 1000 miles* is statistically significant, suggesting that AI helps banks to extend credit primarily to very distant borrowers who are located at least 1,000 miles away.

We conduct two robustness tests. First, we alternatively measure loan growth as loans in *t*+1 less loans in *t*, scaled by average loans in *t* and *t*+1. Second, like Granja et al. (2022), we alternatively control for the interaction between *Ln*(*Distance*) and percentage change in unemployment rate or the interaction between *Ln*(*Distance*) and the net percentage of domestic banks increasing spreads of loan rates over banks' cost of funds to small firms. We find that our results in Table 3 hold, as the coefficient on *AI* × *Ln*(*Distance*) remains positive and significant (untabulated).

4.4 Addressing the endogeneity concern

To address the concern that banks' AI use is endogenous, we conduct an instrumental variable test. Following the recent literature on AI's geographic diffusion (e.g., Hunt et al. 2024), we posit that the presence of AI vendors located close to the bank's *headquarters* significantly increases a bank's probability to use AI. This instrument introduces relatively exogenous variation in AI because banks rarely relocate their headquarters, and this decision is unlikely driven by proximity to AI vendors. Also, it is unlikely that an AI vendor near a bank's headquarters is associated with the bank's distant borrowers, who are defined by their distance from the bank's nearest *branch*.

Specifically, the Census's 2019 technology survey not only collects data on firms' AI use, but also asks whether the surveyed firm sells AI or goods and services that include AI. To identify AI vendors, we restrict the firms to those who answered "yes" to this question, are in the software industry (i.e., NAICS code = 5112), and indicated that their motivation for selling AI is to either "upgrade goods or services" or "enter new markets or adapt existing products to new markets." We construct the instrumental variable, *AI Vendor*, based on whether there is an AI vendor located within a 150-mile radius of the bank's headquarters. We collect banks' headquarters from their Call Reports and measure the geodetic distance between the bank's headquarters' ZIP code and the AI vendor's ZIP code.[9]

Because the variable of interest in the main model is an interaction term (i.e., *AI* × *Ln*(*Distance*)), simply replacing *AI* with predicted *AI* in the second stage will run into the "forbidden regression" problem (see details on p.236, Wooldridge (2010)). To avoid complications from this issue, we restrict our sample to distant borrowers only (i.e., those located 1,000 miles aways from the bank) to remove the interaction term from our main model, resulting in a model with *AI* itself as the main variable. By doing so, we still focus on the same set of borrowers that are affected by AI based on our findings in Table 4 and now can run a traditional two-stage least-squares (2SLS) test.

Table 5 presents the results. In Column (1), we first replicate the main result after restricting the sample to borrowers who are more than 1,000 miles away. Consistent with Table 4, we find that banks with higher AI usage increase lending to these distant borrowers. Column (2) reports the first-stage result, where we regress *AI* on *AI Vendor* and all other bank controls

[9] We use ZIP codes because AI vendors' street addresses are not available in the Census survey.

and fixed effects. We find that having an AI vendor nearby significantly increases banks' AI use. The Kleibergen-Paap F-statistic of 13.2 is higher than the conventional threshold of 10, suggesting that it is not a weak instrument (Stock et al. 2002, p.522). Column (3) reports the second-stage result, where we regress loan growth on predicted AI from the first stage and other bank controls and fixed effects. We find that the instrumented AI is significantly and positively associated with banks' lending to distant borrowers, suggesting that AI's effect on distant lending is likely causal. The coefficient magnitude of 0.889 is larger than but comparable to that in Column (1).

4.5 Do all advanced technologies increase distant lending?

In this section, we examine whether other advanced technologies have a similar effect on distant lending as AI does. Earlier research shows that technological advancements from the 1970s to the 1990s led to an increase in distance between banks and borrowers at the time (e.g., Petersen and Rajan 2002). Therefore, it is possible that any type of technological improvement could lead to more distant lending.

We utilize other questions in the ABS technology module to conduct these tests. Specifically, non-AI technologies in the 2018 ABS include cloud service, augmented reality, automated storage and retrieval system, radio-frequency identification inventory system, robotics, and touchscreens/kiosks for customer interface. Non-AI technologies in the 2019 ABS include cloud-based computing systems, specialized software, robotics, and specialized equipment. Non-AI technologies in the 2020 ABS include cloud computing, computer infrastructure, automation, internet-connected devices, mobile communication technologies, digital technologies for collaboration, communication, enterprise resource planning, and blockchain. See Appendix B for more details.

Among all non-AI technologies, cloud systems is the only one consistently surveyed about during our sample period, albeit with slightly different wording. Therefore, we categorize the non-AI technologies into three groups: cloud systems, other software, and hardware. Other software includes automated storage and retrieval systems, radio-frequency identification inventory system, and touchscreens/kiosks for customer interface for year 2017; specialized software for year 2018; and automation, mobile communication technologies, digital technologies for collaboration, communication, enterprise resource planning, and blockchain for year 2019. Hardware includes robotics in 2017; robotics and specialized equipment in 2018; and computer infrastructure and internet-connected devices in 2019. Similar to how we define *AI*, we construct three continuous measures to capture the extent to which the bank uses cloud systems (*Cloud*), other software (*Other Software*), and hardware (*Hardware*).

The mean of *Cloud* is 2.51, meaning that banks' average use of cloud systems is between low and modest. The mean use of other software (i.e., 2.82) is slightly higher than that of cloud systems, while the mean use of hardware (i.e., 1.89) is slightly lower. The standard deviations in all three technology variables are of similar magnitude as *AI*, suggesting that there is meaningful variation in the use of these technologies among U.S. banks.

Table 6 presents the regression results after we replace *AI* with *Cloud*, *Other Software*, and *Hardware* in Equation (2). We find that the coefficients on the interaction terms between these non-AI technologies and distance are not statistically significant, suggesting that banks with more non-AI technologies do not lend more to distant borrowers. These results contrast with those for AI and highlight that AI's effect on distant lending appears to be unique.

4.6 Are AI's effects stronger for underserved areas?

A possible mechanism for the increase in distant lending by AI is that AI helps banks identify creditworthy borrowers who would otherwise be overlooked by traditional credit assessment models. Testing this mechanism requires identifying borrowers with historically less access to bank loans. Because CRA is aggregated at the bank-county level, we use two measures to capture these borrowers at the county level: the percentage of population living in poverty and the number of banks serving small businesses in the county. If AI helps banks identify new creditworthy borrowers, we predict our results to be stronger among poorer counties and counties with less bank presence.

Table 7 reports the results after we partition the sample by median poverty rate in Panel A and bank density in Panel B. We find that the coefficient on *AI* × *Ln*(*Distance*) is significantly positive across all four subsamples. The coefficient is significantly larger for poorer counties (χ^2 = 4.08, *p*-value = 0.04) and counties with less bank presence (χ^2 = 2.98, *p*-value = 0.08), suggesting that AI enables banks to extend credit to borrowers that traditionally have fewer borrowing opportunities.

4.7 AI's effect on the total credit supply to distant counties

Our finding that AI enables banks to lend more to distant counties leads to the question of how the total credit supply to these counties is affected. If AI merely helps banks take market share from existing local banks, we may not observe an overall increase in loan growth among these counties. To answer this question, we convert the CRA bank-county-year sample into county-year observations to study county-level loan growth. We retain observations with at least one bank lending from over 1,000 miles away. Our independent variable of interest (*AI Bank*) is an indicator variable that identifies the counties to which a bank with AI usage lends from over 1,000 miles away. Online Appendix A1 shows that the coefficient on *AI Bank* is positive and significant, suggesting that these distant counties served by AI banks experience higher loan growth. This finding provides further evidence that AI enables banks to meet unmet credit needs in distant counties.

4.8 AI's effect on loan performance

To examine the effect of AI on loan performance, we rely on the SBA dataset, which is at the individual loan level. Consistent with our other tests, we only include loans where the lenders and borrowers are both located in the 48 contiguous U.S. states. To facilitate calculation of lender-borrower distance, we exclude borrowers without valid street addresses, such as those with addresses listed as post office boxes or lot numbers. We then manually clean borrowers' street addresses before converting the addresses into geographical coordinates using the Census Bureau TIGER/Line shapefiles. Following Brown and Earle (2017) and Granja et al. (2022), we do not include canceled loans. Merging the bank-year sample with AI and bank controls to the SBA dataset yields 38,000 (rounded) loans.

Panel C of Table 1 reports the descriptive statistics of this sample. We find that the average charge-off percentage is 1.7%. The average lending distance is 110 miles and loans with distances between 50 and 250 miles, between 250 and 1,000 miles, and over 1,000 miles account for 5%, 8%, and 4% of the sample. Compared with the CRA sample, the SBA sample has a smaller share of distant loans, likely because the unit of observation is a bank-county-year in CRA but is an individual loan in SBA. The mean portion guaranteed by SBA is 61%, suggesting that banks retain significant "skin in the game" by bearing an average of 39% of the loss if a loan defaults.

We run the following model using this sample:

$$\begin{aligned} ChargeOff_{l,i,t+1} = \; & \beta_1 AI_{i,t} \times Ln(Distance_{i,c}) + \beta_2 AI_{i,t} + \beta_3 Ln(Distance_{i,c}) \\ & + \beta_4 Size_{i,t} \times Ln(Distance_{i,c}) + \beta_5 \Delta GDP_{t+1} \times Ln(Distance_{i,c}) \end{aligned}$$

$$+ \text{Loan Controls}_{l} + \text{Bank Controls}_{i,t} + \text{SBA program FEs}$$
$$+ \text{County} \times \text{Year FEs} + \text{Borrower Industry FEs} + \varepsilon \quad (3)$$

where l indexes loans, i indexes banks, and t indexes origination years. *ChargeOff* equals the amount that has been charged off scaled by the loan's principal amount. If the loan does not have any charge-off, *ChargeOff* equals 0. Because borrowers' addresses are available in the SBA dataset, we calculate *Distance* as the geodetic distance between the borrower's address and the bank's closest branch. If AI helps banks lower the default rate of loans to distant borrowers, we expect β_1 to be negative.

We include the same interaction variables (*Size* × *Ln*(*Distance*) and Δ*GDP* × *Ln*(*Distance*)), the same set of bank controls (*Size*, *Public*, *ROA*, *Leverage*, *Liquid Assets*, *Securities*, *Loans*, *Residential Loans*, *Commercial Loans*, *Nonperforming Loans*, *Derivatives*, *Deposits*), and county × year fixed effects as Equation (2). In addition, we include loan controls (*Loan Amount*, *Maturity*, and *Guaranteed*), borrower industry (3-digit NAICS code) fixed effects, and SBA program fixed effects (e.g., SBA Express and Preferred Lending Program). See variable definitions in Appendix A. We cluster standard errors by bank.

Table 8 presents the regression results. In Column (1), we find that the loan experiences a higher charge-off as the distance between lenders and borrowers increases. This result is consistent with DeYoung et al. (2008) and suggests that distant borrowers are on average riskier. In Column (2) where we include *AI* × *Ln*(*Distance*), we find that the coefficient on *AI* × *Ln*(*Distance*) is significantly negative, suggesting that banks with more AI use experience lower defaults from distant borrowers. The results hold even after controlling for *Size* × *Ln*(*Distance*) in Column (3).

In Column (4), we discretize distance using the same cutoffs as in the credit supply analyses. We find that the coefficients on *AI* × distance bin indicators are all significantly negative and monotonically decrease as lending distance increases, suggesting that AI helps banks improve performance of loans to borrowers located over 50 miles away. A one-standard-deviation increase in AI approximately reduces charge-offs by 0.96% for borrowers located 50 to 250 miles and 250 to 1,000 miles away and by 1.20% for borrowers located over 1,000 miles away. As a benchmark, the coefficients on the distance bin indicators are significantly positive and suggest that relative to the baseline (i.e., borrowers located within 50 miles), those located 50 to 250 miles away on average experience 1.2% higher charge-offs, those located 250 to 1,000 miles experience 1.1% higher charge-offs, and those located over 1,000 miles away experience 1.6% higher charge-offs. Therefore, the 0.96 to 1.20% reduction in charge-offs offsets 75 to 87% of incremental charge-off from these distant borrowers.

We conduct two robustness tests. First, we convert the continuous *ChargeOff* variable to a binary variable that equals one if the loan experienced any charge-off and zero otherwise. Following Breuer and deHaan (2024) and Kielty et al. (2023), we estimate an OLS model to avoid the incidental parameter problem. Second, we measure the continuous *ChargeOff* variable using a three-year performance window for comparability in loan performance across years. We rerun the performance test using these alternative dependent variables. We find that our inference remains the same (i.e., interaction on *AI* × *Ln*(*Distance*) remains negative and statistically significant) (untabulated).

4.9 AI's effect on interest spread

To examine the effect of AI on interest spread, we use the same loan-level SBA sample as the loan performance test and re-run Equation (3) with *Interest Spread* as the dependent variable.

Interest Spread is defined as the interest rate at loan origination minus the concurrent Treasury rate with the closest maturity (i.e., the risk-free rate). SBA sets a maximum rate of the prime rate plus 2.25% (2.75%) for loans with principal amount of more than $50,000 and maturity of less than seven years (seven years or more). The standard deviation of interest spread is 1.537%, representing meaningful variation in interest spreads in our sample despite SBA's ceilings.

Table 9 presents the regression results. In Column (1), we find that lenders charge significantly higher interest rates as lending distance increases, suggesting that lenders perceive distant borrowers as riskier. In Column (2) where we include *AI* × *Ln*(*Distance*), we find that the coefficient on *AI* × *Ln*(*Distance*) is significantly negative, suggesting that banks with more AI charge significantly lower interests to distant borrowers. The results hold even after controlling for *Size* × *Ln*(*Distance*) in Column (3). These results indicate that banks with greater AI usage accurately assessed the distant borrowers they lent to, who demonstrated better loan performance *ex post*, as low risk at loan origination.

In Column (4), we discretize distance using the same cutoffs as in the credit supply and loan performance analyses. We find that the coefficients on *AI* × distance bin indicators are all significantly negative and generally decrease as lending distance increases, corroborating the inference that AI primarily helps banks identify creditworthy borrowers located over 50 miles away at loan origination. A one-standard-deviation increase in AI approximately reduces interest spread by 0.17% for borrowers located 50–250 miles away, by 0.25% for borrowers located 250–1,000 miles away, and by 0.23% for borrowers located over 1,000 miles away. As a benchmark, the coefficients on distance indicators are significantly positive and suggest that relative to the baseline (i.e., borrowers located within 50 miles), those located 50 to 250 miles away on average pay 0.35% higher interests, those located 250 to 1,000 miles pay 0.47% higher interests, and those located over 1,000 miles away pay 0.74% higher interests. Therefore, a 0.17 to 0.25% reduction in interest spread offsets 32 to 53% of incremental interest cost for these distant borrowers. Our results are robust to using the raw interest rates (without taking out the risk-free rate) (i.e., the coefficient on $AI \times Ln(Distance)$ remains negative and significant).

5. Conclusion

In contrast with the emerging importance of AI and the abundant anecdotal evidence of its rapid adoption in the U.S. banking industry, scant empirical evidence exists for its effect on banks. We fill this void by using confidential microdata from Census Bureau's new technology survey. We provide an overall picture of AI adoption in the U.S. banking industry. Specifically, we show an increasing trend in AI use, rising from 14% in 2017 to 43% in 2019. AI use appears to be driven primarily by a bank's overall size, rather than specific business lines (e.g., loans, or securities).

In the context of small business markets, we provide evidence on how AI changes banks' lending practices. We find that banks with greater AI use lend significantly more to distant borrowers. Using an instrumental variable, we mitigate the concern that banks endogenously adopt AI. We contrast AI with other advanced technologies, such as cloud systems, and show that the effect of AI on distance lending appears to be unique. AI's effect is more pronounced for poorer counties and counties with less bank presence. A supplemental test suggests that AI increases the overall credit supply to distant counties.

Moreover, AI helps improve the performance of loans to distant borrowers and lower the interest spread charged to these borrowers, consistent with the notion that AI helps banks identify creditworthy, distant borrowers at loan origination. Overall, our findings suggest that AI reduces banks' reliance on soft information in lending by improving their ability to collect and analyze hard information. The benefit of improved use of hard information appears to outweigh the loss of soft information, leading to better loan performance.

References

Adams, R. M., K. P. Brevoort, and J. C. Driscoll. 2023. Is lending distance really changing? Distance dynamics and loan composition in small business lending. *Journal of Banking & Finance* 156:107006.

Agarwal, S., and I. Ben-David. 2018. Loan prospecting and the loss of soft information. *Journal of Financial Economics* 129 (3):608-628.

Agarwal, S., and R. Hauswald. 2010. Distance and Private Information in Lending. *The Review of Financial Studies* 23 (7):2757-2788.

Aiken, L. S., and S. G. West. 1991. *Multiple regression: Testing and interpreting interactions*. Thousand Oaks, CA, US: Sage Publications, Inc.

Babina, T., A. Fedyk, A. He, and J. Hodson. 2024. Artificial intelligence, firm growth, and product innovation. *Journal of Financial Economics* 151:103745.

Berg, T., A. Fuster, and M. Puri. 2022. FinTech Lending. *Annual Review of Financial Economics* 14 (Volume 14, 2022):187-207.

Berger, A. N., W. S. Frame, and N. H. Miller. 2005. Credit Scoring and the Availability, Price, and Risk of Small Business Credit. *Journal of Money, Credit and Banking* 37 (2):191-222.

Biz2X. *Understanding the Role of Artificial Intelligence in Lending* 2022 [cited. Available from https://www.biz2x.com/blog/understanding-the-role-of-artificial-intelligence-in-lending/.

Blankespoor, E., E. deHaan, and I. Marinovic. 2020. Disclosure Processing Costs, Investors' Information Choice, and Equity Market Outcomes: A Review. *Journal of Accounting and Economics* 70 (2-3):1-46.

Bord, V. M., V. Ivashina, and R. D. Taliaferro. 2021. Large banks and small firm lending. *Journal of Financial Intermediation* 48:100924.

Bowman, M. D. 2024. Speech by Governor Bowman on banking policy developments. *Board of Governors of the Federal Reserve System*.

Breuer, M., and E. deHaan. 2024. Using and Interpreting Fixed Effects Models. *Journal of Accounting Research* 62 (4):1183-1226.

Brown, J. D., and J. S. Earle. 2017. Finance and Growth at the Firm Level: Evidence from SBA Loans. *The Journal of Finance* 72 (3):1039-1080.

Burks, J. J., D. W. Randolph, and J. A. Seida. 2019. Modeling and interpreting regressions with interactions. *Journal of Accounting Literature* 42 (1):61-79.

Business Wire. 2024. New Abrigo small business lending gives financial institutions automation and AI tech to grow their portfolios and communities.

Campbell, D., M. Loumioti, and R. Wittenberg-Moerman. 2019. Making sense of soft information: interpretation bias and loan quality. *Journal of Accounting and Economics* 68 (2):101240.

Census Act, 13 U.S.C. § 224.

Census Bureau. *2019 Annual Business Survey* 2019 [cited. Available from https://www2.census.gov/programs-surveys/abs/information/abs_2019.pdf.

Charoenwong, B., Z. T. Kowaleski, A. Kwan, and A. G. Sutherland. 2024. RegTech: Technology-driven compliance and its effects on profitability, operations, and market structure. *Journal of Financial Economics* 154:103792.

Chen, W., J. K. Kang, and A. Mohan. 2023. Data-driven technologies and the local information advantages in small business lending. *Working Paper*.

Chen, W., and S. Srinivasan. 2024. Going digital: implications for firm value and performance. *Review of Accounting Studies* 29 (2):1619-1665.
Core, F., and F. De Marco. 2024. Information Technology and Credit: Evidence from Public Guarantees. *Management Science* 70 (9):6202–6219.
Cortés, K. R., Y. Demyanyk, L. Li, E. Loutskina, and P. E. Strahan. 2020. Stress tests and small business lending. *Journal of Financial Economics* 136 (1):260-279.
Crosman, P. 2024. Bankwell Bank pilots generative AI in small-business lending. *American Banker*.
Demerjian, P. 2024. Financial statements vs. FinTech: A discussion of Minnis, Sutherland, and Vetter. *Journal of Accounting and Economics*:101716.
DeYoung, R., D. Glennon, and P. Nigro. 2008. Borrower–lender distance, credit scoring, and loan performance: Evidence from informational-opaque small business borrowers. *Journal of Financial Intermediation* 17 (1):113-143.
Di Maggio, M., D. Ratnadiwakara, and D. Carmichael. 2022. Invisible Primes: Fintech Lending with Alternative Data. *National Bureau of Economic Research Working Paper Series* No. 29840.
Dornbrook, J. 2024. Banking on AI. *Kansas City Business Journal*.
Dryer, T. 2018. How machine learning is quietly transforming small business lending. *Forbes*.
Erel, I., and J. Liebersohn. 2022. Can FinTech reduce disparities in access to finance? Evidence from the Paycheck Protection Program. *Journal of Financial Economics* 146 (1):90-118.
Even-Tov, O., A. Su, and K. P. Wang. 2024. Do information processing costs matter to regulators? Evidence from the U.S. mortgage companies' supervision. *The Accounting Review* (Forthcoming).
Federal Reserve Banks. 2022. 2022 Report on Employer Firms: Based on the 2021 Small Business Credit Survey.
Fuster, A., P. Goldsmith-Pinkham, T. Ramadorai, and A. Walther. 2022. Predictably Unequal? The Effects of Machine Learning on Credit Markets. *The Journal of Finance* 77 (1):5-47.
Gallo, L., E. Labro, and J. Omartian. 2023. Overreliance on Data in Forecasting. *Working Paper*.
Gargano, F. 2024. Citi streamlines commercial lending with the help of AI. *American Banker*.
Glennon, D., and P. Nigro. 2005. Measuring the Default Risk of Small Business Loans: A Survival Analysis Approach. *Journal of Money, Credit and Banking* 37 (5):923-947.
Granja, J., C. Leuz, and R. G. Rajan. 2022. Going the Extra Mile: Distant Lending and Credit Cycles. *The Journal of Finance* 77 (2):1259-1324.
He, Z., J. Huang, and C. Parlatore. 2024. Information Span and Credit Market Competition. *National Bureau of Economic Research Working Paper Series* No. 33141.
He, Z., S. Jiang, D. Xu, and X. Yin. 2022. Investing in Lending Technology: IT Spending in Banking. *National Bureau of Economic Research Working Paper Series* No. 30403.
Howell, S. t., T. Kuchler, D. Snitkof, J. Stroebel, and J. Wong. 2024. Lender Automation and Racial Disparities in Credit Access. *The Journal of Finance* 79 (2):1457-1512.
Huang, R. 2024. The Financial Consequences of Online Review Aggregators: Evidence from Yelp Ratings and SBA Loans. *Management Science* (Forthcoming).
Hunt, J., I. M. Cockburn, and J. Bessen. 2024. Is Distance from Innovation a Barrier to the Adoption of Artificial Intelligence? *National Bureau of Economic Research Working Paper Series* No. 33022.
Jia, N., N. Li, G. Ma, and D. Xu. 2024. Corporate Responses to Generative AI: Early Evidence from Conference Calls. *Working Paper*.

Kielty, P. D., K. P. Wang, and D. L. Weng. 2023. Simplifying Complex Disclosures: Evidence from Disclosure Regulation in the Mortgage Markets. *The Accounting Review* 98 (4):191-216.

Lau, A. *How A Community Bank Simplified Small Business Lending with Open Banking Innovation* 2024 [cited. Available from https://www.geekwire.com/sponsor-post/how-a-community-bank-simplified-small-business-lending-with-open-banking-innovation/.

Levine, R., C. Lin, Q. Peng, and W. Xie. 2020. Communication within Banking Organizations and Small Business Lending. *The Review of Financial Studies* 33 (12):5750-5783.

Liberti, J. M., and M. A. Petersen. 2019. Information: Hard and Soft. *The Review of Corporate Finance Studies* 8 (1):1-41.

Liu, M. 2022. Assessing Human Information Processing in Lending Decisions: A Machine Learning Approach. *Journal of Accounting Research* 60 (2):607-651.

Mason, T. 2024. *Smaller banks are using AI too*. S&P Global 2023 [cited 10/16 2024]. Available from https://www.spglobal.com/market-intelligence/en/news-insights/research/smaller-banks-are-using-ai-too.

McElheran, K., J. F. Li, E. Brynjolfsson, Z. Kroff, E. Dinlersoz, L. Foster, and N. Zolas. 2024. AI adoption in America: Who, what, and where. *Journal of Economics & Management Strategy* 33 (2):375-415.

McKinsey. 2021. Building the AI bank of the future.

Minnis, M. 2011. The Value of Financial Statement Verification in Debt Financing: Evidence from Private U.S. Firms. *Journal of Accounting Research* 49 (2):457-506.

Moody's. *Machine learning: Challenges, lessons, and opportunities in credit risk modeling* 2017 [cited. Available from https://www.moodys.com/web/en/us/insights/credit-risk/machine-learning-challenges-lessons-and-opportunities-in-credit-risk-modeling.html.

Muro, M., and S. Liu. 2021. The geography of AI: Which cities will drive the artificial intelligencce revolution?: Brookings Institution.

Nguyen, L. 2022. Banks Tiptoe Toward Their Cloud-Based Future. *New York Times*.

Petersen, M. A., and R. G. Rajan. 2002. Does Distance Still Matter? The Information Revolution in Small Business Lending. *The Journal of Finance* 57 (6):2533-2570.

Puri, M., Y. Qian, and X. Zheng. 2024. From Competitors to Partners: Banks' Venture Investments in Fintech. *Working Paper*.

Seamans, R., and M. Raj. 2018. Artificial Intelligence, Labor, Productivity, and the Need for Firm-Level Data. In *The Economics of Artificial Intelligence: An Agenda*: National Bureau of Economic Research, Inc, 553-565.

Son, H. *JPMorgan Software Does in Seconds What Took Lawyers 360,000 Hours* 2017 [cited. Available from https://www.bloomberg.com/news/articles/2017-02-28/jpmorgan-marshals-an-army-of-developers-to-automate-high-finance.

Stock, J. H., J. H. Wright, and M. Yogo. 2002. A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments. *Journal of Business & Economic Statistics* 20 (4):518-529.

U.S. Chamber of Commerce. *The State of Small Business Now* 2023 [cited. Available from https://www.uschamber.com/small-business/state-of-small-business-now.

Vashishtha, R. 2019. Discussion of "Making sense of soft information: Interpretation bias and loan quality". *Journal of Accounting and Economics* 68 (2):101239.

Wooldridge, J. M. 2010. *Econometric Analysis of Cross Section and Panel Data*: The MIT Press.

Zolas, N., Z. Kroff, E. Brynjolfsson, K. McElheran, D. N. Beede, C. Buffington, N. Goldschlag, L. Foster, and E. Dinlersoz. 2020. Advanced Technologies Adoption and Use by U.S. Firms: Evidence from the Annual Business Survey. *National Bureau of Economic Research Working Paper Series* No. 28290.

Appendix A
Variable Definitions

Variable	Definition
$\Delta Ln(SBL)$	The change in natural logarithm of one plus the total dollar amount (in thousands) of small business loans by bank i in county c from year t to t+1.
Ln(*Distance*)	Natural logarithm of geodetic distance between the centroid of the borrower's county (borrower's address) and the bank's closest branch for CRA (SBA).
AI	Level of AI use; 0 for no use, 1 for testing, 2 for low use, 3 for moderate use, and 4 for high use.
Size	The natural logarithm of total assets (in $10 million).
ROA	Net income divided by total assets.
Leverage	Total liabilities divided by total assets.
Loans	Total loans divided by total assets.
Public	=1 if the bank is publicly traded, 0 otherwise.
Liquid Assets	The sum of cash, federal funds sold, and securities purchased under agreements to resell, scaled by total assets.
Securities	The sum of held-to-maturity and available-for-sale securities, scaled by total assets.
Residential Loans	Loans secured by residential properties, scaled by total loans.
Commercial Loans	Commercial and industrial loans, scaled by total loans.
Nonperforming Loans	The sum of loans past due more than 90 days and nonaccrual loans, scaled by total loans.
Derivatives	Derivatives assets minus derivative liabilities, scaled by total assets.
Deposits	Deposits divided by total liabilities.
Loan Amount	The natural logarithm of the loan's principal amount.
Maturity	Length of the loan term (in months).
Guaranteed	The amount of the loan guaranteed by SBA divided by the loan amount.
Interest Spread	The interest rate at loan origination minus the concurrent Treasury rate with the closest maturity (i.e., the risk-free rate).
ChargeOff	The amount that has been charged off scaled by the loan's principal amount. *ChargeOff* = 0 if the loan does not have any charge-off.
ΔGDP	The percentage change in real GDP from year t to t+1.

Appendix B
Survey Questions on the Use of AI and Other Technologies

We present the questions from each year's technology survey used in the paper, highlighting (in red) those related to AI. Following (Zolas et al. 2020; McElheran et al. 2024), we define AI as any of machine learning, natural language processing, machine vision, and voice recognition in 2018–the only year when AI was not directly surveyed. We code AI usage as follows:

- *AI* = 0 (no use): Bank's response is "No use" in 2018, "Did not use" in 2019, and "Not at all" in 2020.
- *AI* = 1 (testing): Bank's response is "Testing but not using in production or service" in 2018 and "Tested, but did not use in production or service" in 2019.
- *AI* = 2 (low use): Bank's response is "In use for less than 5% of production or service" in 2018, "Low use" in 2019, and "To a small extent" in 2020.
- *AI* = 3 (modest use): Bank's response is "In use for between 5% and 25% of production or service" in 2018, "Moderate use" in 2019, and "To some extent" in 2020.
- *AI* = 4 (high use): Bank's response is "In use for more than 25% of production or service" in 2018, "High use" in 2019, and "A great extent" in 2020.

We define other technologies variables (*Cloud*, *Other Software*, and *Hardware*) following the same procedure. When multiple technologies are involved, we define the variable based on the highest level of usage among them.

2018:

		No use	Testing but not using in production or service	In use for less than 5% of production or service	In use for between 5% - 25% of production or service	In use for more than 25% of production or service	Don't know
	a. Augmented reality	☐	☐	☐	☐	☐	☐
	b. Automated guided vehicles (AGV) or AGV systems	☐	☐	☐	☐	☐	☐
Other Software	c. Automated storage and retrieval systems	☐	☐	☐	☐	☐	☐
AI	d. Machine learning	☐	☐	☐	☐	☐	☐
AI	e. Machine vision software	☐	☐	☐	☐	☐	☐
AI	f. Natural language processing	☐	☐	☐	☐	☐	☐
Other Software	g. Radio-frequency identification (RFID) inventory system	☐	☐	☐	☐	☐	☐
Hardware	h. Robotics	☐	☐	☐	☐	☐	☐
Other Software	i. Touchscreens/ kiosks for customer interface (Examples: self-checkout, self-check-in, touchscreen ordering)	☐	☐	☐	☐	☐	☐
AI	j. Voice recognition software	☐	☐	☐	☐	☐	☐

CLOUD SERVICE PURCHASES

Considering the amount spent on each of these IT functions, how much was spent on cloud services? *(Cloud services are services provided by a third party that this business accesses on-demand via the internet.)* ***Select one for each row.***

		None	Up to 50%	More than 50%	All	Don't know	Don't use IT function
Cloud	a. All IT functions	☐	☐	☐	☐	☐	☐

2019:

AI → **Artificial Intelligence**

- ☐ Did not use
- ☐ Tested, but did not use in production or service
- ☐ Low use
- ☐ Moderate use
- ☐ High use
- ☐ Don't know

Cloud → **Cloud-Based Computing Systems and Applications**

- ☐ Did not use
- ☐ Tested, but did not use in production or service
- ☐ Low use
- ☐ Moderate use
- ☐ High use
- ☐ Don't know

Other Software → **Specialized Software**

- ☐ Did not use
- ☐ Tested, but did not use in production or service
- ☐ Low use
- ☐ Moderate use
- ☐ High use
- ☐ Don't know

Hardware → **Robotics**

- ☐ Did not use
- ☐ Tested, but did not use in production or service
- ☐ Low use
- ☐ Moderate use
- ☐ High use
- ☐ Don't know

Hardware → **Specialized Equipment**

- ☐ Did not use
- ☐ Tested, but did not use in production or service
- ☐ Low use
- ☐ Moderate use
- ☐ High use
- ☐ Don't know

2020:

	Select one for each row.	A great extent	To some extent	To a small extent	Not at all
Hardware →	a. Computer infrastructure (server technologies)	☐	☐	☐	☐
Cloud →	b. Cloud computing	☐	☐	☐	☐
AI →	c. Artificial Intelligence	☐	☐	☐	☐
Other Software →	d. Automation	☐	☐	☐	☐
Hardware →	e. Internet-connected devices	☐	☐	☐	☐
Other Software →	f. Mobile communication technologies	☐	☐	☐	☐
Other Software →	g. The use of digital technologies for collaboration	☐	☐	☐	☐
Other Software →	h. Communication (i.e., through social media)	☐	☐	☐	☐
Other Software →	i. Digital technologies for planning and management (i.e., enterprise resource planning, customer relationship management)	☐	☐	☐	☐
Other Software →	j. Digital technologies for distributed ledgers (blockchain)	☐	☐	☐	☐

Table 1
Descriptive Statistics

Panel A: Determinants analysis

	Mean	S.D.
AI	0.614	1.151
Total Assets (rounded)	7,326,000	22,710,000
ROA	0.013	0.015
Leverage	0.852	0.160
Loans	0.664	0.183
Public	0.345	0.475
Liquid Assets	0.088	0.133
Securities	0.188	0.127
Residential Loans	0.378	0.197
Commercial Loans	0.143	0.107
Nonperforming Loans	0.008	0.010
Derivatives	0.000	0.001
Deposits	0.860	0.250

Panel B: Loan growth analyses

	Mean	S.D.
Small $\Delta Ln(SBL)$	0.462	2.253
Large $\Delta Ln(SBL)$	0.411	3.082
AI	1.141	1.305
Distance	459.6	541.0
1,000 miles	0.144	0.351
250 miles	0.360	0.480
50 miles	0.216	0.412

Panel C: Loan performance and pricing analyses

	Mean	S.D.
ChargeOff	0.017	0.114
Interest Spread	5.024	1.537
AI	0.916	1.202
Distance	110.2	347.0
1,000 miles	0.036	0.187
250 miles	0.077	0.266
50 miles	0.048	0.215
Loan Amount (rounded)	392,400	689,100
Maturity	123.8	71.1
Guaranteed	0.610	0.132

Notes: This table presents descriptive statistics for variables used in our bank-year determinants analysis in Panel A, bank-county-year loan growth analyses in Panel B, and loan-level loan performance and pricing analyses in Panel C. For Census Bureau disclosure avoidance purposes, we round the mean and standard deviation of *Total Assets* and *Loan Amount*. See Appendix A for variable definitions.

Table 2
Determinants of Banks' Adoption of AI

Dependent variable = *AI*	(1)
Size	0.101***
	(3.83)
ROA	3.331
	(1.34)
Leverage	-0.366
	(-0.73)
Loans	0.168
	(0.27)
Public	0.023
	(0.26)
Liquid Assets	0.409
	(0.72)
Securities	-0.061
	(-0.10)
Residential Loans	-0.099
	(-0.41)
Commercial Loans	-0.638
	(-1.63)
Nonperforming Loans	1.148
	(0.38)
Derivatives	46.72
	(0.78)
Deposits	-0.063
	(-0.33)
Year FEs	✓
State FEs	✓
Observations (rounded)	1,500
Adj. R^2	0.11

Notes: This table presents the regression results of estimating the determinants of banks' use of artificial intelligence. *AI* equals banks' level of artificial intelligence use in year *t*, with 0 as no use, 1 as testing, 2 as low use, 3 as moderate use, and 4 as high use. For Census Bureau disclosure avoidance purposes, we round the number of observations. See Appendix A for variable definitions. Standard errors are calculated by clustering observations by bank. *t*-statistics are reported in parentheses below coefficient estimates. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Table 3
The Effect of AI Adoption on Credit Supply

Dependent variable = Δ*Ln*(*SBL*)	(1) Small	(2) Large	(3) Small	(4) Large	(5) Small	(6) Large
AI × Ln(Distance)			**0.055****	**0.037****	**0.058****	**0.039****
			(2.42)	**(2.30)**	**(2.25)**	**(2.35)**
Ln(Distance)	-0.098***	-0.034	-0.157***	-0.075***	-0.149***	-0.071**
	(-3.40)	(-1.58)	(-3.70)	(-2.72)	(-3.38)	(-2.48)
AI	0.047	0.041	-0.217***	-0.140*	-0.235**	-0.148*
	(0.86)	(1.06)	(-2.63)	(-1.76)	(-2.38)	(-1.88)
Bank controls	✓	✓	✓	✓	✓	✓
Δ*GDP × Ln(Distance)*	✓	✓	✓	✓	✓	✓
Size × Ln(Distance)					✓	✓
County × Year FEs	✓	✓	✓	✓	✓	✓
Observations (rounded)	172,000	172,000	172,000	172,000	172,000	172,000
Adj. R^2	0.07	0.01	0.07	0.02	0.07	0.02

Notes: This table presents the regression results of estimating the effect of artificial intelligence on banks' small business lending to distant borrowers. Following Granja et al. (2022), we measure loan growth using the log change of one plus the volume of loans originated by a bank in a county (Δ*Ln*(*SBL*)) and measure distance using the log distance between the borrower's county and bank's nearest branch (*Ln*(*Distance*)). *AI* equals banks' level of artificial intelligence use in year *t*, with 0 as no use, 1 as testing, 2 as low use, 3 as moderate use, and 4 as high use. Following Adams et al. (2023), we define loans as small if their amounts are less than $100,000, and as large if their amounts range between $100,000 and $1,000,000. For Census Bureau disclosure avoidance purposes, we round the number of observations and do not tabulate bank controls, which include all 12 bank characteristics in Table 2. See Appendix A for variable definitions. Standard errors are calculated by clustering observations by bank. t-statistics are reported in parentheses below coefficient estimates. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Table 4
The Effect of AI Adoption on Credit Supply by Distance

Dependent variable = $\Delta Ln(SBL)$	(1) Small	(2) Large
AI × 1,000 miles	**0.349*****	**0.237*****
	(3.14)	**(2.64)**
AI × 250 miles	**0.117**	**0.094**
	(1.17)	**(1.52)**
AI × 50 miles	**0.049**	**0.023**
	(0.79)	**(0.64)**
1,000 miles	-0.750***	-0.394***
	(-3.23)	(-2.99)
250 miles	-0.505***	-0.247**
	(-3.32)	(-2.54)
50 miles	-0.251***	-0.023
	(-3.07)	(-0.33)
AI	-0.041	-0.020
	(-0.89)	(-0.46)
Bank controls	✓	✓
Δ*GDP × Ln(Distance)*	✓	✓
Size × Ln(Distance)	✓	✓
County × Year FEs	✓	✓
Observations (rounded)	172,000	172,000
Adj. R^2	0.07	0.02

Notes: This table presents the regression results of estimating the effect of artificial intelligence on banks' small business lending after discretizing bank-borrower distance into four bins: less than 50 miles (base), between 50 and 250 miles (*50 miles*), between 250 and 1,000 miles (*250 miles*), and over 1,000 miles (*1,000 miles*). Following Granja et al. (2022), we measure loan growth using the log change of one plus the volume of loans originated by a bank in a county ($\Delta Ln(SBL)$). *AI* equals banks' level of artificial intelligence use in year *t*, with 0 as no use, 1 as testing, 2 as low use, 3 as moderate use, and 4 as high use. Following Adams et al. (2023), we define loans as small if their amounts are less than \$100,000, and as large if their amounts range between \$100,000 and \$1,000,000. For Census Bureau disclosure avoidance purposes, we round the number of observations and do not tabulate bank controls, which include all 12 bank characteristics in Table 2. See Appendix A for variable definitions. Standard errors are calculated by clustering observations by bank. *t*-statistics are reported in parentheses below coefficient estimates. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Table 5
Instrumental Variable

Dependent variable =	(1) $\Delta Ln(SBL)$	(2) AI	(3) $\Delta Ln(SBL)$
AI	**0.324****		
	(2.17)		
AI Vendor		**0.876*****	
		(3.63)	
$\widehat{AI}$			**0.889*****
			(2.79)
Bank controls	✓	✓	✓
County × Year FEs	✓	✓	✓
Observations (rounded)	21,500	21,500	21,500
Adj. R^2	0.04	0.48	0.04

Notes: This table presents the regression results of our instrumental variable analyses. Following Granja et al. (2022), we measure loan growth using the log change of one plus the volume of loans originated by a bank in a county ($\Delta Ln(SBL)$) and measure distance using the log distance between the borrower's county and bank's nearest branch ($Ln(Distance)$). *AI* equals banks' level of artificial intelligence use in year t, with 0 as no use, 1 as testing, 2 as low use, 3 as moderate use, and 4 as high use. In Column (1), we confirm our main finding of the relation between AI use and loan growth among distant borrowers (borrower counties located at least 1,000 miles from the lender). Column (2) presents the first stage of our instrumental variable analysis. *AI Vendor* is an indicator variable that equals 1 if there is an AI vendor within a 150-mile radius of the bank's headquarters, and 0 otherwise. Column (3) presents the second stage of our instrumental variable analysis, using predicted AI from the first stage in place of *AI* in Column (1). For Census Bureau disclosure avoidance purposes, we round the number of observations and do not tabulate bank controls, which include all 12 bank characteristics in Table 2. See Appendix A for variable definitions. Standard errors are calculated by clustering observations by bank. t-statistics are reported in parentheses below coefficient estimates. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Table 6
Falsification Tests

Dependent variable = Δ*Ln*(*SBL*)	(1) Small	(2) Large	(3) Small	(4) Large	(5) Small	(6) Large
Cloud* × *Ln(Distance)	**-0.007**	**0.009**				
	(-0.27)	**(0.58)**				
Other Software* × *Ln(Distance)			**0.007**	**0.003**		
			(0.41)	**(0.21)**		
Hardware* × *Ln(Distance)					**-0.033**	**-0.008**
					(-1.46)	**(-0.58)**
Ln(Distance)	✓	✓	✓	✓	✓	✓
Cloud	✓	✓				
Other Software			✓	✓		
Hardware					✓	✓
Bank controls	✓	✓	✓	✓	✓	✓
Δ*GDP* × *Ln(Distance)*	✓	✓	✓	✓	✓	✓
Size × *Ln(Distance)*	✓	✓	✓	✓	✓	✓
County × Year FEs	✓	✓	✓	✓	✓	✓
Observations (rounded)	161,000	161,000	166,000	166,000	169,000	169,000
Adj. R^2	0.06	0.01	0.06	0.02	0.06	0.01

Notes: This table presents the regression results of estimating the effect of cloud computing, other (non-AI) software, and hardware on banks' small business lending to distant borrowers. Following Granja et al. (2022), we measure loan growth using the log change of one plus the volume of loans originated by a bank in a county (Δ*Ln*(*SBL*)) and measure distance using the log distance between the borrower's county and bank's nearest branch (*Ln*(*Distance*)). *Cloud*, *Other Software*, and *Hardware* equal banks' level of cloud, other (non-AI) software, and hardware use in year *t*, with 0 as no use, 1 as testing, 2 as low use, 3 as moderate use, and 4 as high use. Following Adams et al. (2023), we define loans as small if their amounts are less than $100,000, and as large if their amounts range between $100,000 and $1,000,000. For Census Bureau disclosure avoidance purposes, we round the number of observations and do not tabulate bank controls, which include all 12 bank characteristics in Table 2. See Appendix A for variable definitions. Standard errors are calculated by clustering observations by bank. *t*-statistics are reported in parentheses below coefficient estimates. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Table 7
The Mechanism

Panel A: Poorer Counties

Dependent variable = $\Delta Ln(SBL)$	(1)	(2)
AI × Ln(Distance)	**0.087***** **(2.66)**	**0.065**** **(2.25)**
Test of coefficient difference	χ^2 = 4.08**	
Ln(Distance)	✓	✓
AI	✓	✓
Bank controls	✓	✓
ΔGDP × Ln(Distance)	✓	✓
Size × Ln(Distance)	✓	✓
County × Year FEs	✓	✓
Partition	Higher poverty	Lower poverty
Observations (rounded)	86,000	86,500
Adj. R^2	0.04	0.04

Panel B: Underserved Counties

Dependent variable = $\Delta Ln(SBL)$	(1)	(2)
AI × Ln(Distance)	**0.098**** **(2.49)**	**0.059**** **(2.27)**
Test of coefficient difference	χ^2 = 2.98*	
Ln(Distance)	✓	✓
AI	✓	✓
Bank controls	✓	✓
ΔGDP × Ln(Distance)	✓	✓
Size × Ln(Distance)	✓	✓
County × Year FEs	✓	✓
Partition	Fewer lenders	More lenders
Observations (rounded)	84,000	88,500
Adj. R^2	0.04	0.05

Notes: This table presents the cross-sectional regression results of estimating the effect of artificial intelligence on banks' small business lending to distant borrowers based on borrower poverty or lender availability. Following Granja et al. (2022), we measure loan growth using the log change of one plus the volume of loans originated by a bank in a county ($\Delta Ln(SBL)$) and measure distance using the log distance between the borrower's county and bank's nearest branch ($Ln(Distance)$). *AI* equals banks' level of artificial intelligence use in year *t*, with 0 as no use, 1 as testing, 2 as low use, 3 as moderate use, and 4 as high use. We partition the sample by median county poverty percentage in Panel A and by the median number of unique lenders to small businesses in each county in Panel B. In Panel A, Column (1) reports the results using counties with higher poverty, and Column (2) reports the results using counties with lower poverty. In Panel B, Column (1) reports the results using counties with fewer lenders, and Column (2) reports the results using counties with more lenders. For Census Bureau disclosure avoidance purposes, we round the number of observations and do not tabulate bank controls, which include all 12 bank characteristics in Table 2. See Appendix A for variable definitions. Standard errors are calculated by clustering observations by bank. *t*-statistics are reported in parentheses below coefficient estimates. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Table 8
The Effect of AI Adoption on Loan Performance

Dependent variable = *ChargeOff*	(1)	(2)	(3)	(4)
AI × Ln(Distance)		**-0.002***** **(-4.42)**	**-0.002***** **(-4.10)**	
AI × 1,000 miles				**-0.010**** **(-2.43)**
AI × 250 miles				**-0.008***** **(-2.66)**
AI × 50 miles				**-0.008***** **(-2.99)**
Ln(Distance)	0.001* (1.82)	0.002*** (3.33)	0.003*** (3.78)	
1,000 miles				0.016*** (3.03)
250 miles				0.011** (2.42)
50 miles				0.012** (2.50)
AI	0.001 (0.41)	0.004*** (2.82)	0.004*** (2.61)	0.002* (1.66)
Bank controls	✓	✓	✓	✓
Loan controls	✓	✓	✓	✓
ΔGDP × Ln(Distance)	✓	✓	✓	✓
Size × Ln(Distance)			✓	✓
Borrower Industry FEs	✓	✓	✓	✓
SBA Program FEs	✓	✓	✓	✓
County × Year FEs	✓	✓	✓	✓
Observations (rounded)	38,000	38,000	38,000	38,000
Adj. R^2	0.02	0.02	0.02	0.02

Notes: This table presents the regression results of estimating the effect of artificial intelligence on the loan performance of banks' distant borrowers. The dependent variable is the amount that has been charged off scaled by the loan's principal amount. We measure distance using the log distance between the borrower and bank's nearest branch (*Ln*(*Distance*)). *AI* equals banks' level of artificial intelligence use in year *t*, with 0 as no use, 1 as testing, 2 as low use, 3 as moderate use, and 4 as high use. In Column (4), we discretize *Distance* into four bins: less than 50 miles (base), between 50 and 250 miles (*50 miles*), between 250 and 1,000 miles (*250 miles*), and over 1,000 miles (*1,000 miles*). For Census Bureau disclosure avoidance purposes, we round the number of observations and do not tabulate bank controls, which include all 12 bank characteristics in Table 2, or loan controls, which include *Loan Amount*, *Maturity*, and *Guaranteed.* See Appendix A for variable definitions. Standard errors are calculated by clustering observations by bank. *t*-statistics are reported in parentheses below coefficient estimates. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Table 9
The Effect of AI Adoption on Loan Pricing

Dependent variable = *Interest Spread*	(1)	(2)	(3)	(4)
AI × Ln(Distance)		**-0.042*****	**-0.034*****	
		(-4.06)	**(-3.71)**	
AI × 1,000 miles				**-0.195****
				(-2.31)
AI × 250 miles				**-0.209*****
				(-3.43)
AI × 50 miles				**-0.141*****
				(-2.99)
Ln(Distance)	0.060**	0.095***	0.105***	
	(2.22)	(3.47)	(4.21)	
1,000 miles				0.737***
				(4.43)
250 miles				0.470***
				(3.48)
50 miles				0.350***
				(4.43)
AI	0.023	0.112***	0.098**	0.059*
	(0.76)	(2.69)	(2.44)	(1.69)
Bank controls	✓	✓	✓	✓
Loan controls	✓	✓	✓	✓
ΔGDP × Ln(Distance)	✓	✓	✓	✓
Size × Ln(Distance)			✓	✓
Borrower Industry FEs	✓	✓	✓	✓
SBA Program FEs	✓	✓	✓	✓
County × Year FEs	✓	✓	✓	✓
Observations (rounded)	38,000	38,000	38,000	38,000
Adj. R^2	0.46	0.46	0.46	0.46

Notes: This table presents the regression results of estimating the effect of artificial intelligence on loan pricing to banks' distant borrowers. The dependent variable is the interest rate at loan origination minus the concurrent Treasury rate with the closest maturity (i.e., the risk-free rate). We measure distance using the log distance between the borrower and bank's nearest branch (*Ln*(*Distance*)). *AI* equals banks' level of artificial intelligence use in year *t*, with 0 as no use, 1 as testing, 2 as low use, 3 as moderate use, and 4 as high use. In Column (4), we discretize *Distance* into four bins: less than 50 miles (base), between 50 and 250 miles (*50 miles*), between 250 and 1,000 miles (*250 miles*), and over 1,000 miles (*1,000 miles*). For Census Bureau disclosure avoidance purposes, we round the number of observations and do not tabulate bank controls, which include all 12 bank characteristics in Table 2, or loan controls, which include *Loan Amount*, *Maturity*, and *Guaranteed.* See Appendix A for variable definitions. Standard errors are calculated by clustering observations by bank. *t*-statistics are reported in parentheses below coefficient estimates. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Online Appendix A1
The Effect of AI Use on Credit Supply to Distant Counties

Dependent variable = $\Delta Ln(SBL_{county})$	(1)
AI Bank	+***
Year FEs	✓
County FEs	✓
Observations (rounded)	8,000

Notes: This table presents the regression results of estimating the effect of AI use on small business lending to distant counties. We aggregate the CRA bank-county-year sample into county-year observations and retain only counties served by at least one distant bank (at least 1,000 miles away). $\Delta Ln(SBL_{county})$ is the log change of one plus the volume of loans in the county. *AI Bank* equals 1 if a bank using AI (low, moderate, or high use) lends to this county from over 1,000 miles away, and 0 otherwise. For Census Bureau disclosure avoidance purposes, we round the number of observations. Standard errors are calculated by clustering observations by borrower county. ***, **, and * indicate significance at the 1%, 5%, and 10% levels. The table has not gone through the full Census review process, so we are only permitted to present sign and significance.