Finance and Economics Discussion Series

Federal Reserve Board, Washington, D.C.
ISSN 1936-2854 (Print)
ISSN 2767-3898 (Online)

Information Technology in Banking and Entrepreneurship

Toni Ahnert, Sebastian Doerr, Nicola Pierri, Yannick Timmer

2024-083

Please cite this paper as:
Ahnert, Toni, Sebastian Doerr, Nicola Pierri, and Yannick Timmer (2024). "Information Technology in Banking and Entrepreneurship," Finance and Economics Discussion Series 2024-083. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2024.083.

Information Technology in Banking and Entrepreneurship

Toni Ahnert* Sebastian Doerr† Nicola Pierri‡
Yannick Timmer§

August 2024

Abstract

We study the importance of information technology (IT) in banking for entrepreneurship. Guided by a parsimonious model, we establish that job creation by young firms is stronger in US counties more exposed to banks with greater IT adoption. We present evidence consistent with banks' IT adoption spurring entrepreneurship through a collateral channel: entrepreneurship increases by more in IT-exposed counties when house prices rise. Further analysis suggests that IT improves banks' ability to determine collateral values, in particular when collateral appraisal is more complex. IT also reduces the time and cost of disbursing collateralized loans.

JEL classification: D82, G21, L26.

Keywords: technology in banking, entrepreneurship, information technology, collateral, screening.

We would like to thank our discussants Diana Bonfim, Marti Guasch, Daniil Kashkarov, Ralf Meisenzahl, Mikhael Paaso, Mitchell Petersen, Tarun Ramadorai, Glenn Schepens, Consuelo Silva, and Huan Tang, as well as Manuel Adelino, Nigel Chalk, Giovanni Dell'Ariccia, Jon Frost, Davide Malacrino, Andrea Presbitero, Anke Weber, and Wei Xiong for their insightful comments. We are grateful to seminar participants at the Santiago Finance Workshop 2022, FIRS, the Barcelona GSE Summer Forum, the 4th IMF Annual Macro-Financial Research Conference, the 2nd DC Junior Finance Conference, Midwest Finance Association Annual Conference, the Price College of Business Colloquium on The Peril & Promise of Artificial Intelligence for Corporations, International Network for Economic Research Conference, EFiC 2021 Conference in Banking and Corporate Finance, Corporate Finance Day at RSM, The Future of Growth Conference, the IBEFA session at ASSA 2022, EBA Policy Research Workshop, the IMF, University of Bonn, and University of Halle. We thank Chenxu Fu for excellent research assistance. The views expressed in the paper are those of the authors and do not necessarily represent the views of the Bank for International Settlements, the Board of Governors of the Federal Reserve, the European Central Bank, the IMF, its Executive Board, or its Management. Vendor data only used by Board author when working at the IMF.

*European Central Bank and CEPR. Email: toni.ahnert@ecb.europa.eu.

†Bank for International Settlements and CEPR. Email: sebastian.doerr@bis.org.

‡International Monetary Fund. Email: npierri@imf.org.

§Federal Researve Board. Email: yannick.timmer@frb.gov.

1 Introduction

The rise of information technology (IT) in the financial sector has dramatically changed how information is gathered, processed, and analyzed (e.g. Vives 2019). This development has important implications for credit supply, as one of banks' key functions is to screen and monitor borrowers. Financing for young firms is likely to be especially sensitive to such changes in lenders' technology. They have produced limited information available to lenders and often rely on banks as a source of funding (Robb and Robinson, 2014; Babina, 2020).

This paper studies how the rise of IT in the financial sector affects entrepreneurship. Our empirical analysis is guided by a set of hypotheses derived from a parsimonious model in which banks' screening strategy depends on their level of IT adoption. We first establish that US counties in which more IT-intensive banks operate experience more job creation by startups. Guided by the model's predictions, we then present evidence consistent with banks' IT adoption facilitating entrepreneurship through a collateral channel: the presence of IT-intensive banks strengthens the responsiveness of job creation by entrepreneurs to changes in local real estate values. This pattern is especially pronounced in industries that rely more on real estate collateral and in markets where evaluating collateral is more difficult. Further analysis suggests that IT adoption also reduces the cost and time of disbursing collateralized loans.

We measure IT adoption at the bank-branch level as the ratio of PCs per employee, following seminal papers on IT adoption among non-financial firms (Bresnahan et al., 2002; Brynjolfsson and Hitt, 2003; Beaudry et al., 2010; Bloom et al., 2012). This simple measure of IT adoption, which is based only on hardware availability, is a strong predictor of alternative measures, such as the IT budget or adoption of frontier technologies, but has much better data coverage. We purposely focus on banks' general adoption of IT, rather than specific technologies (e.g. ATMs or online banking), because of the multi-purpose nature of IT.

Following Pierri and Timmer (2022), we decompose IT adoption into a bank fixed effect, which constitutes our main measure of IT at the bank level, and a borrower-county fixed effect. This additive model allows us to purge bank-level IT adoption from unobservable local factors that could affect both the IT intensity of local branches and local economic conditions. A large literature has adopted similar approaches to disentangle supply and demand forces (Khwaja and Mian, 2008; Amiti and Weinstein, 2018).

We use banks' IT adoption and their historical geographic footprint to compute county-level *exposure* to IT in the banking sector. It is computed as the weighted average bank-level IT adoption of banks operating in a given county. Weights are given by the initial share of local branches. Constructing local IT exposure from banks' historical footprint ameliorates concerns about banks' selecting into counties based on unobservable county characteristics, such as economic dynamism or growth trajectories.

The empirical analysis begins with establishing that higher county-level IT exposure is associated with significantly higher entrepreneurial activity. Entrepreneurship is measured as the employment share of firms of age 0 to 1, as in Adelino et al. (2017). Economically, our estimates imply that a one-standard-deviation higher IT exposure is associated with a 0.4 percentage points (pp) higher employment share in young firms. In light of the steady decline in the employment share of young firms – which fell by around 3 pp since the 1990s – the economic magnitude is sizeable. Consistent with a credit-supply channel, we find that job creation by startups in counties more exposed to IT is relatively larger in industries that depend more on external financing (Rajan and Zingales, 1998). This pattern remains robust to including industry and county fixed effects.

The positive relationship between IT exposure and startup activity could be explained by reverse causality or omitted variable bias. Reverse causality is unlikely to be a major concern in our empirical setting: lending to startups represents only a small fraction of banks' overall lending, which makes it unlikely that banks' overall IT adoption is driven by an expected increase in startup activity in specific counties. Yet confounding factors could drive the association between IT and entrepreneurship. For instance, a better-educated workforce may make it easier for banks to hire IT-savvy staff and also create more business opportunities for startups. To mitigate these concerns, we perform a large number of robustness tests. For example, we show that including a wide set of county-level controls, including the IT adoption of non-financial firms, does not affect the results.

Another possible concern is that IT could capture other bank-specific factors, such as differences in business models, quality, or profitability. Therefore, we compute counties' exposure to several bank characteristics, e.g. size, funding or assets structure, capitalization, and profitability. The relationship between entrepreneurship and local IT exposure remains unchanged when such controls are included.

We then investigate the channels underlying the relationship between county exposure and entrepreneurship. Our theoretical framework predicts that IT adoption improves banks' ability to lend against collateral, or hard information more generally. This prediction follows from the assumption that IT improves banks' ability to screen via collateral (either through better or cheaper/faster screening). This assumption is motivated by reinforcing trends. For one, advances in technology reduce the costs and time of several real estate-related processes. For example, IT expedites appraisal, research, and sales (Jud et al., 2002; Kummerow and Lun, 2005; Sawyer et al., 2005). It also improves banks' ability to evaluate collateral through hedonic pricing (Hill, 2013; Wei et al., 2022). Moreover, IT facilitates the flow of information, and in particular hard information such as collateral values, between banks' headquarters and local branches (Petersen and Rajan, 2002).

We test the prediction that IT exposure affects the relationship between collateral values and startup activity by exploiting variation in house price growth across counties. We thereby follow the literature showing that entrepreneurs often pledge their home equity as collateral (Adelino et al., 2015; Bahaj et al., 2020). Consistent with the model's predictions, we find that job creation by startups increases by more when collateral values rise, and especially so in IT-exposed counties. And that the amplifying effect of IT exposure is strongest in industries where home equity is of high importance to startups – measured either by firms' propensity to use home equity or the amount of startup capital required to start a business (Hurst and Lusardi, 2004; Adelino et al., 2015; Doerr, 2021).

To support the argument that IT spurs entrepreneurship by facilitating bank lending against collateral, we investigate different dimensions of the collateral channel. First, we show that IT spur entrepreneurship in particular in areas where house price volatility is high and assessing collateral values more difficult. This finding suggests that IT helps banks in assessing the value of collateral. Second, we find the effects of IT on entrepreneurship to be stronger in banks with more complex organizational structures. IT adoption may hence be particularly helpful in transferring information about collateral values between a bank branch and the headquarters. Consistent with this argument, we also show that the sensitivity of banks' small business lending to a local income shock declines in the distance of the borrower county to banks' headquarters, but less so for IT-intensive banks.

We further present evidence that banks' IT adoption affects the cost of originating secured loans and the time it takes to process loan applications. For this, we use data on small business loans. First, we introduce a model extension. It predicts that an IT-induced decline in the cost of making a secured loan has a stronger effect on bank lending the smaller the loan size. We find empirical support for this prediction: during periods of rising house prices (and hence collateral values), high-IT banks' growth in small business lending increases, and substantially more so for loans with smaller amounts. Second, drawing on data from the Small Business Administration, we show that banks with higher IT adoption process collateralized loan applications significantly faster than banks with lower IT adoption. This evidence is, to the best of our knowledge, the first to show that banks' IT adoption improves loan processing times. However, the magnitude of the relationship between IT and disbursement time is economically smaller compared to other studies (see e.g. Fuster et al. (2019)), suggesting this mechanism is of secondary importance in explaining the relationship between banks' IT adoption and entrepreneurship.

Performing analyses with granular bank-county level data on small business lending brings two advantages. First, it allows us to measure IT adoption at the bank-level directly, so that we can use an instrumental variable to obtain exogenous variation. Second, we can include granular county*year fixed effects that control for potentially confounding factors at the county level, including loan demand.

Our instrumental variable is based on the distance between a bank's headquarters (HQ) and the nearest land-grant colleges. Students of these colleges, established at the end of the nineteenth century to provide technical education, are significantly more likely to major in technical subjects than other students. The establishment of these colleges thus fosters the local presence of technical knowledge and expertise. Importantly, the location of land-grant colleges within a state is partly due to historical accidents and it is practically random from today's perspective (Moretti, 2004). Moreover, the location of banks' HQ is mostly explained by historical heritage. It usually predates the IT revolution by decades and it is uncorrelated with the presence of land-grant colleges in a county. The instrumental variable addresses the concern that banks' IT adoption could be correlated with other (unobservable) bank characteristics that also drive lending to small or young businesses (He et al., 2021; Pierri and Timmer, 2022).

The distance from land-grant colleges, by fostering technical knowledge, could also affect local economic conditions, businesses, and entrepreneurship. Our empirical setting allows us to address this concern directly: we include county*year fixed effects to effectively compare lending to borrowers in the same county by different banks (with headquarters that have different distances to land-grant colleges). Our instrumental variable estimates are therefore unaffected by potentially confounding factors that may determine local credit demand, including human capital, technology adoption, and the distance from land-grant colleges.

Our findings do not preclude that banks' IT adoption affects startups through additional channels. For example, IT could also affect how hard information beyond the quality and value of real estate collateral is transmitted and processed. In the context of IT-intensive fintech lenders, Di Maggio and Yao (2021) show that hard information embodied in credit reports explains most of the variation in interest rates. That being said, borrowing against real estate collateral constitutes an important source of financing for entrepreneurs, and asymmetric information is more prevalent in lending to entrepreneurs compared to e.g. in the mortgage market. Our results thus suggest that banks' IT adoption has improved entrepreneurs' access to bank credit. As we find no deterioration in startup quality, the rise of IT in the financial sector may have partly offset the overall decline in firm formation.

Related literature. Our paper contributes to the literature on financial technology and banking. Banks' increasing technological sophistication could enable them to more efficiently screen and monitor new clients (Hauswald and Marquez, 2003; Ahnert and Kuncl, 2023), as well as process hard information (Petersen and Rajan, 2002; Liberti and Mian, 2009), and consequently be more resilient to shocks (Pierri and Timmer, 2022).[1] Another implication is that IT adoption by banks leads to greater lender-borrower distance (Petersen, 1999; Berger and Udell, 2002; Hauswald and Marquez, 2006) and a change in banks' branch network structure (Lin et al., 2021; Amberg and Becker, 2024). Empirical evidence on the effects of IT on credit supply is still scarce. D'Andrea and Limodio (2019) show how high-speed internet promoted credit provision by African banks. Core and De Marco (2023) document the role of IT in providing credit to small firms in Italy during the COVID pandemic. We provide novel evidence that banks' IT adoption can spur bank lending against collateral, and thereby increase employment among startups.

We also relate to papers that highlight the importance of housing collateral for corporate investment (Catherine et al., forthcoming; Chaney et al., 2012) and entrepreneurial activity (Hurst and Lusardi, 2004; Adelino et al., 2015; Corradin and Popov, 2015; Schmalz et al., 2017; Bahaj et al., 2020; Doerr, 2021). Problems of asymmetric information about the quality of new borrowers are especially acute for young firms that are costly to screen and monitor (Degryse and Ongena, 2005; Agarwal and Hauswald, 2010). Banks thus often require collateral until they acquired sufficient information (Jiménez et al., 2006; Hollander and Verriest, 2016; Prilmeier, 2017). Our results suggest that the rise of IT in the financial sector has further increased the importance of housing collateral to entrepreneurs.

Finally, we speak to the recent literature that investigates how the rise of financial technology changes information processing and the resulting consequences for households (Berg et al., 2019; Fuster et al., 2021; Allen et al., 2022) and firms (Hau et al., 2018;

[1] Similar to Pierri and Timmer (2022) we document a benefit of banks' IT adoption. However, Pierri and Timmer (2022) study the impact of IT on banks and financial stability, while this paper's main focus is on the impact of IT on local entrepreneurship and thus on the real economy. Importantly, Pierri and Timmer (2022) show that IT benefits banks through lower NPLs during the system-wide financial crisis of 07/08, while this paper highlights the benefits of IT for lending to startups during a boom.

Erel and Liebersohn, 2020; Beaumont et al., 2021; Gopal and Schnabl, 2022; Kwan et al., 2021; Cumming et al., 2022). Di Maggio and Yao (2021) argue that fintechs rely disproportionately on hard information in the process of granting loans. Fuster et al. (2019) find evidence that fintechs are faster at processing mortgage applications. Our results suggest that the same is true for IT-savvy banks. In addition, we document material effects for firms' access to credit and employment, reflecting the importance of bank lending to young and small firms. An advantage of focusing on variation in IT adoption among banks is that our results are unlikely to be explained by regulatory arbitrage or privacy concerns. Both factors have been shown to be drivers of the growth of fintechs (Buchak et al., 2018; Doerr et al., 2023).

The remainder of the paper proceeds as follows. Section 2 provides an overview of our data. Section 3 presents a set of hypotheses derived from a simple theoretical framework of bank lending and screening to guide the empirical analysis. Section 4 presents evidence on the relationship between IT in banking, entrepreneurship, and collateral at the county level. Section 5 zooms in on the IT collateral channel and investigates its dimensions. Section 6 provides extensions and robustness tests, while Section 7 concludes.

2 Data and Variable Construction

This section explains the construction of the main variables and reports summary statistics. The analysis focuses on the years from 1999 to 2007. While banks continued to adopt IT in more recent years, the post-crisis period saw substantial financial regulatory reform (such as the Dodd-Frank Act and regular stress tests), which has affected banks' ability to lend to young and small firms. The absence of major financial regulatory changes during our sample period makes it well-suited to identify the effects of banks' IT adoption on entrepreneurship.

IT adoption and exposure. Data on banks' IT adoption come from an establishment-level survey on personal computers per employee in establishments across the US. It is provided by CiTBDs Aberdeen (previously known as "Harte Hanks") for the years 1999, 2003, 2004, and 2006. We focus on establishments in the banking sector (based on the SIC2 classification and excluding savings institutions and credit unions). We end up with 143,607 establishment-year observations.

We first hand-merge the CiTBD Aberdeen data with data on bank holding companies (BHCs) collected by the Federal Reserve Bank of Chicago. We use the Financial Institution Reports, which provide consolidated balance sheet information and income statements for domestic BHCs. We then compute a BHC-level measure of IT adoption from a regression of the share of personal computers per employee in each bank branch on a bank fixed effect, while controlling for the size and the location of the establishment, as well other characteristics through fixed effects at the level of the establishment county. Specifically, we estimate $PCs/Emp_{est,t} = IT_b + \theta_{est\ type} + \theta_c + \theta_t + \gamma \cdot log(emp_{est}) + \epsilon_{est,t}$. The variation captured by the bank fixed effects, denoted as IT_b, is our main measure of IT adoption at the bank level. The focus on BHCs rather than local branches or banks is due to the facts that (a) most of the variation in branch-level IT adoption is explained by variation at the BHC-level, (b) technology adoption at individual branches could in principle be influenced by unobservable county-level factors, which we account for through branch-location fixed effects, and (c) using a larger pool of observations per bank reduces measurement error.

To compute county exposure to IT in the financial sector, we then merge the resulting Aberdeen-BHC data set to the Federal Deposit Insurance Corporation (FDIC) Summary of Deposits (SOD) data. These data provide information on the number of branches of each bank in a county. We combine IT_b with the branch network of each bank in 1999.

The average IT adoption of all banks present in a county is defined as:

$$IT\ exposure_c = \sum_{b=1}^{N} IT_b * \frac{No.\ branches_{b,c}}{No.\ branches_c}, \tag{1}$$

where $No.\ branches_{b,c}$ is the number of branches of bank b in county c in 1999 and $No.\ branches_c$ is the total number of branches across all banks in a county in 1999 for which IT_b is available. For the ease of interpretation, $IT\ exposure_c$ is standardized to a mean of zero and a standard deviation of one. Higher values indicate that banks with branches in a given county have adopted relatively more IT.

Our measure of IT adoption is based on the use of personal computers across bank branches in the United States. The ratio of PCs per employee has not only the most comprehensive coverage, but has also been used extensively in the literature (Bresnahan et al., 2002; Brynjolfsson and Hitt, 2003; Beaudry et al., 2010; Bloom et al., 2012). We purposely focus on banks' general adoption of IT, rather than specific technologies (e.g. ATMs or online banking as in Hannan and McDowell (1987) or Hernández-Murillo et al. (2010)), because of the multi-purpose nature of IT. To examine the validity of our measure, we exploit additional information on banks' IT budget available in the 2016 vintage. The correlation between the IT budget of an establishment and the number of computers as a share of employees is 0.65 in 2016. There is also a strong positive correlation between PCs per employee and the probability of the adoption of cloud computing. These correlations provide assurance that the number of PCs per employee is a valid measure of IT adoption.

County and industry data. Data on young firms are obtained from the Quarterly Workforce Indicators (QWI). The QWI provide detailed data on end-of-quarter employment at the county-two-digit NAICS industry-year level, with a breakdown by firm age.

QWI are the only publicly available data set that provides information on county employment by firm age and industry.

We follow the literature and define young firms or entrepreneurs as firms aged 0–1 (Adelino et al., 2017; Curtis and Decker, 2018; Doerr, 2021). For each two-digit industry in each county we use 4^{th} quarter values. Note that the employment of young firms is a flow and not a stock variable, as it measures the number of jobs created by new firms in a given year. In our baseline specification, we scale the job creation of young firms by total employment in the same county-industry cell, but results are unaffected by other normalization choices. There is significant variation in job creation rates by startups both across and within states, and entrepreneurial activity is high also outside of e.g. tech hubs such as the Silicon Valley.

The 2007 Public Use Survey of Business Owners (SBO) provides firm-level information on sources of business start-up and expansion capital. For each two-digit NAICS industry we compute the fraction of young firms out of all firms that reports using home equity financing or personal assets (*home equity* henceforth) to start or expand their business (Doerr, 2021; Doerr et al., 2022). In addition, we collect information on the reported capital required to start a company in each industry (Hurst and Lusardi, 2004; Adelino et al., 2015). Following Rajan and Zingales (1998), we use Compustat data to measure industry-level dependence on external finance as capital expenditure minus cash flow over capital expenditure, averaged over the decade prior to our sample period.

The US Department of Agriculture provides a list of land-grant colleges and universities that were established in 1862 and 1890. Data on enrolment by major and test scores are obtained from from the Integrated Postsecondary Education Data System survey for 1996.

Other county-level variables include the total population, the share of the black population, the share of the population of age 65 and older, the unemployment rate, house price growth, and per capita income. The respective data sources are: Census Bureau Population Estimates, Bureau of Labor Statistics Local Area Unemployment Statistics, Federal Housing Finance Agency (FHFA) repeat sales House Price Index (HPI), and Bureau of Economic Analysis Local Area Personal Income.

Bank data. The FDIC provides detailed bank balance sheet data in its Statistics on Depository Institutions (SDI). To construct bank-level controls, we collect second quarter data for each year on banks' total assets, Tier 1 capital ratio, non-interest and total income, total investment securities, overhead costs (efficiency ratio), non-performing loans, return on assets, and total deposits.

We further use Community Reinvestment Act (CRA) data on loan origination at the bank-county-year level, collected by the Federal Financial Institutions Examination Council. CRA data contain information on loans with commitment amounts below $1 million originated by financial institutions with more than $1 billion in assets. We aggregate the data to the bank-county level and then compute loan growth as log differences. We also compute loan growth for loans with an origination amount smaller than $100,000.

Finally, we collect data on the processing time of small business loans from the Small Business Administration (SBA)'s 7(a) loan program, which contains predominantly secured loans. To select the sample, we drop revolver loans and loans with missing or zero loan size. We also exclude loans disbursed under the SBA Express program, as these follow a different process and do not require collateral. The SBA data provide information on the date the loan was approved and the date of the first loan disbursement. We take the difference between the two variables to compute the loan processing time. It is likely a reasonable proxy for the time between application date and approval date, or processing times more generally: once a loan is approved the lender starts the closing process, which includes securing collateral, preparing loan documents, and fulfilling any other authorization requirements. Note that the SBA Lending Programs are guarantee programs where the SBA guarantees a portion of loans originated by commercial lending institutions against losses from defaults, rather than lending directly to qualifying borrowers.[2]

Descriptive statistics. In the average county, the employment share of entrepreneurs out of total employment equals 5.4%, with a standard deviation of 1.4%. These numbers are in line with the aggregate employment share of young firms from 1999 to 2007, which stands at 4.7%. At the county-industry level, mean and standard deviation average 5.7% and 3.5%. The IT variable has by definition a mean of 0 and a standard deviation of 1. Table 1 reports further summary statistics of variables at the county and bank level.

Table 2 reports the balancedness in terms of county-level covariates. We split the sample into counties in the bottom and top tercile of IT exposure. Except for population, we do not find any significant differences across counties. Counties with high and low exposure are similar in terms of their industrial structure, demographics, and IT adoption of non-financial firms.

3 Hypotheses Development

In this section we develop a simple theoretical framework to derive testable hypotheses that guide our subsequent empirical analysis. A key building block of the model is asymmetric information: firms' quality is initially unobserved by banks. To mitigate the arising adverse selection problem, banks screen by either acquiring information about

[2]To qualify for a 7(a) loan, a borrower must meet various requirements. In particular, to qualify, borrowers must exhaust other funding sources, including personal sources, before seeking financial assistance, and be willing to pledge collateral for the loan. In sum, SBA 7(a) loans are intended as a last resort (Bachas et al., 2021).

firms to learn their type (unsecured lending) or requesting collateral (secured lending).

The agents in the economy are banks and firms. There are two dates $t = 0, 1$, no discounting, and universal risk-neutrality. There are two goods: a good for consumption or investment and collateral that can back borrowing at date 0.

Firms have a new project at date 0 that requires one unit of investment. They are penniless in terms of the investment good but have pledgeable collateral C at date 0. Firms are heterogeneous at date 0 along two publicly observable dimensions. First, a firm's collateral is drawn from a continuous distribution G. The market price of collateral at date 1 (in terms of consumption goods) is P, so the collateral value is $P\,C$ at date 1. Second, firms are either old (O) or young (Y), where we refer to young firms as entrepreneurs. The mass of firms is normalized to one and the share of young firms is $y \in (0, 1)$. For expositional clarity, firm age and collateral are independent.

The key friction is asymmetric information about a firm's type, that is the quality of the project. The project yields a contractible payoff $x > 1$ at date 1 if successful and 0 if unsuccessful. Projects of good firms are more likely to be successful: the probability of success is p_G for good firms and p_B for bad ones, where $0 < p_B < p_G < 1$ and only good projects have a positive NPV,

$$p_B\, x < 1 < p_G\, x. \tag{2}$$

Project quality (type G or B) is privately observed by the firm but not by banks. The share of good projects at date 0 is $q > 0$, which is independent of bank or firm characteristics. We assume that the share of good projects is low,

$$[q p_G + (1 - q) p_B]\, x < 1, \tag{3}$$

so the adverse selection problem is severe enough for banks to choose to screen all borrowers in equilibrium. As a result, all loans granted are made to good firms.

There is a unit mass of banks endowed with one unit of the investment good at date 0 to grant a loan. An exogenous fraction $h \in (0, 1)$ of banks adopted IT in the past and is therefore a high-IT bank, while the remainder is a low-IT bank.

Each bank has two tools to screen borrowers. First, the bank can pay a fixed cost F to learn the type of the project (screening by information acquisition). This cost can be interpreted as the time cost of a loan officer identifying the quality of the project. We assume that this cost is lower for old firms than for young firms:

$$F_O < F_Y, \tag{4}$$

which captures that old firms have (i) a longer track record and thus lower uncertainty about future prospects; or (ii) larger median loan volumes in practice, so the fixed cost is relatively less important.

Second, the bank can screen by asking for collateral at date 0 that is repossessed and sold at date 1 if the firm defaults on the loan. In this case, the bank does not directly learn the firm's type, but the self-selection by firms – whereby only firms with good projects choose to seek funding from banks – reveals their type in equilibrium. We assume that the cost of screening via collateral is lower for high-IT banks than for low-IT banks:

$$v_{HighIT} < v_{LowIT}, \tag{5}$$

which captures that it is easier or cheaper for a high-IT bank to verify the existence of collateral, determine its market value, or document and convey these pieces of information to its headquarter.

This important assumption builds on literature examining the ways IT facilitates real-estate related process. IT can expedite appraisal, research, and sales, thereby reducing the costs and time of originating secured loans (Jud et al., 2002; Kummerow and Lun, 2005; Sawyer et al., 2005).[3] IT also stands to improve banks' ability to evaluate collateral through hedonic pricing (Hill, 2013; Wei et al., 2022).[4] Finally, IT facilitates the flow of information, and in particular hard information such as on collateral values, between banks' headquarters and local branches (Petersen and Rajan, 2002).[5]

For expositional clarity the fixed costs F are independent of the bank's type and the costs of screening via collateral v are independent of firm age. Our results generalize as long as the high-IT bank has a comparative advantage in screening via collateral.

We assume that banks and firms are randomly matched. The lending volume maximizes joint surplus, where banks receive a fraction $\theta \in (0, 1)$ of the surplus generated. This assumption simplifies the market structure because it implies that a startup does not make a loan application with multiple banks, thus excluding competitive interaction between lenders.

In what follows, we assume a ranking of screening costs relative to the expected surplus of good projects:

$$v_{HighIT} < F_O < p_G x - 1 < \min\{F_Y, v_{LowIT}\}. \tag{6}$$

In equilibrium, only good firms may receive credit because all firms are screened in

[3]For example, Kummerow and Lun (2005) argue that "firms [used to] access sales data on microfiche, a tedious, time-consuming search process. [...] The result of being able to obtain sales information more quickly by fax or email was to increase the number of valuations completed per day. [...] A process that used to take several days could be compressed to a few hours. Valuers who used to do 3–4 valuations a day, now can complete 7–8 per day, including property inspections".

[4]As discussed in Hill (2013), early attempts at hedonic pricing only had access to rather limited data sets and computing power, and hence required manual input and models were quite costly to estimate. As a result of the combination of the development of new data sets and increased computing power, computer intensive nonparametric methods have become popular. It also became standard to make use of the longitude and latitude data of individual properties in hedonic regressions, as well as their distance to e.g. public parks, schools, or public transportation. These developments have greatly improved banks' ability to assess real estate values (Wei et al., 2022).

[5]In the Online Appendix we present evidence supporting these arguments: high IT banks are more likely to originate secured loans in the syndicate loan market; and IT improves the flow of information between banks' headquarters and local branches.

some way to detect lemons. Young firms with a good project cannot receive credit from a low-IT bank because the information cost is too high, as implied by the assumption in (6). (For a relaxation of this assumption, see Extension 2 below.) Young firms with a good project receive credit when matched with a high-IT bank and when possessing enough collateral, $C > C_{min}$, which applies to a fraction $1 - G(C_{min})$ of these firms. The bound on the collateral C_{min} ensures that young firms of the bad type do not pretend to be of good type, so the binding incentive compatibility constraint is

$$p_B(x - r) \equiv (1 - p_B)PC_{min}, \tag{7}$$

where r is the bank's lending rate.[6] Equation 7 has an intuitive interpretation: its left-hand side is the benefit of pretending to be a good type and receiving a loan from a bank, keeping the surplus $x - r$ whenever the project succeeds, which happens with the success probability of the bad type p_B. The right-hand side is the cost of forgoing the market value of collateral when the project fails. Equation (7) makes clear that the minimum level of collateral depends negatively on its price, $C_{min} = C_{min}(P)$ with $\frac{dC_{min}}{dP} < 0$. In sum, sufficient collateral ensures that only good firms receive loans in equilibrium.

Old firms with a good project always receive credit. When matched to a high-IT bank, lending is backed by collateral if the firm has enough of it, otherwise the high-IT bank ensures the project quality via information acquisition. When matched with a low-IT bank, screening via information acquisition is exclusively used.

Together, these points allows us to state the model's predictions about the share of expected lending to young firms s_Y (out of total expected lending) and how it depends on the share of high-IT banks h and the collateral price P. See Appendix A1.1 for proofs.

[6] When the bank has adopted IT, its cost of lending is $1 + v_{HighIT}$ and the surplus from lending is $p_G x - (1 + v_{HighIT})$ in equilibrium because only firms with a good project are funded. Since the bank keeps a fraction θ of this surplus, the equilibrium lending rate is $r^*_{HighIT} = \theta p_G x + (1 - \theta)(1 + v_{HighIT})$.

Proposition 1 *The share of lending to young firms is*

$$s_Y \equiv \frac{yh[1-G(C_{min})]}{1-y+yh[1-G(C_{min})]}. \tag{8}$$

The first three predictions describe the comparative statics.

Prediction 1. A higher share of high-IT banks increases the share of lending to young firms, $\frac{ds_Y}{dh} > 0$.

Prediction 2. Higher collateral values increases the share of lending to young firms, $\frac{ds_Y}{dP} > 0$.

Prediction 3. Higher collateral values increase the share of lending to young firms by more when the share of high-IT banks is higher, $\frac{d^2 s_Y}{dhdP} > 0$.

To gain intuition for these predictions, note that a higher share of high-IT banks implies that good young firms with sufficient collateral can receive funding more often (because they are matched with a bank that lends to them more often). A higher value of collateral, in turn, lowers the minimum collateral requirement C_{min} and thus increases expected lending along the extensive margin (more young firms have sufficient collateral).

In equilibrium, all potential borrowers are screened and only good projects are financed, regardless of the screening choice or the bank type. Thus, the model implies that IT adoption does not affect the quality of firms who are funded by banks, as summarized in the following prediction.

Prediction 4. Bank IT adoption does not affect the quality (default rate) of firms receiving funding in equilibrium.

Some of our model's implications are related to evidence documented in other work. The positive impact of real estate values on entrepreneurship reflects that higher real estate prices increase collateral values and thereby mitigate informational frictions and relax borrowing constraints (Rampini and Viswanathan, 2010; Adelino et al., 2015; Schmalz et al., 2017; Bahaj et al., 2020). Moreover, young firms use collateral more extensively than old firms in equilibrium. Since firm age and size are correlated in the data, this implication is consistent with recent evidence on the greater importance of collateral for lending to small businesses (Chodorow-Reich et al., 2021; Custodio et al., 2021).

4 IT Exposure and Entrepreneurship

This sections proposes a set of empirical tests at the county level for the hypotheses presented in Section 3. It then presents results.

4.1 IT exposure and entrepreneurship

Prediction 1 implies a positive relationship between the share of high-IT banks in a county and local entrepreneurial activity. To test this prediction, we estimate the following cross-sectional regression at the county-industry level:

$$\begin{aligned} \text{startups}_{c,i} = {} & \beta_1\ \textit{IT exposure}_c + \beta_2\ \textit{constraint}_i \\ & + \beta_3\ \textit{IT exposure}_c \times \textit{constraint}_i + \textit{controls}_c + \theta_c + \phi_i + \varepsilon_{c,i}. \end{aligned} \tag{9}$$

The dependent variable is the employment share of firms of age 0-1 (startups) out of total employment in each county (c) and 2-digit industry (i), averaged over 1999-2007. Scaling young firm employment by total employment has the benefit that county- or industry-specific shocks common to all firms within a county and/or sector will be cancelled out. *IT exposure* denotes county exposure to IT-intensive banks as of 1999. The variable *constraint* captures industry-level dependence on external finance. Standard errors are clustered at the county level and regressions are weighted by the total county population.

The relationship between IT exposure and local entrepreneurship could be driven by observable or unobservable local characteristics. To mitigate this concern, we include a rich set of county-level controls, all as of 1999. By controlling for the log of the total population we avoid comparing smaller rural counties to larger urban ones. We further control for the share of the population of age 65 and older, as younger individuals may be more likely to start companies and also have higher IT proficiency (Ouimet and Zarutskie, 2014; Bernstein et al., 2021). Similarly, we control for the share of adults with a bachelor degree or higher. Other socio-demographic controls, such as the share of the black population, the unemployment rate, and household income, purge our estimates from a potential correlation between local income or investment opportunities and the variables of interests. We also control for differences in the industrial structure of counties by including the employment shares in the major 2-digit SIC industries 23, 31, 44, 62, and 72. Finally, we control for the average PCs per employee in non-financial firms to address the concern that startup activity may thrive in location where IT is more readily available in general. As discussed further below, we also enrich the specification with granular fixed effects.

Abstracting from interaction terms, **Prediction 1** implies that $\beta_1 > 0$. Before moving to the regression analysis, panel (a) in Figure 1 shows a significant positive relationship between IT exposure and startup employment. It provides a binned scatterplot, with the share of employment among firms age 0–1 on the vertical axis and county IT exposure on the horizontal axis. We now investigate this pattern in greater detail.

Table 3 shows a positive relationship between county IT adoption and startup activity. Column (1) shows that counties with higher levels of IT exposure also have a significantly higher share of employment among young firms. Column (2) shows that the coefficient declines only slightly in magnitude when we add county-level controls, while the R-squared increases more than ten-fold. Column (3) adds industry fixed effects to control for unobservable confounding factors at the industry level. Including these fixed effects does not change the coefficient of interest in a statistically or economically meaningful way, despite a sizeable increase in the R-squared by 20 pp. The stability of the coefficient in light of the increase in R-squared suggests that the effect of counties' IT exposure on job creation by startups is orthogonal to observable county and unobservable industry characteristics, reducing potential concerns about self-selection and omitted variable bias (Altonji et al., 2005; Oster, 2019).

The economic magnitude of the estimated effect is sizeable: in column (3), a one standard deviation higher IT exposure is associated with a 0.38 pp increase in the share of young firm employment (7% of the mean). While the employment share of young firms has declined steadily (Decker et al., 2016) – by around 3 pp since the 1990s – these results suggest that banks' IT adoption partly offset this trend.

The model suggests that IT spurs entrepreneurship through a relaxation in firms' borrowing constraints. We thus expect the positive correlation in columns (1)–(3) to be stronger in industries that depend more on external finance. We augment the regression with an interaction term between IT adoption and industry-level dependence on external finance (β_3 in Equation (9)). In column (4), the coefficient on the interaction term between IT exposure and external financial dependence is positive and statistically significant. Counties with higher IT exposure have a higher share of employment among young firms precisely in those industries that depend more on external finance. This pattern is consistent with the notion that the effect is driven by the impact of banks' IT on startups' financing. In column (5), we further enrich our specification with county fixed effects to control for any observable and unobservable confounding factors at the local level. Results are near-identical to column (4). Unobservable county factors are thus unlikely to explain the relationship between entrepreneurship and IT exposure.

To interpret the magnitude, column (6) replicates column (4), but uses a dummy with a value of one if an industry lies in the top tercile in terms of external financial dependence. A one standard deviation higher IT exposure is associated with a 0.22 pp increase in the share of young firm employment in industries that depend less on external finance, but a 0.69 pp increase in industries that depend more on external finance.

Taken together, Table 3 provides support for **Prediction 1**: a larger local presence of IT-intensive banks is associated with more startup activity. This is especially so in sectors that depend more on external financing, suggesting that the relationship is driven by better access to credit.

4.2 IT, collateral, and entrepreneurship

Predictions 2 & 3 state that *i*) higher collateral values increase startup activity and *ii*) that they do so especially in counties with higher IT exposure.

We test these predictions by examining how local IT exposure affects the sensitivity of entrepreneurship to changes in house prices. We use a county-industry-year panel from 1999 to 2007 and estimate the following regression:

$$\begin{aligned} \text{startups}_{c,i,t} = &\ \gamma_1\ IT\ exposure_c + \gamma_2\ \Delta HPI_{c,t} \\ &+ \gamma_3\ IT\ exposure_c \times \Delta HPI_{c,t} \\ &+ controls_{c,t-1} + \tau_t + \varepsilon_{c,i,t}. \end{aligned} \tag{10}$$

The dependent variable is the employment share of firms of age 0-1 out of total employment in county (c) and 2-digit industry (i) in year (t). *IT exposure* denotes counties' IT exposure. ΔHPI is the yearly county-level change in the house price index. County-level controls include county size (log of total the population), the share of the population of age 65 and older, the share of the black population, education, the unemployment rate, the industrial structure, and IT adoption among non-financial firms. We also include year fixed effects to absorb common trends. Standard errors are clustered at the county level and regressions are weighted by total county population.

Table 4, column (1) confirms that higher IT exposure is associated with a higher share of young firm employment. This is in line with the cross-sectional results in Table 3 on **Prediction 1** ($\gamma_1 > 0$). Column (2) then tests **Prediction 2**, which implies that $\gamma_2 > 0$. It shows that a rise in house prices is associated with an increase in entrepreneurship at the local level.

We then test **Prediction 3** by including the interaction term between changes in local house prices and county exposure to IT in the banking sector (γ_3). We expect $\gamma_3 > 0$, i.e. that an increase in house prices increases startup activity especially in counties more-exposed to IT. Column (3) shows that higher house price growth spurs entrepreneurship in areas with more IT, consistent with **Prediction 3**. To illustrate this finding Figure 2 plots the estimated effect of house price growth on entrepreneurship for counties with low, medium, and high IT exposure. In counties with average exposure, there is a positive relationship between house price growth and entrepreneurship (indicated by the green line), consistent with the positive coefficient on IT adoption itself. The red line plots the same relationship for a county with exposure two standard deviations above the mean. As a result of the positive interaction term between IT and house price growth, the slope is steeper relative to the average county. For counties with IT exposure is two standard deviations below the mean, changes in house prices do not have a significant effect on entrepreneurship (blue line). Moreover, the graph illustrates that even when zeroing in on counties where house prices fell during our sample period, job creation by young firms was stronger in high IT compared to lower IT regions.

In terms of magnitude, suppose house price growth increases by 4 pp (corresponding roughly to an increase from the 25^{th} percentile to the 75^{th} percentile). In counties with an IT exposure of one, the share of employment at young firms would increase by 0.276

pp more per year than in counties with an IT exposure of zero (based on column (3)).

To isolate the variation of interest and control for confounding factors at the local or industry level, we include granular fixed effects. Column (4) of Table 4 adds industry*year fixed effects that account for unobservable changes at the industry level. The interaction coefficient remains similar in terms of sign, size, and significance to column (3).

Finally, we provide complementary evidence on the role of collateral, building on previous work demonstrating that the importance of real estate collateral differs across industries. Specifically, young firms have been shown to be more responsive to changes in collateral values in industries in which the average required start-up capital is lower, or in industries in which a larger share of entrepreneurs relies on home equity to start or expand their business (Adelino et al., 2015; Doerr, 2021). Focusing on differences between industries within the same county and year allows us to additionally include county*year fixed effects. We thus purge our estimates from the impact of any time-varying county-level shocks, in addition to controlling for industry-specific trends. Columns (5) and (6) show that the positive effect of rising house prices on startups in more IT exposed counties is especially pronounced in industries whose financing is expected to be more sensitive to changes in collateral values, as indicated by the positive and significant coefficient on the triple interaction term. Note that the remaining coefficients are absorbed by fixed effects.

In sum, Table 4 provides evidence in line with **Predictions 2 & 3**: entrepreneurship increases when local collateral values rise, and in particular so in counties with higher exposure to IT-intensive banks.

4.3 IT exposure and startup quality

Prediction 4 states that higher startup activity due to IT exposure does not lower the quality of the average firm receiving funding in equilibrium. As IT improves the screening process, there is no trade off between the quantity of credit and the marginal quality of the borrower.

In the model firm quality is disciplined by the probability of default, which is unobservable in the data. Instead, we proxy startup quality with the average growth rate of employment of startups during their first years of life. To this end, we construct 'transition rates' (Adelino et al., 2017). As the QWI report employment of firms of age 0–1, 2–3, or 4–5 in a given year, we can substract the employment of startups (firms age 0 or 1 year) two years earlier from employment of firms of age 2–3 to obtain the change in jobs created by continuing startups during that period. The transition rate in a county-industry cell is thus defined as $transition^{2-3}_{c,i,t} = \frac{Employment\ Age\ 2-3_{c,i,t+2} - Employment\ Startup_{c,i,t}}{Total\ Employment_{c,i,t}}$ in year t. We construct similar transition rates for firms transitioning from age 2–3 to 4–5. The averages of the variables $transition^{2-3}$ and $transition^{4-5}$ are -0.37 and -0.51, with standard deviations of 1.66 and 1.36.

We then estimate a cross-sectional regression similar to Equation 9, where the dependent variable is the average transition rate between 1999 and 2007. Columns (1)-(3) in Table 5 show that there is no systematic correlation between a county's IT exposure and the transition rates of local startups to age 2–3, neither on average nor in industries that are more dependent on external finance. We obtain similar results for the transition rates of firms of age 2–3 years to 4–5 years in columns (4)-(6). These results lend support to **Prediction 4**.[7]

[7]In the Online Appendix, we further show that there is no economically or statistically significant relationship between IT exposure and startups' patenting activity. This result is consistent with the argument that greater firm entry does not come at the cost of lower average startup quality.

The absence of any significant relationship between IT exposure and local startup quality could suggest that IT adoption by banks has aggregate implications. The formation of more startups, without a decline in quality, could bring benefits in terms of aggregate business dynamism, employment and productivity growth (Haltiwanger et al., 2014; Klenow and Li, 2020).

5 IT and the Dimensions of the Collateral Channel

In this section, we shed light on different dimensions through which IT could improve banks' ability to lend against collateral. First, we examine banks' ability to evaluate collateral in complex markets and communicate that information within the organization. Second, we analyze whether IT lowers the cost of issuing collateralized loans and the time it takes to disburse them.

5.1 Complexity of evaluating collateral

An important dimension relates to the complexity of determining the market value of collateral. In particular, greater computing power and improvements in hedonic pricing models allow for non-linearity, variable interaction, or other complex valuation situations (Herath and Maier, 2010; Hill, 2013). Hedonic pricing also allows lenders to consider factors beyond standard variables such as property size, age, or location. IT may thus enable banks to better assess real estate values, predict future performance, and price risks. The benefits of information technology adoption may therefore be especially high in circumstances where the value of collateral is more complex to assess or transmit.

We first investigate whether the effects of IT exposure vary with the complexity of evaluating real estate collateral. Large swings in house prices may complicate the evaluation of real estate and prediction of its future value for the lender in case of foreclosure. Therefore, we calculate the standard deviation of yearly house price growth over the sample period for each county and split the sample into high/low volatility counties based on the median volatility. We then estimate the empirical relationship between IT and local entrepreneurship (i.e. Equation (9)) separately for each sub-sample. In columns (1) and (2) in Table 6 we find a substantially larger coefficient for IT exposure in more volatile markets. This pattern is consistent with the argument that IT spurs entrepreneurship through collateralized lending especially when evaluating collateral is more difficult.

Another dimension of complexity is the organizational structure of banking groups. In more complex organizational structures, IT adoption may be particularly helpful in transferring information about collateral values (or hard information more generally) between a bank branch and the headquarters. Following Correa and Goldberg (2022), we measure the complexity of each bank in the FDIC dataset by counting the number of entities within the bank. We then use the FDIC deposits data to project the bank data onto US counties to obtain a measure of organizational complexity of locally active banks. We split the sample into counties above and below the median of complexity. The relationship between IT exposure and entrepreneurship is significantly stronger in counties that have more complex banks (columns (3) vs (4) in Table 6). This result is in line with the argument that IT fosters entrepreneurship through ameliorating the flow of information within organizations.

Consistent with better information transmission within banks, we show in the Online Appendix that the distance between a bank's headquarters and its borrowers matters more for low-IT than for high-IT banks, further suggesting that IT reduces frictions in the flow of information within a bank. The intuition is formalized in a model extension we provide in the appendix.

5.2 Cost and speed of originating secured loans

Beyond the ease and accuracy of determining collateral values, IT adoption could affect the cost and amount of time it takes to evaluate collateral and originate loans to startups. To investigate these dimension, we turn to data on small business loans.[8]

We start with investigating the link between banks' IT adoption and loan origination costs. Since we have no direct data on origination costs, we use our model to derive a prediction. In the baseline model, we have considered a fixed loan size, normalized to 1. Our result is based on the assumption that high-IT banks have a lower cost of making a loan $v_{HighIT} < F_O < p_G x - 1$. Here we generalize this result, examining loans of different sizes. In particular, consider loans of size L, where the loan size is independent of other firm characteristics and identically distributed across banks. Thus, the average cost of a high-IT bank, that is the cost per unit of lending, is

$$1 + \frac{v_{HighIT}}{L},$$

where the high-IT bank again experiences a reduction in the cost v, as in the baseline scenario. It is immediate from the average cost that banks with a small loan (low values of L) benefit more than banks with a large loan. Thus, the model predicts that the effect of a reduction in the cost of lending based on hard information is expected to be stronger for smaller loans.

To test this prediction, we estimate the following regression from 1999 to 2007 at the

[8] To the best of our knowledge, no data exists with a separate breakdown of lending to young businesses. However, the vast majority of young firms is small, making small business lending a reasonable proxy.

bank-county-year level for loans of different sizes:

$$\begin{aligned} \Delta loans_{b,c,t} &= \beta_1\ IT_b + \beta_2\ \Delta\ HPI_{c,t} + \beta_3\ IT_b \times \Delta HPI_{c,t} \\ &+ bank\ controls_{b,t-1} + \tau_{c,t} + \varepsilon_{b,c,t}. \end{aligned} \tag{11}$$

The dependent variable is the growth in total CRA small business lending by bank b to borrower county c in year t. We distinguish between loans with an amount below \$1,000,000 and loans with an amount below \$100,000. The main explanatory variable IT measures the use of IT at the bank level, as described in Section 2. ΔHPI measures the yearly change in the house price index in the borrower county. Regressions further include county*year fixed effects, effectively comparing lending by different banks to borrowers in the same county. Bank-level controls are the log of total assets, deposits over total liabilities, the share non-interest income, securities over total assets, return on assets, the equity ratio (Tier 1), and the wholesale funding ratio. We also control for the share of mortgage loans out of total loans. We thus hold the allocation of credit *across* lending segments constant. We cluster standard errors at the county level.

Based on **Prediction 3**, IT-intensive banks have an advantage in lending against collateral (as indicated by the county-level analysis in Section 4). We thus expect their lending to be more sensitive to changes in local collateral values, i.e. house prices ($\beta_3 > 0$). Moreover, according to our model extension we expect β_3 to be larger in magnitude for small loans.

Column (1) in Table 7, panel (a) shows a larger responsiveness of small business lending by high-IT banks to rising house prices, as indicated by the significant positive coefficient on the interaction term. Note that the specification includes time-varying fixed effects at the county level. We thus essentially compare small business lending by two banks that differ in their IT intensity to borrowers in the *same* county, mitigating concerns that the relationship between bank lending and house prices is due to unobservable local factors. Column (2) repeats the exercise for loans of an amount of $100,000 or less. It shows that patterns are qualitatively similar, but the estimated coefficient on the interaction term is substantially larger in magnitude (0.267 vs 0.159). These results are consistent with the hypothesis that IT reduces the fixed cost of collateralized lending.

Next, we investigate whether IT affects the processing time of collateralized small business loans. For this, we use SBA data and estimate the following loan-level (i) regression:

$$ln(processing\ time)_{i(b)} = \beta\ IT_b + controls_{b/i} + \theta_{c,t} + \varepsilon_i. \tag{12}$$

The dependent variable is the log of the days between approval and disbursement of loan i by lender b to a small business. The average (median) days equal 83 (54). IT is the lender's IT adoption. Controls include the log of the loan amount, the loan maturity, the share of the loan amount that is guaranteed, and dummies for the borrower type (whether the business is a corporation, partnership, or individual); we also include baseline bank-level controls. We further control for the delivery method, as processing times vary greatly across methods, and include county*time fixed effects.

Columns (3) and (4) in panel (a) of Table 7 report results. Column (3) shows a negative and significant relationship between IT and processing time. A one standard deviation increase in IT leads to a decline in processing times by around 2.5%. When we further include industry*time fixed effects in column (4), the estimated coefficient remains similar in magnitude. A one standard deviation increase in banks' IT adoption reduces processing times by 2.7%, or 2.2 days. These results suggest high-IT banks are faster at processing loan applications/disbursing loans than low-IT banks.

All in all, this section provides evidence for two dimensions of the IT collateral channel that spurs entrepreneurship. First, by improving banks' ability to evaluate collateral and transmit such information within the organization. And second, by reducing the time and cost of disbursing collateralized loans. Assessing the relative importance of these channels is challenging in our empirical setting, as it would require a more structural approach. Nonetheless, the empirical evidence suggests that improvements in banks' ability to evaluate collateral and transmit such information are the more important dimensions of the collateral channel. The effect of IT adoption on entrepreneurship is substantially stronger in more volatile markets or within complex organizations. Meanwhile, while statistically significant, IT adoption's effect on processing times averages just 2.7% of the mean processing time. This number is comparatively small. For example, Fuster et al. (2019) find that fintechs are between 8 and 10 days (15%–18% of the mean processing time) faster at processing loans than non-fintech lenders in the US mortgage market.

To more directly compare the two channels, the relevant parameter is the elasticity of entrepreneurship to loan processing times. If this elasticity is high, even minor reductions in processing times could significantly boost entrepreneurship. To provide a suggestive estimate of this parameter, in the Online Appendix we estimate Equation (9), but regress the employment share of young firms on the average processing time of SBA loans in each county. We obtain a coefficient of -0.016 (see Table A12). As discussed above, a one standard deviation increase in banks' IT adoption reduces processing time by 2.2 days. Combining these numbers translates into a 0.0352 pp (-0.016×-2.2) higher share of employment among young firms due to IT's effect on processing times. This implies that faster processing times explain around 10% (0.0352/0.361) of the overall effect of IT on entrepreneurship we find. In contrast, a one standard deviation higher IT adoption by banks is associated with a 0.54 pp to 0.7 pp increase in employment among young firms in complex markets, and a substantially smaller 0.22 pp to 0.35 pp increase in less complex markets.

While these back-of-the-envelope calculations suggest that IT improving banks' ability to evaluate collateral is economically more important, a more rigorous assessment of the relative importance of the different channels is a promising avenue for future research.

6 Extensions and Robustness

This section presents a number of additional tests. In particular, we confirm our bank-level results in IV regressions and discuss robustness exercises to our county-level analysis.

6.1 Instrumental variable analysis

The relationship between IT adoption and bank lending could be driven by unobservable factors. For instance, banks may invest more in IT to serve certain costumers or local productivity shocks may drive both banks' IT adoption and credit demand. Studying the relationship between a bank's IT adoption and its lending directly allows us to obtain exogenous variation in IT-adoption through an instrumental variable.

We exploit the quasi-random allocation of land-grant colleges, which acted as a shift in the availability of local technical expertise (Moretti, 2004) and predicts banks' IT adoption (He et al., 2021; Pierri and Timmer, 2022). The Morrill Act of 1862, and its follow-up in 1890, endowed states with federal land to found universities, with a focus on teaching science, agriculture, and other technical subjects. The presence of a land-grant college remains an important determinant of the supply of skilled labour in a city even today, especially for the IT sector. Their exact location, however, is largely due to historical accidents and close to random from today's perspective (Moretti, 2004). It is also unrelated to current local economic factors (Kantor and Whalley, 2019), as well as to the presence of banks' HQ in the same county (Pierri and Timmer, 2022), reflecting that the formation of banks' headquarters usually predates the IT revolution by many decades.

Land-grant colleges could spur banks' IT adoption through different channels. They directly increase the supply of tech-inclined graduates that banks could hire, which could incentivize their IT adoption. Additionally, a shorter distance to campuses could lead to knowledge spillovers and the diffusion of ideas and technology (Keller, 2002), making bank mangers more likely to invest in IT.

As decision-making at a firm is often driven by its HQ, we base our instrument on the distance of a bank's HQ to the nearest land-grant colleges. In a first step, we compute the distance in log miles (plus one) between the county of each land-grant college j and a bank's HQ county, weighted by the size of the college in terms of STEM enrollment. In a second step, we compute a measure of the average distance to land-grant colleges. There is no clear economic reason to expect why the distance to only the nearest, second- or third-nearest college should matter. In addition, distances to the nearest colleges are positively correlated. We thus take an agnostic approach and take the first principal component of the distance to the nearest two land-grant colleges as our baseline instrumental variable, so the IV captures only the salient variation in distances.

The identification assumption underlying our instrument is that the distance to the nearest land-grant colleges affects the ability to lend to small businesses through banks' IT adoption, and not through other bank-specific channels or changes in the demand for credit. Students of land-grant colleges are significantly more likely to major in technical subjects and less likely to major in business and management sciences (Pierri and Timmer, 2022). The introduction of these colleges is thus akin to a shifter of the availability of local technical knowledge for banks, rather than overall managerial capabilities. It mitigates the concern that the impact of IT adoption is capturing better managerial practices (Bresnahan et al., 2002).

Land-grant colleges could affect non-financial firms in close proximity by increasing the availability of technical knowledge, thereby directly affecting entrepreneurship and credit demand. Therefore, we compare the lending of *different banks* – whose HQs have different distances to land-grant colleges – to borrowers *in the same county.* To do so, we include in our specifications time-varying fixed effects at the borrower-county level.

Results in panel (b) of Table 7 confirm our OLS results. Column (1) shows a larger responsiveness of small business lending by high-IT banks to rising house prices, as indicated by the significant coefficient on the interaction term. Column (2) repeats the exercise for loans with amounts of $100,000 or less and shows that estimated coefficients are larger for smaller loans, consistent with our model prediction. Columns (3)–(4) confirm the negative effect of IT adoption on processing times. Across specifications, the first-stage F-statistics for all instrumented variables safely exceeds the critical value of 10. IV coefficients are, however, larger in magnitude than their OLS counterparts, which could reflect measurement error in the potentially endogenous IT adoption variable (Pancost and Schaller, 2022).[9]

6.2 Further tests

To show that the relationship between IT exposure and entrepreneurship is robust, we perform a series of additional tests. They are reported in the Online Appendix. We show that our results are insensitive to an alternative construction of IT exposure based on either the unweighted average of the IT adoption of banks that operate in a county, or the share of local deposits. Excluding firms in the financial and education industries, or individual regions that have particularly high IT exposure or entrepreneurial activity, does not affect our results. Excluding the top 20 counties in terms of venture capital (VC) funding activity (which receive almost 80% of total VC funding) yields results similar to

[9] As our measure of IT adoption is built from the ratio of computers per employee, it does not take into account differences across IT applications, it does not capture heterogeneous quality of IT equipment, nor does it considers workers' different ability of using such equipment. We therefore expect a significant amount of noise, and it is unsurprising that IV coefficients are larger than OLS ones.

our baseline. Similarly, normalizing the share of employment in startups by the previous year's total employment to rule out our results are driven by the denominator leaves our conclusion unaltered. We also show that our main findings are present in tradable industries, which are less affected by local economic conditions. We also investigate the increase in IT adoption over time. We find that counties more exposed to the *increase* in IT in banking also experienced relatively higher *growth* in startup rates.

We further augment our baseline regression with a set of controls for counties' exposure to other bank characteristics. That is, for all banks in our IT data, we compute a set of characteristics to capture bank profitability, funding and assets structure, size, and capitalization (capital ratio, return on assets, return on equity, wholesale funding ratio, deposit ratio, log assets, securitization ratio). We then project these characteristics on US counties by relying on banks' historical footprint, as we do for IT. We find that the relationship between exposure to bank IT and entrepreneurship is unchanged by the inclusion of these controls, mitigating the concern that this finding is spuriously driven by correlation with other bank factors. Looking at aggregate bank lending, we show that higher IT adoption is associated with a significant increase in the share of small business lending, but no decline in banks' overall C&I lending. These patterns suggest that the increase in lending to young firms with collateral does not crowd out other lending to corporates that is based on soft information.

We furthermore provide evidence consistent with two model extensions. First, we find that a higher share of high-IT banks increases the share of lending to young firms by significantly less in recourse states than in non-recourse states. Second, we show that the geographic distance between lenders and borrowers matters more for low-IT than for high-IT banks' loan growth when local loan demand increases. We conclude by presenting suggestive evidence of a potential downside of the beneficial effect of banks' IT adoption on entrepreneurship through the collateral channel: it may magnify pre-existing disparities in housing wealth due to racial discrimination in the mortgage market.

7 Conclusion

Over the last decades, banks have invested in information technology at a grand scale. However, only scant evidence exists on the effects of banks' IT adoption on lending and the real economy. In this paper we focus on how banks' IT adoption affects startups. We do so for two reasons. First, startups matter greatly for aggregate employment, innovation, and growth; and second, they are opaque borrowers and hence likely to be especially sensitive to technologies that affect lenders' information acquisition.

We find that IT adoption in the financial sector has spurred entrepreneurship. In regions where banks with higher IT-adoption have a larger footprint, job creation by startups was relatively stronger. This relationship is particularly pronounced in industries that rely more on external finance. We show – both theoretically and empirically – that collateral plays an important role in explaining these patterns. As IT likely improves banks' ability to assess and transmit the value and quality of collateral, banks with higher IT adoption lend more when the value of entrepreneurs' collateral increases.

Our results have implications for policy. Banks have been ardent adopters of technology during the last years. Meanwhile the role of fintech companies that rely on technology and algorithms, rather than loan officers, to provide credit to small businesses has been steadily increasing (Gopal and Schnabl, 2022). These developments have triggered a debate on the impact of IT adoption in the financial sector on the real economy, for example through its impact on the relative importance of soft and hard information, or the need for collateral (Gambacorta et al., 2020). Our findings suggest that IT adoption can spur job creation by young firms by making lending against collateral, or hard information more general, easier. This finding raises the prospect that the rising adoption of financial technology in the financial sector eases financial constraints for young and dynamic firms.

References

Adelino, Manuel, Song Ma, and David Robinson (2017) "Firm age, investment opportunities, and job creation", *Journal of Finance*, 72 (3), pp. 999–1038.

Adelino, Manuel, Antoinette Schoar, and Felipe Severino (2015) "House prices, collateral, and self-employment", *Journal of Financial Economics*, 117 (2), pp. 288–306.

Agarwal, Sumit and Robert Hauswald (2010) "Distance and private information in lending", *Review of Financial Studies*, 23 (7), pp. 2758–2788.

Ahnert, Toni and Martin Kuncl (2023) "Government loan guarantees, market liquidity, and lending standards", *Management Science, Accepted.*

Allen, Linda, Yu Shan, and Yao Shen (2022) "Do fintech mortgage lenders fill the credit gap? evidence from natural disasters", *Journal of Financial and Quantitative Analysis.*

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber (2005) "Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools", *Journal of Political Economy*, 113 (1), pp. 151–184.

Amberg, Niklas and Bo Becker (2024) "Banking without branches ", *Working Paper.*

Amiti, Mary and David E. Weinstein (2018) "How Much Do Idiosyncratic Bank Shocks Affect Investment? Evidence from Matched Bank-Firm Loan Data", *Journal of Political Economy*, 126 (2).

Babina, Tania (2020) "Destructive creation at work: How financial distress spurs entrepreneurship", *Review of Financial Studies*, 33 (9), pp. 4061–4101.

Bachas, Natalie, Olivia S Kim, and Constantine Yannelis (2021) "Loan guarantees and credit supply", *Journal of Financial Economics*, 139 (3), pp. 872–894.

Bahaj, Saleem, Angus Foulis, and Gabor Pinter (2020) "Home values and firm behaviour", *American Economic Review*, 110 (7), pp. 2225–2270.

Beaudry, Paul, Mark Doms, and Ethan Lewis (2010) "Should the personal computer be considered a technological revolution? Evidence from US metropolitan areas", *Journal of Political Economy*, 118 (5), pp. 988–1036.

Beaumont, Paul, Huan Tang, and Eric Vansteenberghe (2021) "The role of fintech in small business lending: Evidence from France", *Working Paper*.

Berg, Tobias, Valentin Burg, Ana Gombović, and Manju Puri (2019) "On the rise of fintechs–credit scoring using digital footprints", *Review of Financial Studies*, 33 (7), pp. 2845—-2897.

Berger, Allen N. and Gregory F. Udell (2002) "Small Business Credit Availability and Relationship Lending: The Importance of Bank Organisational Structure", *Economic Journal*, 112 (477), pp. 31–53.

Bernstein, Shai, Emanuele Colonnelli, Davide Malacrino, and Tim McQuade (2021) "Who creates new firms when local opportunities arise?", *Journal of Financial Economics*.

Bloom, Nicholas, Raffaella Sadun, and John Van Reenen (2012) "Americans do it better: US multinationals and the productivity miracle", *American Economic Review*, 102 (1), pp. 167–201.

Bresnahan, Timothy F, Erik Brynjolfsson, and Lorin M Hitt (2002) "Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence", *Quarterly Journal of Economics*, 117 (1), pp. 339–376.

Brynjolfsson, Erik and Lorin M Hitt (2003) "Computing productivity: Firm-level evidence", *Review of Economics and Statistics*, 85 (4), pp. 793–808.

Buchak, Greg, Gregor Matvos, Tomasz Piskorski, and Amit Seru (2018) "Fintech, regulatory arbitrage, and the rise of shadow banks", *Journal of Financial Economics*, 130 (3), pp. 453–483.

Catherine, Sylvain, Thomas Chaney, Zongbo Huang, David Alexandre Sraer, and David Thesmar (forthcoming) "Quantifying reduced-form evidence on collateral constraints", *Journal of Finance.*

Chaney, Thomas, David Sraer, and David Thesmar (2012) "The collateral channel: How real estate shocks affect corporate investment", *American Economic Review*, 102 (6), pp. 2381–2409.

Chodorow-Reich, Gabriel, Olivier Darmouni, Stephan Luck, and Matthew Plosser (2021) "Bank liquidity provision across the firm size distribution", *Journal of Financial Economics.*

Core, Fabrizio and Filippo De Marco (2023) "Information technology and credit: evidence from public guarantees", *Management Science.*

Corradin, Stefano and Alexander Popov (2015) "House prices, home equity borrowing, and entrepreneurship", *Review of Financial Studies*, 28 (8), pp. 2399–2428.

Correa, Ricardo and Linda S Goldberg (2022) "Bank complexity, governance, and risk", *Journal of Banking & Finance*, 134, p. 106013.

Cumming, Douglas, Hisham Farag, Sofia Johan, and Danny McGowan (2022) "The digital credit divide: Marketplace lending and entrepreneurship", *Journal of Financial and Quantitative Analysis.*

Curtis, E Mark and Ryan Decker (2018) “Entrepreneurship and state taxation”, *Working Paper*.

Custodio, Claudia, Diana Bonfim, and Clara C Raposo (2021) “The sensitivity of SME's investment and employment to the cost of debt financing”, *Working Paper*.

D'Andrea, Angelo and Nicola Limodio (2019) “High-speed internet, financial technology and banking in Africa”, *Working Paper*.

Decker, Ryan A, John Haltiwanger, Ron S Jarmin, and Javier Miranda (2016) “Declining business dynamism: Implications for productivity”, *Brookings Institution, Hutchins Center Working Paper*.

Degryse, Hans and Steven Ongena (2005) “Distance, Lending Relationships, and Competition”, *Journal of Finance*, 40 (1), pp. 231–266.

Di Maggio, Marco and Vincent Yao (2021) “Fintech borrowers: Lax screening or cream-skimming?”, *The Review of Financial Studies*, 34 (10), pp. 4565–4618.

Doerr, Sebastian (2021) “Stress tests, entrepreneurship, and innovation”, *Review of Finance*, 25 (5), pp. 1609—1637.

Doerr, Sebastian, Thomas Drechsel, and Donggyu Lee (2022) “Income inequality and job creation”.

Doerr, Sebastian, Leonardo Gambacorta, Luigi Guiso, and Marina Sanchez del Villar (2023) “Privacy regulation and fintech lending”, *Working Paper*.

Erel, Isil and Jack Liebersohn (2020) “Does fintech substitute for banks? Evidence from the paycheck protection program”, *Working Paper*.

Fairlie, Robert W, Alicia Robb, and David T Robinson (2020) “Black and white: Access to capital among minority-owned startups”, *Working Paper.*

Fuster, Andreas, Paul Goldsmith-Pinkham, Tarun Ramadorai, and Ansgar Walther (2021) “Predictably unequal? The effects of machine learning on credit markets”, *Journal of Finance.*

Fuster, Andreas, Matthew Plosser, Philipp Schnabl, and James Vickery (2019) “The role of technology in mortgage lending”, *Review of Financial Studies*, 32 (5), pp. 1854–1899.

Gambacorta, Leonardo, Yiping Huang, Zhenhua Li, Han Qiu, and Shu Chen (2020) “Data vs collateral”, *BIS Working Paper*, 881.

Ghent, Andra C and Marianna Kudlyak (2011) “Recourse and residential mortgage default: Evidence from US states”, *Review of Financial Studies*, 24 (9), pp. 3139–3186.

Gopal, Manasa and Philipp Schnabl (2022) “The rise of finance companies and fintech lenders in small business lending”, *Review of Financial Studies.*

Haltiwanger, John, Ian Hathaway, and Javier Miranda (2014) “Declining business dynamism in the US high-technology sector”, *Available at SSRN 2397310.*

Hannan, Timothy H and John M McDowell (1987) “Rival precedence and the dynamics of technology adoption: An empirical analysis”, *Economica*, pp. 155–171.

Hau, Harald, Yi Huang, Hongzhe Shan, and Zixia Sheng (2018) “Fintech credit, financial inclusion and entrepreneurial growth”, *Working Paper.*

Hauswald, Robert and Robert Marquez (2003) “Information technology and financial services competition”, *Review of Financial Studies*, 16 (3), pp. 921–948.

Hauswald, Robert and Robert Marquez (2006) "Competition and strategic information acquisition in credit markets", *Review of Financial Studies*, 19 (3), pp. 967–1000.

He, Zhiguo, Sheila Jiang, Douglas Xu, and Xiao Yin (2021) "Investing in lending technology: IT spending in banking", *Working Paper*.

Herath, Shanaka and Gunther Maier (2010) "The hedonic price method in real estate and housing market research: a review of the literature", *Working Paper*.

Hernández-Murillo, Rubén, Gerard Llobet, and Roberto Fuentes (2010) "Strategic online banking adoption", *Journal of Banking & Finance*, 34 (7), pp. 1650–1663.

Hill, Robert J (2013) "Hedonic price indexes for residential housing: A survey, evaluation and taxonomy", *Journal of Economic Surveys*, 27 (5), pp. 879–914.

Hollander, Stephan and Arnt Verriest (2016) "Bridging the gap: The design of bank loan contracts and distance", *Journal of Financial Economics*, 119 (2), pp. 399–419.

Hurst, Erik and Annamaria Lusardi (2004) "Liquidity constraints, household wealth, and entrepreneurship", *Journal of Political Economy*, 112 (2), pp. 319–347.

Ivashina, Victoria and David Scharfstein (2010) "Bank lending during the financial crisis of 2008", *Journal of Financial Economics*, 97 (3), pp. 319–338.

Jiménez, Gabriel, Vicente Salas, and Jesús Saurina (2006) "Determinants of collateral", *Journal of Financial Economics*, 81 (2), pp. 255–281.

Jud, G, Daniel Winkler, and Stacy Sirmans (2002) "The impact of information technology on real estate licensee income", *Journal of Real Estate Practice and Education*, 5 (1), pp. 1–16.

Kantor, Shawn and Alexander Whalley (2019) "Research proximity and productivity: long-term evidence from agriculture", *Journal of Political Economy*, 127 (2), pp. 819–854.

Keller, Wolfgang (2002) "Geographic localization of international technology diffusion", *American Economic Review*, 92 (1), pp. 120–142.

Khwaja, Asim Ijaz and Atif Mian (2008) "Tracing the Impact of Bank Liquidity Shocks: Evidence from an Emerging Market", *American Economic Review*, 98 (4), pp. 1413–1442.

Klenow, Peter J and Huiyu Li (2020) "Innovative growth accounting", *NBER Macroeconomics Annual 2020, volume 35*, University of Chicago Press.

Kummerow, Max and Joëlle Chan Lun (2005) "Information and communication technology in the real estate industry: Productivity, industry structure and market efficiency", *Telecommunications Policy*, 29 (2-3), pp. 173–190.

Kwan, Alan, Chen Lin, Vesa Pursiainen, and Mingzhu Tai (2021) "Stress testing banks' digital capabilities: Evidence from the COVID-19 pandemic", *Working Paper*.

Liberti, José María and Atif Mian (2009) "Estimating the effect of hierarchies on information use", *Review of Financial Studies*, 22 (10), pp. 4057–4090.

Liberti, José María and Mitchell A Petersen (2019) "Information: Hard and soft", *Review of Corporate Finance Studies*, 8 (1), pp. 1–41.

Lin, Chen, Chicheng Ma, Yuchen Sun, and Yuchen Xu (2021) "The telegraph and modern banking development, 1881–1936", *Journal of Financial Economics*, 141 (2), pp. 730–749.

Mian, Atif and Amir Sufi (2014) "What explains the 2007–2009 drop in employment?", *Econometrica*, 82 (6), pp. 2197–2223.

Moretti, Enrico (2004) "Estimating the social return to higher education: Evidence from longitudinal and repeated cross-sectional data", *Journal of Econometrics*, 121 (1-2), pp. 175–212.

Munnell, Alicia H, Geoffrey MB Tootell, Lynn E Browne, and James McEneaney (1996) "Mortgage lending in Boston: Interpreting HMDA data", *The American Economic Review*, pp. 25–53.

Oster, Emily (2019) "Unobservable selection and coefficient stability: Theory and evidence", *Journal of Business & Economic Statistics*, 37 (2), pp. 187–204.

Ouimet, Paige and Rebecca Zarutskie (2014) "Who works for startups? The relation between firm age, employee age, and growth", *Journal of Financial Economics*, 112 (3), pp. 386–407.

Pancost, N Aaron and Garrett Schaller (2022) "Measuring measurement error", *Working Paper*.

Petersen, Mitchell A. (1999) "Banks and the Role of Lending Relationships: Evidence from the US Experience", *Working Paper*.

Petersen, Mitchell A and Raghuram G Rajan (1995) "The effect of credit market competition on lending relationships", *Quarterly Journal of Economics*, 110 (2), pp. 407–443.

Petersen, Mitchell A. and Raghuram G. Rajan (2002) "Does Distance Still Matters? The Information Revolution in Small Business Lending.", *Journal of Finance*, 57 (6), pp. 2533–2570.

Pierri, Nicola and Yannick Timmer (2022) "The importance of technology in banking during a crisis", *Journal of Monetary Economics*.

Prilmeier, Robert (2017) "Why do loans contain covenants? Evidence from lending relationships", *Journal of Financial Economics*, 123 (3), pp. 558–579.

Rajan, Raghuram G and Luigi Zingales (1998) "Financial dependence and growth", *American Economic Review*, 88 (3), pp. 559–586.

Rampini, Adriano A and S Viswanathan (2010) "Collateral, risk management, and the distribution of debt capacity", *Journal of Finance*, 65 (6), pp. 2293–2322.

Robb, Alicia M and David T Robinson (2014) "The capital structure decisions of new firms", *Review of Financial Studies*, 27 (1), pp. 153–179.

Sawyer, Steve, Rolf T Wigand, and Kevin Crowston (2005) "Redefining access: Uses and roles of information and communication technologies in the US residential real estate industry from 1995 to 2005", *Journal of Information Technology*, 20 (4), pp. 213–223.

Schmalz, Martin C, David A Sraer, and David Thesmar (2017) "Housing collateral and entrepreneurship", *Journal of Finance*, 72 (1), pp. 99–132.

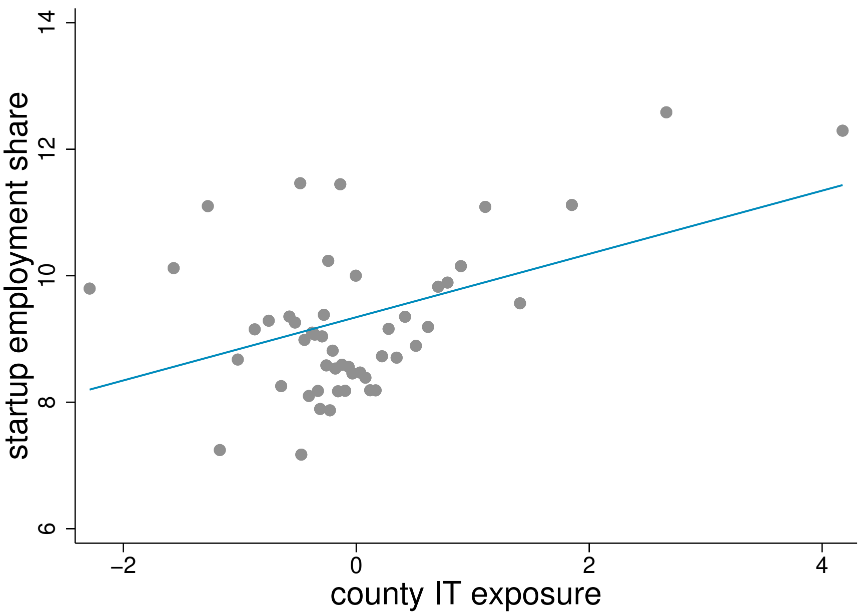
Vives, Xavier (2019) "Digital disruption in banking", *Annual Review of Financial Economics*, 11, pp. 243–272.

Vives, Xavier and Zhiqiang Ye (2020) "Information technology and bank competition", *Working Paper*.

Wei, Cankun, Meichen Fu, Li Wang, Hanbing Yang, Feng Tang, and Yuqing Xiong (2022) "The research development of hedonic price model-based real estate appraisal in the era of big data", *Land*, 11 (3), p. 334.

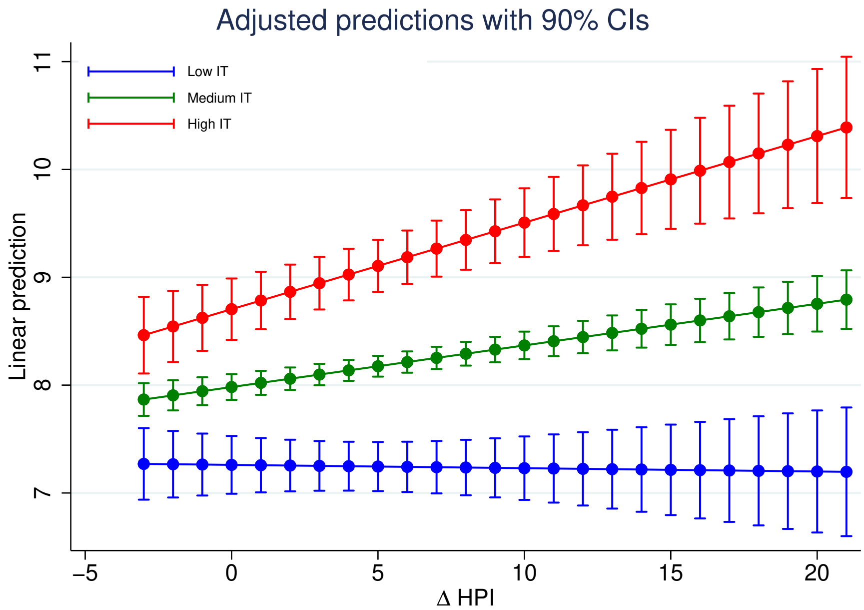
Figures and Tables

Figure 1: **Job creation by young firms and banks' IT adoption**



This figure shows a binned scatterplot of the share of employment by young firms over total employment in a county-industry cell, averaged over the sample period, on the vertical axis and county-level exposure to banks' IT adoption on the horizontal axis.

Figure 2: **Job creation by young firms and house price changes**



This figure plots the estimated effect of house price changes on job creation by young firms as a function of high (two standard deviation above mean) medium (mean) and low (two standard deviation below the mean) county IT exposure.

Table 1: **Descriptive statistics**

Panel (a): County level

Variable	Obs	Mean	Std. Dev.	Min	Max	P25	P50	P75
IT exposure	1774	0	1	-2.393	4.114	-.458	-.17	.289
bachelor or higher	1774	.183	.083	.06	.605	.122	.16	.223
log(income pc)	1774	10.791	.217	10.199	11.595	10.639	10.771	10.911
log(pop)	1774	10.995	1.135	8.501	16.06	10.186	10.774	11.651
share pop old	1774	.138	.037	.029	.349	.114	.137	.158
share pop black	1774	.091	.133	0	.855	.006	.03	.114
unemployment rate	1774	4.671	2.388	.7	29.7	3.1	4.1	5.8
employment share NAICS 23	1774	.059	.03	.004	.369	.04	.052	.071
employment share NAICS 31	1774	.216	.131	.003	.685	.114	.194	.297
employment share NAICS 44	1774	.158	.04	.052	.512	.131	.155	.181
employment share NAICS 62	1774	.137	.052	.01	.448	.101	.132	.165
employment share NAICS 72	1774	.097	.045	.02	.568	.072	.088	.111
PCs per employee (non-fin)	1774	.497	.094	.166	.951	.44	.499	.553

Panel (b): Bank level

Variable	Obs	Mean	Std. Dev.	Min	Max	P25	P50	P75
IT adoption	1189	0	1	-2.47	2.319	-.561	-.084	.553
log(assets)	1189	14.583	1.62	10.786	20.958	13.424	14.192	15.344
deposit ratio	1189	.831	.142	.007	.996	.783	.864	.925
non-interest income	1189	.188	.101	.022	.704	.124	.163	.226
secured assets	1189	.216	.107	0	.682	.145	.203	.279
return on assets	1189	.013	.005	-.021	.038	.011	.013	.015
equity ratio	1189	.092	.023	.049	.294	.078	.087	.099

This table reports summary statistics at the county and bank level.

Table 2: **Balancedness at the county level**

	low IT		**high IT**		**mean diff.**
	mean	sd	mean	sd	t
bachelor or higher	0.18	(0.09)	0.18	(0.08)	1.24
log(income pc)	10.79	(0.23)	10.77	(0.21)	1.64
log(pop)	10.94	(1.11)	10.82	(1.10)	2.00
share pop old	0.14	(0.04)	0.14	(0.04)	-1.63
share pop black	0.09	(0.14)	0.09	(0.13)	0.47
unemployment rate	4.71	(2.31)	4.60	(2.25)	0.84
employment share NAICS 23	0.06	(0.03)	0.06	(0.03)	-0.20
employment share NAICS 31	0.22	(0.13)	0.21	(0.13)	0.12
employment share NAICS 44	0.16	(0.04)	0.16	(0.04)	-0.13
employment share NAICS 62	0.14	(0.05)	0.14	(0.05)	-0.12
employment share NAICS 72	0.09	(0.04)	0.10	(0.05)	-1.62
PCs per employee (non-fin)	0.50	(0.10)	0.49	(0.09)	1.12
Observations	592		591		1183

This table compares the means of county-level control variables in counties in the bottom and top tercile of the distribution of IT exposure. The column *mean diff* denotes the t-value for the difference in means.

Table 3: **County IT exposure and entrepreneurship**

VARIABLES	(1) share 0-1	(2) share 0-1	(3) share 0-1	(4) share 0-1	(5) share 0-1	(6) share 0-1
IT exposure	0.462***	0.405***	0.378***	0.378***		0.220**
	(0.118)	(0.100)	(0.100)	(0.099)		(0.110)
IT exposure $\times$ ext. fin. dep				0.167***	0.162***	0.471***
				(0.043)	(0.042)	(0.159)
Observations	25,779	25,779	25,779	25,779	25,779	25,779
R-squared	0.003	0.046	0.248	0.248	0.350	0.249
County Controls	-	✓	✓	✓	-	✓
NAICS FE	-	-	✓	✓	✓	✓
County FE	-	-	-	-	✓	-

This table reports results for regressions at the county-industry level (see Equation 9). The dependent variable is the employment share of firms of age 0-1. *IT Exposure* denotes the IT adoption of banks in the county, standardized to a mean of zero and a standard deviation of one. *Ext. fin. dep* measures the dependence on external finance in an industry. Standard errors are clustered at the county level. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: **County IT exposure, entrepreneurship, and collateral**

VARIABLES	(1) share 0-1	(2) share 0-1	(3) share 0-1	(4) share 0-1	(5) share 0-1	(6) share 0-1
IT exposure	0.303***	0.296***	0.127	0.105		
	(0.087)	(0.083)	(0.089)	(0.087)		
Δ HPI		0.038***	0.040***	0.042***		
		(0.011)	(0.011)	(0.010)		
IT exposure × Δ HPI			0.029**	0.031**		
			(0.014)	(0.013)		
IT exposure × Δ HPI × Low SU capital					0.044**	
					(0.021)	
IT exposure × Δ HPI × home equity						0.271*
						(0.155)
Observations	134,499	134,499	134,499	134,499	168,598	168,598
R-squared	0.040	0.041	0.041	0.235	0.331	0.331
County Controls	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	-	-	-
NAICS × Year FE	-	-	-	✓	✓	✓
County × Year FE	-	-	-	-	✓	✓

This table reports results for regressions at the county-industry-year level (see Equation 10). The dependent variable is the employment share of firms of age 0-1. *IT Exposure* denotes the IT adoption of banks in the county, standardized to a mean of zero and a standard deviation of one. Δ *HPI* is the yearly change in house prices in county *c*. *Low SU capital* is a dummy that takes on a value of one for industries that require low amounts of capital required to start a company. *Home equity* measures the dependence on home equity of young firms in an industry as a source to start or expand operations. Standard errors are clustered at the county level. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: **County IT exposure and transition rates**

VARIABLES	(1) tr 0/1-2/3	(2) tr 0/1-2/3	(3) tr 0/1-2/3	(4) tr 2/3-4/5	(5) tr 2/3-4/5	(6) tr 2/3-4/5
IT exposure	-0.030	-0.030		-0.019	-0.019	
	(0.035)	(0.035)		(0.026)	(0.026)	
IT exposure × ext. fin. dep		-0.023	-0.025		-0.015	-0.014
		(0.019)	(0.019)		(0.014)	(0.014)
Observations	23,729	23,729	23,729	22,675	22,675	22,675
R-squared	0.068	0.068	0.138	0.046	0.046	0.119
County Controls	✓	✓	-	✓	✓	-
NAICS FE	✓	✓	✓	✓	✓	✓
County FE	-	-	✓	-	-	✓

This table reports results for regressions at the county-industry level (see Equation 9). The dependent variable is the transition rate of firms of age 0–1 to 2–3 (columns 1–3) and of age 2–3 to 4–5 (columns 4–6). *IT Exposure* denotes the IT adoption of banks in the county, standardized to a mean of zero and a standard deviation of one. *Ext. fin. dep* measures the dependence on external finance in an industry. Standard errors are clustered at the county level. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: **County IT exposure and entrepreneurship – Complexity**

VARIABLES	(1) share 0-1	(2) share 0-1	(3) share 0-1	(4) share 0-1
IT exposure	0.537***	0.352***	0.697***	0.220
	(0.193)	(0.134)	(0.199)	(0.136)
Observations	11,943	11,943	13,236	12,506
R-squared	0.003	0.002	0.004	0.001
Measure	High Vol.	Low Vol.	Complex	Not Complex

his table reports results for regressions at the county-industry level (see Equation 9). The dependent variable is the employment share of firms of age 0-1. *IT Exposure* denotes the IT adoption of banks in the county, standardized to a mean of zero and a standard deviation of one. *Ext. fin. dep* measures the dependence on external finance in an industry. Each pair of columns report the regressions in a different sub-sample. Columns (1) and (2) report the results of the regression estimated separately in counties with house price volatility above and below the median. Columns (3) and (4) report the results of the regression estimated separately in counties where banks have organizational complexity – measured as the number of entities within each BHC – above or below the median. Standard errors are clustered at the county level. *** p<0.01, ** p<0.05, * p<0.1.

Table 7: **Banks' IT, origination costs, and processing times**

Panel (a): OLS

VARIABLES	(1) Δ loans	(2) Δ loans (sm)	(3) ln(days)	(4) ln(days)
IT	0.005	0.001	-0.025***	-0.027***
	(0.004)	(0.004)	(0.009)	(0.008)
IT × Δ house prices	0.159**	0.267***		
	(0.069)	(0.060)		
Observations	120,495	120,495	59,408	59,352
R-squared	0.201	0.208	0.234	0.256
Bank Controls	✓	✓	✓	✓
Year FE	C*T	C*T	C*T	C*T + I*T

Panel (b): IV

VARIABLES	(1) Δ loans	(2) Δ loans (sm)	(3) ln(days)	(4) ln(days)
IT	-0.109***	-0.168***	-0.330***	-0.331***
	(0.024)	(0.021)	(0.096)	(0.095)
IT × Δ house prices	2.824***	3.788***		
	(0.338)	(0.310)		
Observations	120,495	120,495	59,408	59,352
Bank Controls	✓	✓	✓	✓
Year FE	C*T	C*T	C*T	C*T + I*T
F stat	174.4	174.4	11.19	11.16

In each panel, columns (1) and (2) report results for Equation (11). The dependent variable is the change in total CRA loans by bank b to borrowers in county c in year t in column (1); and CRA loans with an amount of less than $100,000 in column (2). IT measures the IT adoption of bank b, ΔHPI is the yearly change in house prices in county c. Columns (3) and (4) report results for Equation (12). The dependent variable is the log of the days between approval and disbursement of loan i. Panel (b) instruments bank-level IT adoption with the land-grant colleges instrument. Standard errors are clustered at the county level. *** p<0.01, ** p<0.05, * p<0.1.

A1 Online Appendix

A1.1 Model proofs

Recall from the discussion in the main text that only projects of high quality are funded in equilibrium irrespective of the type of bank, so Prediction 4 follows immediately. Thus, we can henceforth limit attention to firms with a good project. Next, we construct the share of expected lending to young firms as a fraction of total expected lending, s_Y.

All old firms with a good project are funded, which are of quantity $q(1-y)$. Young firms with a good project, which are of measure qy, are funded when they meet a high-IT bank, which occurs with probability h, and when they have enough collateral $C > C_{min}$, which holds for a fraction $1-G(C_{min})$ of these firms (all characteristics are independent). Thus, the measure of lending to young firms is $qyh[1-G(C_{min})]$. Taken these points together, we obtain the share s_Y stated in Proposition 1.

The derivatives follow, where we use $\frac{dC_{min}}{dP} = -\frac{p_B}{1-p_B}(x-r)\frac{1}{P^2} < 0$ to sign them:

$$\frac{ds_Y}{dh} = \frac{y(1-y)[1-G(C_{min})]}{(1-y+yh[1-G(C_{min})])^2} > 0 \tag{13}$$

$$\frac{ds_Y}{dP} = \frac{y(1-y)h}{(1-y+yh[1-G(C_{min})])^2}\, g(C_{min}) \left(-\frac{dC_{min}}{dP}\right) > 0 \tag{14}$$

$$\frac{d^2 s_Y}{dhdP} = \frac{y(1-y)\big(1-y-yh[1-G(C_{min})]\big)}{(1-y+yh[1-G(C_{min})])^3}\, g(C_{min}) \left(-\frac{dC_{min}}{dP}\right) > 0, \tag{15}$$

where the sign of $\frac{d^2 s_Y}{dhdP}$ arises from observing that $1-y-yh[1-G(C_{min})] \geq 1-y-yh = 1-y(1-h) > 0$.

We turn to the case of recourse, where the no-recourse derivatives are unchanged:

$$\frac{ds_Y^R}{dh} = \frac{y(1-y)(1-i)[1-G(C_{min})]}{(1-y+y\{i+(1-i)h[1-G(C_{min})]\})^2} > 0 \tag{16}$$

We have $\frac{ds_Y^R}{dh} \to \frac{ds_Y}{dh}$ for $i \to 0$; since $i > 0$, this reduces the numerator and increases the denominator of $\frac{ds_Y^R}{dh}$ relative to $\frac{ds_Y}{dh}$, so $\frac{ds_Y^R}{dh} < \frac{ds_Y^{NR}}{dh} = \frac{ds_Y}{dh}$.

A1.2 Model extensions

This section presents two extensions to the baseline model.

Extension 1: Recourse. Recourse – i.e. lenders' ability to possess other borrower assets or future income through a deficiency judgment – can substitute for the need of screening borrowers through collateral. To study the role of recourse, we assume that a fraction $i \in (0, 1)$ of firms generate an additional external income I at date 1. Banks may have recourse to this income, depending on whether they are located in states with recourse (R) or with no recourse (NR). In recourse states, all banks can obtain this external income, while only high-IT banks have the comparative advantage in lending via collateral. For expositional clarity, we assume that the external income is independent of other firm characteristics and that it suffices to back the loan, $I \geq PC_{min}$.

Nothing changes in no-recourse states, so the share of lending to young firms is $s_Y^{NR} = s_Y$ given in Equation 8. In recourse states, by contrast, young firms now also receive funding when they have additional income (a fraction i of them do). Because their future income is no smaller than the collateral value, no additional incentive problems arise and only young firms of high quality seek funding. Thus, the share of lending to young firms in recourse states is

$$s_Y^R \equiv \frac{y\left\{i + (1-i)h[1 - G(C_{min})]\right\}}{1 - y + y\left\{i + (1-i)h[1 - G(C_{min})]\right\}}. \tag{17}$$

The next prediction compares recourse to no-recourse states.

Prediction 5. A higher share of high-IT banks increases the share of lending to young firms by less in recourse states than in non-recourse states, $\frac{ds_Y^{NR}}{dh} > \frac{ds_Y^R}{dh}$. Quite intuitively, this result arises because recourse to future income mitigates the effective comparative advantage of high-IT banks in using collateral.

Extension 2: Geographical distance. A large literature in banking highlights the importance of geographical distance between lenders and borrowers and how it affects the relative values of hard and soft information. In our model, high-IT banks have a comparative advantage in screening based on collateral, which can be interpreted as hard-information lending (and is thus unaffected by distance). Low-IT banks lend based on information acquisition instead. To allow for a role of distance, we assume that low-IT banks can screen some young firms, namely those that are close. Hence, we relax Assumption 6 by assuming

$$v_{HighIT} < F_Y^{close} < p_G x - 1 < \min\{F_Y^{distant}, v_{LowIT}\}, \tag{18}$$

where the cost of information acquisition is low enough relative to the expected surplus of a good project when the firm is close to the bank. Let $d \in (0, 1)$ be the fraction of young firms that is distant and the remainder is close.

Thus we can express for each type of bank the share of credit to young firms as a proportion of total credit, ϕ, and how it depends on the bank's distance to the borrower. For a high-IT bank, this share is invariant to distance:

$$\phi_{HighIT} = \frac{y[1 - G(C_{min})]}{y[1 - G(C_{min})] + 1 - y} = \phi_{HighIT}^{distant} = \phi_{HighIT}^{close}, \tag{19}$$

because all young firms with sufficient collateral are funded (irrespective of distance). For a low-IT bank, by contrast, this share depends on distance:

$$\phi^{distant}_{LowIT} = 0 < \frac{y(1-d)}{y(1-d)+1-y} = \phi^{close}_{LowIT}, \tag{20}$$

because no distant young firms are funded, but geographically close ones are. Note that when most young firms are distant (a high d), we have $\phi_{HighIT} > \phi^{close}_{LowIT}$. Also note that the shares of low-IT banks are independent of the price of collateral, so $\frac{d\phi_{LowIT}}{dP} = 0$.

Prediction 6. The geographic distance between lenders and borrowers matters more for low-IT than high-IT banks. Specifically, the share of lending to young firms varies more with distance for low-IT banks than for high-IT banks:

$$\phi^{close}_{LowIT} - \phi^{distant}_{LowIT} > \phi^{close}_{HighIT} - \phi^{distant}_{HighIT}. \tag{21}$$

The advantage of high-IT banks in hard information lending makes their lending less sensitive to the lender-borrower distance. Of particular relevance for the empirical analysis is how the distance between borrowers and lenders impacts the sensitivity of credit to local economic conditions. Adelino et al. (2017) document that startups strongly respond to changes in economic opportunities and are responsible for a larger share of job creation when local opportunities arise thanks to a positive income shock. As the responsiveness of startup activity to local shocks is larger than for older firms, the more a bank lends to startups in a market, the larger its credit supply should respond to local economic conditions. Therefore, **Prediction 6** implies that low IT banks' credit responds less to local economic conditions in counties that are more-distant from the banks' HQ, while distance does not matter for the responsiveness of lending by high IT banks.

A1.3 Additional empirical evidence

Recourse default. Recourse can partially substitute for the need of screening borrowers through collateral. The ability to recourse in the case of foreclosure or default thus diminishes the misalignment of interests (Ghent and Kudlyak, 2011). **Prediction 5** thus implies that the positive relationship between IT exposure and entrepreneurship is more pronounced in non-recourse states. To test this prediction, we exploit heterogeneity across US states in terms of legal and practical considerations that make obtaining a deficiency judgment more or less difficult for lenders. We follow Ghent and Kudlyak (2011) to classify recourse and non-recourse states according to whether they allow, at least in some cases, deficiency judgment. We then estimate the cross-sectional relationship between IT and entrepreneurship (i.e. Equation 9) for counties in recourse versus non-recourse states.[10]

Columns (1) and (2) in Table A1 highlight that the positive relationship between IT exposure and job creation by startups is stronger in non-recourse states. We confirm this finding in interaction specifications in columns (3) and (4). Columns (3) shows that in recourse states the relationship between IT adoption and entrepreneurship is significantly weaker. Column (4) confirms the finding when we exclude North Carolina, as its classification presents some ambiguity.

The role of distance in lending. In the model, IT lowers the cost of banks to verify the existence and market value of collateral, and transmit the information to their (distant) HQ. This mechanism is consistent with work that suggests that IT adoption by banks reduces the importance of distance in lending decisions, as it enables a more effective transmission of hard information (Petersen and Rajan, 2002; Vives and Ye, 2020).

[10] Ghent and Kudlyak (2011) relies on recourse / non-recourse classifications of states from the 21st edition (2004) of the National Mortgage Servicer's Reference Directory to show that recourse clauses impact borrowers' behavior.

Prediction 6 thus states that with increasing IT adoption, lending should become more responsive to new investment opportunities in more distant counties. Following a large literature that shows that informational frictions increase with lender-borrower distance (Liberti and Petersen, 2019), we test whether the relationship between local investment opportunities and lender-borrower distance varies with banks' use of IT. We consider the following specification from 1999 to 2007 at the bank-county-year level:

$$\begin{aligned} \Delta loans_{b,c,t} = \beta_1 \ log(distance)_{b,c} + \beta_2 \ \Delta income \ p.c._{c,t} \\ + \beta_3 \ log(distance)_{b,c} \times \Delta income \ p.c._{c,t} \\ + controls_{c/b,t-1} + \theta_{c,t} + \varepsilon_{b,c,t}, \\ \text{if IT} = \text{low/high}. \end{aligned} \tag{22}$$

The dependent variable is the log difference in total CRA small business lending by bank b to borrower county c in year t. The variable $log(distance)$ measures the log of the distance between banks' HQ county and the county of the borrower. We proxy investment opportunities in borrower countries with the log change in county-level income per capita (Adelino et al., 2017). Regressions further include standard county controls, as well as year or county*year fixed effects. Bank-level controls are the log of total assets, deposits over total liabilities, the share non-interest income, securities over total assets, return on assets, the equity ratio (Tier 1), and the wholesale funding ratio. We also control for the share of mortgage loans out of total loans. We thus hold the allocation of lending of credit *across* lending segments constant. Standard errors are clustered at the county level. An increase in local investment opportunities is expected to increase local lending ($\beta_1 > 0$), especially in borrower counties nearer to the HQ ($\beta_3 < 0$). If banks' IT adoption reduces the importance of distance, then β_3 should be significantly smaller in magnitude for *high IT* banks.

Results in Table A2 are in line with the hypotheses. Column (1) shows that rising local incomes are associated with higher local loan growth. Greater distance reduces the sensitivity of banks' small business lending in response to local investment opportunities, as the interaction terms between changes in income and distance is negative. This findings holds when we include county*year fixed effects to control for any unobservable time-varying borrower-county characteristics in column (2). Columns (3) and (4) show that the lower responsiveness of bank lending in counties located further away is present only among low IT banks; for high IT banks, distance has no dampening effect.

An interaction specifications in column (5) confirms this finding: While distance reduces the sensitivity of lending to changes in local investment opportunities for low IT banks, among high IT banks distance matters significantly less. Results are similar when we focus on total lending through loans with origination amounts below $100,000, which are usually granted to smaller companies. Note that coefficients increase in magnitude, which is consistent with the common finding that informational frictions are more severe among smaller firms.

Finally, columns (7)–(8) replicate columns (5)–(6), but instrument banks' IT, as well as the associated interaction terms, with the IV based on distance to the nearest two land-grant colleges. The main coefficients are similar in terms of sign and significance, but again larger in magnitude.

IV: first stage. Figure A1 shows a strong negative association between the distance to the nearest land-grant colleges and banks' IT adoption (panel a). Further, in regressions we control for an extensive set of bank-level controls – most importantly bank size, which is commonly associated with economies of scale that could facilitate IT adoption. As panel (b) of Figure A1 shows, the strong relationship between distance to the nearest land-grant colleges and IT adoption remains when we condition on bank size (log assets).

IT and the use of collateral. A key assumption of the model is that high IT banks have a relative cost advantage in screening through collateral. While we do not have loan-level information on collateralized lending to startups, we can provide empirical evidence on the presence of collateral for large corporate loans with data from DealScan (Ivashina and Scharfstein, 2010). Figure A2 shows that the share of loans that are collateralized is positively correlated with bank IT adoption. To ensure that this correlation is not driven by (unobservable) borrower heterogeneity, we estimate the following linear probability model:

$$secured_{b,i,t} = \beta\ IT_b + \tau_t + \theta_i + \varepsilon_{b,i,t}, \tag{23}$$

where b denotes a bank that granted a loan in year t to corporate borrower i and $secured_{b,i,t}$ is a dummy equal to one whenever the loan is collateralized. Results in Table A3 confirm that more IT-intensive banks are more likely to require collateral than other banks, even when controlling for borrower characteristics through borrower fixed effects.

The role of local competition. Our theoretical framework abstracts from interactions between local competition and IT adoption in the banking sector. Instead, banks and borrowers share the surplus from lending if a loan is granted. To ensure that local competition does not affect our key empirical results, we re-estimate Equation 9, but control for market concentration (measured through the HHI) and its interaction with IT. Results are presented in Table A4, where columns (1)–(2) construct the HHI from CRA loan shares and columns (3)–(4) from deposit shares. In general, higher concentration is associated with higher startup activity. This could reflect that lenders in less competitive markets have a sufficiently high surplus to acquire costly soft information or that they might be more prone to lend to startups because know they expect to extract more surplus in the future as young firms grow (Petersen and Rajan, 1995). However, there is no significant interaction between concentration and local IT adoption in banking, and the positive impact of IT on startups remains largely unaffected when we account for the local market structure. This result supports the assumption to abstract from local competition.

IT adoption over time. An alternative approach to test **Prediction 1** is to analyze the relationship between the *increase* in IT adoption and *changes* in entrepreneurship at the county-level. To do so, we compute the change in county exposure as

$$\Delta IT_c = \sum_{b=1}^{N} \Delta IT_b * \frac{No.\ Branches_{b,c}}{No.\ Branches_c}, \tag{24}$$

where ΔIT_b is the increase of IT adoption between 1999 and 2006 of bank b. We find that counties more exposed to an increase in IT in banking also experienced stronger performance of startups, as illustrated by panel (b) in Figure A3. The positive correlation between changes in IT adoption in banking and changes in startup rates is also confirmed by more formal regression analysis presented in Table A5. Note that this first-difference approach implicitly controls any county-level (time invariant) observable and unobservable characteristics.

Minority entrepreneurship. Our results indicate that IT increases the importance of real estate collateral in lending decisions, which could suggest that entrepreneurs with insufficient personal or family wealth may not be able to benefit to the same extent as others. Previous research has shown that some communities, such as racial and ethnic minorities, have experienced long lasting discrimination in the mortgage market (Munnell et al., 1996) and have thus been accumulated less real estate wealth. Minority entrepreneurs also face more hurdles in access to capital (Fairlie et al., 2020).

The QWI report employment by race, but not the race of the entrepreneur. To the extent that entrepreneurs are likely to hire from their personal networks or job referral are more likely among people of the same ethnic or racial group, startups with a larger share of black employees are more likely to be owned by a black entrepreneur. We therefore investigate the relationship between IT in banking and the share of startups' employees that are black within a county, normalized by subtracting the same share for white employees. Table A6 reveals that counties more exposed to IT in banking have a lower share of black employees among startups. This result suggests that IT adoption in banking fosters entrepreneurship and business dynamism in general, but may perpetuate inequality across demographic groups.

Other tests. Table A7 presents robustness tests to our main results at the county level. Column (1) replicates the baseline result for comparison (see column (3) in Table 3). In column (2), IT exposure is the unweighted average of the IT adoption of banks that operate in a county and in column (3) exposure is weighted by the share of local deposits (rather than the number of branches). The positive association between IT exposure and entrepreneurial activity remains, highlighting that it is not driven by any specific choice of the construction of the IT exposure measure. Column (4) excludes employment in startups in the financial and education industries and column (5) excludes Wyoming, the state with the highest exposure to banks' IT adoption. Results remain unaltered. Column (6) includes state fixed effects and shows that results are also present when we exploit within-state variation only. Column (7) normalizes the share of employment in startups by the previous year's total employment to rule out that our results are driven by a decline in total employment instead of an increase in young firms' employment. In fact, column (8) shows that IT positively but insignificantly affects total employment, suggesting that the employment growth in young firms can promote total job growth, but as young firms are mostly small when founded, the effect is likely to weak to make a strong contemporaneous economic and statistical impact on total employment. our results are not driven by a decline in total employment. Column (9) focuses in firms in tradable industries, which are less affected by local economic conditions.[11] Finally, columns (10) and (11) address the concern that the availability of other forms of external financing, venture capital (VC) in particular, may be correlated with IT exposure. As VC funding is highly concentrated in a small fraction of the US territory, we exclude the top 20 counties (representing almost 80% of VC funding at the time) or seven states with the highest VC activity,[12] and find results similar to baseline.

Table A8 shows that our results are robust to using the number of young firms, rather than their employment, as dependent variable.

To assess the effects on firm quality, we use the number of patents by firms of age 0–5 per employee in firms of age 0–5 as dependent variable in county-level regressions. We estimate county-level regressions with IT exposure as independent variable. Table A9 shows results: neither for patents per employee nor per firm in columns (1) and (2) is there a statistically or economically significant correlation with IT exposure. This picture does not change when we weight patents by citation in columns (3)–(4). We find identical results when focus on patents (and patents per employee or number of firms) by startups of age 0 in Table A10.

Table A11 looks at overall bank lending in bank-level regressions. Column (1) uses very small business loans (below $100,000) over assets as dependent variable. A one standard deviation increase in banks IT adoption is associated with a significant 0.3% increase in the share of small business lending (18% of the sample mean). For loans below $1 million in column (2), which make up a larger share of total assets, the increase is

[11] We rely on the tradable classification of 4 digit industries by Mian and Sufi (2014), which we aggregate to the 2 digit level.

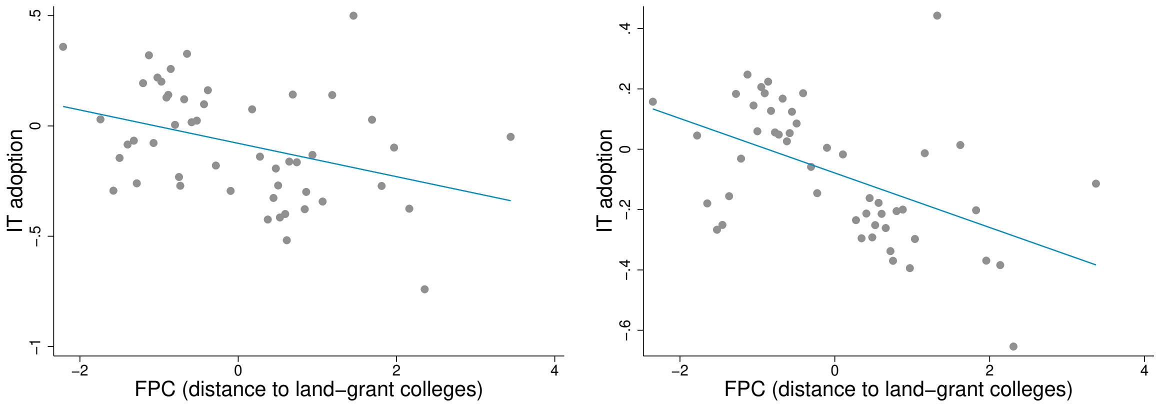
[12] See e.g. https://pitchbook.com/newsletter/28-counties-account-for-80-of-vc-investment-in-the-us.

also significant and equals 1% for a 1sd increase in IT (16% of the mean). Column (3) thus uses the log of total C&I lending as dependent variable and finds a positive, albeit insignificant effect of IT. We obtain similar findings when we use C&I lending over assets in column (4). These results suggest that the increase in lending to young firms (proxied with small business lending) does not appear to crowd out C&I lending overall.

Table A12 controls for the average processing time of SBA loans in each county.

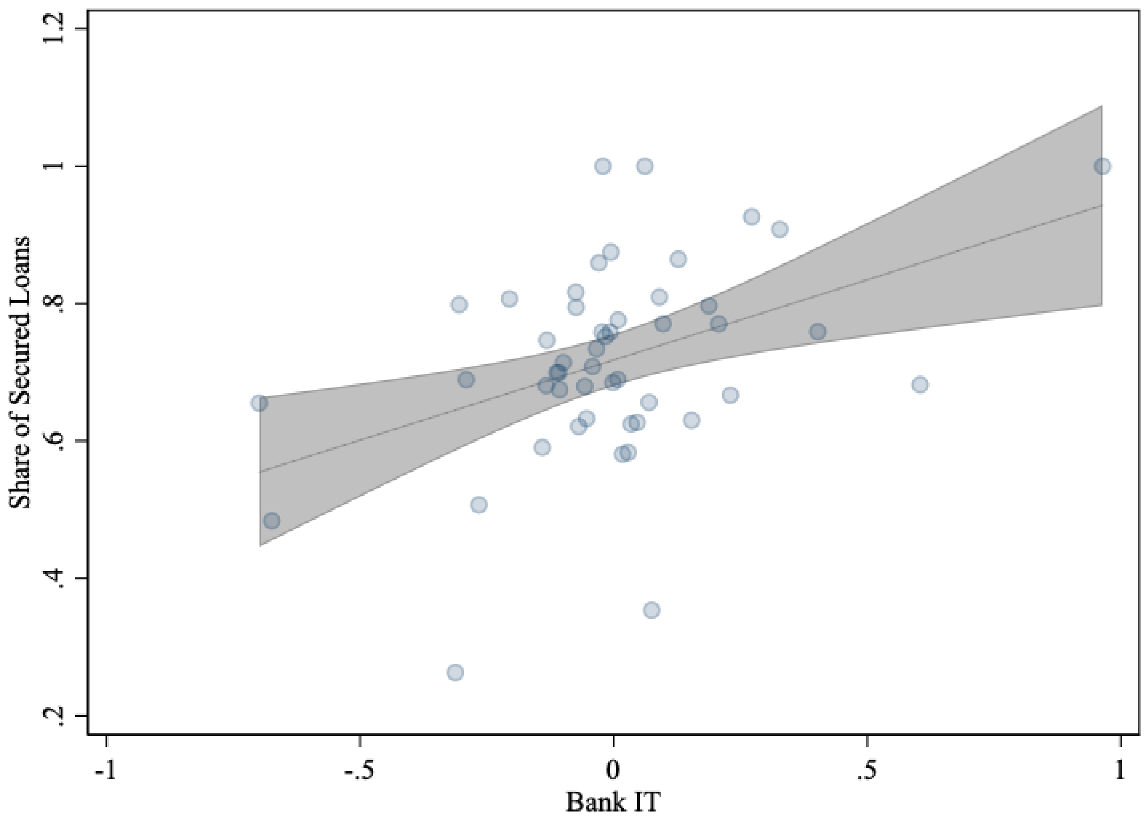
Figure A1: **Distance to land-grant colleges and IT adoption**

(a) Unconditional (b) Conditional on bank size



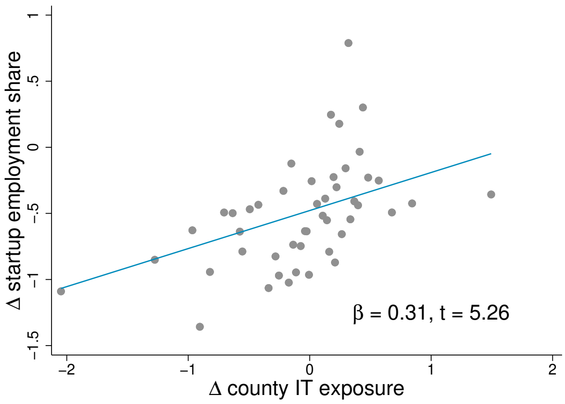
Panel (a) shows a binned scatterplot of banks' IT adoption on the vertical axis against the first principal component (FPC) of the distance of banks' HQ to the nearest two land-grant colleges on the horizontal axis. Panel (b) shows the same binned scatterplot but conditional on bank size, measured via the log of total bank assets.

Figure A2: Share of Secured Loans



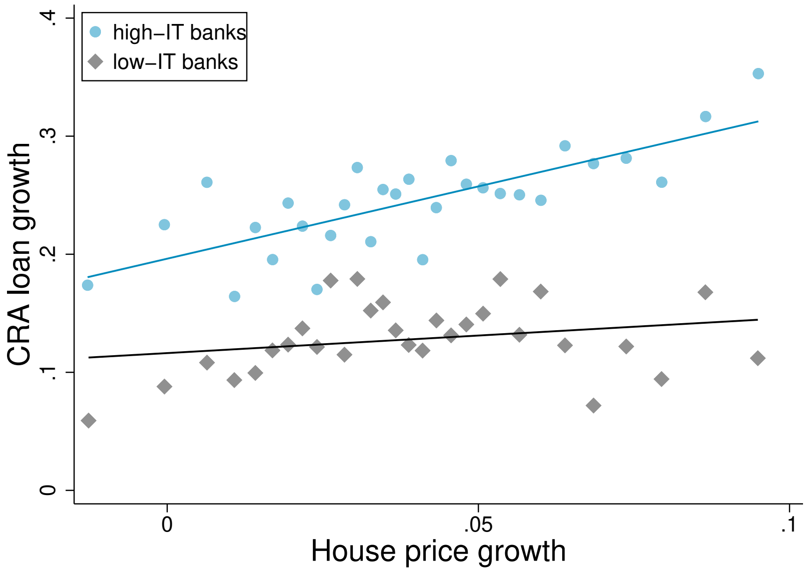
This figure shows the share of secured loans in the Dealscan syndicated loan data and banks' IT adoption.

Figure A3: **Job creation by young firms and banks' IT adoption – changes**



This figure shows a binned scatterplot of the *change* in the startup rate in a county-industry between 2000 and 2007 (in percentage points) on the y-axis and the exposure of a county to banks' *change* in IT adoption between 2000 and 2007 (standardized) on the x-axis.

Figure A4: **Banks' IT, house prices, and loan growth**



This figure shows a binscatter of CRA loan growth on the vertical axis and county-level house price growth on the horizontal axis. The sample is split into banks above and below the median along the IT distribution. In a regression of CRA loan growth on house price growth ($\Delta CRA_{b,c,t} = \Delta$ *house price growth*$_{c,t} + \varepsilon_{b,c,t}$), the respective coefficients (t-values) for high- and low-IT banks are 1.22 (5.93) and 0.30 (1.77).

Table A1: **IT exposure and recourse**

VARIABLES	(1) share 0-1	(2) share 0-1	(3) share 0-1	(4) share 0-1
IT exposure	0.305***	0.471***	0.700***	0.673***
	(0.0966)	(0.176)	(0.203)	(0.204)
Recourse State × IT exposure			-0.463**	-0.434**
			(0.220)	(0.220)
Observations	20,046	5,696	25,742	24,630
R-squared	0.275	0.359	0.272	0.273
County Controls	✓	✓	✓	✓
NAICS FE	✓	✓	✓	✓
Specification	Recourse	Non-Recourse	Interaction	No NC

This table reports results from cross-sectional regressions at the county-industry level (see Equation 9). The dependent variable is the share of the employment in firms of age 0-1 in county c and industry i. $IT\ Exposure_c$ is the IT adoption of banks in the county, measured by the IT adoption of banks historically present in the county, and standardized with mean zero and a standard deviation of one. $Recourse\ State_s$ a dummy that is one if the state is a recourse state. Column (1) shows the baseline specification only for recourse states. Column (2) shows the baseline specification only for non-recourse states. Column (3) and (4) show the regression with an interaction between a $Recourse\ State_s$ and $IT\ Exposure_c$. Column (4) excludes North Carolina, as its classification presents some ambiguity. Standard errors are clustered at the county level *** p<0.01, ** p<0.05, * p<0.1.

Table A2: **Banks' IT, distance, and lending**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
			low IT	high IT			IV	IV
VARIABLES	Δ loans	Δ loans	Δ loans	Δ loans	Δ loans	Δ loans (sm)	Δ loans	Δ loans (sm)
log(distance)	0.015***	0.019***	0.048***	-0.002	0.018***	0.016***	0.001	-0.002
	(0.003)	(0.003)	(0.005)	(0.006)	(0.003)	(0.003)	(0.007)	(0.006)
Δ income	0.019***							
	(0.004)							
Δ income × log(distance)	-0.003***	-0.004***	-0.009***	0.003***	-0.004***	-0.005***	0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
IT					0.049***	0.040***	1.157***	0.884***
					(0.017)	(0.015)	(0.225)	(0.184)
Δ income × IT					-0.017***	-0.021***	-0.192***	-0.164***
					(0.004)	(0.004)	(0.039)	(0.033)
log(distance) × IT					-0.006*	-0.002	-0.204***	-0.158***
					(0.003)	(0.003)	(0.037)	(0.030)
Δ income × log(distance) × IT					0.004***	0.005***	0.041***	0.037***
					(0.001)	(0.001)	(0.007)	(0.006)
Observations	144,722	144,144	73,865	47,146	144,144	125,756	144,144	125,756
R-squared	0.025	0.167	0.267	0.302	0.168	0.199		
Bank Controls		✓	✓	✓	✓	✓	✓	✓
County × Year FE		✓	✓	✓	✓	✓	✓	✓

This table reports results for regressions at the bank-county-year level (see Equation 22). The dependent variable is the change in total CRA loans by bank b to county c in year t or in CRA loans with an amount of less than \$ 100,000. IT_b is the IT adoption of bank b. $\Delta\ Income_{c,t}$ is the change in per capita income in county c between year $t-1$ and t. $log(distance)_{b,c}$ is the log of the number of miles between bank b's headquarters and county *low/high IT* refers to banks in the bottom/top tercile of the IT distribution. Standard errors are clustered at the bank and county level. The Kleibergen-Paap Wald F-statistics for all instrumented variables considered in columns (7) and (8) jointly equal 8.17 and 7.80. *** p<0.01, ** p<0.05, * p<0.1.

Table A3: **Secured Loans and Bank IT adoption**

VARIABLES	(1) Secured	(2) Secured	(3) Secured	(4) Secured	(5) Secured
Bank IT	0.230***	0.279***	0.039*	0.046**	0.033*
	(0.051)	(0.057)	(0.022)	(0.019)	(0.017)
Observations	211,796	211,795	207,889	207,888	147,212
R-squared	0.018	0.049	0.820	0.824	0.822
Borrower FE	-	-	✓	✓	✓
Year FE	-	✓	-	✓	✓
Cluster	Bank	Bank	Bank	Bank	Bank
Sample	All	All	All	All	Pre-GFC

This table reports results from syndicated loan-level regression using data from Dealscan. The dependent variable is a dummy that equals one if the loan is secured and 0 otherwise. Standard errors are clustered at the bank-level *** p<0.01, ** p<0.05, * p<0.1.

Table A4: **The role of local competition**

VARIABLES	(1) share 0-1	(2) share 0-1	(3) share 0-1	(4) share 0-1
IT exposure	0.393***	0.415***	0.372***	0.372***
	(0.110)	(0.100)	(0.113)	(0.113)
HHI	2.439***	2.483***	4.895***	4.893***
	(0.910)	(0.906)	(1.019)	(1.017)
HHI × IT exposure		0.646		-0.015
		(0.603)		(0.954)
Observations	25,779	25,779	25,779	25,779
R-squared	0.249	0.249	0.252	0.252
County Controls	✓	✓	✓	✓
NAICS FE	✓	✓	✓	✓
Cluster	County	County	County	County
HHI	CRA lending	CRA lending	FDIC deposits	FDIC deposits

This table reports results for the following regression: $\text{startups}_{c,i} = \beta\ IT\ exposure_{c,99} + \delta\ HHI_{c,99} + \gamma\ IT\ exposure_{c,99} \times HHI_{c,99} + controls_{c,99} + \phi_i + \varepsilon_{c,i}$, where $\text{startups}_{c,i}$ is defined as the share of the employees in county c and industry t which is employed at a firm with at most 1 year of life. The share is then averaged across the years 2000 and 2007. IT_c is the IT adoption of banks in the county, measured by the IT adoption of banks historically present in the county, and standardized with mean zero and a standard deviation of one. $HHI_{c,99}$ is the Herfindahl-Hirschman Index in county c, where market shares are computed from either small business lending in 1999 (from CRA data) or deposits in 1999 (from FDIC data). Standard errors are clustered at the county level. *** p<0.01, ** p<0.05, * p<0.1.

Table A5: **County IT exposure and Entrepreneurship - Long Differences**

VARIABLES	(1) Δ share 0-1	(2) Δ share 0-1	(3) Δ share 0-1	(4) Δ share 0-1	(5) Δ share 0-1
Δ IT exposure	0.153*	0.241***	0.248***	0.210**	
	(0.084)	(0.085)	(0.085)	(0.088)	
Δ IT exposure × ext. fin. dep				0.258*	0.201
				(0.142)	(0.136)
Observations	15,952	15,952	15,952	15,952	15,952
R-squared	0.000	0.007	0.021	0.014	0.144
County Controls	-	✓	✓	✓	-
NAICS FE	-	-	✓	✓	✓
County FE	-	-	-	-	✓
Cluster	County	County	County	County	County

This table reports results from cross-sectional regressions at the county-industry level. The dependent variable is the change in the share of the employment in firms of age 0-1 in county c and industry i between 2006 and 2000. $\Delta IT\ Exposure_b$ is the change in the IT adoption of banks in the county, measured by the change in IT adoption of banks historically present in the county (between 2006 and 2000), and standardized with mean zero and a standard deviation of one. $ext.fin.dep_i$ the dependence on external finance in an industry. Standard errors are clustered at the county level *** p<0.01, ** p<0.05, * p<0.1.

Table A6: **Black Entrepreneurship**

	(1)	(2)	(3)
VARIABLES	Share of startup employees who are black (minus share of white)		
IT exposure	-0.259***	-0.257***	-0.245***
	(0.098)	(0.094)	(0.094)
Observations	21,714	21,714	21,714
R-squared	0.001	0.013	0.047
County Controls	-	✓	✓
NAICS FE	-	-	✓
Cluster	County	County	County

The left hand side variable is defined as the difference between the minority young employment share and non-minority young employment share, where young employment share is the share of employees in young firms in a demographic group relative to total employees in demographic group in a county sector. Standard errors are clustered at the county level *** p<0.01, ** p<0.05, * p<0.1.

Table A7: **County-level robustness**

VARIABLES	(1) share 0-1	(2) share 0-1	(3) share 0-1	(4) share 0-1	(5) share 0-1	(6) share 0-1	(7) share 0-1 (lagged)	(8) Δ Employment	(9) share 0-1	(10) share 0-1	(11) share 0-1	(12) share 0-1	(13) share 0-1
IT exposure	0.378***	0.163**		0.398***	0.375***	0.333***	0.418***	0.054	0.809*	0.247***	0.349***	0.344***	0.405***
	(0.100)	(0.073)		(0.106)	(0.099)	(0.092)	(0.126)	(0.065)	(0.421)	(0.088)	(0.095)	(0.097)	(0.103)
IT exposure (deposit weighted)			0.342***										
			(0.094)										
Observations	25,779	25,779	25,779	21,735	25,544	25,779	25,440	25,774	2,105	21,150	25,519	24,900	18,652
R-squared	0.248	0.252	0.248	0.252	0.248	0.268	0.208	0.215	0.279	0.283	0.247	0.251	0.242
County Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
NAICS FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Spec	Baseline	No Weights	Deposit Share	No Finance	NoWyoming	State FE	Lagged Denominator	Δ Total Employment	Only Tradable	No High-VC States	No High-VC Counties	Coverage: control	No Low Coverage Counties

This table reports results for the following regression: $\text{startups}_{c,i} = \beta\ IT\ exposure_{c,99} + controls_{c,99} + \theta_c + \phi_i + \varepsilon_{c,i}$, where $\text{startups}_{c,i}$ is defined as the share of the employees in county c and industry t which is employed at a firm with at most 1 year of life. The share is then averaged across the years 2000 and 2007. IT_c is the IT adoption of banks in the county, measured by the IT adoption of banks historically present in the county, and standardized with mean zero and a standard deviation of one. The Table report results from a set of robustness exercises. (1) Is the baseline regression. Column (2): local IT adoption is the unweighted average of the IT adoption of banks present in the county. In Column (3) we project bank IT adoption by the deposit share rather than the number of branches on the county. In column (4) we exclude finance and education as a sector. In (5) We exclude Wyoming. (6) We include state FE. (7) We divide employment creation of young firms by lagged total employment in the county sector cell. In Column (8) we use the change in total employment as a dependent variable. Standard errors are clustered at the county level. In (9) we restrict our sample to firms in tradable industries. In (10) and (11) we exclude high venture capital states and counties, respectively. In column (12) we control for the coverage. In (13) we exclude low coverage counties. *** p<0.01, ** p<0.05, * p<0.1.

Table A8: **County IT exposure and entrepreneurship – BDS**

VARIABLES	(1) share E 0-1	(2) share E 0	(3) share F 0	(4) Δ F 0	(5) log(F 0)
IT exposure	0.463***	0.113***	0.183***	0.002*	0.036***
	(0.112)	(0.025)	(0.069)	(0.001)	(0.013)
Observations	1,771	1,771	1,771	1,771	1,771
R-squared	0.017	0.426	0.481	0.254	0.982
County Controls	✓	✓	✓	✓	✓
Cluster	County	County	County	County	County

This table reports results for regressions at the county level. The dependent variable is the employment share of firms of age 0-1 in column (1); the employment share of firms of age 0 in column (2); the share of firms of age 0 out of all firms in column (3); the change in the number of firms of age 0 in column (4); and the log of the number of firms of age 0 in column (5). *IT exposure* denotes the IT adoption of banks in the county, standardized to a mean of zero and a standard deviation of one. Standard errors are clustered at the county level. *** p<0.01, ** p<0.05, * p<0.1.

Table A9: **Startup quality – patenting (0–5 years)**

VARIABLES	(1) pat/emp	(2) pat/firm	(3) pat/emp (w)	(4) pat/firm (w)
IT exposure	-0.024	-0.186	0.021	0.234
	(0.091)	(0.491)	(0.102)	(0.663)
Observations	1,771	1,771	1,771	1,771
R-squared	0.071	0.128	0.088	0.136
County Controls	✓	✓	✓	✓
Cluster	County	County	County	County

This table reports results from cross-sectional regressions at the county-level. The dependent variable is the number of patents by firms of age 0–5 over employment by firms of age 0–5 (in thousands) or the number of firms of age 0–5 (in thousands) in county *c*. *IT Exposure* is the IT adoption of banks in the county, measured by the IT adoption of banks historically present in the county, and standardized with mean zero and a standard deviation of one. Standard errors are clustered at the county level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A10: **Startup quality – patenting (0 year)**

VARIABLES	(1) pat/emp	(2) pat/firm	(3) pat/emp (w)	(4) pat/firm (w)
IT exposure	-0.039	-0.454	-0.035	-0.232
	(0.225)	(0.853)	(0.135)	(0.600)
Observations	1,771	1,771	1,771	1,771
R-squared	0.053	0.097	0.119	0.162
County Controls	✓	✓	✓	✓
Cluster	County	County	County	County

This table reports results from cross-sectional regressions at the county-level. The dependent variable is the number of patents by firms of age 0 over employment by firms of age 0 (in thousands) or the number of firms of age 0 (in thousands) in county *c*. *IT Exposure* is the IT adoption of banks in the county, measured by the IT adoption of banks historically present in the county, and standardized with mean zero and a standard deviation of one. Standard errors are clustered at the county level. *** p<0.01, ** p<0.05, * p<0.1.

Table A11: **Aggregate bank lending and IT**

VARIABLES	(1) SB loans (sm)/assets	(2) SB loans/assets	(3) log(CI loans)	(4) CI loans/assets
IT	0.003**	0.010***	0.012	0.002
	(0.002)	(0.003)	(0.035)	(0.003)
Observations	542	542	542	542
R-squared	0.057	0.053	0.881	0.065
Bank Controls	✓	✓	✓	✓

This table reports results from cross-sectional regressions at the bank-level. The dependent variable are different measures of bank lending. *IT* is the IT adoption of bank *b*. Standard errors are clustered at the bank level. *** p<0.01, ** p<0.05, * p<0.1.

Table A12: **County IT exposure and entrepreneurship – processing time**

VARIABLES	(1) share 0-1	(2) share 0-1	(3) share 0-1
IT exposure	0.361***		0.380***
	(0.100)		(0.097)
processing time		-0.016***	-0.016***
		(0.004)	(0.004)
Observations	25,376	25,376	25,376
R-squared	0.248	0.249	0.251
County Controls	✓	✓	✓
NAICS FE	✓	✓	✓

This table reports results for regressions at the county-industry level (see Equation 9). The dependent variable is the employment share of firms of age 0-1. *IT Exposure* denotes the IT adoption of banks in the county, standardized to a mean of zero and a standard deviation of one. *processing time* measures the average processing time of SBA loans in each county. Standard errors are clustered at the county level. *** p<0.01, ** p<0.05, * p<0.1.