Search Frictions in Credit Markets

Clément Mazet-Sonilhac

SciencesPo DEPARTMENT OF ECONOMICS

Discussion Paper

SEARCH FRICTIONS IN CREDIT MARKETS

Clément Mazet-Sonilhac

SCIENCES PO ECONOMICS DISCUSSION PAPER

No. 2024-14

SEARCH FRICTIONS IN CREDIT MARKETS*

Clément Mazet-Sonilhac†
Bocconi University and IGIER

Using data on over 3.5 million firm-bank relationships, I document new facts suggesting that search frictions are pervasive in the market for small business loans. Motivated by this evidence, I develop a theory of firm-bank matching where firms incur costs to search for banking partners. The model predicts that a reduction in search frictions increases (i) inter-city credit flows and (ii) firm-bank distance. I confirm these predictions by estimating my model on French data, using the staggered roll-out of broadband internet as an instrument. Using the structurally estimated parameters, I find that internet access reduced the cost of debt for small businesses by 4.9%

Keywords: Search frictions, Broadband internet, Firm-bank matching.

*Winner of the USC Marshall School of Business Trefftzs Best PhD Paper Award at the WFA (2022) and WFA Brattle Group PhD Candidate Award for Outstanding Research (2022). I am grateful to Philippe Aghion, Antonin Bergeaud, Johannes Boehm, Cecilia Bustamante, Thomas Chaney, Federico Ciliberto, Anne Duquerroy, Michele Fioretti, Nicola Gennaioli, Stéphane Guibaud, Basile Grassi, Samuel Kortum, Mathias Lé, Nicola Limodio, Clément Malgouyres, Robert Marquez, Thierry Mayer, Isabelle Méjean, Christopher Palmer, Daniel Paravisini, Jonathan Parker, Rafael Repullo, Jean-Marc Robin, Xavier Ragot, Sandra Poncet, Edouard Schaal, Nicolas Serrano-Velarde, David Thesmar, Adrien Verdelhan, Xavier Vivès, and participants at MIT Sloan, Cemfi, Bocconi University, CREi, SciencesPo, GSIE, INSEAD/Collège de France, and Banque de France seminars for helpful comments.

†Email: clement.mazetsonilhac@unibocconi.it

1 Introduction

In corporate credit markets, banks do not post contract terms, and no central organization is responsible for setting prices or conducting transactions. In this decentralized environment, firms devote time and resources searching for the right banking partner.[1] An entrepreneur seeking bank credit, for example, must visit several bank branches to meet with loan officers, present her business plan, and undergo a lengthy application process before possibly securing a loan.[2]

While the effect of information asymmetries on credit allocation has received considerable academic attention (e.g., Akerlof, 1970; Stiglitz and Weiss, 1981; Petersen and Rajan, 1995; Udell, 2015), little is known about how search frictions affect firm-bank matching and firm access to credit. Yet, these two frictions interact and complement each other: search frictions imply costly search before the two parties meet, while information asymmetries shape credit allocation after they come into contact. Put simply, an environment of high search frictions prevents firms from meeting the loan officer best able to understand and finance their business and hence from being offered the best contractual credit terms. Understanding the specific role of search frictions is therefore particularly important for policymakers who aim to facilitate firm access to credit, not only since recent developments in information and communication technology (ICT) are likely to affect search frictions but also because policies that reduce search costs differ substantially from policies that target information asymmetries.

In this paper, I study the role of search frictions in the context of corporate credit markets for SMEs in France. I ask the following questions: How do search frictions shape firm-bank matching? Does a decrease in search frictions induced by ICT diffusion affects credit allocation and the cost of debt for small firms? To address these questions, this paper proceeds in four steps. First, I document novel empirical facts about local credit markets suggesting a key role for search frictions. Second, I build a model of the matching between firms and bank branches that combines search frictions and costly screening, and generates predictions about how such frictions affect the allocation of credit to SMEs. Third, I use rich bank-firm matched data over the period 1998-2005 to test and estimate the model structurally. In order to dig into the causal impact of search frictions, I propose a novel instrument variable strategy that exploits a natural experiment, the staggered diffusion of Broadband Internet in France, as a shock to search

[1] den Haan et al. (2003); Wasmer and Weil (2004); Duffie et al. (2005).

[2] According to recent surveys, small- and medium-size enterprises (SMEs) contact 2.7 financial institutions on average, submit 3 credit applications, and undergo a time-consuming application process: over 33 hours are spent, on average, on loan request paperwork (Federal Reserve Bank, 2014; Mills and McCarthy, 2014). Overall, about one-third of SMEs experience a difficult and lengthy credit application process (Infosys, 2018; Federal Reserve Bank, 2022). Honka et al. (2017) shows that retail consumers are, on average, aware of 6.8 banks and consider 2.5 banks when shopping for a new bank account.

frictions. Lastly, I use the structure of the model to quantify the impact of reducing search frictions on the cost of debt for small businesses.

In the first part of the paper, I leverage unique regulatory data that contain, for the universe of firm-bank matches in France, the identity and location of both the firm and the bank branch providing credit. I establish novel empirical facts suggesting that search frictions are pervasive in corporate credit markets. Because search frictions are inherently unobservable, I follow the empirical literature on consumer search (Gavazza and Lizzeri, 2021) focusing on trader heterogeneity, trading delays, and, more importantly price dispersion. I find that (i) bank branches are heterogeneous in size, the number of markets in which they operate, and the distance to their customers; (ii) the average time between the creation of a firm and its first bank match is 15 months, but is shorter in submarkets with more banking options, measured by the number of bank branches per square kilometer; and (iii) corporate loan markets exhibit a high level of price dispersion, for the same credit product and for similar firms, even across branches of the same bank. More than 30% of the variance in interest rates within branches remains unexplained after controlling for loan and firm characteristics.

Motivated by these findings, I develop a partial equilibrium model of firm-bank matching that features two-sided heterogeneity—bank branches and firms—and two information frictions. First, search frictions hinder firms' ability to locate and match with the right banking partner. Formally, I build on Lenoir et al. (2018), where the search process is random and depends on the level of search frictions between two submarkets.[3] Second, upon meeting firms, bank branches incur a cost to screen projects and learn about firms' quality. Screening costs are a function of branch productivity and the physical distance between a branch and a firm, which ultimately determines the interest rate quote.

This model serves three purposes. First, it highlights the mechanisms through which search frictions affect credit allocation. Second, by adding structure to the search process, it generates a number of testable predictions linking the variation in the level of search frictions to (i) the cost of debt for small businesses, (ii) credit flows between cities, and (iii) firm-bank matching. In particular, the model predicts a gravity structure for credit flows between cities. This equation states that bilateral credit flows between cities are proportional to city size and are inversely proportional to geographic distance and search costs. When search costs decrease, this gravity structure is distorted: firms meet with more potential lenders and eventually borrow credit at a lower rate. Third, the model allows me to perform a counterfactual analysis on the cost of debt for small firms.

[3]In this paper, search frictions encompasses both search and transactions costs. It represents the costly and time-consuming process of searching for available banking options, learning about the banks' characteristics and offers, contacting and meeting with loan officers, pitching the business project, preparing the applications files and exchanging information with the banker, and finally waiting for a decision and, eventually, for the funds.

Identifying the causal effect of search frictions on firm-bank matching and credit outcomes is difficult because of its endogeneity and also because the level of search costs is not observed. I address both challenges by using a natural experiment that exploits the staggered diffusion of broadband internet in France, from 1999 to 2007. In particular, I propose a novel IV strategy for the timing of broadband internet diffusion. Broadband internet operators relied on pre-existing infrastructures—local copper loops networks used for phone call transmission and large optic fiber cables used by railway and highway companies—to gradually deploy the the asymmetric digital subscriber line (ADSL) technology, with one objective: to connect the maximum number of customers while minimizing the cost of investment. Accordingly, I solve a simple maximization program to predict the optimal timing of broadband availability that exploits variations in population density interacted with distance to pre-existing infrastructures. For this purpose, I use a dataset on broadband internet availability at the city-level compiled by Malgouyres, Mayer and Mazet-Sonilhac (2021) and combine it with data on population density (measured ex-ante) and telecommunication infrastructures: i) exact location of local copper loops and ii) optic fiber cables installed along highways and railways. Thus, the interaction between pre-existing infrastructures with the density of population, controlling for firm density, predicts the timing of the arrival of broadband internet. In other words, I extract the part of the broadband internet shock that is orthogonal to local economic conditions. This setup rules out the possibility that the timing of technology diffusion is endogenous to bank branches' strategy, investment, and credit supply to small businesses (or, conversely, to small firms' demand for credit). It also provides a natural ground for an event-study identifying how ICT affects credit markets, through a reduction in search frictions.

Next, I structurally estimate the model and test its main predictions. Using my IV for the staggered diffusion of broadband internet in France, from 1999 to 2005, as a shock that reduced search frictions, I can estimate parameters θ and γ—respectively, the elasticity of credit flows to distance and search frictions—that govern how bilateral credit flows between cities and the cost of debt for small businesses react to a variation in the level search frictions. I first verify that credit flows between cities follow a gravity equation that is distorted by the staggered rollout of broadband internet. I provide causal evidence that this technology-induced reduction in search frictions triggers an average increase by 6% of the share of credit exchanged between pairs of interconnected cities. Consistent with the model's predictions, this effect varies dramatically with the initial level of search frictions. It is higher when two very distant cities are connected but is negative when two neighboring cities, already very closely tied economically, are connected by internet.

My main results rely on regressions carried out at the city level.[4] This geographical level simply matches the treatment level and allows me to use an estimation sample of a manageable size (24 million observations). However, some model predictions at a less aggregated level require leveraging bank branch-level data. I further document that broadband internet diffusion allows banks to match with new firms located in remote submarkets. Connected banks increase their share of credit lent to firms located outside their city by 10% and their share of remote clients by almost 12%. As a result, the average distance between a bank and its customers increases by 10% in the medium term after broadband internet access. These results are robust to several potential threats to identification. I find no evidence of pre-expansion differential trends in branch-level outcomes, and I show that adding city-level controls does not affect my estimates.

Finally, I assess the implications of my findings on the cost of debt for small firms, using the structure of my model. I plug my empirical estimates for θ and γ, as well as parameters calibrated from the data, into the equation linking search frictions to loan prices. Interpreted within my model, the estimates imply that the reduction in search frictions triggered by the diffusion of broadband internet lowered the cost of debt for small businesses by 4.9% on average. This reduction in the cost of debt displays an interesting spatial heterogeneity: it is stronger in rural areas and in medium-sized cities than in the largest French cities. Firms initially located far from bank branches, or with limited access to a variety of potential banking contacts, benefited most from the reduction in search frictions, as it enabled them to connect with new or more suitable banking partners. In this regard, the spread of broadband internet reduced spatial credit inequalities and fostered local bank competition in rural and isolated submarkets.

These results are subject to a main identification concern: my empirical strategy might not be able to separately identify search frictions from other first-order frictions (e.g., asymmetric information) that are also changing as a result of the technology shock. While I cannot definitively rule out potential violations of the exclusion restriction, I address this concern through multiple angles. First, my main identification strategy utilizes a gravity equation with bilateral fixed effects (city x time fixed effects, for both the city of the firm and the city of the bank). This approach effectively captures any productivity-enhancing impact of broadband internet on both firms and banks. If the introduction of broadband internet increases bank branch productivity - thereby improving their ability to screen and reducing asymmetric information - this effect would be absorbed by the city of the bank × time fixed effects. Conversely, firm-level changes in information availability are absorbed by the city of the firm × time fixed effects. Consequently, my analysis isolates the specific impact of a drop in bilateral search frictions resulting from broadband internet adoption. In other words, only if broadband internet affects screening in a city-pair-specific way would this identification strategy fail. Second, broadband internet is a

[4]In this paper, city refers to the French "Unité Urbaine", an administrative concept that empirically designates a territory defined by the continuity of its habitat or buildings. See Appendix B for more details

communication technology that facilitates faster exchanges of information. This distinguishes it from technologies designed to mitigate asymmetric information, such as screening algorithms (e.g., Machine Learning and alternative data) and risk assessment models, which have been adopted at a later stage. For the adoption of these screening technologies to confound my results, their implementation would need to be perfectly aligned with the timing of broadband internet adoption, which is highly unlikely. Finally, my empirical results point toward search frictions rather than asymmetric information. If the primary issue were asymmetric information, we would expect to see effects on the intensive margin of credit - specifically, the amount of money that incumbent firms can borrow. However, my results highlight changes at the extensive margin, as evidenced by the formation of new bank-firm relationships. At the bank branch level, I find no reduction in the average borrower probability of default or increase in total credit granted—outcomes that would be expected with reduced asymmetric information. Notwithstanding these arguments, I cannot exclude that some residual effects may still stem from reductions in asymmetric information.

Literature review. This paper contributes to the literature on corporate finance and search and matching. There is a vast theoretical and empirical literature in corporate finance on the role that informational frictions plays in hampering firms' access to credit, starting with seminal papers by Akerlof (1970) and Stiglitz and Weiss (1981).[5] This literature focuses on asymmetric information and has long highlighted the role of firm-bank relationships in alleviating agency frictions that shape credit supply (for surveys, see, e.g., Boot, 2000; Degryse et al., 2009; Udell, 2015 and Crawford et al., 2018 for recent related structural work). Yet, the role of search frictions, in particular in the formation of firm-bank relationships, has been overlooked despite its potentially equal importance.[6] I contribute to this literature by proposing a novel theory of firm-bank matching and SME access to credit that formally introduces search and contracting frictions. I first provide empirical evidence that emphasizes the importance of search and transactions costs in local credit markets. Then, I write a partial equilibrium model of firm-bank matching subject to both search and costly screening that I structurally take to the data. To the best of my knowledge, this paper is the first to account for search frictions in the formation of firm-bank relationships from a micro perspective and to provide causal evidence of how a reduction in search frictions affect credit allocation. This contribution has important policy implications, as an environment characterized by search frictions may not only amplify and propagate shocks (den Haan et al., 2003; Wasmer and Weil, 2004) but also generate slow recoveries (Boualam, 2018).

[5] See also Sharpe (1990); Rajan (1992); Berger and Udell (1995, 2002); Degryse and Ongena (2005); Agarwal and Hauswald (2010); Drexler and Schoar (2014); Nguyen (2019).

[6] Recently, Argyle et al. (2019) and Allen et al. (2019) show how search frictions affect mortgage and consumer credit markets. Closer to my approach, Agarwal et al. (2020) study the interaction of search and application approvals in mortgage credit markets and build a model combining search and screening. They conclude by calling for the study of other credit markets, such as the small business loan market.

My model delivers a gravity equation for inter-regional credit flows, which echoes the nascent literature in finance and macroeconomics that studies gravity equations for cross-border equity flows (Portes and Rey, 2005), bonds and bank holdings (Coeurdacier and Martin, 2009), and cross-border asset holdings (Okawa and van Wincoop, 2012). Brei and von Peter (2018) estimate gravity equations on international banking flows using a PPML approach and find a substantial role of distance for banking, even with immaterial transport cost, pointing to the role of information frictions. My paper complements and extends these findings by estimating a gravity equation for within-country bank credit flows with two types of informational asymmetries. It confirms the prominent role for distance, even for local credit markets, but also underlines the relevance of search frictions.

This paper also contributes to the vast literature in labor (for a survey, see, e.g., Rogerson et al., 2005) and trade (Chaney, 2014; Allen, 2014; Lenoir et al., 2018) that studies how search and matching frictions affect firms' ability to produce. This literature has long highlighted that it takes time and resources for a worker to land a job, for a firm to fill a vacancy, for an exporter to find customers abroad, or, symmetrically, for an importer to match with the right supplier remotely. I contribute to this literature by showing that search and matching frictions also affect firms' ability to raise external finance. While the theoretical approach developed in my paper is in the spirit of trade models, where firms undergo a random search process, it incorporates screening costs that replace traditional iceberg transport costs. Finally, some recent papers explore how the diffusion of ICTs—including broadband internet—affects such frictions (Allen, 2014; Lendle et al., 2016; Steinwender, 2018; Akerman et al., 2018; Malgouyres et al., 2021; Bhuller et al., 2019; D'Andrea and Limodio, 2019; Jiang et al., 2022; Bergeaud et al., 2024). My contribution to this literature is twofold. First, I structurally estimate the impact of a technology-induced reduction in search frictions on credit markets and credit allocation and provide a causal interpretation of recent structural changes observed in banking markets (Kroszner and Strahan, 1999; Petersen and Rajan, 2002; Vivès and Ye, 2021) among which the erosion of the dependence on geographical proximity between customers and banks. Second, I develop a novel IV for the timing of broadband internet diffusion. As the observed pace of the expansion of broadband internet may be endogenous, I show how to generate a connection timing that only depends on ex-ante city-level characteristics as population density and distance to infrastructure.[7]

[7] In a contemporaneous and independent paper, D'Andrea et al. (2022) study the impact of broadband internet diffusion on Italian banks. The authors adopt a reduced-form approach to document the global effect of broadband internet on both credit volume and price, driven by an increase in banks' productivity and competition. Their approach—focused on the bank side—is complementary to my paper, which uses broadband internet expansion as a shock to structurally identify the impact of a decrease in search frictions faced by SMEs, i.e., the firm side of credit markets.

The rest of the paper is organized as follows. Section 2 presents the data, and documents new facts about corporate credit markets. Section 3 presents the model and the main predictions that guide the empirical analysis. Section 4 presents the empirical context and details the IV strategy. Section 5 describes the empirical methodology, and Section 6 details the results. Section 7 studies the implications for the cost of debt through the model's lens, and Section 8 concludes.

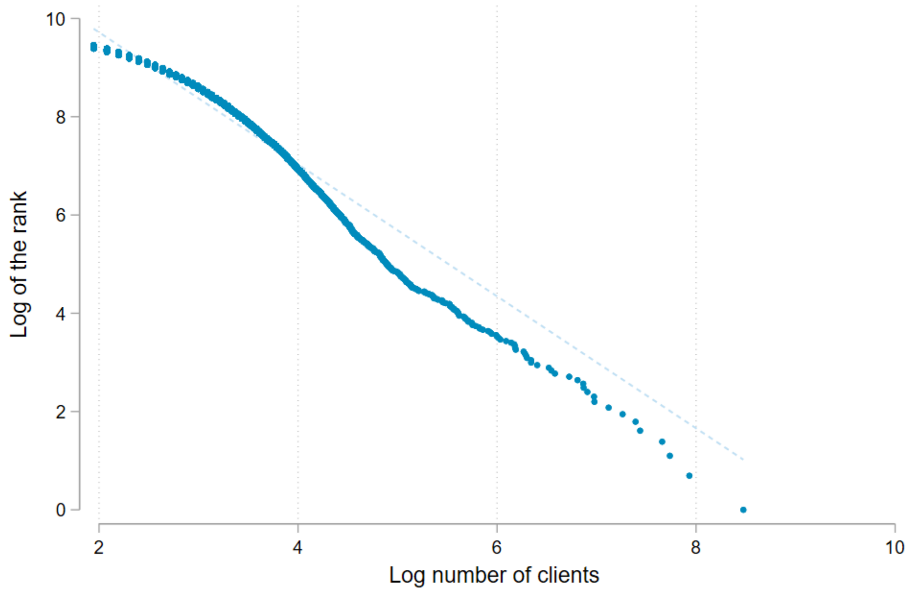
2 Data

In this section, I describe the data I use to study local credit markets in France and how they were affected by broadband internet diffusion over the 1998–2005 period. I combine multiple proprietary datasets from the Banque de France about firm-bank relationships, branch-level credit exposure, and interest rates of new loans, with unique data on broadband internet expansion at the city level. I establish novel empirical facts about search frictions in corporate credit markets. Because search frictions are inherently unobservable, I follow the vast empirical literature on consumer search (Gavazza and Lizzeri, 2021) and focus on trader heterogeneity, trading delays, and price dispersion.

2.1 The French Credit Register

The French Credit Register is a large dataset of firm-bank linkages available at the Banque de France over the 1998–2005 period, and it collects data on bank exposures to residents on a monthly basis to monitor and control systemic risk. The monthly data come from reports by credit institutions that are mandatory provided that their commitments or risk exposures on a company, as defined by a legal unit and referenced by a national identification number (SIREN), reach a total of EUR 75,000. I use a yearly version of the credit register for convenience as I already deal with an oversized dataset. Monthly reports encompass the funds made available or drawn credits, banks' commitments on credit lines and guarantee, and medium- and long-term leasing, factoring, and securitized loans at the branch level. Recipients are single businesses, corporations, and sole proprietorships engaged in professional activities, who may be registered in France or abroad. Reporting financial intermediaries include all resident credit institutions, investment firms, and other public institutions. In 2005, this raw dataset (which excludes individual entrepreneurs) covers information on more than 1.9 million firm-bank relationships, corresponding to more than 1.4 million unique firms or corporations (SIREN), 341 unique banks, and 15,925 bank branches serving firms. The smallest banks in my sample have only one branch. In contrast, the largest one (i.e., Crédit Lyonnais) owns more than 2,800 branches all across the country.

Figure 1: Bank Branch Size



Notes: This figure displays the distribution of bank branch size for the last quarter of 2005, for all banks with at least five clients. Formally, it shows the scatter plot of the log (size) against the log (rank). I compute the size of a branch as its number of clients. Bank branches are ranked by size, with #1 being the largest, #2 the second largest, and so on.

Bank branch heterogeneity. These data allow me to document novel facts on bank branch heterogeneity that guide my theoretical assumptions. Figure 1 displays the distribution of bank branch size for the last quarter of 2005, and shows that bank branches are very heterogeneous in size. The relationship between branch log size—measured as its total number of clients—and log rank is close to a straight line with a slope close to 1. This indicates that the distribution of the branch size follows a power law (i.e., Pareto distribution): very few large branches grant credit to many firms ($\geq$ 10,000 clients), while a vast majority of small offices only finance 10 to 20 clients. This fact is robust when branch size is measured as total credit (see Figure E1). In the Appendix E.1, I show that bank branches also differ markedly from each other with respect to (i) their average distance to clients and (ii) the number of markets in which they operate. I document a positive correlation between branch size and the average geographic distance between the branch and its clients. While small branches focus on a single very local market (branches with less than 10 clients are, on average, 15.5 km away from them), the largest branches grant credit to firms located in other cities or regions (branches with more than 100 clients are, on average, 80 km away from them). Overall, this vast bank branch heterogeneity is consistent with a high level of search frictions: an entrepreneur needs to identify the adequate banking partner as branch characteristics matter for credit allocation.

Trading delays. In decentralized corporate credit markets, new credit arrangement are not immediate as firms must devote time searching for the right banking partner. I document the time needed to secure a loan for new firms (i.e., the time to credit), from their creation to the first banking relationship established. To do so, I rely on an exhaustive dataset of firm creations (SIRENE[8]) from 2000 to 2005 for mono-establishment firms that I combine with information from the Credit Register about firms' first realized banking match. The number of firm creations in my sample range from 18,675 in 2000 to 26,865 in 2005.

Table 1: Trading delay

	Mean	P10	P25	Median	P75	P90	N
Time to credit (Number of months)	15.4	1.0	2.0	8.0	26.0	44.0	137,175

Notes: The table gives the distribution of the variable Time to credit, defined as the number of months elapsed between firm creation and the first match with a bank branch.

After their establishment, firms match with a single bank branch: the period of time running from firm entry to first banking match is 15 years on average. This average time to credit masks a substantial heterogeneity. A quarter of new firms finds their banking partner in less than 2 months. However, more than 25% of those new firms find their banking partner after two years of existence. As this delay in obtaining credit could also be due to factors other than search frictions, I further document that the time to credit is correlated with the density of the banking offer, which I use as a proxy for the cost of meeting several loan officers and making multiple applications. Figure A1 shows that the time to credit is shorter in submarkets with a greater number of banking options, measured by the number of bank branches per square kilometer. The substantial time elapsed between setting up a business and accessing bank credit — particularly for businesses located in cities with a low density of bank branches — echoes recent surveys highlighting that search frictions are high in such markets Mills and McCarthy (2014) requiring start-ups to submit multiple applications to different bankers before securing a loan (Federal Reserve Bank, 2014; Infosys, 2018).

2.2 Loan Rates (M-Contran)

The M-Contran database gathers quarterly surveys about all new loans granted to French firms from 2006 Q1 to 2016 Q4. This information is collected by the Banque de France to compute quarterly aggregate statistics on the interest rates of new loan contracts, with breakdowns by types of loans, borrowing sectors, and types of credit institutions. It also enables Banque de France to estimate and publish usury interest rates, an upward limit on lending rates set by French law. All main credit institutions report exhaustive information for all new individual

[8]The SIRENE data is available at this link

loans from their reporting branches granted during the first month of each quarter. The initial dataset reports, on average, about 100,000 new loans in each quarter. In addition to interest rates, the surveys provide rich information on a wide range of relevant loans characteristics, such as the loan size, the loan's precise purpose (investment, treasury, leasing etc.), its maturity at issuance, whether it is a fixed or adjustable rate, and whether it is secured or not.

Table 2: Explaining Price Dispersion

R^2	(1) Time + Bank FE	(2) (1) + Dep. FE	(3) (2) + Sector FE	(4) (3) + Branch FE	(5) (4) + Loan	(6) (5) + Firm
Equipment Loans	0.628	0.647	0.652	0.672	0.689	0.699
Credit Lines	0.559	0.574	0.579	0.606	0.651	0.657
Leasing	0.491	0.511	0.533	0.533	0.538	0.553

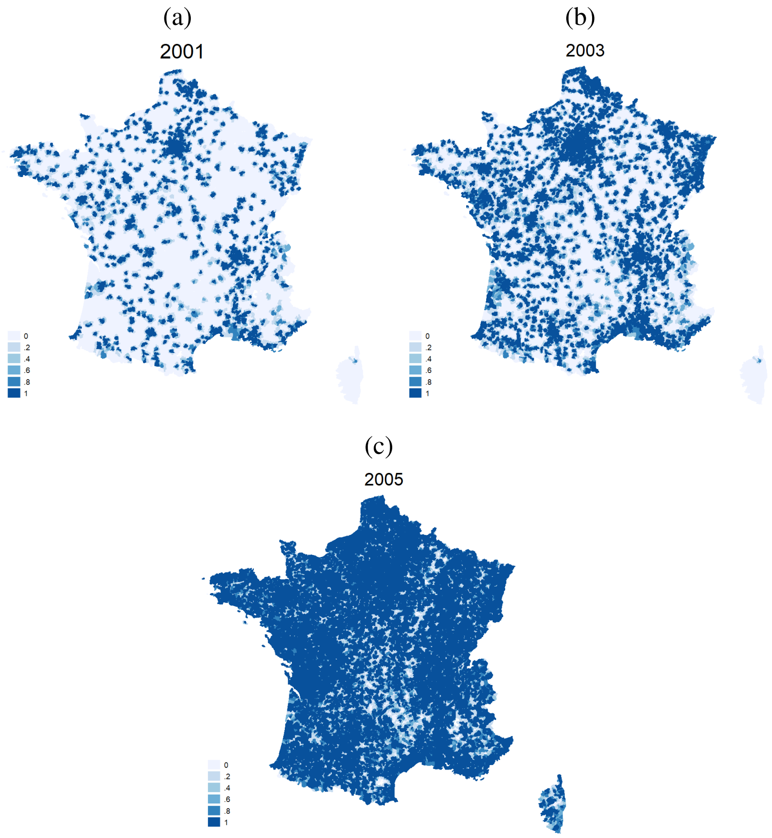
Notes: The table shows R^2 from OLS estimations of equation (30): Interest rate$_{ijtu}$ = Loan$_{ijtu} \cdot \rho_1$ + Firm$_{ijtu} \cdot \rho_2$ + FE$_{s(u)}$ + FE$_j$ + FE$_t$ + FE$_u$ + ϵ_{ijtu}. The dependent variable is the bank interest rate. Loan$_{ijtu}$ is a vector of loan characteristics (term in months, amount, type of rate: fixed or variable), and Firm$_{ijtu}$ is a vector of firm characteristics (age, size, debt, investment grade, turnover). Column (1) includes time and bank fixed effects. In columns (2) to (6), I sequentially add a county fixed effect (i.e., French "départements"), a sector (NACE Rev. 2 French classification) fixed effect, and a quarter fixed effect.

To examine the observed variation in loan prices, I estimate a price equation (see details in Appendix E.3) using these data. Table 2 reports the R^2 of each regression and suggests that credit rates exhibit a substantial dispersion within a time-bank branch-industry-department quadruplet, consistent with recent evidence on mortgage and consumer credit markets (see Argyle et al., 2019; Allen et al., 2019), and also suggests a high level of search frictions in the French credit market for SMEs. The results indicate that, at best, the model accounts for 70% of the observed variance in credit prices, letting more than 30% (40% for leasing) of the variance unexplained even when the model is saturated. Similarly, Cerqueiro et al. (2011) find substantial dispersion in loan rates for seemingly identical borrowers, using confidential Belgian data. The authors attribute this dispersion to information imperfections and asymmetries affecting credit markets and, among them, search costs.

2.3 Broadband Internet Data

I use the unique data of Malgouyres et al. (2021) on the date of upgrade to ADSL for each local exchange (LE) in mainland France. The historical operator (France Télécom, henceforth FT) made these data available to other operators as well as websites so that consumers could gauge the quality of their line for regulatory reasons. The authors also gathered data from the regulatory agency (ARCEP) regarding the geographic coverage of each LE.

Figure 2: Broadband internet rollout in France



Notes: This figure shows the rollout of broadband internet for all cities in mainland France for the years 2001, 2003, and 2003. The dark blue areas represent a large degree of coverage (Z_{ut} close to 1), while the light blue areas are cities with no internet connection (Z_{ut} equal to 0). These data contain the date of upgrade to ADSL for each local exchange in mainland France. The historical operator (France Télécom) had to make these data available to other operators as well as websites to allow consumers to gauge the quality of their line for regulatory reasons.

Combining both datasets, Malgouyres et al. (2021) construct a continuous measure of broadband access of municipality i at year t, denoted $Z_{it}^{\text{municipality level}}$, which is a time-weighted percentage of area covered in municipality i.[9] I construct an additional continuous measure of connectivity at the city level, Z_{ut}^{city} which is simply the (weighted) sum of $Z_{it}^{\text{municipality level}}$ for each $i \in u$:

$$Z_{ut}^{\text{city}} = \sum_{i \in u} w_i \cdot Z_{it}^{\text{municipality level}}, \tag{1}$$

where $i \in u$ denotes the municipalities included in city u and w_i is the weight of municipality i in the total population of city u. Equation 1 implies that Z_{ut}^{city} is continuous between 0 and 1;

[9]Formally, the measure used by Malgouyres et al. (2021) is written as:

$$Z_{it} = \sum_{b \in i} \frac{\text{\# days with access in } b \text{ since Jan 1st of } t}{\text{\#days in year } t} \times \frac{\text{area}_{bt}}{\sum_{b \in i} \text{area}_{bt}}$$

$\widetilde{Z}_{it}$ will be equal to 1 if all of its areas have had access for the entire year and equal to 0.5 if the entire municipality has had access to broadband over half the year t.

0 implies that no firm located in u enjoys an ADSL connection during year t, while 1 indicates that all firms benefit from it the entire year. Figure 2 shows the rollout of broadband internet for all cities in mainland France, from 1998 to 2005. The dark areas represent a large degree of coverage (a high Z_{ut}). In 2000, those areas are confined to the few major cities of France, surrounded by a large majority of no-ADSL territories. By 2003, the treatment has largely spread to lower-scale municipalities, although large parts of France remain dependent on the old technology. The national territory is essentially hole-free in 2005, when almost all LEs have been dealt with.

3 Model

Motivated by the empirical evidence that I document on search frictions in French credit markets, in this section I present a partial equilibrium model of firm-bank matching and inter-regional credit flows based that incorporates realistic geographic aspects. The model features two-sided heterogeneity—bank branches and firms—and information frictions of two kinds: informational asymmetries affect banks' ability to screen projects and search frictions hinder firms' ability to locate and match with the right financing partner, as in Eaton et al. (2022) and Lenoir et al. (2018).[10] The model captures the key empirical evidence presented in Section 2 and serves three purposes. First, it highlights the mechanisms through which search frictions affect the allocation of credit. Second, the model is tractable and yields testable predictions that I further bring to the data. Third, it allows me to perform counterfactual analysis on the cost of debt for small firms in Section 7.

3.1 Setup

There are a large number of local submarkets in the economy, indexed by $u = 1,...,N$, each inhabited by an exogenous mass of entrepreneurs (SMEs) and bank branches. In what follows, I use u to refer to the submarket in which a bank branch is located (the origin submarket) and v to refer to the submarket in which the branch customer is located (the destination submarket). In this economy, a single good is consumed by entrepreneurs and provided by bank branches into perfectly substitute varieties: bank credit.

Supply Side. There is a continuum of bank branches in each submarket u, of measure $N_u = S_u \cdot z_{min}^{-\theta}$, with S_u indicating the size of the submarket. Bank branches produce and provide credit

[10]This model is inspired by the recent trade literature that emphasizes the role of search frictions in the international goods market. Here, I model bank credit as a special kind of good that requires buyer-supplier search and matching and involves no traditional transportation cost.

with a single-factor constant returns-to-scale production function.[11] For the sake of simplicity, I make no distinction between short- and long-term loans and do not model more complex credit types as leasing or factoring. Bank branches operating in submarket u incur an exogenous input unit cost c_u that encompasses the branch office rent, loan officer wage, or marketing expenses, among others. The productivity of bank branch b_u located in submarket u is independently drawn from a Pareto distribution of parameter θ and support $[z_{min}, +\infty[$:

$$z_{b_u} \sim \text{Pareto}\ (z_{min}, \theta). \tag{2}$$

This Pareto assumption is data driven: Figure 1 shows that the relationship between branch log size—measured as its total number of clients—and log rank is close to a straight line with a slope close to 1. Thus, the number of bank branches located in submarket u that can provide credit with efficiency above z is written as $N_u(z) = S_u \cdot z^{-\theta}$. Bank branches located in u additionally incur a variable cost d_{uv} when lending to a firm located in a remote submarket v, which is a function of the physical distance between submarkets u and v. This variable cost encompasses the fact that branches located closer to borrowing firms enjoy a local comparative advantage, stemming from a better knowledge of the local economic environment and actors. Thus, branches' market power arises from this proximity to local borrowers and erodes over distance (Degryse and Ongena, 2005).

Demand Side. Each submarket u is populated with a continuum of ex-ante heterogeneous entrepreneurs (or SMEs) with an investment project I, of size normalized to one, and no cash. Entrepreneurs differ in their productivity z_e. The productivity of an entrepreneur e_u located in submarket u, is independently drawn from a Pareto distribution of parameter γ and support $[z^f_{min}, +\infty[$:

$$z_{e_u} \sim \text{Pareto}\ (z^f_{min}, \gamma), \tag{3}$$

such as the number of firms in a submarket written as $F_u = S_u \cdot {z^f_{min}}^{-\gamma}$. To start their investment project, entrepreneurs need to raise external finance from banks—which is the only source of external finance available for small firms. Because of search frictions, it is difficult to locate the right banking partner; as a result, an entrepreneur has to undergo costly search process.

Search and Matching. I build on Eaton et al. (2022), where matching between buyers and sellers is random, and search is biased geographically toward submarkets in which search frictions are low. Each entrepreneur meets with a discrete number of bank branches, some located in their own submarket and some located remotely. This search process is a reduced form for the

[11] Constant returns-to-scale technology for bank branches seems to be a reasonable assumption as it has been documented for large financial institutions (McAllister and McManus, 1993). However, increasing returns-to-scale appears to better fit the data for most US banks, a fact further documented in Wheelock and Wilson (2012). Using increasing rather than constant returns-to-scale does not affect the predictions of the model.

active search for banking partners: entrepreneurs need to gather information about bank branch characteristics, contact loan officers, and physically meet with them to get a price quote.

Formally, the discrete number of branches met in submarket u is drawn into the distribution $N_u = S_u \cdot z_{min}^{-\theta}$. This implies that the number of branches met with efficiency higher than z is drawn in $N_u(z)$. As a consequence, the set of potential lenders drawn by entrepreneur e_u is the random variable Θ_{e_u}, which is the sum of potential banking partners met in each of the N submarkets. Θ_{e_u} reflects the strength of search frictions affecting the submarket u; in a frictionless world, each entrepreneur from u would meet with all bank branches in the economy and, in turn, would end up applying for credit from their optimal banking partner (i.e., the first best match). In decentralized credit markets with search frictions hindering the number of meetings and price quotes, the first best match is not always feasible as an entrepreneur may never meet with the right loan officer.

In Eaton et al. (2022) and Lenoir et al. (2018), there is no firm heterogeneity and the likelihood to meet with a supplier from v is the same for all the firms in u. I instead assume that entrepreneur heterogeneity matters: it reflects the fact that more productive entrepreneurs incur lower search costs. I model the search process as independent draws in the distribution of bank branches; each bank branch b_u located in u has the probability $z_{e_v}\kappa_{uv}$ to be drawn by entrepreneur e_v located in v. z_{e_v} stands for the firm productivity (normalized such that $z_{e_v} \in [0, 1]$). κ_{uv} (also $\in [0, 1]$) is a pair-specific u-v inverse measure of the strength of search frictions. While I remain agnostic about the determinants of κ_{uv}—such as distance (home bias), network effects, etc.—this heterogeneity across submarkets suggests that entrepreneurs' search will be geographically biased toward submarkets with lower search frictions. Formally, $\mathbb{P}[b_u \in \Theta_{e_v}] = z_{e_v}\kappa_{uv}$ and $\Theta_{e_v}(u)$, the number of bank branches from u met by an entrepreneur from v, follows a binomial law such that

$$\text{Card}(\Theta_{e_v}(u)) = z_{e_v}\kappa_{uv} \cdot S_u z_{min}^{-\theta}. \tag{4}$$

Under the Poisson limit theorem, the binomial law of parameters $(z_{e_v}\kappa_{uv}, S_v z_{min}^{-\theta})$ can be approximated by a Poisson law of parameter $z_{e_v}\kappa_{uv} \cdot S_v z_{min}^{-\theta}$; this approximation is used in the rest of the analysis. This modeling has two major implications. First, more productive firms will, mechanically, meet with more bank branches, not only locally but also in distant submarkets. De facto, productive entrepreneurs are more likely to find a good match among Θ_{e_v}, while entrepreneurs with a low productivity may end up with only a few bad quotes.

Second, heterogeneity in κ_{uv} across submarkets implies that entrepreneurs' search will be biased geographically toward submarkets in which search frictions are low. An important feature of the search process is that bank branch heterogeneity does not affect the probability of meeting; in particular, there is no directed search toward the most productive branches. I argue that this is a reasonable assumption given that branch characteristics (specialization, growth rate, etc.) as well as loan officer background, preferences, and bargaining ability are difficult to assess from an outsider. This echoes Cerqueiro et al. (2011)'s notion of loan officer discretion in the loan rate setting process, especially for small and opaque businesses. Gathering information about how much a bank branch will be a good fit turns out to be costly and complex, so I assume that branch characteristics are ex-ante unobserved and do not affect the meeting probability.

Conditional on meeting with a loan officer in u, entrepreneur e_v pitches their investment project and gets a price quote. I assume a simplistic bank pricing strategy (bank branches always price at their marginal cost, as in a perfect competition framework) and a reduced-form cost function that captures costly screening through both branch and entrepreneur productivity, exogenous unit cost c_u, and transportation cost d_{uv}. The interest rate offered by bank branch b_u to lend to entrepreneur e_v writes $r_{b_u,e_v} = \frac{c_u d_{uv}}{z_{e_v} z_{b_u}}$.[12] The interest rate increases with d_{uv}, a function of branch-firm distance, and the unit cost of production c_u. In addition, the loan rates are negatively correlated to both branch and entrepreneur productivity.

This formulation accounts for costly screening, that play a key role in firm-bank matching models (Crawford et al., 2018). Importantly, there is no adverse selection: once the bank screens a project, it perfectly observe entrepreneurs' heterogeneous default risk z_{e_v}. The assumption of bank branches pricing at their marginal cost is strong, particularly in a context of credit markets subject to informational asymmetries. An alternative is to assume that bank branches compete à la Bertrand: the branch that offers the best contract terms does not price at its marginal cost but instead equals the marginal cost of the branch with the second best offer. Another alternative is to assume a Nash bargaining equilibrium in which the branch with the best offer and the entrepreneur share the surplus of the match. Under the assumption of inelastic demand, competition à la Bertrand and Nash bargaining do not affect the model predictions about firm-branch matching. Since firm-branch matching is the main outcome of my model, I keep the marginal cost pricing assumption for the sake of simplicity.

After meeting with all the loan officers in Θ_{e_v}, entrepreneur e_v decides to match with the one offering the best contract terms (i.e., lowest interest rate). The interest rate paid by entrepreneur e_v writes:

$$r_{e_v} = \underset{b_u \in \Theta_{e_v}}{\operatorname{argmin}} \{ \frac{c_u d_{uv}}{z_{b_u} z_{e_v}} \}. \tag{5}$$

The Poisson search process combined with the Pareto distribution of bank branch size allows an analytical formula for r_{e_v}. Eaton et al. (2022) demonstrates that the assumption of Poisson

[12] The bank branch b_u program writes: $\Pi_u = z_{match} Kr - K c_u d_{uv}$ with $z_{match} = f(z_{b_u}, z_{e_v})$. I assume a multiplicative reduced-form expression for $f(z_{b_u}, z_{e_v})$, which is common in the labor literature where the productivity of a worker-firm match is simply the product of both productivities.

draws into a Pareto distribution delivers a Weibull distribution for the minimum interest rate introduced in equation 5. Formally,

$$\mathbb{P}(r_{e_v} \leq r) = W_{e_v}(r) = 1 - \exp\left(- r^{\theta} z_{e_v}^{\theta+1} \sum_{u=1}^{N} S_u \cdot (c_u d_{uv})^{-\theta} \kappa_{uv} \right). \tag{6}$$

Conditional on r fixed effects, entrepreneurs in submarket v obtain, on average, a best offer if the level of competition is high (i.e., $\sum_{u=1}^{N} S_u \cdot (c_u d_{uv})^{-\theta}$ is large) due to the proximity with vast and crowded submarkets. In the same vein, the lower the search frictions faced by entrepreneurs in v (i.e., the greater κ_{uv}), the better the contract terms will be, on average. Finally, the entrepreneur's productivity directly impacts the likelihood to be matched with a branch that offers a low interest rate because more productive entrepreneurs draw a larger set Θ_{e_v} of potential lenders that somehow compensates for the adverse effect of search frictions. Note that a larger θ also alleviates the negative impact of search frictions but is mostly in favor of less productive entrepreneurs: indeed, less heterogeneity between bank branches advantages the entrepreneur with the least price quotes.

3.2 Predictions

In this section, I derive a number of theoretical predictions about i) the magnitude of aggregate credit flows between any two submarkets, ii) firm-branch matching, and iii) the number of clients by bank branch along the distribution of the branch's productivity. I then investigate how a shock on search frictions modifies those predictions.

3.2.1 Aggregate Credit Flows

Let Π_{uv} be the share of credit granted in submarket v by bank branches located in submarket u (over the total credit borrowed by firms in v). As all investment projects share the same size, normalized to one, the expected share of credit distributed in v by branches located in u is the sum over all entrepreneurs in v of a dummy variable equal to 1 if the entrepreneur is matched with a branch in u and 0 otherwise:

$$\Pi_{uv} = \frac{\sum_{e_v=1}^{F_v} \mathbb{I}\{\mathrm{M}(e_v) = u\}}{\sum_{e_v=1}^{F_v} 1}, \tag{7}$$

where $\mathrm{M}(e_v) = u$ indicates that entrepreneur e_v decides to match with a branch from u.

Lenoir et al. (2018) show that using the law of large numbers, Π_{uv} is equal to the expected value of $\mathbb{I}\{\mathrm{M}(e_v) = u\}$ across entrepreneurs in v, which is the probability that the best contract terms offered to any entrepreneur in v comes from branch u. Here, a crucial condition is that random variables $\mathbb{I}\{\mathrm{M}(e_v) = u\}$ are independent and identically distributed, which is straightforward if entrepreneurs are ex-ante identical. In my case, with entrepreneur heterogeneity, I show that this condition holds as the likelihood to ultimately match with submarket u does not depend on z_{e_v}. Thus, after some calculations, equation 7 rewrites as follows (see Appendix C):

$$\mathbb{P}\Big[\{\ \mathrm{M}(e_v) = u\}|r\Big] = \frac{z_{e_v}^{\theta+1}\kappa_{uv}S_u(c_u d_{uv})^{-\theta}}{\sum_{k=1}^{N} z_{e_v}^{\theta+1}\kappa_{kv}S_k(c_k d_{kv})^{-\theta}}. \tag{8}$$

Equation 8 indicates that the likelihood for an entrepreneur located in v to match with a bank branch from u does not vary along the distribution of productivity. When z_{e_v} increases, the number of branches drawn by entrepreneur e_v in u mechanically increases but so do the number of branches drawn in other competing submarkets ($\forall k \neq u$), which results in a constant probability of matching with u. Entrepreneur productivity only impacts the contract terms, not the destination of the match. Second, under the assumption that Pareto distributions of bank branch productivity share the same shape parameter θ across submarkets, $\mathbb{P}[\{\ \mathrm{M}(e_v) = u\}|r]$ is the same for each price quote r. Thus, the structural expression for the share of credit distributed in v by branches located in u is

$$\Pi_{uv} = \mathbb{E}_{ev}\Big[\{\mathbb{I}\ \mathrm{M}(e_v) = u\}\Big] = \frac{\kappa_{uv}S_u(c_u d_{uv})^{-\theta}}{\sum_{k=1}^{N}\kappa_{kv}S_k(c_k d_{kv})^{-\theta}}. \tag{9}$$

Two forces are at stake: i) the relative magnitude of search frictions between submarkets u and v with respect to the magnitude of search frictions affecting all the other potential submarkets $k \neq u$ and ii) the relative size and efficiency of the submarket u compared to submarkets' sizes and efficiencies in the rest of the economy. From, equation 9, I derive two predictions about aggregate credit flows and the impact of a shock on search frictions κ_{uv}.

Prediction 1: Gravity Equation for Bank Credit. As shown in Lenoir et al. (2018), a log-linearization of equation 9 delivers a gravity equation for the share of credit distributed in v by branches located in u:

$$\log \Pi_{uv} = \log \kappa_{uv} - \theta \cdot \log d_{uv} + \mathrm{FE}_v + \mathrm{FE}_u, \tag{10}$$

where FE_u stands for $\log S_u c_u^{-\theta}$ and FE_v equals $-\log \sum_{k=1}^{N}\kappa_{kv}S_k(c_k d_{kv})^{-\theta}$. Gravity equations are not common in the finance literature, with notable exceptions for cross-border equity flows (Portes and Rey, 2005), bonds, and bank holdings (Coeurdacier and Martin, 2009). Okawa and van Wincoop (2012) proposes a theoretical foundation of a gravity equation for cross-border asset holdings' gravity including financial frictions in the form of informational asymmetries about assets' future returns. To the best of my knowledge, this paper is the first to propose a structural gravity equation for within-country bank credit flows with two types of informational asymmetries.

Prediction 2: A Shock on Search Frictions. I investigate the effect of a reduction of bilateral search frictions—e.g., the development of broadband internet and online banking services—on aggregate credit flows. The first-order condition of equation 9 with respect to κ_{uv} leads to

$$\frac{\partial \ln \Pi_{uv}}{\partial \kappa_{uv}} = \underbrace{\frac{\partial \ln \kappa_{uv}}{\partial \kappa_{uv}}}_{(a)} + \Big[- \underbrace{\frac{\partial \ln \sum_{k=1}^{N} \kappa_{kv} S_k (c_k d_{kv})^{-\theta}}{\partial \kappa_{uv}}}_{(b)} \Big]. \tag{11}$$

Two opposite mechanisms are at stake. First, a reduction in search frictions has a direct and strictly positive effect (a) on bilateral credit flows. This connectivity effect reflects the fact that it becomes less costly for entrepreneurs located in v to gather information about bank branches and loan officers in u and to meet with them. Formally, the likelihood of meeting with a bank branch from u increases for each entrepreneur in v; i.e., there will be, on average, more bank branches from u in Θ_{e_v}.

Second, the competition effect (b) captures the increasing competition between bank branches, conditional on being met, induced by the higher number of potential banking partners. The expression (b) is rewritten as $\frac{S_u (c_u d_{uv})^{-\theta}}{\sum_{k=1}^{N} \kappa_{kv} S_k (c_k d_{kv})^{-\theta}} \geq 0$ such that the competition effect is negative and may compensate for the direct effect of connectivity. Equation 11 simplifies to

$$\frac{\partial \ln \Pi_{uv}}{\partial \kappa_{uv}} = \frac{1}{\kappa_{uv}} - \frac{S_u (c_u d_{uv})^{-\theta}}{\sum_{k=1}^{N} \kappa_{kv} S_k (c_k d_{kv})^{-\theta}} \geq 0. \tag{12}$$

Note that the effect of a reduction in search frictions κ_{uv} is heterogeneous across submarkets u. The larger or closer a submarket, the smaller the total impact. This captures the fact that large and nearby submarkets already benefit from a visibility advantage; entrepreneurs in v easily meet with bankers from those very accessible and visible submarkets, which, in turn, are able to offer attractive contract terms. In contrast, small and remote submarkets benefit more for a reduction in search frictions.

3.2.2 Branch-Entrepreneur Matching

In this section I investigate the matching process between an entrepreneur located in v and any bank branch located in the N submarkets. I derive predictions about (i) the number and the quality of entrepreneurs who ultimately match with a particular bank branch and (ii) the impact of a reduction in search frictions on the matching equilibrium. Both predictions can be tested using firm-branch data and confronted to motivating empirical work presented in Appendix E.

Prediction 3: Positive Assortative Matching. I consider a bank branch located in u and its likelihood $F_{b_u}(e_v)$ to lend to entrepreneur e_v located in v as a result of the search and matching process. $F_{b_u}(e_v)$ can be decomposed as the likelihood for entrepreneur e_v to draw and meet with b_u and the likelihood for b_u to be the lowest cost supplier. Formally,

$$\begin{aligned} F_{b_u}(e_v) &= \mathbb{P}\Big(b_u \in \Theta_{e_v}\Big) \times \mathbb{P}\Big(\underset{\Theta_{e_v}}{\operatorname{argmin}}\{\frac{c_u d_{uv}}{z_{b_u} z_{e_v}}\} = b_u\Big) \\ &= \mathbb{P}\Big(b_u \in \Theta_{e_v}\Big) \times \Big(1 - \mathbb{P}(r_{e_v} < r_{bu})\Big). \end{aligned} \tag{13}$$

From equation 6 for minimum price distribution, I have an analytical formula for $\mathbb{P}(r_{e_v} \leq r_{bu})$.[13] By definition of the random search process, $\mathbb{P}(b_u \in \Theta_{e_v}) = z_{e_v}\kappa_{uv}$. Thus, the model delivers the following expression for the likelihood that entrepreneur e_v ultimately decides to borrow from b_u, namely $F_{b_u}(e_v)$:

$$F_{b_u}(e_v) = z_{e_v}\kappa_{uv} \times \exp\Bigg(- (c_u d_{uv})^{\theta} z_{b_u}^{-\theta} z_{e_v} \sum_{u=1}^{N} S_u \cdot (c_u d_{uv})^{-\theta}\kappa_{uv}\Bigg). \tag{14}$$

The likelihood of a match between branch b_u and entrepreneur e_v strictly increases in z_{bu} and z_{e_v}. The branch productivity has a simple and direct effect via the attractiveness of the price quote, while the entrepreneur productivity has two distinct impacts: a direct positive effect on the meeting likelihood that appears in the first part of equation 14 and a negative but smaller effect—in the exponential term—that captures a competition effect (conditional on meeting, the likelihood of being the lowest cost bank branch decreases with the number of other branches met). This indicates a positive assortative matching between very productive branches and entrepreneurs. On the contrary, low-productivity bank branches are likely to match only with unproductive entrepreneurs who do not enjoy a large set of potential partners. This is consistent with the empirical evidence from the probit regression presented in Appendix E: SMEs are more likely to match with branches that ex-ante exhibit a high level of complementarity in term of distance, size or specialization.

Prediction 4: Firm-Branch Matching and Search Frictions. The first-order condition of equation 14 with respect to κ_{uv} indicates how the matching process is affected by a shock on bilateral search frictions between any two submarkets u and v.

$$\frac{\partial \ln F_{b_u}(e_v)}{\partial \kappa_{uv}} = \underbrace{\frac{\partial \ln (z_{e_v}\kappa_{uv})}{\partial \kappa_{uv}}}_{(a)} \underbrace{- (c_u d_{uv})^{\theta} z_{b_u}^{-\theta} z_{e_v} \sum_{k=1}^{N} S_k \cdot (c_k d_{uk})^{-\theta} \frac{\partial \kappa_{kv}}{\partial \kappa_{uv}}}_{(b)}.$$

[13] I make the usual approximation that $\mathbb{P}(r_{e_v} \leq r_{bu}) \approx \mathbb{P}(r_{e_v} < r_{bu})$.

Similar to Prediction 2, the impact of a reduction in search frictions is twofold, with a direct and positive connectivity effect (a) and an indirect and negative competition effect (b). The connectivity effect captures the enhanced visibility of the bank branch, and the competition effect reflects the fact that, conditional on being drawn, it becomes more difficult to offer the lowest price quote. Those two antagonistic effects are analogous to the visibility and competition channels described by Lenoir et al. (2018) in the context of international trade. Equation 15 offers a reduced-form expression that summarizes both effects:

$$\frac{\partial \ln F_{b_u}(e_v)}{\partial \ \kappa_{uv}} = \frac{1}{\kappa_{uv}} - \frac{S_u z_{e_v}}{z_{b_u}^{\theta}}. \tag{15}$$

The impact of search frictions varies along the distribution of branch productivity. High-productivity branches (high z_{b_u}) benefit more from the reduction in search costs as the direct impact dominates the competition effect: those large and efficient branches located in v now meet with much more entrepreneurs from v and, conditional on being met, enjoy a dominant position compared to smaller and less productive branches when it comes to offering attractive contact terms. Their likelihood to be the lowest price quote and, then to ultimately be chosen, increases. On the contrary, low-productivity branches that previously benefited from the low level of competition (i.e., few other branches drawn from u) are now exposed to a tougher rivalry.

3.3 Discussion

In the model, banks play a relatively passive role, contrary to what is usually assumed in the literature where search frictions are absent (Hauswald and Marquez, 2003, 2006). Closer to my paper, Chang et al. (2023) develops a matching model with heterogeneous banks (in terms of costs) and firms (in terms of risk), similar to the different bank and firm productivities in my model, but focuses on banks' strategy: banks can either lend to a firm or or opt for alternative investment opportunities. My approach omits these strategic choices, instead introducing search frictions and a geographical dimension. I argue that the assumption of passive banks both enhances model tractability and aligns well with the context of the market for small business loans. Small firms, as studied here, are rarely actively approached by bankers; indeed, surveys indicate that it is often small firms that bear the cost of search (Federal Reserve Bank, 2014). Additionally, part of bank competition may occur at a national level in France (advertising and marketing)[14] rather than locally, where bank branches serve as meeting points for firm credit demand and bank supply. However, the model could be enriched with a more sophisticated pricing strategy. A solution would be to assume that bank branches compete à la Bertrand: the branch that offers the best contract terms does not price at its marginal cost but instead equals

[14]In contrast to Honka et al. (2017), which highlights substantial variation in advertising intensity across U.S. states, France's banking sector is largely centralized, with advertising and marketing typically managed at the national level by a few large banks, rather than locally by individual branches.

the marginal cost of the branch with the second best offer. Under the assumption of inelastic demand, Bertrand competition would not alter the model's predictions regarding firm-branch matching (i.e., who matches with whom); instead, it would impact the predicted interest rate, which would then depend on the screening cost of the bank offering the second-best rate.

Another key aspect of my theoretical approach is that the bank branching strategy is not modeled: the number and location of bank branches is exogenous and stable over time. The latter is an important concern, as recent work has shown that digital disruption reduced the size of bank branch networks in the US (Jiang et al., 2022). In the French context, however, this assumption seems reasonable as the bulk of widespread branch closures and consolidations in France occurred post-2008, largely in response to the Great Financial Crisis and technological shifts that accelerated later. For the period covered here, branch closures were minimal, with the number of bank branches in my sample (13,000 in 2005) actually increasing slightly overtime. Second, while bank branch management is out of the scope of this paper, I take this into account in my empirical analysis, adding a dyadic measure of the supply of branches belonging to the same bank in each pair of cities.

4 Empirical Context

In this section, I provide a description of my empirical strategy. I exploit the staggered deployment of broadband internet in France between 1998 and 2007 as a large-scale quasi-natural experiment to study the impact of a reduction in search frictions on credit markets. I first describe the technological and institutional context of broadband internet diffusion in France. Then, I discuss the identification strategy and propose a new IV for broadband internet deployment.

4.1 The Diffusion of Broadband Internet in France

ADSL is a data communication technology that enables fast data transmission over copper telephone lines (much faster than what a conventional voiceband modem could provide). In ADSL technology, bandwidth and bit rate are said to be asymmetric, which means greater toward the customer premises (downstream) than the reverse (upstream). Eligibility for ADSL depends on the distance between the final customer (e.g., home or office) and the LE since the intensity and the quality of the analog signal decreases as it is routed over the copper lines. LEs are the telephone exchanges owned by the incumbent operator FT (later renamed Orange) into which subscribers' telephone lines end. As of 2008, there were about 17,000 LEs spread throughout the country. Initially dedicated to telephone networks, LEs are essential for internet users who subscribe to ADSL as they aggregate local traffic and then direct it via the so-called backbone (i.e., higher levels of the network) toward the world wide web. A key feature of ADSL technology is that one can supply high-speed internet by upgrading the LE while relying on the existing (copper) local loop to connect the premises of the final customers. The upgrading involves installing an equipment inside the LE called a DSLAM (digital subscriber line access multiplexer) that is required to recover the data transmitted via ADSL on the local copper loop and adapt it so that it can be transmitted to the higher levels of the network (which are typically relying on optical fiber). The upgrading of local LEs is the key source of variation that I use in my empirical analysis.

ADSL technology became popular during the 1990s, as many OECD countries were planning to expand ICT-related services. In the early 2000s in France, the deployment of the technology beyond the country's largest cities was slow, for multiple reasons. First, FT, the monopolistic telecom supplier at the time and still the main supplier today, was unsure as to whether it was going to be able to make the upgraded infrastructure available to new competitors with a positive markup or not. The uncertainty regarding the wholesale price FT was going to be able to charge made the firm reluctant to upgrade LEs beyond the largest cities (see Sénat, 2002, p. 232). This uncertainty was lifted after a series of decisions by the regulatory agency set the conditions of that wholesale market (Arcep, 2002).

Moreover, while FT had to invest massively in upgrading its LEs to ADSLs, it went through a debt crisis that ended with what was essentially a government bailout in 2002. One can find anecdotal evidence of the impatience of the French government in accounts of Parliamentary debate (at the Senate) regarding the excessively slow expansion of broadband internet (Tregouet, 2001) and the difficult cooperation between the French government (the ministry in charge of the industry) and FT. Under the impulse of the government—which increased its stake in the firm during its 2002 bailout—FT pledged in 2003 to cover 90% of the French (metropolitan) population by the end of 2005, i.e., all LEs with more than 1,000 lines, for a total investment of 600 million euros (750 million euros in 2018 prices) (Telecom, 2003).

Overall, the account of the broadband expansion in France over the period suggests that it was gradual due to uncertainty regarding the capacity of FT to undergo the investment until 2002. After 2002, with the strong impulse of the government, FT started covering more secondary areas with a focus on the overall number of lines per LE, with only limited attention paid to local economic potential. While accelerated, the coverage remained gradual due to operational limits on the part of FT and took about two more years than anticipated in 2003. Further evidence on the adoption of ICTs by the French banking sector, and on how high-speed internet reduced search frictions for SMEs is given in Appendix D.

4.2 Identification and IV

My identification strategy is based on the gradual diffusion of this new ADSL technology in different LEs over space and time. Because this staggered diffusion is likely to be endogenous to local economic outcomes, I propose a novel IV for the timing of broadband expansion that addresses this challenge.[15] My IV strategy is based on a theoretical optimal investment plan for infrastructure upgrading. As previously mentioned, ADSL technology combines local copper loops and a large optic fiber network. Thus, when FT connected a specific city with broadband internet, the total cost of the project was twofold: the cost of upgrading the LE and the connection cost between the LE and the global optical cable network, which depends on the physical distance between the LE and the closest optic fiber cable. On the other hand, the gain for the internet supplier to upgrade an LE was proportional to the number of newly connected inhabitants. My optimal theoretical investment plan opposes the connection costs to the connection gains for each city. The gains are the number of potential clients (measured before 1999) reached consequent to an LE upgrading, and I use the distance (in kilometers) to the closest optic fiber as a proxy of the connection costs.

A key feature of this IV strategy lies in the exogeneity of the distance between an LE and the closest optical fiber cable, as the optic fiber network construction was anterior to the ADSL expansion. The network has been built, in part, by other economic actors, before 1998 and for a completely different purpose. Indeed, highway firms and the French railway company (SNCF) installed optic fiber cables along the lines (respectively, the roadsides) for fast data transmission: surveillance videos, internal communications, etc. FT leased the existing infrastructure to those companies in order to expedite broadband expansion. Figure A2 displays the location of around 13,000 LEs, highways, and railroads that pre-existed before the beginning of broadband internet expansion in France. For each LE, I compute the shortest geodesic distance to a highway or a railroad and use this as a proxy of connection costs.

Formally, I predict the optimal connection rank $\hat{R}_i$ for the LE i, only taking into account two presumably exogenous measures of costs and gains. I use this optimal connection rank $\hat{R}_i$ in place of the observed rank R_i to predict the theoretical year of connection and thus $\hat{Z}_{it}$. As a consequence, the optimal connection rank is not polluted by concomitant or correlated economic shocks that may affect the connection timing and only depends on pre-existing and

[15]The question of what the criteria were for deciding to treat one LE before another has been studied by Malgouyres, Mayer and Mazet-Sonilhac (2021). Empirical evidence shows that the main determinant of broadband internet expansion was the city-level population density, with no role for levels or trends in the economic patterns of the city, and was slowed down by the sunk cost of upgrading the infrastructure, consistent with statistical analysis. In particular, the authors highlight the fact that broadband expansion occurring to maximize population coverage with no special consideration for economic potential is strongly supported by a statistical analysis of the determinants of broadband coverage that they carry out in their paper.

time-unvarying city characteristics:

$$R_i = \alpha + \beta_1 \cdot \text{Density}_{i,1998} + \beta_2 \cdot \text{Shortest Distance}_{i,1998} + \underbrace{\epsilon_i}_{\text{Endogenous part}}$$

$$\hat{R}_i = \hat{\alpha} + \hat{\beta}_1 \cdot \text{Density}_{i,1998} + \hat{\beta}_2 \cdot \text{Shortest Distance}_{i,1998}.$$

Figure A3 shows the rank correlation between optimal versus observed connection ranks. The combination of exogenous connection gains and costs have a strong predictive power, with an R^2 close to 0.70. Finally, the mapping between the predicted connection rank $\hat{R}_i$ and the connectivity variable $\hat{Z}_{it}$ follows the correspondence between R_i and Z_{it} observed in the data such that $\hat{Z}_{it}$ and Z_{it} also display a strong positive correlation. Finally, I define the degree of connection between two cities as $\mathbb{C}_{uvt} = Z_{vt} \times Z_{ut}$. $\mathbb{C}_{uvt}$ belongs to $[0, 1]$, and this measure captures the ability for firms located in u to locate and communicate with bank branches located in v, using the world wide web. My IV strategy delivers a similar measure of connectivity between cities named $\hat{\mathbb{C}}_{uvt} = \hat{Z}_{vt} \times \hat{Z}_{ut}$.

5 Empirical Approach

In this section, I describe my empirical approach. I test the three main predictions from the model: (i) inter-regional credit flows follow a gravity equation, (ii) the broadband internet rollout affects inter-regional flows through a reduction in search frictions, and (iii) the broadband internet rollout affects firm-branch matching.

5.1 Gravity Equation for Aggregate Credit Flows

I use data on bilateral credit exposure from the Credit Register, aggregated at the city level for all single-city SMEs in my sample, to compute bilateral credit shares Π_{uvt}. Π_{uvt} is the amount of credit granted by bank branches located in u to SMEs located in v, over the total credit stock of SMEs in v. Π_{uvt} lies in $[0, 1]$, and 0 indicates that none of the firms located in v borrow credit from a bank branch located in u at date t (i.e., the aggregate credit flow from v to u is null). On the contrary, Π_{uvt} equals 1 and implies that firms in v are fully financed by branches in u. In an economy without inter-regional exchanges, all Π_{uvt} would equal 0, except Π_{uut} as all firms would be financed by local bank branches. Alternatively, I use other measures of flows to distinguish between the extensive and the intensive margin: the total amount of credit granted to firms in v by branches in u, the average loan granted, the number of firms in v financed by u, or the share of firms financed by u.

Baseline Specification. A very broad literature in international trade studies the gravity equation and its estimation (see Head and Mayer, 2014 for an overview). Santos Silva and Tenreyro (2006) and Santos Silva and Tenreyro (2011) show that the PPML estimator is a promising workhorse for estimating gravity equations, particularly in the presence of many zeros. It is perfectly suited for estimating multiplicative models, without the log-linearization of the dependent variable.[16] I adopt this standard approach and rely on the new estimator for Poisson pseudo-regression models with multiple high-dimensional fixed effects developed by Correia et al. (2019).

Formally, I estimate equation 9 in its multiplicative form:

$$Y_{uvt} = \exp\left[\ln S_{vt} + \ln M_{ut} + \beta_1 \ln dist_{uv} + \beta_2 \ln X_{uvt}\right] + \epsilon_{uvt}, \tag{16}$$

where Y_{uvt} is the bilateral dependent variable (credit flows, shares, number of clients served, etc.) aggregated at the city level. u, v, and t are indices for origin (the city from which the bank branches operate), destination (the city in which the borrowing firms are located), and time. S_{vt} and M_{ut} are the origin city $\times$ year and the destination city $\times$ year fixed effects; these fixed effects ensure that the theoretical restrictions implied by structural gravity are satisfied. X_{uvt} is a vector of time-varying pair characteristics (e.g., trade of goods between cities u and v, dummy variable for belonging to the same department, region, etc.) that may affect firm-bank matching and financial decisions. Finally, I do not include a pair-specific fixed effects as my goal is to identify the coefficient β_1 associated with physical distance $dist_{uv}$ between u and v. Here, the physical distance captures both the monitoring costs and the search frictions: according to my model and consistent with the gravity estimation in international trade and finance, β_1 should be negative.

Impact of Broadband Internet Diffusion. In this subsection I study the impact of a reduction in search frictions on aggregate credit flows using the staggered diffusion of broadband internet, i.e., prediction 11 of the model. I impose the following functional form for $\log \kappa_{uvt} = \gamma\, \mathbb{C}_{uvt} + \varrho X_{uvt} + \varepsilon_{uvt}$. Inspired by the international trade literature that investigates the impact of time-varying trade policies as trade agreements, I consistently identify the effect of time-varying connectivity between cities using a dynamic PPML estimator with fixed effects in a difference-in-difference setting. Namely, I add to the baseline equation the time-varying variable of interest $\mathbb{I}_{uvt}$ that captures the pair-specific variation in online connectivity (the reduction in search frictions):

$$Y_{uvt} = \exp\left[\ln S_{ut} + \ln M_{vt} + \beta_1 \ln dist_{uv} + \beta_2 \ln X_{uvt} + \beta_3\, \mathbb{C}_{uvt}\right] + \epsilon_{uvt}. \tag{17}$$

[16] The PPML estimator identifies the coefficients using the same first-order conditions that are used by the ML estimator derived from the Poisson distribution. However, it does not require the dependent variable to be Poisson distributed (Fally, 2015).

The goal is to consistently estimate the average effect of $\mathbb{C}_{uvt}$, a continuous variable indicating the degree of internet connectivity between u and v, using a structural gravity specification derived from the model. The origin × year and destination × year fixed effects—S_{vt} and M_{ut}—are crucial as they absorb all the time-varying impacts of broadband internet, which are not pair specific.[17] Another widely used specification includes dyadic fixed effects, namely origin-destination fixed effects, to absorb all time-invariant pair characteristics that may be correlated with the likelihood of being mutually connected. I propose an alternative version of specification 17 including pair-specific fixed effects: a direct consequence is that I cannot estimate the coefficient β_1 relating to the physical distance $dist_{uv}$ in this specification.[18]

5.2 Firm-Branch Matching

I then test how broadband internet diffusion affects the likelihood of firm-branch matching and the number of remote firms from u financed by a bank branch located in v. For this purpose, I leverage bank branch-level data to verify Prediction 4 in a dynamic event-study approach, similar to Malgouyres et al. (2021). I estimate a dynamic specification where I allow the effect on a branch b located in city u at year t to vary with time-to-treatment.

With the level of observation being a branch b located in city u, I can discretize the treatment status by setting the treatment status to 1 after the city experienced its highest increase in the predicted treatment variable $\hat{Z}_{ut}^{\text{city level}}$. Formally, I define the year of treatment as $t_{u0} = \text{argmax}_t\{\Delta\hat{Z}_{ut}^{\text{city level}}\}$ and the discretized treatment status as $\mathbb{1}\{t \geq t_{u0}\}$. The year of treatment for each city is denoted as t_{u0}. I index time-to-treatment with d (negative before treatment, positive after). The sample covers the years 1998 to 2005, and I restrict the set of observations where $d \in \{-6, -5, \ldots, +4, +5\}$. The main estimating equation is as follows:

$$Y_{b(u)t} = \sum_{\substack{d=-5 \\ d\neq -1}}^{d=5} \beta_d \times \mathbb{1}\{t = d + t_{0u}\} + \mathbf{x}'_{ut}\delta + \alpha_{b(u)} + \psi_{r(u),t} + \varepsilon_{ut}, \tag{18}$$

where $\alpha_{b(u)}$ and $\psi_{r(u),t}$ are fixed effects for branch b located in city u and for the department (of the city)-year, and $\mathbf{x}'_{ut}$ is a vector of time-dependent city-level covariates. I drop two indicator

[17] Broadband internet not only affects search frictions but might also impact city sizes, exogenous production costs or branch, and firm productivity. In this setting, that does not bias the β_3 point estimate because of the urban-unit-time fixed effects.

[18] A key aspect of the empirical strategy is based on the performance of the PPML estimator with multiple high-dimensional fixed effects. Santos Silva and Tenreyro (2011) show that the PPML estimator is well behaved (and outperforms the OLS) when the dependent variable displays a large proportion of zeros, using a Monte-Carlo approach. In Appendix F, I extend and adapt their simulation exercise to the exact case of my empirical setting. I present simulation evidence on the performance of the PPML estimator when the panel data are generated by a constant elasticity model, with (i) a large proportion of zeros, (ii) with a time-varying shock, and (iii) when all units are not simultaneously treated. My results confirm and extend the findings of Santos Silva and Tenreyro (2006) and Santos Silva and Tenreyro (2011), showing that both the PPML and the gamma pseudo-maximum likelihood (GPML) estimators are well behaved in the cases considered.

variables for $d = -5$ and $d = -1$. This restriction is necessary to avoid multi-collinearity and to identify the fully dynamic underlying data-generating process in the staggered design (Borusyak and Jaravel, 2017; Gross et al., 2018). To ensure that this restriction is not influential in the results, I display the results with alternative normalizations in the robustness section. The specification presented in equation 18 includes leads and lags, and including leads allows me to assess the presence of pre-trends. I also estimate a simpler "semi-dynamic" specification where only the lags of the treatment are included, as presented in equation 19):

$$Y_{b(u)t} = \sum_{d=0}^{d=5} \beta_d \times \mathbb{1}\{t = d + t_{0u}\} + \mathbf{x}'_{ut}\delta + \alpha_v + \psi_{r(u),t} + \varepsilon_{ut}. \tag{19}$$

The event-study coefficients $\widehat{\beta_d}$ estimated from equation 18 can be interpreted causally under the identifying assumption that, conditional on receiving broadband internet over the period considered and conditional on bank branch and city fixed effects, the timing of broadband internet rollout is unrelated to the outcome. The presence of systematic local factors that would drive both broadband and credit would be cause for concern. This potential issue is investigated by assessing the sensitivity of the coefficients to including a large set of controls and fixed effects meant to account for city and local labor market shocks. Finally, the outcome variable $Y_{b(u)t}$ measures several aspects of the branch-lending activity: i) the average distance to clients, ii) the share of remote clients (located outside the branch's city), and iii) the share of credit granted remotely (i.e., to remote clients).

6 Results

In this section I first show that inter-regional credit flows follow a gravity equation by estimating equation (16), using a PPML approach. I then estimate equation (17) and document an increase in inter-regional credit flows triggered by the staggered rollout of broadband internet, associated with a reduction in search frictions. Finally, I estimate equation (18) to document that broadband internet diffusion allows banks to match with new firms located remotely. I show the results for different specifications and assess the robustness of the results.

Gravity Equation for Aggregate Trade Flows. Table 3 shows the results for estimating equation (16), using a PPML approach described in Section 5.1. The dependent variable is the credit share Π_{uvt}, defined as the amount of credit granted by bank branches located in u to SMEs located in v, over the total credit stock of SMEs in v. Column (1) displays the results with no control, origin city × year fixed effects, and destination city × year fixed effects. In columns (2) to (5), I sequentially add pair-level controls: a dummy variable equal to 1 if both the firm and the bank are located in the same region (*région*), in the same county (*départment*), and the lagged log of trade flows between counties.

The coefficient of interest β_1, associated with the log of the distance between city u and v, is negative and close to –2. This magnitude implies that firms borrow credit from banks located in a nearby city nearly four times more than from similar banks located at twice the distance. The distance coefficient decreases but remains close to –2 when I consider banks and firms located within the same region or county or when I control by lagged trade flows in column (4), which is the baseline specification. I cannot compare this magnitude with results found in the

Table 3: Gravity Equation for Inter-Regional Credit Flows

	Share of credit in v borrowed from u: Π_{uvt}				
	(1)	(2)	(3)	(4)	(5)
Log distance $dist_{uv}$	-2.197***	-1.966***	-1.829***	-1.765***	-1.741***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Same region		1.482***			
		(0.018)			
Same county			1.722***		0.803***
			(0.016)		(0.023)
log (Trade Flows)				0.473***	0.304***
				(0.004)	(0.006)
Origin (u) × Year FE	Yes	Yes	Yes	Yes	Yes
Destination (v) × Year FE	Yes	Yes	Yes	Yes	Yes
Origin-Destination FE	No	No	No	No	No
R^2	0.66	0.66	0.67	0.66	0.67
Observations	24,545,811	24,545,811	24,545,811	24,545,640	24,545,640

Notes: The table shows the PPML estimation of equation (16). $dist_{uv}$ = bilateral distance. Same region is a dummy variable equal to 1 if both the firm and the bank are located in the same region. Same county is a dummy variable equal to 1 if both the firm and the bank are located in the same county. Log trade flows is the lagged log of bilateral trade flows between counties. Columns (1) to (5) include fixed effects for origin × years and destination × years. The sample period is 1997–2005, and the sample consists of all origin-destination-year combinations where at least one firm is located with positive credit. Robust standard errors are in in parentheses.

literature since, to the best of my knowledge, this paper is the first to estimate gravity equations for inter-regional credit flows. Portes and Rey (2005) document the negative impact of distance for international cross-border equity flows and find coefficients twice as low, between –0.529 and –0.881. Brei and von Peter (2018) run a similar estimation on international banking flows and adopt a PPML approach. They find an estimate close to but smaller than one. A coefficient of magnitude –2 is also larger than what is documented by the vast trade literature, comprising more than 2,500 estimates of the distance effect (Head and Mayer, 2014). Therefore, my results indicate that distance-related frictions are likely to be more important at a very local level than for international credit flows. This confirms the local nature of credit markets for SMEs and echoes the first-order importance of distance as a determinant of access to credit, well documented in the banking literature (e.g., Petersen and Rajan, 1995; Agarwal and Hauswald, 2010).

For all $u - v$ pairs with no bilateral credit flow, the share of credit is set to zero, which represents the vast majority of the observations. For robustness, I run a similar estimation but only keep pairs of cities with positive credit flows. Table A1 shows the results. The coefficient of interest is still negative and significant but is of a smaller magnitude: β_1 is now close to -1 and remains stable when adding the control variables.

Impact of a Reduction in Search Frictions. I now test how the gravity equation for inter-regional credit flows is distorted by a technology-induced reduction in search frictions. I formally test the prediction (11) of my model by running a PPML estimation of the augmented gravity equation (17). The main results are presented in Table 4, where the first column reproduces column (4) of Table 3 for comparison purposes. In column (2), I include the continuous variable $\mathbb{C}_{uvt}$ for observed internet interconnection between two cities. Similar to Table 3, to estimate an effect on distance, origin and destination × year fixed effects are included but not pair-specific fixed effects.

The estimate is positive and significant, in line with the model prediction, and it suggests that a technology-induced reduction in search frictions distorts the gravity equation for credit flows. The amount of credit exchanged between firms and banks located in connected cities increases relative to other non-connected banking partners. The magnitude of the estimate implies that the share of credit granted to firms located in v from banks located in a remote city u increases by 44% on average when u and v are connected. The estimate for distance is not affected by the inclusion of the ADSL variable nor the control for bilateral trade flows.

My model predicts a heterogeneous effect of broadband internet with respect to distance, which is the result of two opposing forces: a connectivity effect (positive impact) and a competition effect (negative impact). While the former dominates overall, the competition effect could prevail for markets that are geographically close, as banks located in these close markets already benefit from a visibility advantage. In contrast, remote submarkets benefit more from a reduction in search frictions. I formally test this prediction by including the interaction variable for the bilateral distance (expressed as a deviation from the sample average) and the treatment variable in column (3). I find a positive and statistically significant effect of the interaction variable with distance, which means that the elasticity of credit flows with respect to distance decreases in magnitude with broadband internet. In other words, the positive impact of a reduction in search frictions on credit flows is higher when two very distant cities are connected, which verifies the intuition of the model. On the contrary, the effect is almost null or even negative when two neighboring cities, already economically very closely tied, are interconnected by the internet.

Table 4: TECHNOLOGY-INDUCED REDUCTION IN SEARCH FRICTIONS

	Share of credit in v borrowed from u: Π_{uvt}				
	(1)	(2)	(3)	(4)	(5)
Log distance $dist_{uv}$	-1.765*** (0.004)	-1.760*** (0.004)	-1.844*** (0.006)	-1.764*** (0.004)	-1.865*** (0.006)
$\mathbb{C}_{uvt}$		0.370*** (0.058)	0.800*** (0.059)		
Log distance $dist_{uv}$ x $\mathbb{C}_{uvt}$			0.230*** (0.008)		
$\hat{\mathbb{C}}_{uvt}$				0.058 (0.059)	0.399*** (0.062)
Log distance $dist_{uv}$ x $\hat{\mathbb{C}}_{uvt}$					0.199*** (0.008)
log (Trade Flows)	0.473*** (0.004)	0.475*** (0.004)	0.470*** (0.004)	0.474*** (0.004)	0.468*** (0.004)
Origin (u) × Year FE	Yes	Yes	Yes	Yes	Yes
Destination (v) × Year FE	Yes	Yes	Yes	Yes	Yes
Origin-Destination FE	No	No	No	No	No
R^2	0.66	0.67	0.68	0.67	0.68
Observations	24,545,640	24,545,640	24,545,640	24,545,640	24,545,640

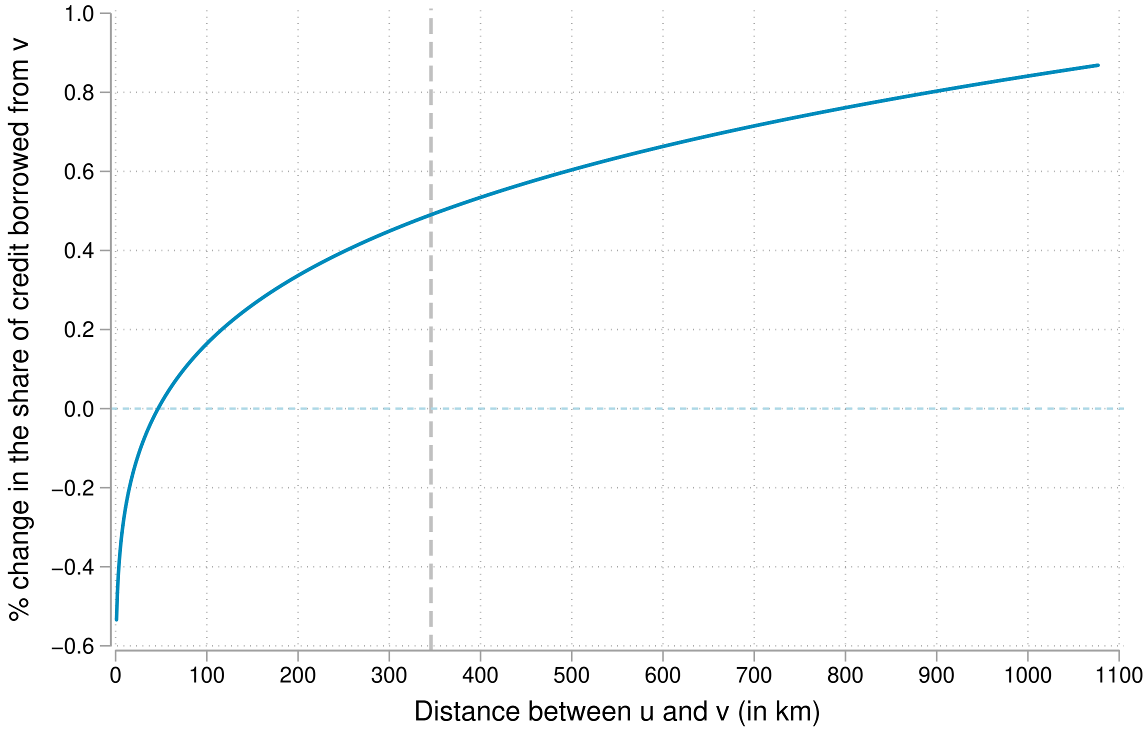
Notes: The table shows the PPML estimation of equation (16). $dist_{uv}$ = bilateral distance. $\mathbb{C}_{uvt} = Z_{vt} \times Z_{ut}$ is a continuous variable that indicates the degree of broadband internet inter-connectivity between two cities. $\mathbb{C}_{uvt}$ belongs to $[0, 1]$. This measure captures the ability for firms located in u to locate and communicate with bank branches located in v, using the world wide web. My IV strategy delivers a similar measure of connectivity between cities named $\hat{\mathbb{C}}_{uvt} = \hat{Z}_{vt} \times \hat{Z}_{ut}$. Log trade flows is the lagged log of bilateral trade flows between counties. Columns (1) to (5) include fixed effects for origin × years and destination × years. The sample period is 1997–2005, and the sample consists of all origin-destination-year combinations where at least one firm is located with positive credit. Robust standard errors are in in parentheses.

An important concern with these results is that the potential endogeneity of internet take-up biases these estimates. To address this concern, in columns (4) and (5), I estimate equation (17) using the IV for broadband internet interconnection, namely $\hat{\mathbb{C}}_{uvt}$, instead of the observed connection variable used in columns (2) and (3). The sign of the coefficient of interest remains unchanged, although the magnitude of the effect declines. The overall effect in column (4) is imprecisely estimated and implies that the share of credit granted to firms located in v from banks located in a remote city u now increases by only 6% when u and v are connected.

The results in column (5) documenting the heterogeneity of the effect with respect to distance are very comparable to those in column (3) that do not use the IV. In particular, the interaction term is positive and significant. In economic terms, these results mean that an increase in internet availability of 10 percentage points increases credit flows for a city at the 25th distance percentile by 24% less than for a city at the 75th distance percentile. Figure 3 illustrates this heterogeneity. While the effect of being interconnected is negative when cities are geographically nearby (the competition effect dominates), it increases sharply and because positive after the 50th kilometer.

After the 100th kilometer, the curve's slope is much flatter. Appendix Figure A6 maps the heterogeneous effect of being connected to Paris, showing that Marseille (second biggest French city) benefit more than Lyon (third biggest city). I then test for robustness along different

Figure 3: Heterogeneity of the effect with respect to distance



Notes: This figure plots the marginal effect of broadband internet with respect to the distance between two interconnected cities u and v, as estimated in equation (17). The x-axis represents the distance in kilometers, and the y-axis shows the total effect of broadband internet connection on the share of credit borrowed from v by firms located in u, in percentages. The vertical gray dashed line represents the average distance between two cities in my sample and correspond to the overall effect estimated in column (4) of Table 4.

dimensions. I first estimate the augmented gravity equation (11) by adding bilateral pair fixed effects that control for any unobserved characteristics of the city pair that are constant over time (see Head and Mayer, 2014). Table A2, columns (2) and (4) show the results. Although less significant and of lower magnitude, they are virtually unchanged and confirm the positive but heterogeneous effect of the technology-induced reduction in search frictions. To account for the dynamic nature of credit flows, I then replicate the baseline results by adding the lag dependent variable to the regressors in Table A3. This implies that credit relationships existing in the previous year provide a basis for the credit flows observed in the current year.[19] The main findings are confirmed by this alternative specification. Finally, I estimate models (16) and (17) on the extensive margin of credit flows. I use the share of relationships between banks

[19]However, given that the time dimension is much lower than N, the estimates are likely to suffer from the Nickell bias in the dynamic model.

located in u and firms located in v divided by the number of banking relationships in v rather than the share of credit used above. The results, presented in Tables A4 and A5, are consistent with previous my findings. This indicates that the margin of adjustment is mainly the extensive margin (creation of new firm-bank relationships) rather than the intensive margin (increase in the average loan), consistent with the model predictions.

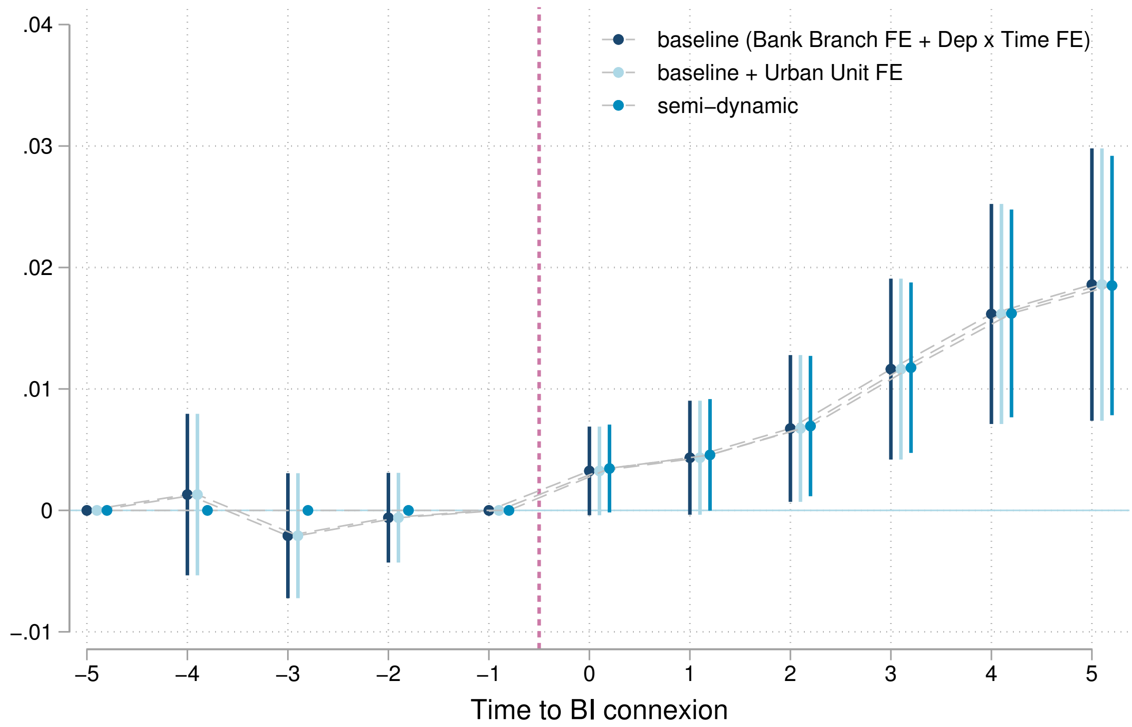
Firm-Bank Matching. The model predicts that broadband internet diffusion affects the firm-branch matching process: the reduction in search frictions allows very distant firms and banks to match, and in turn the share of remote firms (located outside u) financed by a bank located in u increases. I formally test prediction (15) by leveraging micro bank branch-level data in the dynamic event-study setting described above in Section 5.2. The main variable I consider is the share of credit given by bank branches located in a given city to firms located outside that city. With a high level of search frictions, matches between firms and banks occur only locally and the share of remote credit is close to zero. On the contrary, if searching is free, firms can meet with and borrow from banks located far away and the share of credit lent remotely is large.

Figure 4 displays the results, plotting estimated coefficients from equation (18). The dark blue dots report the results from the baseline specification including bank branch and county-year fixed effects. The estimates exhibit a flat trend before the event (i.e., the normalizing measure of time since access $d = -1$) and a break in the trend after. The coefficient for $d = 5$ in that specification is 0.018, suggesting that the expansion of access to broadband internet increases the bank branch-level share of credit lent to distant firms by about 10%, five years after the period of the largest expansion. Not only is this in line with the model predictions, but it is also consistent with the aggregate impact on credit flows documented above.

My second specification adds city fixed effects to the regression and aim to control for time-invariant city characteristics. The light blue dots in Figure 4 show the estimates. Here again, I find no sign of a pre-trend before broadband expansion contrasting with a steady growth afterwards. The estimated effect after five years is virtually the same as in the baseline case. Finally, the last set of coefficients, plotted in royal blue, represent a semi-dynamic version of the baseline specification (see equation 19). The regression should in theory be more efficiently estimated—as the number of parameters to be estimated is lower—however, both the estimates and the standard errors stay very stable and close to the fully dynamic specification in practice.

I then explore the possible mechanisms behind the bank-level increase in the share of credit lent to remote clients documented above. I study how the extensive margin of credit, i.e, the number of banking relationships, is affected by the technology-induced reduction in search frictions (as opposed to the average credit, the intensive margin). For this purpose, the dependent variable I consider is the share of remote firms financed by bank branches located in a given city, which is

Figure 4: Share of credit to remote firms



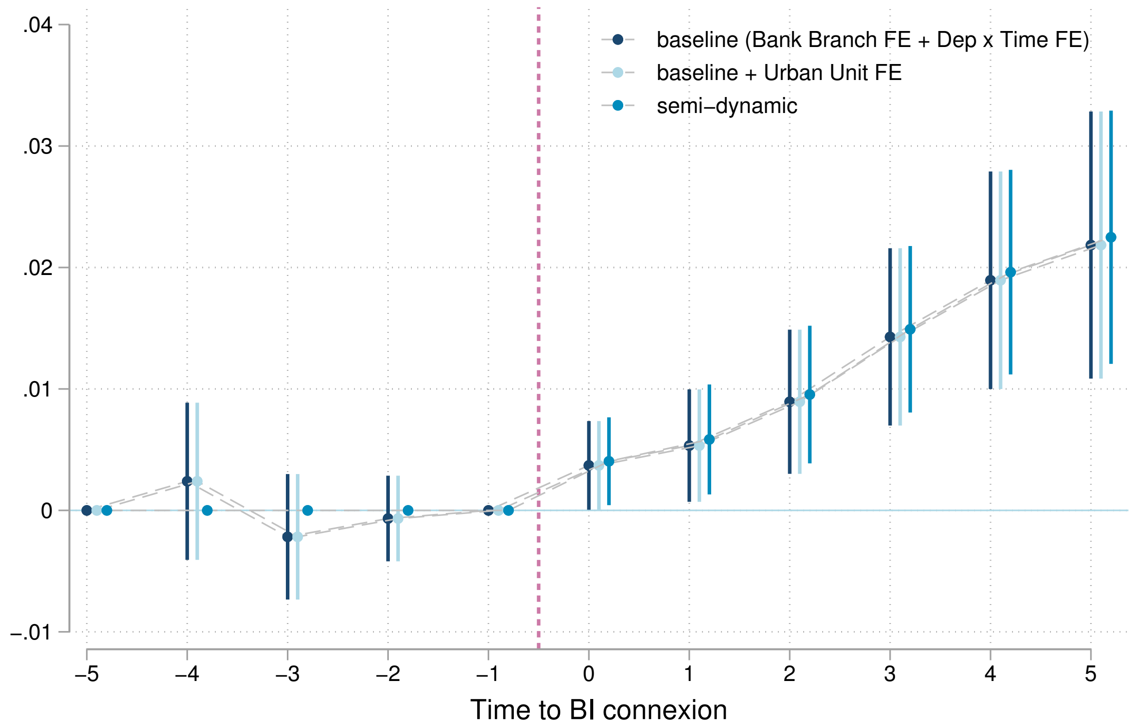
Notes: This figure plots estimates for the specification in equation (18)—fully dynamic—and in equation (19)—semi-dynamic. The dependent variable is the share of credit granted to firms located outside the bank's city. The sample includes all bank branches with a positive credit exposure. The baseline specification (navy blue) includes bank branch fixed effects and year × county fixed effects. City fixed effects are included in the alternative fully dynamic specification (light blue). Ninety-five percent confidence intervals are presented. Standard errors are clustered at the city level.

defined as the number of clients located outside the bank's city divided by the total number of clients.

Figure 5 shows the results. Similar to the previous estimation, I find a flat pre-trend and a positive effect afterwards. The coefficient for $d = 5$ now equals 0.023, implying that the technology-induced reduction in search frictions increases the bank branch-level share of remote clients by about 12%, five years after the period of the largest expansion. Interestingly, the effect on the extensive margin is comparable but slightly higher than the overall effect. This suggests that (i) the increase in between city credit flows is mainly driven by the creation of new relationships (new matches between banks and firms located in different cities), and (ii) those new credit relationships are, on average, smaller than the existing ones.

A direct consequence of these results is that firm-bank distance should increase following broadband internet expansion since there are more matches between firms and banks located in different cities. I directly verify this hypothesis by running a similar dynamic event-study regression with the (weighted) firm-bank distance as the dependent variable. Figure 6 displays

Figure 5: Share of remote clients

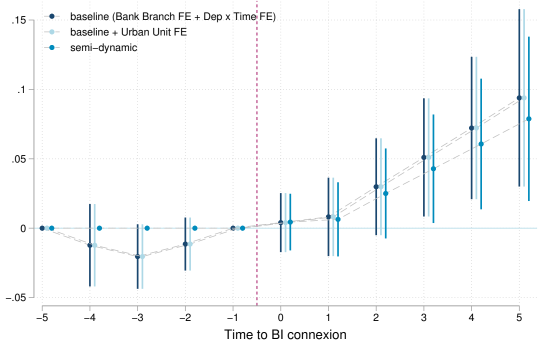


Notes: This figure plots estimates for the specification in equation (18)—fully dynamic—and equation (19)—semi-dynamic. The dependent variable is the share of clients financed outside the bank's city. The sample includes all bank branches with a positive credit exposure. The baseline specification (navy blue) includes bank branch fixed effects and year × county fixed effects. City fixed effects are included in the alternative fully dynamic specification (light blue). Ninety-five percent confidence intervals are presented. Standard errors are clustered at the city level.

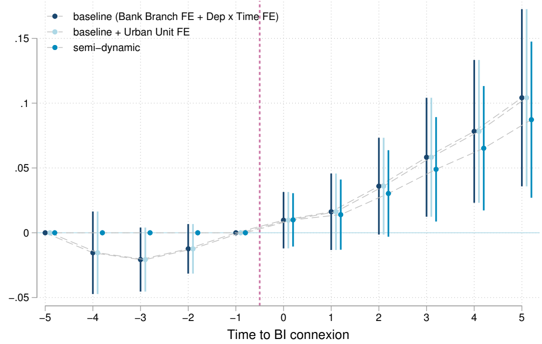
the results. Figure 6(a) documents a positive and significant effect of broadband internet on firm-bank distance. The magnitude implies that connected banks match with firms that are located, on average, 10% further away, five years after the shock. Figure 6(b) shows that the result is not significantly affected if the average distance is weighted by credit exposure. These results are in line with Kroszner and Strahan (1999), who find that innovation in information technology has reduced the dependence on geographic proximity between customers and banks in the US since the 1970s. These results also align with Petersen and Rajan (2002), who document a decline in the local nature of small business lending, as the average distance between small firms and their lenders grew from 51 miles in the 1970s to 161 miles in the 1990s—an increase of over 200% over 20 years. The smaller change in distance observed in my study is likely due to the shorter timeframe and the fact that I isolate the causal impact specifically attributed to internet adoption. To the best of my knowledge, this paper is the first to provide a causal interpretation for these facts, suggesting that innovations in information technology—namely broadband internet diffusion—reduced the role of both transaction and search costs in shaping credit outcomes, allowing firms to search for credit further and leading to structural changes

Figure 6: Firm-bank distance

(a) Unweighted



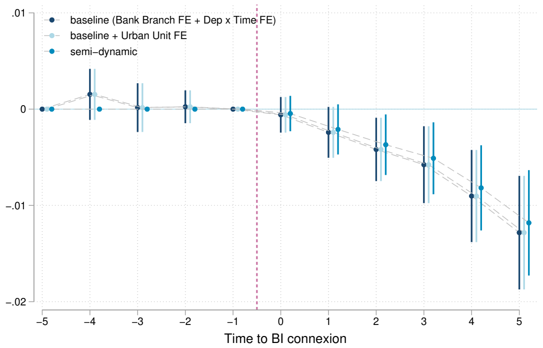
(b) Weighted by Credit Exposure



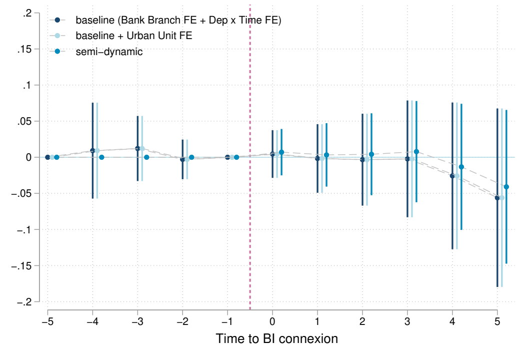
Notes: These figures plot estimates for the specification in equation (18)—fully dynamic—and equation (19)—semi-dynamic. In the left panel, the dependent variable is the (log) average firm-bank distance measured at the branch level. In the right panel, the (log) distance is weighted by credit exposure. The sample includes all bank branches with a positive credit exposure. The baseline specification (navy blue) includes bank branch fixed effects and year × county fixed effects. City fixed effects are included in the alternative fully dynamic specification (light blue). Ninety-five percent confidence intervals are presented. Standard errors are clustered at the city level.

Figure 7: Bank risk

(a) Share of Risky Borrowers



(b) Average Borrower Risk



Notes: These figures plot estimates for the specification in equation (18)—fully dynamic—and equation (19)—semi-dynamic. In the left panel, the dependent variable is the ratio of risky relationships to the total number of relationships measured at the branch level. In the right panel, the portfolio average investment grade. The sample includes all bank branches with a positive credit exposure. The baseline specification (navy blue) includes bank branch fixed effects and year × county fixed effects. City fixed effects are included in the alternative fully dynamic specification (light blue). Ninety-five percent confidence intervals are presented. Standard errors are clustered at the city level.

in local credit markets. Additionally, Figures A4 and A5 in Appendix show no impact on bank branch size, measured as total credit or the number of clients, nor on the average loan size (intensive margin). These findings are consistent with the gravity results presented above, which emphasize the extensive margin channel, and confirm that broadband internet allows new firm-bank relationships to be established between submarkets rather than cause an increase in the size of existing credit relationships. Finally, I study the consequences for bank branch risk. Enhanced competition, access to new markets, and new customers may increase bank risk-taking substantially. I proxy the risk-taking behavior of banks using the ratio of risky relationships to the total number of relationships and the portfolio average investment grade. The results in Figure 7 indicate no risk-taking by connected branches. The average borrower risk is not significantly affected, and the share of risky relationships (i.e., credit relationships with a Banque de France grade higher than five) decreases slightly, by 10% five years after the connection.

7 Implications for the Cost of Debt

In this section, I use the empirical results from Section 6 to quantify the impact of the technology-induced reduction in search frictions on loan prices through the lens of my model. Mapping the model prediction (9) into the gravity equation I estimate gives the following equivalence between the model parameters and the empirical estimates: $\beta_1 = \nu\,\theta$, $\beta_2 = \varrho$, and $\beta_3 = \gamma$. Thus, the technology-induced reduction in search frictions κ_{uv} for a pair of connected cities is formally written as:

$$\Delta \ln\kappa_{uv} = \hat{\gamma}\,\Delta\,\mathbb{C}_{uv} = -0.058. \tag{20}$$

The distribution of the minimum loan price (equation 6) obtained by an entrepreneur located in v is rewritten as follows:

$$\mathbb{P}\big(r_{e_v} \leq r\big) = W_{e_v}(r) = 1 - \exp\Bigg(- r^{\theta} z_{e_v}^{\theta+1} \sum_{u=1}^{N} S_u \cdot (c_u d_{uv})^{-\theta}\kappa_{uv}\Bigg),$$

which delivers the following expected lowest rate P_{e_v} for the entrepreneur e_v:

$$P_{e_v} = \mathbb{E}[r|e_v] = z_{e_v}^{-\frac{\theta+1}{\theta}} (\sum_{u=1}^{N} S_u \cdot (c_u d_{uv})^{-\theta}\kappa_{uv})^{-\frac{1}{\theta}} \times \Gamma\big[1+\frac{1}{\theta}\big], \tag{21}$$

where Γ stands for the Gamma function. By taking the log of equation 21, I can isolate the effect on P_{e_v} of the technology-induced reduction in search frictions caused by the diffusion of broadband internet:

$$\Delta ln P_{e_v} = ln\,\frac{P_{e_v}(1)}{P_{e_v}(0)} = -\frac{1}{\theta}\,ln\Bigg(\frac{\sum_{u=1}^{N}\hat{\kappa}_{uv}(1)\cdot S_u(0)\cdot(c_u d_{uv})^{-\theta}(0)}{\sum_{u=1}^{N}\hat{\kappa}_{uv}(0)\cdot S_u(0)\cdot(c_u d_{uv})^{-\theta}(0)}\Bigg), \tag{22}$$

where (1) indicates the state of the economy after broadband internet access and (0) before. As the origin fixed effect FE_u equals $S_u \cdot c_u^{-\theta}$ and $d_{uv}^{-\theta}$ equals $dist_{uv}^{-\nu\theta}$, equation (22) simplifies as follows:

$$\Delta ln \hat{P}_{e_v} = -\frac{1}{\theta}\,ln\Bigg(\frac{\sum_{u=1}^{N}\hat{\kappa}_{uv}(1)\cdot F\hat{E}_u(0)\cdot dist_{uv}^{-\hat{\nu}\hat{\theta}}}{\sum_{u=1}^{N}\hat{\kappa}_{uv}(0)\cdot F\hat{E}_u(0)\cdot dist_{uv}^{-\hat{\nu}\hat{\theta}}}\Bigg). \tag{23}$$

While broadband internet may affect all variables in this equation as the firm productivity z_{e_v}, the cost c_u, and the θ parameter of the bank branches' size distribution, I keep them constant to conduct counterfactual exercises. I plug into this equation my empirical estimates for $\beta_1 = -\hat{\nu}\hat{\theta}$, FE_u and $\hat{\kappa}_{uv}(1) = \exp(\hat{\gamma} + \hat{\rho}X_{uv(1)})$ as well as parameters calibrated from the data (θ, $dist_{uv}$) to compute $\Delta \widehat{lnP_{e_v}}$, the change in the lower expected cost of debt triggered by the broadband internet-induced reduction in search frictions. My model predicts an average decline of –4.9% in 2005 compared to what it would have been without any lowering of search and contracting costs. This decline in the cost of debt would have been higher if all French cities were connected at the end of 2005, with an average value of –5.8%. This result echoes the conclusions of Hauswald and Marquez (2003), who show how improved access to information, under some conditions, makes markets more competitive so that customers benefit from technological progress. The mechanism herein departs from theirs as I focus on firms' search and do not model the process by which banks search for customers, but the intuition of an easier dissemination of information is similar. In a recent theoretical approach, Vivès and Ye (2021) also study how the diffusion of information technology affects competition in the lending market. Here again, they do not model search frictions (in their model, ICTs reduce the effect of bank–borrower distance on monitoring/screening cost) but share similar conclusions: ICT progress may trigger an increase in banks' competition intensity and, as a consequence, the loan rates offered to entrepreneurs decline.[20]

Figure A7 illustrates the spatial dimension of the decline in firms' cost of debt. Dark blue cities are the ones in which this decline is the strongest (larger than the 75th percentile), while the light gray ones indicates a reduction lower than the median. It is noteworthy that all the largest French cities (Paris, Marseille, Bordeaux, or Lyon) benefit less from this decline than suburban or rural areas. This result is in line with the model intuition that a reduction in search frictions precipitates an increase in competition due to the ability for firms to search further and meet with more bank branches. In already crowded markets with a lot of active banks and a high level of competition—typically in large city centers—firms are not highly constrained by the search frictions and the decline is low. On the contrary, in isolated and rural submarkets, where firms must make costly efforts to meet with different bankers and eventually match with the right one, the reduction in search frictions triggers a substantial decline in loan prices. Several other dimensions of heterogeneity are explored in Appendix (see Figure A8), where I show that if population skills and average income do not matter, the age structure of the workforce is negatively correlated with the strength of the decline: a city populated with youth is more likely to benefit from a technology-induced reduction of the cost of debt for SMEs.

[20] I interpret the observed reduction in the cost of debt as a direct consequence of the reduction in search costs, along the lines of my model, but the impact of ADSL diffusion could also be due to an increase of banks' ability to screen projects (through productivity z_b or d_{uv}). This effect is absorbed by the c × time fixed effects in my empirical analysis.

8 Conclusion

In this paper, I develop a new theory of firm-bank matching subject to search frictions and provide causal evidence on how these frictions affect firm-bank matching and the allocation of bank credit, using the staggered rollout of broadband internet in France as a shock on transaction and search costs. I show that this technology-induced reduction in search frictions triggers an increase of 6% of the share of credit exchanged between interconnected cities. This positive effect varies with the initial level of search frictions: it is higher when two very distant cities are connected and is almost null when two neighboring cities, already very closely tied economically, are interconnected by internet.

Leveraging bank branch-level data, I document that broadband internet diffusion allows banks to match with new firms located remotely. Connected banks increase their share of credit lent to firms located outside their city by 10% and their share of remote clients by almost 12%. As a result, the average distance between a bank and its customers increases by 10% in the medium run after broadband internet access. Finally, I plug these estimates into the equation linking search frictions to loan prices. Interpreted within my model, the reduced-form estimates imply that the reduction in search frictions due to the large diffusion of broadband internet lowered the cost of debt for small businesses by 4.9% on average.

Overall, this paper highlights the role of transactions and search costs in shaping firms' access to credit. Credit markets with high search frictions make financing by bank credit both difficult, time-consuming, and onerous, especially for small businesses. These findings call for a variety of economic policies that aim to make the process of searching and applying for credit more fluid, efficient, and less burdensome, particularly in a period of pandemic marked by the disappearance of face-to-face interactions and the consequent surge of digitalization.

References

Agarwal, Sumit and Robert Hauswald, "Distance and Private Information in Lending," *The Review of Financial Studies*, 04 2010, *23* (7), 2757–2788.

__ , **John Grigsby, Ali Hortaçsu, Gregor Matvos, Amit Seru, and Vincent Yao**, "Searching for Approval," Working Paper 27341, National Bureau of Economic Research June 2020.

Akerlof, George A., "The Market for Lemons: Quality Uncertainty and the Market Mechanism," *The Quarterly Journal of Economics*, 1970, *84* (3), 488–500.

Akerman, Anders, Edwin Leuven, and Magne Mogstad, "Information Frictions, Internet and the Relationship between Distance and Trade," *mimeo*, 2018, p. 48.

Allen, Jason, Robert Clark, and Jean-François Houde, "Search Frictions and Market Power in Negotiated-Price Markets," *Journal of Political Economy*, 2019, *127* (4), 1550–1598.

Allen, Treb, "Information Frictions in Trade," *Econometrica*, 2014, *82* (6), 2041–2083.

Arcep, "Accés haut débit via l'ADSL: historique des décisions de l'ART et de l'ARCEP," Technical Report 2002.

Argyle, Bronson, Taylor Nadauld, and Christopher Palmer, "Real Effects of Search Frictions in Consumer Credit Markets," *MIT Sloan School Working Paper 5242-17*, 01 2019.

Bergeaud, Antonin, Clément Malgouyres, Clément Mazet-Sonilhac, and Sara Signorelli, "Technological change and domestic outsourcing," *Journal of Labor Economics*, 2024.

Berger, Allen and Gregory Udell, "Relationship Lending and Lines of Credit in Small Firm Finance," *The Journal of Business*, 1995, *68* (3), 351–81.

__ **and** __ , "Small Business Credit Availability and Relationship Lending: The Importance of Bank Organisational Structure," *Economic Journal*, 2002, *112* (477), F32–F53.

__ , **John H. Leusner, and John J. Mingo**, "The efficiency of bank branches," *Journal of Monetary Economics*, 1997, *40* (1), 141–162.

Bhuller, Manudeep, Andreas Kostøl, and Trond Christian Vigtel, "How Broadband Internet Affects Labor Market Matching," Memorandum, Oslo University December 2019.

Boot, Arnoud W.A., "Relationship Banking: What Do We Know?," *Journal of Financial Intermediation*, 2000, *9* (1), 7–25.

Borusyak, Kirill and Xavier Jaravel, "Revisiting Event Study Designs with an Application to the Estimation of the Marginal Propensity to Consume," *mimeo*, 2017, p. 33.

Boualam, Yasser M., "Credit Markets and Relationship Capital," *Working Paper at University of Pennsylvania,*, 2018.

Brei, Michael and Goetz von Peter, "The distance effect in banking and trade," *Journal of International Money and Finance*, 2018, *81*, 116–137.

Cerqueiro, Geraldo, Hand Degryse, and Steven Ongena, "Rules versus discretion in loan rate setting," *Journal of Financial Intermediation*, 2011, *20* (4), 503 – 529.

Chaney, Thomas, "The Network Structure of International Trade," *American Economic Review*, November 2014, *104* (11), 3600–3634.

Chang, Briana, Matthieu Gomez, and Harrison Hong, "Sorting out the effect of credit supply," *Journal of Financial Economics*, 2023, *150* (3), 103719.

Coeurdacier, Nicolas and Philippe Martin, "The geography of asset trade and the euro: Insiders and outsiders," *Journal of the Japanese and International Economies*, 2009, *23* (2), 90–113.

Correia, Sergio, Paulo Guimaraes, and Thomas Zylkin, "ppmlhdfe: Fast Poisson Estimation with High-Dimensional Fixed Effects," *arXiv e-prints*, Mar 2019, p. arXiv:1903.01690.

Crawford, Gregory S, Nicola Pavanini, and Fabiano Schivardi, "Asymmetric information and imperfect competition in lending markets," *American Economic Review*, 2018, *108* (7), 1659–1701.

D'Andrea, Angelo and Nicola Limodio, "High-Speed Internet, Financial Technology and Banking in Africa," BAFFI CAREFIN Working Papers 19124 2019.

Degryse, Hand and Steven Ongena, "Distance, Lending Relationships, and Competition," *The Journal of Finance*, 2005, *60* (1), 231–266.

Degryse, Hans, Moshe Kim, and Steven Ongena, *Microeconometrics of Banking Methods, Applications, and Results*, Oxford University Press, 2009.

den Haan, Wouter J., Garey Ramey, and Joel Watson, "Liquidity flows and fragility of business enterprises," *Journal of Monetary Economics*, 2003, *50* (6), 1215–1241.

Drexler, Alejandro and Antoinette Schoar, "Do Relationships Matter? Evidence from Loan Officer Turnover," *Management Science*, 2014, *60* (11), 2722–2736.

Duffie, Darrell, Nicolae Gârleanu, and Lasse Heje Pedersen, "Over-the-counter markets," *Econometrica*, 2005, *73* (6), 1815–1847.

Duquerroy, Anne, Clément Mazet-Sonilhac, Jean-Stéphane Mésonnier, and Daniel Paravisini, "Branch Closures and Specialization," *Mimeo*, 2019.

D'Andrea, Angelo, Marco Pelosi, and Enrico Sette, "Broadband and bank intermediation," Technical Report, Working Paper. Bocconi University, Milan, Italy 2022.

Eaton, Jonatahn, Samuel Kortum, and Francis Kramarz, "Firm-to-firm trade: Imports, exports, and the labor market," Technical Report, National Bureau of Economic Research 2022.

Ernst and Young, "The future of SME banking," *EY Financial Services Report*, 2018.

Fally, Thibault, "Structural gravity and fixed effects," *Journal of International Economics*, 2015, *97* (1), 76–85.

Federal Reserve Bank, "Small Business Credit Survey," 2014.

_ , "Small Business Credit Survey. 2022 Report on Employer Firms," 2022.

Gavazza, Alessandro and Alessandro Lizzeri, "Frictions in Product Markets, Handbook of Industrial Organization," 2021, *4, Issue 1*, 433–484.

Gross, Tal, Matthew J Notowidigdo, and Jialan Wang, "The Marginal Propensity to Consume Over the Business Cycle," 2018.

Hauswald, Robert and Robert Marquez, "Information Technology and Financial Services Competition," *The Review of Financial Studies*, 06 2003, *16* (3), 921–948.

_ **and** _ , "Competition and Strategic Information Acquisition in Credit Markets," *The Review of Financial Studies*, 02 2006, *19* (3), 967–1000.

Head, Keith and Thierry Mayer, "Chapter 3 - Gravity Equations: Workhorse,Toolkit, and Cookbook," in Gita Gopinath, Elhanan Helpman, and Kenneth Rogoff, eds., *Handbook of International Economics*, Vol. 4 of *Handbook of International Economics*, Elsevier, 2014, pp. 131 – 195.

Honka, Elisabeth, Ali Hortaçsu, and Maria Ana Vitorino, "Advertising, consumer awareness, and choice: evidence from the U.S. banking industry," *The RAND Journal of Economics*, 2017, *48* (3), 611–646.

Infosys, "Banks! It's time to change your game in SMEs lending," 2018.

Jiang, Erica Xuewei, Gloria Yang Yu, and Jinyuan Zhang, "Bank competition amid digital disruption: Implications for financial inclusion," *USC Marshall School of Business Research Paper Sponsored by iORB*, 2022.

Kroszner, Randall S. and Philip E. Strahan, "What Drives Deregulation? Economics and Politics of the Relaxation of Bank Branching Restrictions," *The Quarterly Journal of Economics*, 1999, *114* (4), 1437–1467.

Lendle, Andreas, Marcelo Olarreaga, Simon Schropp, and Pierre-Louis Vézina, "There Goes Gravity: eBay and the Death of Distance," *Economic Journal*, 2016, *126* (591), 406–441.

Lenoir, Clemence, Isabelle Mejean, and Julien Martin, "Search Frictions in International Good Markets," 2018 Meeting Papers 878, Society for Economic Dynamics 2018.

Malgouyres, Clément, Thierry Mayer, and Clément Mazet-Sonilhac, "Technology-Induced Trade Shocks? Evidence from Broadband Expansion in France," *Journal of Internation Economics*, July 2021, *123* (103520).

McAllister, Patrick H. and Douglas McManus, "Resolving the scale efficiency puzzle in banking," *Journal of Banking & Finance*, 1993, *17* (2-3), 389–405.

Mills, Karen and Brayden McCarthy, "The state of small business lending: Credit access during the recovery and how technology may change the game," *Harvard Business School General Management Unit Working Paper*, 2014, (15-004).

Nguyen, Hoai-Luu Q., "Are Credit Markets Still Local? Evidence from Bank Branch Closings," *American Economic Journal: Applied Economics*, January 2019, *11* (1), 1–32.

Okawa, Yohei and Eric van Wincoop, "Gravity in International Finance," *Journal of International Economics*, 2012, *87* (2), 205 – 215.

Petersen, Mitchell A. and Raghuram Rajan, "The Effect of Credit Market Competition on Lending Relationships," *The Quarterly Journal of Economics*, 1995, *110* (2), 407–443.

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

Portes, Richard and Hélène Rey, "The determinants of cross-border equity flows," *Journal of International Economics*, 2005, *65* (2), 269 – 296.

Rajan, Raghuram G, "Insiders and outsiders: The choice between informed and arm's-length debt," *The Journal of finance*, 1992, *47* (4), 1367–1400.

Rogerson, Richard, Robert Shimer, and Randall Wright, "Search-Theoretic Models of the Labor Market: A Survey," *Journal of Economic Literature*, December 2005, *43* (4), 959–988.

Sénat, "Le bilan de la loi n 96-659 de réglementation des télécommunications," *Rapport d'information du Sénat*, 2002.

Sharpe, Steven A., "Asymmetric Information, Bank Lending, and Implicit Contracts: A Stylized Model of Customer Relationships," *The Journal of Finance*, 1990, *45* (4), 1069–1087.

Silva, João Santos and Silvana Tenreyro, "The Log of Gravity," *The Review of Economics and Statistics*, 2006, *88* (4), 641–658.

_ **and** _ , "Further simulation evidence on the performance of the Poisson pseudo-maximum likelihood estimator," *Economics Letters*, 2011, *112* (2), 220–222.

Steinwender, Claudia, "Real effects of information frictions: When the states and the kingdom became united," *American Economic Review*, 2018, *108* (3), 657–96.

Stiglitz, Joseph E. and Andrew Weiss, "Credit Rationing in Markets with Imperfect Information," *The American Economic Review*, 1981, *71* (3), 393–410.

Telecom, France, "Conférence de presse du 10 juin 2003 "Internet Haut Débit pour tous : France Télécom s'engage"," Technical Report 2003.

Tregouet, Rene, "Question écrite n. 30844 (Sénat, 01/02/2001)," *Senat*, 2001.

Udell, Gregory F, "SME Access to Intermediated Credit: What Do We Know and What Don't We Know?," in Angus Moore and John Simon, eds., *Small Business Conditions and Finance*, RBA Annual Conference Volume (Discontinued), Reserve Bank of Australia, December 2015.

Vivès, Xavier and Zhiqiang Ye, "Information Technology and Bank Competition," *CEPR Working Paper*, 2021.

Wasmer, Etienne and Philippe Weil, "The Macroeconomics of Labor and Credit Market Imperfections," *American Economic Review*, September 2004, *94* (4), 944–963.

Wheelock, David and Paul Wilson, "Do Large Banks Have Lower Costs? New Estimates of Returns to Scale for U.S. Banks," *Journal of Money, Credit and Banking*, 2012, *44* (1), 171–199.

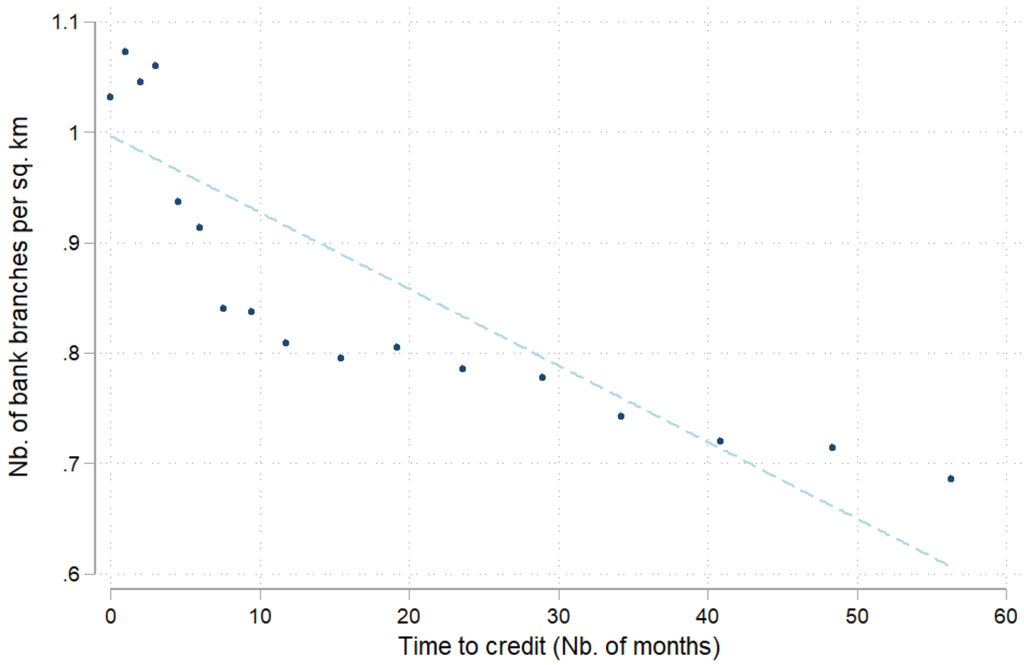
Online Appendix

Outline

- Appendix A presents additional empirical results
- Appendix B presents the geographic units in more details
- Appendix C presents the details of the derivation of Eq. 7
- Appendix D documents BI adoption by French banks
- Appendix E details empirical facts about credit markets
- Appendix F presents a simulation exercise for the PPML estimator

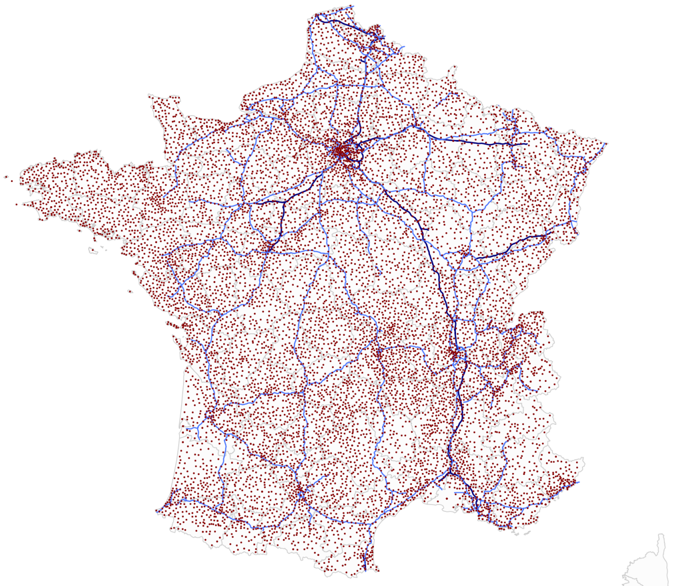
A Additional Results

Figure A1: Correlation between trading delay and bank branch density



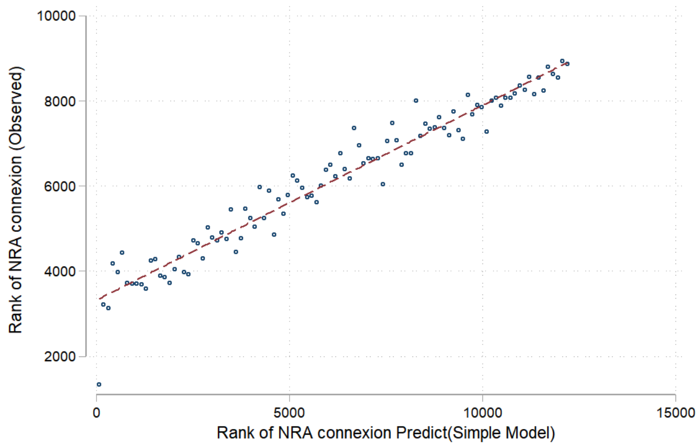
Notes: This figure plots the binned scatter plot for the correlation between the variable Time to credit, defined as the number of months between firm creation and the first match with a bank branch, and the local bank branch density, defined as the number of bank branches per square kilometer in the firm's city.

Figure A2: Local Exchanges, Highways, and Railroads before 1999



Notes: This figure shows the location of around 13,000 local exchanges (red dots), highways (light blue lines), and railroads (dark blue lines) pre-existing before the beginning of the broadband internet expansion in France.

Figure A3: Optimal connection rank predicted versus observed connection rank



Notes: This figure shows the rank correlation between optimal versus observed connection ranks. The combination of exogenous connection gains and costs have a strong predictive power, with an R^2 close to 0.70. The optimal connection rank $\hat{R}_i$ is predicted for each local exchange i, only taking into account two presumably exogenous measures of costs (shortest distance to existing infrastructure) and gains (population density).

Table A1: Gravity Equation for Inter-Regional Credit Flows

	Share of credit in v borrowed from u: $\Pi_{uvt} > 0$				
	(1)	(2)	(3)	(4)	(5)
Log distance $dist_{uv}$	-1.001***	-0.992***	-0.934***	-0.971***	-0.955***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Same region		0.081***			
		(0.013)			
Same county			0.419***		0.669***
			(0.010)		(0.016)
log (Trade Flows)				0.047***	-0.094***
				(0.003)	(0.005)
Origin (u) × Year FE	Yes	Yes	Yes	Yes	Yes
Destination (v) × Year FE	Yes	Yes	Yes	Yes	Yes
Origin-Destination FE	No	No	No	No	No
R^2	0.27	0.27	0.28	0.28	0.28
Observations	143,747	143,747	143,747	143,747	143,747

Notes: The table shows the PPML estimation of equation (16). $dist_{uv}$ = bilateral distance. Same region is a dummy variable equal to 1 if both the firm and the bank are located in the same region. Same county is a dummy variable equal to 1 if both the firm and the bank are located in the same county. Log trade flows is the lagged log of bilateral trade flows between counties. Columns (1) to (5) include fixed effects for origin × years and destination × years. The sample period is 1997–2005, and the sample consists of origin-destination-year combinations with positive credit flows, where at least one firm is located with positive credit.

Table A2: TECHNOLOGY-INDUCED REDUCTION IN SEARCH FRICTIONS: PAIR FIXED EFFECTS

	Share of credit in v borrowed from u: Π_{uvt}			
	(1)	(2)	(3)	(4)
Log distance $dist_{uv}$	-1.844***		-1.865***	
	(0.006)		(0.006)	
$\mathbb{C}_{uvt}$	0.800***	-0.009		
	(0.059)	(0.033)		
Log distance $dist_{uv}$ x $\mathbb{C}_{uvt}$	0.230***	0.035***		
	(0.008)	(0.004)		
$\hat{\mathbb{C}}_{uvt}$			0.399***	0.037***
			(0.062)	(0.004)
Log distance $dist_{uv}$ x $\hat{\mathbb{C}}_{uvt}$			0.199***	0.108***
			(0.008)	(0.039)
log (Trade Flows)	0.470***	-0.018**	0.468***	-0.018**
	(0.004)	(0.009)	(0.004)	(0.009)
Origin (u) × Year FE	Yes	Yes	Yes	Yes
Destination (v) × Year FE	Yes	Yes	Yes	Yes
Origin-Destination FE	No	Yes	No	Yes
R^2	0.68	0.41	0.68	0.42
Observations	24,545,640	250,780	24,545,640	250,780

Notes: The table shows the PPML estimation of equation (17). $dist_{uv}$ = bilateral distance. $\mathbb{C}_{uvt} = Z_{vt} \times Z_{ut}$ is a continuous variable that indicates the degree of broadband internet interconnectivity between two cities. $\mathbb{C}_{uvt}$ belongs to $[0, 1]$, and this measure captures the ability for firms located in u to locate and communicate with bank branches located in v, using the world wide web. My IV strategy delivers a similar measure of connectivity between cities named $\hat{\mathbb{C}}_{uvt} = \hat{Z}_{vt} \times \hat{Z}_{ut}$. Log trade flows is the lagged log of bilateral trade flows between counties. Columns (1) and (3) include fixed effects for origin × years and destination × years, and columns (2) and (4) add pair fixed effects. The sample period is 1997–2005, and the sample consists of all origin-destination-year combinations where at least one firm is located with positive credit.

Table A3: TECHNOLOGY-INDUCED REDUCTION IN SEARCH FRICTIONS WITH LAGS

	Share of credit in v borrowed from u Π_{uvt}					
	(1)	(2)	(3)	(4)	(5)	(6)
Log distance $dist_{uv}$	-1.561***	-1.764***	-1.175***	-1.265***	-1.174***	-1.274***
	(0.003)	(0.004)	(0.004)	(0.006)	(0.004)	(0.006)
$\mathbb{C}_{uvt}$			0.097*	0.416***		
			(0.053)	(0.053)		
Log distance $dist_{uv} \times \mathbb{C}_{uvt}$				0.214***		
				(0.007)		
$\hat{\mathbb{C}}_{uvt}$					0.107**	0.286***
					(0.052)	(0.053)
Log distance $dist_{uv} \times \hat{\mathbb{C}}_{uvt}$						0.168***
						(0.007)
log (Trade Flows)		0.473***	0.482***	0.478***	0.482***	0.477***
		(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Origin (u) × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Destination (v) × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Origin-Destination FE	No	No	No	No	No	No
Lag dependant variable	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.65	0.66	0.68	0.69	0.68	0.69
Observations	24,545,640	24,545,640	20,633,055	20,633,055	20,633,055	20,633,055

Notes: The table shows the PPML estimation of equation (17) with the lag dependent variable included. $dist_{uv}$ = bilateral distance. $\mathbb{C}_{uvt} = Z_{vt} \times Z_{ut}$ is a continuous variable that indicates the degree of broadband internet inter-connectivity between two cities. $\mathbb{C}_{uvt}$ belongs to $[0, 1]$. This measure captures the ability for firms located in u to locate and communicate with bank branches located in v, using the world wide web. My IV strategy delivers a similar measure of connectivity between cities named $\hat{\mathbb{C}}_{uvt} = \hat{Z}_{vt} \times \hat{Z}_{ut}$. Log trade flows is the lagged log of bilateral trade flows between counties. Columns (1) to (5) include fixed effects for origin × years and destination × years. The sample period is 1997–2005, and the sample consists of all origin-destination-year combinations where at least one firm is located with positive credit.

Table A4: TECHNOLOGY-INDUCED REDUCTION IN SEARCH FRICTIONS: EXTENSIVE MARGIN

	Share of firms in v borrowing from u: Π_{uvt}				
	(1)	(2)	(3)	(4)	(5)
Log distance $dist_{uv}$	-2.246***	-2.024***	-1.882***	-1.827***	-1.825***
	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
Same region		1.498***			
		(0.015)			
Same county			1.738***		
			(0.014)		
log (Trade Flows)				0.480***	0.481***
				(0.004)	(0.004)
$\mathbb{C}_{uvt}$					0.200***
					(0.049)
Origin (u) × Year FE	Yes	Yes	Yes	Yes	Yes
Destination (v) × Year FE	Yes	Yes	Yes	Yes	Yes
Origin-Destination FE	No	No	No	No	No
R^2	0.66	0.66	0.67	0.66	0.67
Observations	24545811	24545811	24545811	24545640	24545640

Notes: The table shows the PPML estimation of equation (16). $dist_{uv}$ = bilateral distance. $\mathbb{C}_{uvt} = Z_{vt} \times Z_{ut}$ is a continuous variable that indicates the degree of broadband internet inter-connectivity between two cities. $\mathbb{C}_{uvt}$ belongs to $[0, 1]$. This measure captures the ability for firms located in u to locate and communicate with bank branches located in v, using the world wide web. My IV strategy delivers a similar measure of connectivity between cities named $\hat{\mathbb{C}}_{uvt} = \hat{Z}_{vt} \times \hat{Z}_{ut}$. Log trade flows is the lagged log of bilateral trade flows between counties. Columns (1) to (5) include fixed effects for origin × years and destination × years. The sample period is 1997–2005, and the sample consists of all origin-destination-year combinations where at least one firm is located with positive credit.

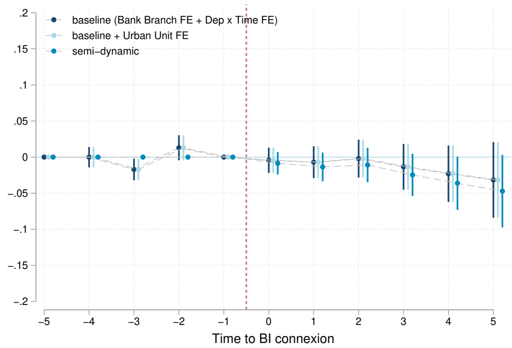
Table A5: Technology-Induced reduction in search frictions: extensive margin

	Share of firms in v borrowing from u Π_{uvt}				
	(1)	(2)	(3)	(4)	(5)
Log distance $dist_{uv}$	-2.246***	-1.825***	-1.905***	-1.927***	
	(0.003)	(0.004)	(0.005)	(0.006)	
$\mathbb{C}_{uvt}$		0.200***	0.624***		
		(0.049)	(0.050)		
Log distance $dist_{uv} \times \mathbb{C}_{uvt}$			0.220***		
			(0.007)		
$\hat{\mathbb{C}}_{uvt}$				0.176***	0.103***
				(0.054)	(0.031)
Log distance $dist_{uv} \times \hat{\mathbb{C}}_{uvt}$				0.191***	0.035***
				(0.007)	(0.004)
log (Trade Flows)		0.481***	0.476***	0.474***	-0.010
		(0.004)	(0.004)	(0.004)	(0.008)
Origin (u) × Year FE	Yes	Yes	Yes	Yes	Yes
Destination (v) × Year FE	Yes	Yes	Yes	Yes	Yes
Origin-Destination FE	No	No	No	No	Yes
R^2	0.65	0.66	0.68	0.69	0.69
Observations	24545811	24545640	24545640	24545640	250780

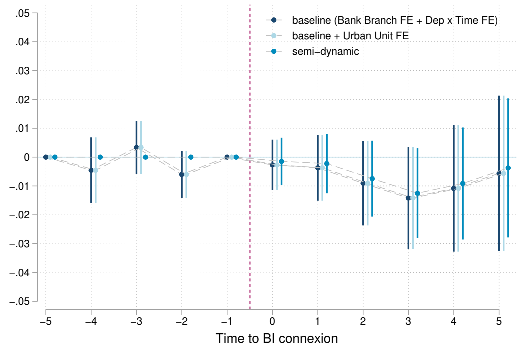
Notes: The table shows the PPML estimation of equation (16). $dist_{uv}$ = bilateral distance. $\mathbb{C}_{uvt} = Z_{vt} \times Z_{ut}$ is a continuous variable that indicates the degree of broadband internet inter-connectivity between two cities. $\mathbb{C}_{uvt}$ belongs to $[0, 1]$. This measure captures the ability for firms located in u to locate and communicate with bank branches located in v, using the world wide web. My IV strategy delivers a similar measure of connectivity between cities named $\hat{\mathbb{C}}_{uvt} = \hat{Z}_{vt} \times \hat{Z}_{ut}$. Log trade flows is the lagged log of bilateral trade flows between counties. Columns (1) to (5) include fixed effects for origin × years and destination × years. The sample period is 1997–2005, and the sample consists of all origin-destination-year combinations where at least one firm is located with positive credit.

Figure A4: Bank branch total and average credit

(a) Total Credit



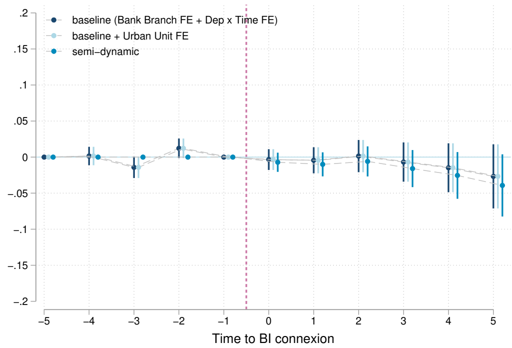
(b) Average Credit



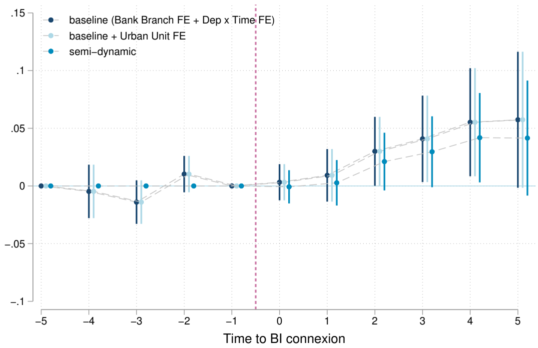
Notes: These figures plot estimates for the specification in equation (18)—fully dynamic—and equation (19)—semi-dynamic. In the left panel, the dependent variable is the (log) total credit measured at the branch level. In the right panel, the dependent variable is the (log) average credit per loan. The sample includes all bank branches with a positive credit exposure. The baseline specification (navy blue) includes bank branch fixed effects and year $\times$ county fixed effects. City fixed effects are included in the alternative fully dynamic specification (light blue). Ninety-five percent confidence intervals are presented. Standard errors are clustered at the city level.

Figure A5: Bank branch total and remote clients

(a) Number of Clients

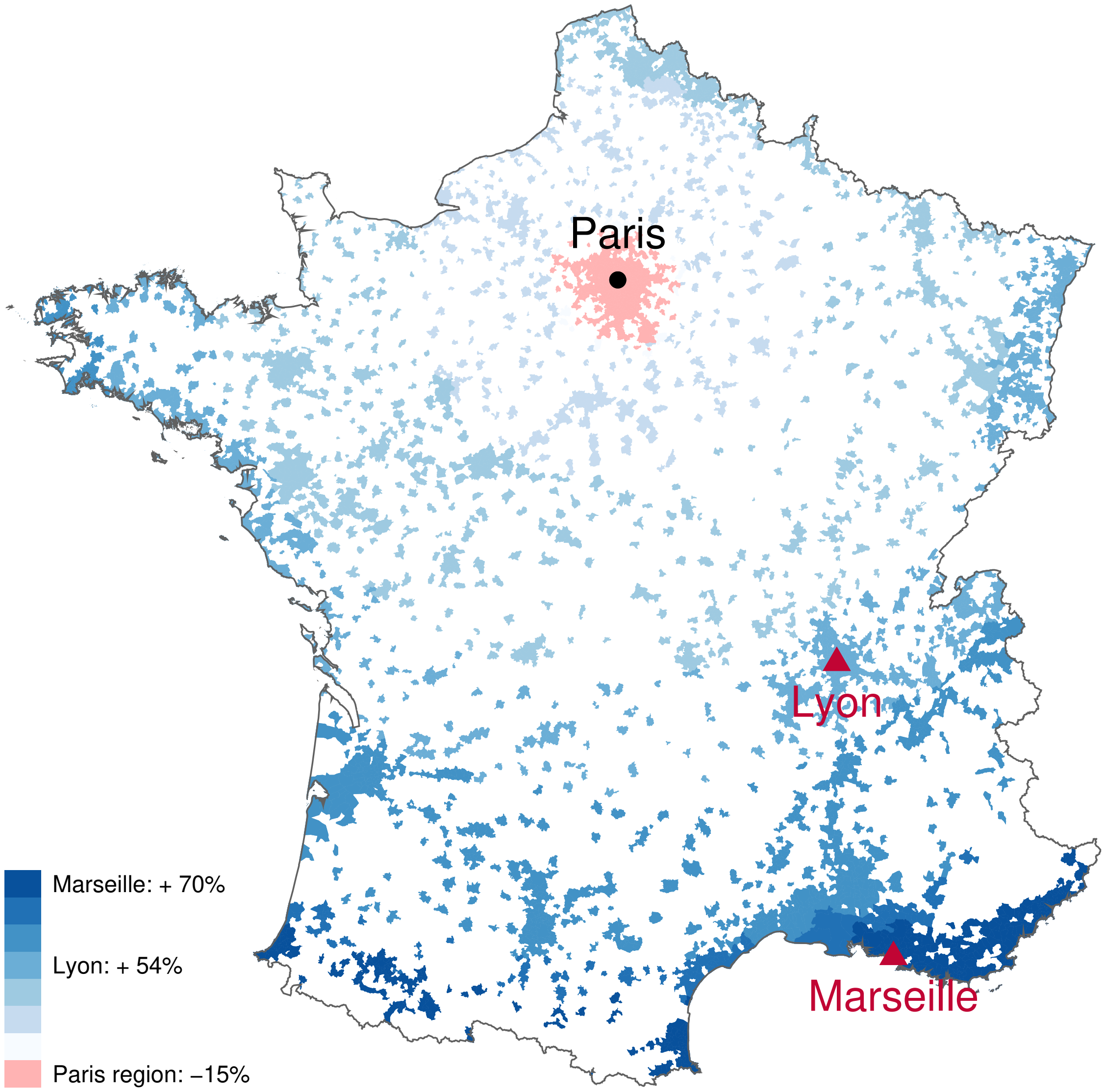


(b) Number of Remote Clients



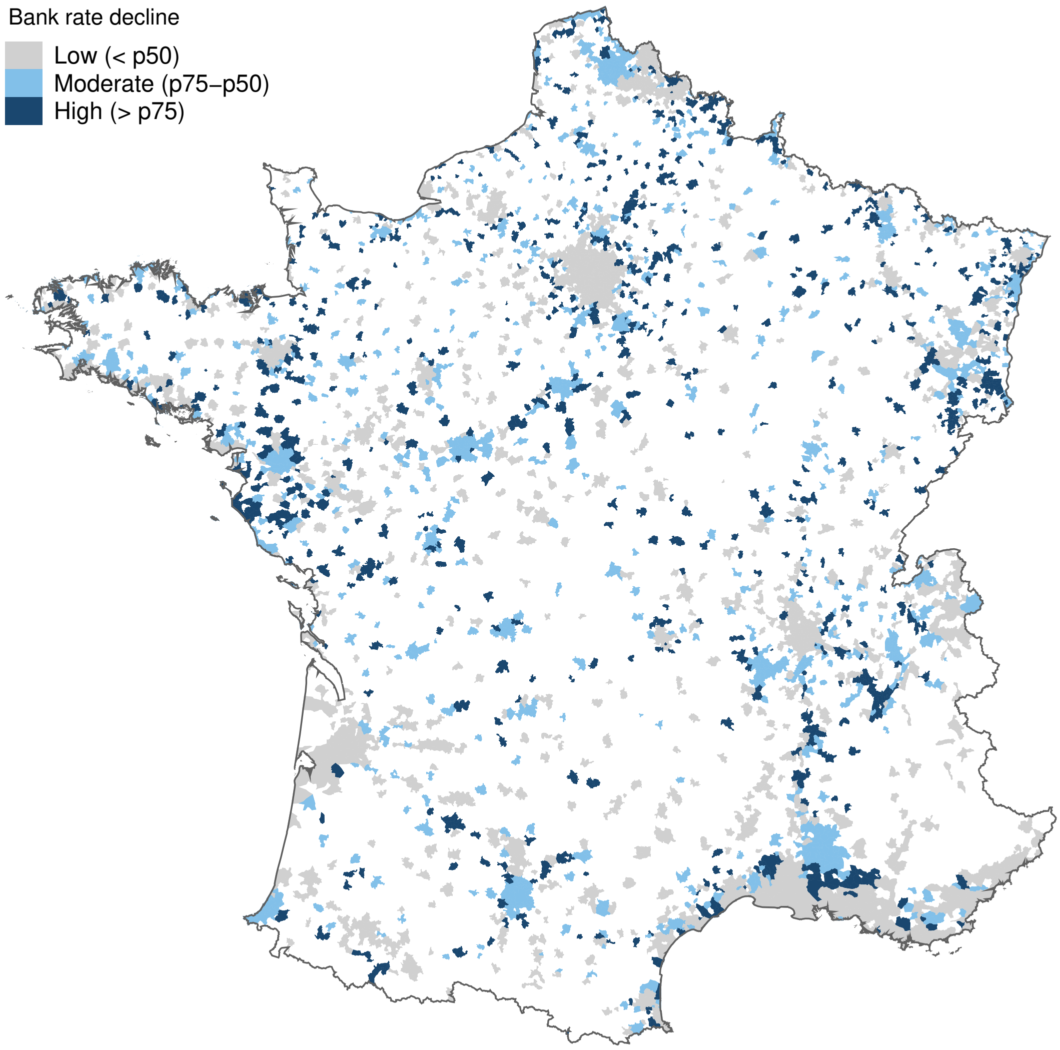
Notes: These figures plot estimates for the specification in equation (18)—fully dynamic—and equation (19)—semi-dynamic. In the left panel, the dependent variable is the (log) number of clients measured at the branch level. In the right panel, the dependent variable is the (log) number of remote clients. The sample includes all bank branches with a positive credit exposure. The baseline specification (navy blue) includes bank branch fixed effects and year $\times$ county fixed effects. City fixed effects are included in the alternative fully dynamic specification (light blue). Ninety-five percent confidence intervals are presented. Standard errors are clustered at the city level.

Figure A6: Heterogeneous effect of broadband internet with respect to distance when Paris is connected



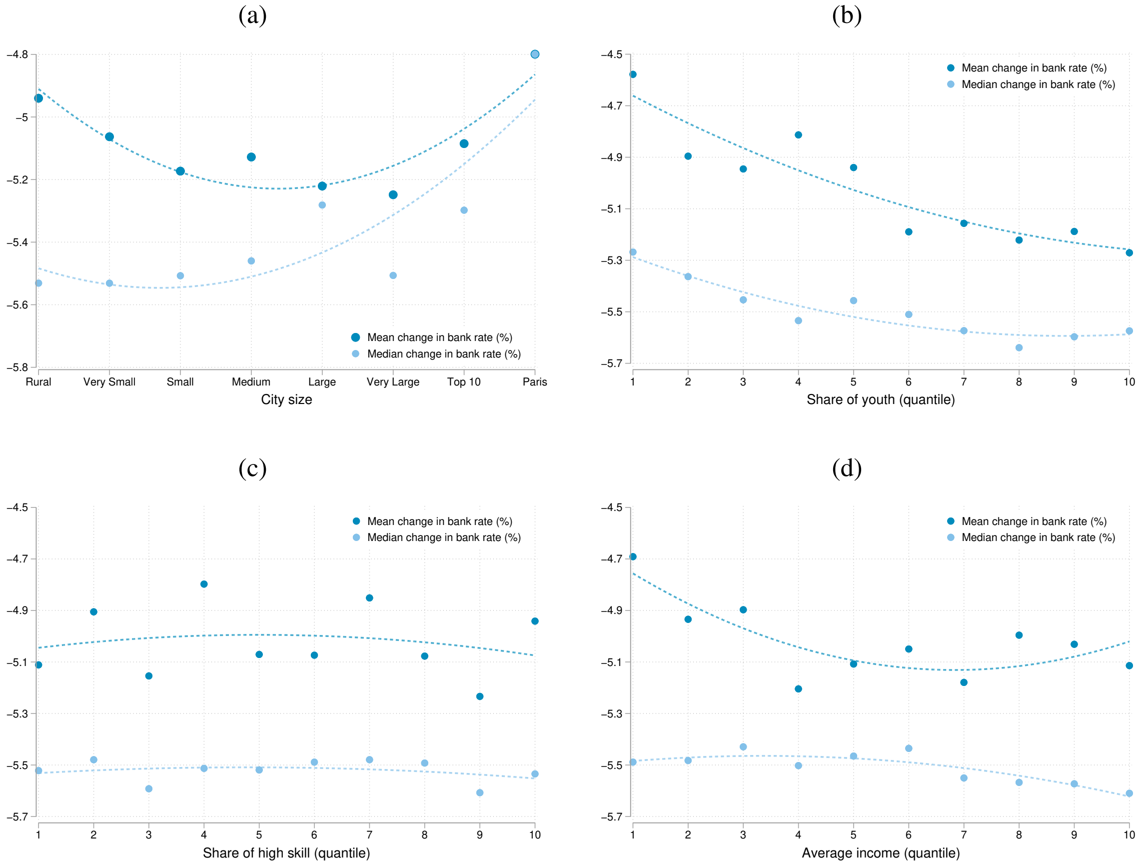
Notes: This figure plots the effect of broadband internet connection on the share of credit borrowed by firms located in any city in France to banks located in Paris. The black dot indicates the Paris location, while the red triangles show Marseille and Lyon, respectively, the second and the third biggest French cities. Dark blue indicates an effect at the 90th percentile, while light red indicates the negative effect of being connected to Paris for cities located nearby.

Figure A7: Reduction of the cost of debt triggered by a reduction in search frictions: spatial heterogeneity



Notes: This map illustrates the spatial dimension of the decline in firms' cost of debt in France. Dark blue cities are the ones in which this decline is the strongest, larger than the 75th percentile. The light blue areas undergo a decline of the cost of debt higher than the median (but lower than the 75th percentile). The light gray ones indicate a reduction lower than the median.

Figure A8: Decline in the cost of debt: heterogeneity by city size, age, skills, and income



Notes: These figures illustrates the heterogeneity in the average (dark blue) and median (light blue) decline in firms' cost of debt in France along several dimensions. In panel (a), cities are grouped by size; in (b) by quantiles of the share of youth (under 30) in the workforce, measured in 1990; in (c) by quantiles of the share of high-skill workers; and (d) by quantiles of income.

B Firm Location, Creation, and Geography

To conduct different empirical exercises, I gather a rich set of (i) administrative data on the creation date and the location of French firms' establishments (from the exhaustive SIRENE database). Firms' establishment locations allow me to identify mono-establishment or mono-city firms, i.e., firms that have only one establishment, or several establishments located in a single city. This is a key feature of my identification strategy, as it allows to identify precisely the time of BI connection. The year of firm creation is useful to study firm entry, time to credit, and first banking partner choice (see Appendix E.2).

I gather geographic data for mainland France on more than 36,000 municipalities, 2,000 urban units ("Unité urbaine"), and 762 urban areas. An urban unit (or city) is a commune or group of communes that includes in its territory a built-up area of at least 2,000 inhabitants where no dwelling is separated from the nearest one by more than 200 meters. In addition, each municipality has more than half of its population in this built-up area. The largest geographic unit that I consider in this paper is the urban area, which is defined as a group of municipalities, all in one piece and without enclaves, consisting of an urban pole with more than 10,000 jobs and rural municipalities or urban units where at least 40% of the employed resident population works in the pole or in municipalities nearby. Finally, municipalities are the finest unit of measurement that I use for distance computation and firm and branch locations. In the city-level part of my empirical analysis, I aggregate branch credit exposure, distance to clients, etc., at the urban unit level.

C Derivation of the gravity equation for credit flows

Lenoir et al. (2018) show that using the law of large numbers, Π_{uv} is equal to the expected value of $\mathbb{I}\{\mathrm{M}(e_v) = u\}$ across entrepreneurs in v, which is the probability that the best contract terms offered to any entrepreneur in v comes from branch u. Here, a crucial condition is that random variables $\mathbb{I}\{\mathrm{M}(e_v) = u\}$ are independent and identically distributed, which is straightforward if entrepreneurs are ex-ante identical. In my case, with entrepreneur heterogeneity, I show that this condition holds as the likelihood to ultimately match with submarket u does not depend on z_{e_v}. Thus, equation 7 writes:

$$\Pi_{uv} = \mathbb{E}_{e_v}\Big[\mathbb{I}\{\mathrm{M}(e_v) = u\}\Big] = \mathbb{P}\Big[\{\mathrm{M}(e_v) = u\}\Big]. \tag{24}$$

I consider a level of price r (and a level of productivity z_{e_v}) fixed effects. Let $D_{u,e_v}(r)$ be the number of branches from u met by entrepreneur e_v that proposes an interest rate below r, formally $D_{u,e_v}(r) = z_{e_v}^{\theta+1} r^\theta S_u (c_u d_{uv})^\theta \kappa_{e_v}$. Then,

$$\mathbb{P}\Big[\{\, \mathrm{M}(e_v) = u\}|r\Big] = \frac{\int_{D_{e_v}(r)>0} \mathbb{P}\Big[\{\, \mathrm{M}(e_v) = u\}|r, D_{e_v}(r)\Big] \,\mathrm{dF}(D_{e_v}(r))}{\int_{D_{e_v}(r)>0} 1 \,\mathrm{dF}(D_{e_v}(r))}. \tag{25}$$

In equation 25, the numerator can be interpreted as the discrete sum over all the total possible number of draws $D_{e_v}(r) = n$, of all the possible combinations of draws from u, i.e., n_u, among n. This leads to

$$\begin{aligned}
\mathbb{P}\Big[\{\, \mathrm{M}(e_v) = u\}|r\Big] = \\
&\sum_{n=1}^{+\infty} \sum_{n_u=0}^{n} \Bigg[\mathbb{P}\Big[\{\, \mathrm{M}(e_v) = u\}|r, D_{u,e_v}(r) = n_u, D_{k\neq u,e_v}(r) = n - n_u\Big] \\
&\times \mathbb{P}\Big[D_{u,e_v}(r) = n_u\Big] \times \mathbb{P}\Big[D_{k\neq u,e_v}(r) = n - n_u\Big]\Bigg] \\
&\times \mathbb{P}\Big[D_{e_v}(r) > 0\Big]^{-1}.
\end{aligned} \tag{26}$$

After some calculations, I obtain the following equation for Π_{uv}, at r fixed:

$$\begin{aligned}
\mathbb{P}\Big[\{\, \mathrm{M}(e_v) = u\}|r\Big] = &\frac{z_{e_v}^{\theta+1} \kappa_{uv} S_u (c_u d_{uv})^{-\theta}}{z_{e_v}^{\theta+1} \sum_{k=1}^{N} \kappa_{kv} S_k (c_k d_{kv})^{-\theta}} \\
&\times 1 - \exp\Bigg(- r^\theta z_{e_v}^{\theta+1} \sum_{u=1}^{N} S_u \cdot (c_u d_{uv})^{-\theta} \kappa_{uv}\Bigg) \\
&\times \mathbb{P}\Big[D_{e_v}(r) > 0\Big]^{-1}.
\end{aligned} \tag{27}$$

Note that $\mathbb{P}[D_{e_v}(r) > 0] = \mathbb{P}[r_{e_v} < r]$, the probability for the minimum price quote to be lower than p, for which an analytical formula is given by equation 6, leading to the following simplification for equation 26:

$$\mathbb{P}\Big[\{\, \mathrm{M}(e_v) = u\}|r\Big] = \frac{z_{e_v}^{\theta+1} \kappa_{uv} S_u (c_u d_{uv})^{-\theta}}{\sum_{k=1}^{N} z_{e_v}^{\theta+1} \kappa_{kv} S_k (c_k d_{kv})^{-\theta}}. \tag{28}$$

D The Rise of Online Banking

In the early 2000s, the large diffusion of ICTs caused a profound change for the banking industry, and broadband internet was the catalyst for this numerical transformation. As digitization proceeded apace, the dimensions of banks' digital evolution varied among segments and products. Digitization became the norm for retail credit processes; personal loan applications could be submitted with a few swipes on a mobile phone, and individuals could obtain cash in as little as a few minutes. Not only did transaction costs decrease, but the rise of online price comparison services and brokers also allowed individuals to search for the best banking partner in a faster and more efficient way.[21] The combination of both transaction and search cost reductions resulted in a severe disruption of search frictions for individuals.

For corporate credit and SME lending, the situation has been mixed. The loan officer remains the most relevant intermediary and the ultimate decision maker in SME lending, as he is ideally suited to understand clients' specific needs and characteristics as well as local market and industry performance, leaving little room for automation. The complexity of loan pricing for SMEs also prevents the use of online brokers or interest rate comparison websites. However, ICTs affected many aspects of the firm-bank relationship, especially for SMEs: a recent survey on UK SMEs (Ernst and Young, 2018) shows that financial services used by SMEs are mainly online banking (85%) followed by branch-based banking (43%), emphasizing the growing importance of digitization. Among other examples, emails allow firms to easily contact a loan officer, and online data and document sharing speeds up the meeting process and reduces transaction costs. Bank websites are showcases designed to attract new customers and provide information about financial products. Finally, customer portals facilitate communication.

To document that French banks had started their digital transition process at the beginning of the 2000s, I gather new data on the large French bank adoption of ICTs. I first check the existence of the bank website with a customer portal before 2000 (using the waybackmachine.com website). I then collect the exact date of the domain name creation (available on nom-domaine.fr). The six largest French banks, which represent around 90% of the total amount of credit granted to firms, were already active online at the beginning of the 2000s, with a sophisticated website, while their domain names were registered on average in the mid-90s, showing an early preoccupation for online visibility. In particular, the average website shows the willingness of banks to improve the accessibility of basic but crucial information for their future SME clients: in a few clicks, it was possible for a new client to get an appointment, find all the bank branches in the area, and ask for information about financial products and services. For each bank branch, the phone number, contact email, and physical address were immediately available. This represents a sharp reduction in search costs borne by the entrepreneur.

[21] Mortgage lending is more complex due to regulatory constraints, yet banks in many developed markets have managed to digitize large parts of their mortgage journey. More than one bank has aspired to automate 95% of retail underwriting decisions.

E Empirical Facts about Corporate Credit Markets

In this appendix section, I detail the body of novel facts about French credit markets that suggest the presence of search frictions. I first document bank branch heterogeneity and endogenous firm-branch matching. I then describe the geography of credit flows and provide evidence of substantial price dispersion. Finally, I mirror those empirical facts with new survey evidence.

E.1 Bank Branch Heterogeneity

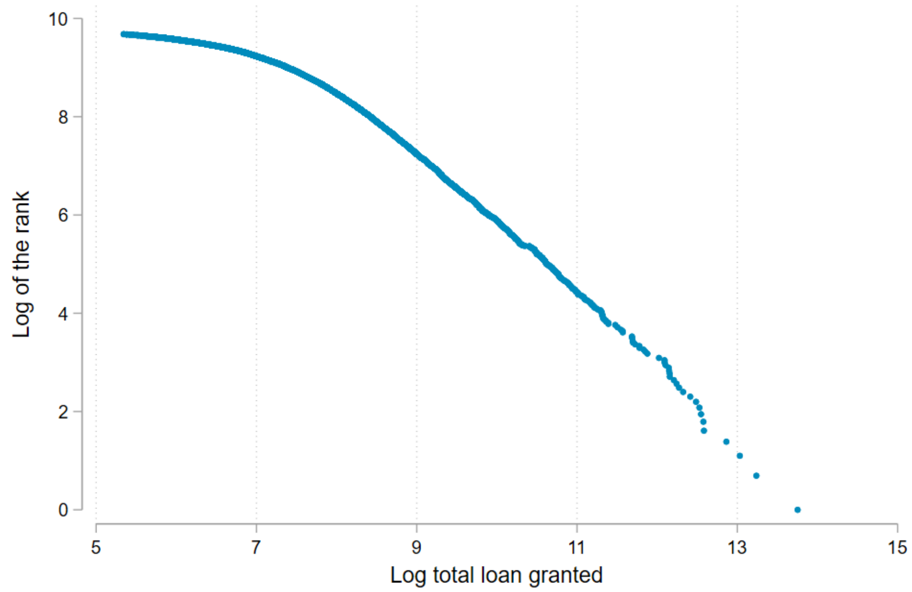
In a highly competitive and decentralized banking sector, large national banks compete locally through their branch networks, across multiple geographic submarkets. Local bank branches and loan officers are therefore the main contact point for entrepreneurs searching for the right banking partner, and branch office characteristics appear as critical factors to firms, especially SMEs, when choosing their financial services providers. While prior literature (Berger et al., 1997) shows that the matching of firms and banks is endogenous and depends on firm and bank characteristics, little evidence exists on the importance of branch characteristics. In this section, I focus on bank branch—rather than bank—heterogeneity. I document four important facts: branches differ markedly from each other with respect to (i) their total credit exposure, (ii) their average distance to clients, (iii) the number of markets in which they operate, and (iv) their portfolio specialization.

Branch Size. Figures 1 and E1 display the distribution of the branch size for the last quarter of 2005. I compute the size of a branch as its total credit exposure (1) and, alternatively, its number of clients (E1). I then rank branches by size,[22] with #1 being the largest branch, #2 the second largest, and so on. Finally, I plot the log (size) versus the log (rank).
The relationship between log size and log rank is close to a straight line, and the slope is very close to 1. This indicates that the distribution of the branch size follows a power law (i.e., Pareto distribution): very few large branches grant credit to many firms ($\geq$ 10,000 clients), while a vast majority of small offices only finance 10 to 20 clients.

Branch-Firm Distance. Figure E2 shows the positive correlation between branch size (measured as total credit exposure and, alternatively, as the number of clients) and the average square geographic distance between the branch and its clients, in kilometers, for the last quarter of 2005. The average squared distance from a branch to its clients is a power function of the branch size: the bigger the branch, the larger the geographic distance. Bank branches differ in their role and ability to finance remote clients. Smaller branches focus on their very local market and

[22] I regress the log rank on log size and find $\log(\text{Rank}) = 12.29 \underset{[0.002]}{\mathbf{-1.04}} \cdot \log(\text{Size})$. The relationship is close to a straight line (R^2=0.95), and the slope is very close to 1 (the standard deviation of the estimated slope is 0.02). This means that the rank of a bank branch is essentially proportional to the inverse of its size.

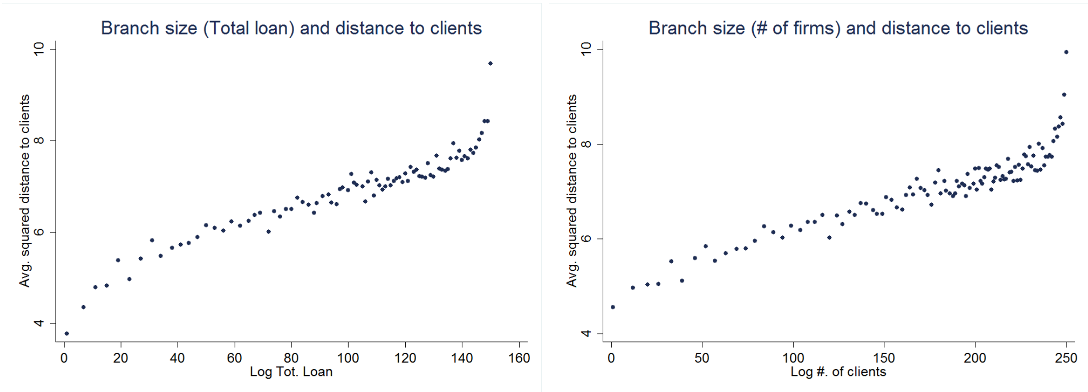
Figure E1: Branch Rank versus Size (Total credit)



Notes: This figure displays the distribution of bank branch size for the last quarter of 2005, for all banks with at least five clients. Formally, it shows the scatter plot of the log (size) against the log (rank). I compute the size of a branch as its number of clients. Bank branches are ranked by size, with #1 being the largest, #2 the second largest, and so on.

concentrate on proximity lending (branches with less than 10 clients are, on average, 15.5 km away from them).

Figure E2: Branch Size and Average Distance to Clients

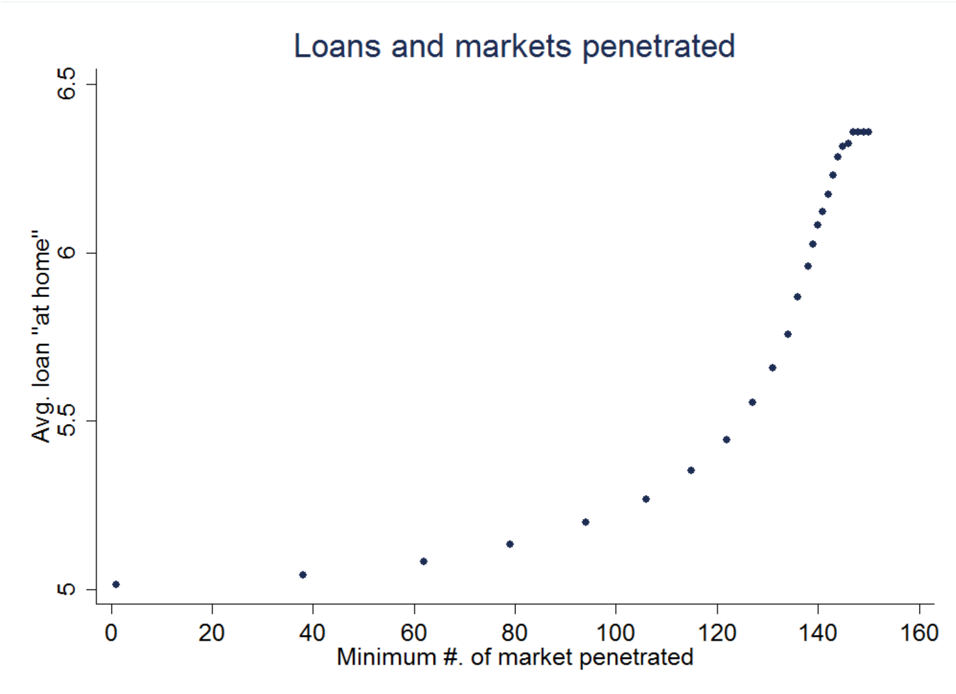


Notes: This figure shows the positive correlation between branch size measured as total credit exposure (left panel) and, alternatively, as the number of clients (right panel) and average square geographic distance between the branch and its clients, in kilometers, for the last quarter of 2005.

On the contrary, the largest branches grant credit to firms located in other cities or regions (branches with more than 100 clients (top 1%) are, on average, 80 km away from them). This heterogeneity may be the result of banks' internal organization and may simply reflect the hierarchy between local branch offices and large business centers. It may also be driven by branch productivity (larger branches being able to screen and monitor more efficiently remote clients), specialization, or local economic conditions, but this question remains beyond the scope of this paper.

Branch Size and Market Entry. To examine how a bank branch's ability to finance distant clients relate to its local activity, I rank and group branches based on the number of remote submarkets (i.e., cities) where they operate. All the branches lend at least to one submarket, while none are active in all. Figure E3 depicts the average credit size in local branch submarkets ("at home") for the group of branches that operate at least in *k* remote submarkets, with *k* on the x-axis.

Figure E3: Average Local Loan and the Number of Distant Markets Penetrated



Notes: This figure displays the log average credit size of a branch in its local submarket ("at home") for the group of branches that operate at least in *k* remote submarkets, with *k* on the x-axis. Bank branches are ranked and grouped based on the number of remote submarkets (i.e., cities) where they operate: all the branches lend at least to one submarket, while none are active in all.

The relationship between the average credit exposure in the local branch submarket and the number of remote submarkets penetrated is strongly positive and corroborates the findings of Figure E2. The larger a branch is locally, the higher the number of distant submarkets served. This striking regularity reveals another facet of the bank branch heterogeneity that relates first to the likelihood, for a firm, of meeting with a bank branch located outside its local submarket and, second, to the ability of a loan officer to remotely screen and monitor this firm.

Bank Branch Specialization. Using similar data and methodology as Duquerroy et al. (2019), I show that within a bank and a city, branches specialize in several dimensions: industry, type of loans, and type of businesses. In particular, some branches appear to heavily finance SMEs compared to other branches located in the same submarket and belonging to the same banking network. I define a branch to be specialized in an industry (respectively, size category) if its portfolio share of lending to firms in an industry (respectively, a size category) is a right-tail outlier in the distribution of portfolio shares of lending by all branches within the same city.

E.2 Endogenous Firm-Branch Matching

In Section E.1, I showed that bank branches markedly differ from each other in several dimensions. Thus, search frictions may arise if firms pay attention to those branch characteristics and devote time and resources to locate and match with the right loan officer. To examine if this observed heterogeneity affects how firms search for the right banking partner and if it leads to endogenous matching, I investigate the importance of branch characteristics in new firm matching decisions. I use a new dataset of firm creations from 2000 to 2005 for mono-establishment firms that I combine with information from the Credit Register about firms' first realized banking match. The number of firm creations in my sample range from 18,675 in 2002 to 26,865 in 2005. After their establishment, firms match with a single bank branch: the period of time running from firm entry to first banking match is 15 months on average. However, close to 50% of those new firms find their banking partner within their first 8 months of existence. For the sake of simplicity, I keep in my final sample of firms that match with their first banking partner in less than five years, which represents 90% of the observations. I end up with 137,175 new firms entering 711 distinct submarkets and matching with 12,952 bank branches. For each of those new firms, I build the universe of possible banking partners, which I define as all the existing branches located in the same urban area that are active the year of entry. Over all possible pairs, I observe only a single realized match.

I present reduced-form evidence of endogenous firm-branch matching: a new firm i, when entering a local credit submarket, is more likely to match with a bank branch j that ex-ante shares some common specialization and size characteristics with the firm. In particular, conditional on distance, a newly created SME from sector s is more likely to borrow from a branch that specializes in lending to small firms or to firms belonging to sector s. Additionally, I find that larger and growing branches appear to easily attract and match with new clients.
Formally, I estimate a probit regression of different specifications of the following equation:

$$\begin{aligned}\mathbb{P}(\text{ Match}_{i,j,t} \mid \text{observables }) = \Phi\Big(&\alpha \cdot \text{Distance}_{i,j} \\ &+ \beta_1 \cdot \text{Size}_{j,t-1} + \beta_2 \cdot \text{Size Growth}_{j,t-1} \\ &+ \beta_3 \cdot \text{Spec. Sector}_{j,t-1} + \beta_4 \cdot \text{Spec. Size}_{j,t-1} \\ &+ \text{Controls}_{i,j,u,t}\Big), \end{aligned} \tag{29}$$

where Φ is the c.d.f. of the standard normal distribution and Spec. Sector$_{j,t-1}$ (respectively, Spec. Size$_{j,t-1}$) equals 1 if branch j is specialized in lending to firms from the same sector (respectively, same size) as firm i, the year before firm i's entry. Size$_{j,t-1}$ is the log size of branch j, measured as total credit exposure and, alternatively, as the number of clients. Size Growth$_{j,t-1}$ is the growth rate of branch j, both measured the year before firm i's entry. Distance$_{i,j}$ is simply the log geodesic distance in kilometers between branch j and firm i. All explanatory variables are lagged so that the branch j characteristics are measured before firm i's entry and are not contaminated by the realized firm-branch matching. The dependent variable is an indicator function that equals 1 if firm i is borrowing from branch j at time t. The coefficient α controls for the direct impact of proximity on the likelihood of firm-branch matching: physical proximity is likely to alleviate informational frictions affecting banks' screening and monitoring costs as being close to clients eases the acquisition and the use of private information in informationally opaque credit markets. Coefficients β_1 and β_2 control for heterogeneity in the size and dynamics of bank branches documented in Section E.1. $\beta_1 > 0$ and $\beta_2 > 0$ suggest that large and fast-growing branches are likely to be more efficient or visible and, conditional on distance, to offer better contract terms. Finally, coefficients β_3 and β_4 control for the direct impact of branch specialization (industry or size category) on firms' choice of banking partner. $\beta_3 > 0$ (respectively, $\beta_4 > 0$) means that a given firm i is more likely to match with a bank branch j that has developed ex-ante a specific advantage in lending to firm i's industry (alternatively, the firm' i size category).

Table E1 shows the marginal effects from the probit estimation of different specifications of equation 29. Standard errors are clustered at the city level. In every specification, all coefficients are statistically significant (at the 1% confidence level) and of the expected signs. Physical proximity (α), branch size (β_1), and branch growth (β_2) increase the likelihood that a firm matches with a bank branch. More interestingly, the actual existing portfolio of a branch shapes its future matches: a branch that specializes in lending to industry a is more likely to be chosen by a firm from industry a when it enters the credit market. To conclude, these results suggest bank branch heterogeneity matters for small firms' financing decisions; SMEs are more likely to search for and then match with branches that ex-ante exhibit a high level of complementarity.

E.3 Price Dispersion

Inspired by the labor market literature, I find that the law of one price does not hold in French credit submarkets. Using rich quarterly micro data on new loans to SMEs from the Sirius/M-Contran database, I find that credit rates exhibit a substantial dispersion within a time-bank branch-industry-department quadruplet. Controlling for loans and borrower characteristics does not affect my result. Formally, I estimate equation 30, which aims to explain the observed

Table E1: Endogenous Firm-Branch Matching

	(1)	(2)	(3)	(4)	(5)
Log Branch-Firm Distance (km)	-0.311***		-0.298***	-0.300***	-0.317***
	(0.001)		(0.001)	(0.001)	(0.001)
Branch size (t-1) × Firm size		0.126***	0.083***	0.081***	0.137***
		(0.001)	(0.001)	(0.001)	(0.001)
Branch growth (t-1)			0.108***	0.197***	0.233***
			(0.004)	(0.005)	(0.005)
Branch Industry spec. (t-1)				0.312***	
				(0.009)	
Branch Size spec. (t-1)					0.875***
					(0.009)
Observations	9,645,373	9,645,373	9,645,373	9,645,373	9,645,373
McFadden R-square	0.15	0.04	0.17	0.17	0.19

Notes: The table shows a probit estimation of different specifications of equation (29). The dependent variable is an indicator function that equals 1 if firm i is borrowing from branch j at time t. Spec. Sector$_{j,t-1}$ (respectively, Spec. Size$_{j,t-1}$) equals 1 if branch j is specialized in lending to firms from the same sector (respectively, same size) as firm i, the year before firm i's entry. Size$_{j,t-1}$ is the log size of branch j, measured as total credit exposure and, alternatively, as the number of clients; and Size Growth$_{j,t-1}$ is the growth rate of branch j, both measured the year before firm i's entry. Distance$_{i,j}$ is simply the log distance between branch j and firm i. All explanatory variables are lagged. Standard errors are clustered at the city level.

variation in loan prices:

$$\text{Interest rate}_{ijtu} = \text{Loan}_{ijtu} \cdot \rho_1 + \text{Firm}_{ijtu} \cdot \rho_2 + \text{FE}_{s(u)} + \text{FE}_j + \text{FE}_t + \text{FE}_u + \epsilon_{ijtu}, \qquad (30)$$

where i stands for the borrower, j for the bank branch, t for the quarter, and u for the urban area in which both the firm and the branch operate. Loan$_{ijtu}$ is a vector of loan characteristics (term in months, amount, type of rate), Firm$_{ijtu}$ is a vector of firm i characteristics (age, size, debt, investment grade, turnover). I sequentially add a bank (alternatively, a branch) fixed effect, county fixed effect (i.e., French *départments*), a sector (NACE Rev. 2 French classification) fixed effect, and a quarter fixed effect. Table 2 shows the R^2 of an OLS regression for different specifications of equation 30 and for three categories of credit: equipment loans, credit lines, and leasing. The results indicate that, at best, the model accounts for 70% of the observed variance in credit prices, causing more than 30% (40% for leasing) of the variance to remain unexplained even when the model is saturated.

F PPML: Difference-in-Difference with Many Zeros

In this Appendix, I test the performance of the PPML estimator in the exact case of my empirical setting. Not only does my estimating sample contain a vast majority of city pairs that do not exchange credit over the entire period (more than 95% of the credit shares equal zero), but I use panel data and a difference-in-difference approach. I present simulation evidence on the performance of the PPML estimator when the panel data are generated by a constant elasticity model, with (i) a large proportion of zeros, (ii) a time-varying shock and, (iii) when all units are not simultaneously treated. In these simulations, the non-negative dependent variable Y_{uvt} is generated so that $\mathbb{P}(Y_{uvt} = 0)$ is substantial and $\mathbb{E}(Y_{uvt}|X_{uvt})$=exp($\beta X'_{uvt}$). To the best of my knowledge, this is the first simulation evidence of the performance of the PPML estimator in this particular setting, which echoes a wide range of papers in international trade that investigate the impact of trade policies within a gravity framework. Following Santos Silva and Tenreyro (2011), the dependent variable Y_{uvt}—which is the total credit granted by branches in v to firms in u—is generated as a finite mixture model of the form Y_{uvt} =$\sum_{j=1}^{m_{uvt}} z_{juvt}$, where m_{uvt} is the number of components of the mixture, and z_{juvt} a continuous random variable with support in $\mathbb{R}$+, distributed independently of m_{uvt}. This data generation process has a direct economic interpretation in my framework. m_{uvt} is the number of branches located in v that serve firms in u, and z_{juvt} is the amount of credit that each of those banks lent their clients located in u. Because m_{uvt} and z_{juvt} are independent, $\mathbb{E}(Y_{uvt}|X_{uvt}) = \mathbb{E}(m_{uvt}|X_{uvt}) \times \mathbb{E}(z_{juvt}|X_{uvt})$. As in Santos Silva and Tenreyro (2011), z_{juvt} is obtained from a gamma distribution with mean 1 and variance 2, which is equivalent to a $\chi^2_{(1)}$ random variable. This implies that conditional on m_{uvt}, Y_{uvt} follows a $\chi^2_{m_{uvt}}$ and then $\mathbb{E}(Y_{uvt}|X_{uvt}) = \mathbb{E}(m_{uvt}|X_{uvt})$. m_{uvt} will be generated as a negative binomial random variable with conditional mean exp($\beta X'_{uvt}$) and a variance equal to $a\mathbb{E}(m_{uvt}|X_{uvt}) + b\mathbb{E}(m_{uvt}|X_{uvt})^2$. I propose the following functional form for $\mathbb{E}(Y_{uvt}|X_{uvt})$:

$$\mathbb{E}(Y_{uvt}|X_{uvt}) = \mathbb{E}(m_{uvt}|X_{uvt}) = exp(\beta_0 + \beta_1 x_{1uv} + \beta_2 x_{2uvt} + \beta_3 x_{3uvt}),$$

where x_{1uv} is the product of the sum of two time-invariant variables drawn from a standard normal distribution x_{1u} and x_{1v}. x_{2uvt} is a time-varying variable equal to 1 with a probability close to 0.4. Formally, I impose a dynamic structure for x_{2uvt} by introducing the underlying variable $w_{2uvt} = \gamma w_{2uvt-1} + \rho\epsilon_{uvt}$, with $\gamma = 1.05$, $\rho = 1.5$, and ϵ_{uvt} is drawn from a standard normal distribution such that $x_{2uvt} = \mathbb{I}(w_{2uvt} > 0.6)$. x_{3uvt} is a city pair-specific treatment dummy variable that equals 1 after the time of treatment t and 0 before. The city pair-specific treatment time is drawn from a discrete uniform distribution over the support $[t_0, T]$. Formally, I generate city-specific treatment date $x_{3u} \sim U[t_0, T]$ and define $x_{3uvt} = \mathbb{I}(t > \max(x_{3u}, x_{3v}))$. Finally, I impose β_0 = 0, β_1 = -1, β_2 = 1, and β_3 = 0.1. This functional form has, again, a direct economic interpretation. The time-invariant variable x_{1uv} (which is symmetrical by definition, i.e., x_{1uv} = x_{1vu}) represents the distance between u and v, x_{2uvt} models the time-varying determinants of credit flows between cities, and x_{3uvt} is analogous to my broadband internet interconnection shock. To complete my simulation setting, I need to define the conditional variance of m_{uvt} and Y_{uvt}. I follow Santos Silva and Tenreyro (2011) and consider the quadratic specification

Table F1: PPML WITH MANY ZEROS IN A DYNAMIC SETTING: SIMULATION RESULTS

Data Generating Process I $\mathbb{P}(Y_{uvt} = 0) = 0.97$	Obs. = 200 x 200 x 10, a = 50 b = 0					
Estimator	β_1		β_2		β_3	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
PPML	-1.003	0.029	1.001	0.037	0.226	0.033
PPML w. D×T + O×T FE	-1.006	0.029	1.006	0.084	0.121	0.099
GPML	-0.874	0.034	1.008	0.039	0.185	0.035
GPML w. D×T + O×T FE	-1.019	0.03	0.995	0.086	0.093	0.101
Data Generating Process II $\mathbb{P}(Y_{uvt} = 0) = 0.88$	Obs. = 200 x 200 x 10, a = 1 b = 5					
Estimator	β_1		β_2		β_3	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
PPML	-1.001	0.013	1.001	0.017	0.226	0.015
PPML w. D×T + O×T FE	-1.001	0.013	0.999	0.038	0.115	0.045
GPML	-0.872	0.018	0.998	0.019	0.187	0.017
GPML w. D×T + O×T FE	-1.005	0.015	0.996	0.041	0.11	0.04

Notes: This table presents the results obtained with 1,000 replicas of the simulation procedure described above. The number of cities N is set to 200 and T equals 10. The estimation sample is therefore composed of 200 × 200 × 10 = 400,000 observations. In the top panel, $a = 50$ and $b = 0$, while a is set to 1 and b to 5 in the bottom panel. The table displays the average point estimate and the standard errors obtained with different estimators, namely PPML and GPML with and without origin × year and destination × year fixed effects. Robust standard errors are in parentheses.

$\mathbb{V}ar(m_{uvt}) = a\mathbb{E}(m_{uvt}|X_{uvt}) + b\mathbb{E}(m_{uvt}|X_{uvt})^2$ so that:

$$\mathbb{V}ar(Y_{uvt}|X_{uvt}) = (1 + 2a)\mathbb{E}(m_{uvt}|X_{uvt}) + 2b\mathbb{E}(m_{uvt}|X_{uvt})^2.$$

Picking the value of a and b allows me to generate a high probability of zeros and different heteroskedasticity patterns. Table F1 presents the results obtained with 1,000 replicas of the simulation procedure described here, in which the number of cities N is set to 200 and T equals 10. The estimation sample is therefore composed of 200 × 200 × 10 = 400,000 observations. In the top panel, $a = 50$ and $b = 0$, while a is set to 1 and b to 5 in the bottom panel. The table displays the point estimate and the standard errors obtained with the different estimators, namely PPML and GPML,[23] with and without origin × year and destination × year fixed effects. These results confirm and extend the findings of Santos Silva and Tenreyro (2006) and Santos Silva and Tenreyro (2011), showing that both the PPML and the GPML estimators are well behaved in the two cases considered. In particular, the coefficient of interest β_3 is consistently estimated with two-way unit × fixed effects. These findings justify my empirical approach and validate my estimation procedure.

[23]GPML is a PPML-like estimation procedure in which the dependent variable is a share instead of the value in level.