# Information Frictions in Credit Markets *

Clément Mazet-Sonilhac †

[Latest version here]

January 6, 2022

**Abstract**

Motivated by empirical evidence I document on local credit markets using data on more than 3.5 million bank-firm relationships, I develop a theory of bank-firm matching, subject to search frictions. Firms undergo a costly search process to locate and match with the right banking partner. Upon matching, banks incur a cost to screen projects. I estimate structurally my model on French data using the staggered roll-out of broadband internet, from 1997 to 2007, as a shock to search frictions. I confirm the model predictions that a reduction in search frictions affects the allocation of credit and firm-bank matching. Finally, I use the structure of my model to quantify the impact of this technology-induced reduction in search frictions on loan prices. I find that broadband access reduced the cost of debt for small businesses by 4.9%.

**JEL classification**: L22, L23, D83
**Keywords:** Search frictions, Broadband internet, Firm-bank matching.

*I am grateful to Philippe Aghion, Antonin Bergeaud, Johannes Boehm, Cecilia Bustamante, Thomas Chaney, Federico Ciliberto, Anne Duquerroy, Michele Fioretti, Stéphane Guibaud, Basile Grassi, Michele Fioretti, Mathias Lé, Clément Malgouyres, Thierry Mayer, Isabelle Méjean, Daniel Paravisini, Jean-Marc Robin, Xavier Ragot, Sandra Poncet, Xavier Vivès, and participants at SciencesPo, GSIE, INSEAD/Collège de France and Banque de France seminars for helpful comments. The views expressed in this paper are my own and do not necessarily reflect those of the Banque de France or the Eurosystem.

†Sciences Po, Banque de France, Collège de France. Email: clement.mazet@sciencespo.fr

# 1 Introduction

It is costly for firms to locate and match with the right banking partner, especially for small and medium-sized enterprises (SMEs) who devote time and resources to this search process. Small firms commonly multiply loan applications (2.7 on average) and undergo a time-consuming application process: over 33 hours are spent on average on loan request paperwork.[1] Overall, about one third of SMEs deplore a difficult and lengthy credit application process.[2] While the effect of information asymmetries on credit allocation is well documented (e.g., Akerlof, 1970; Stiglitz and Weiss, 1981; Petersen and Rajan, 1995), little is known about how search frictions affect bank-firm matching and access to credit. Understanding the role of search frictions is particularly important to policy makers, not only as recent developments in information and communication technology and digitisation in the banking industry are likely to affect them, but also as policies that reduce search costs may differ substantially from policies that reduce traditional agency frictions.

This paper builds 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. It uses 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 frictions. Finally, I use this unique data and the structure of the model to quantify the impact of reducing search frictions on the degree of misallocation in the credit market for SMEs. I find that Broadband Internet access increased bank competition and reduced the cost of debt for small businesses by 4.9%.

The first contribution of the paper is to offer a theory of firm-bank matching subject to informational frictions. I develop a partial equilibrium model of the matching between firms and banks that features two-sided heterogeneity – bank branches and firms – and information frictions. First, search frictions hinder firms' ability to locate and match with the right financing partner. Formally, I build on Lenoir et al. (2018) where the search and matching process is random and depends on the level of search frictions between two submarkets.[3] Second and upon matching, banks incur a cost to screen projects and learn about firms' quality. Screening costs are a function of

[1] see the report *Small Business Credit Survey* from the Federal Reserve System (FED, 2014)
[2] From the survey by Infosys *Banks! It's time to change your game in SMEs lending* (Infosys, 2018)
[3] In this paper, the concept of *search frictions* encompasses both search and transactions costs.

branch productivity and the physical distance between a bank branch and a firm, which ultimately determines the interest rate quote. Those assumptions are motivated by novel empirical evidence on the corporate credit market in France I document using using bank-firm matched data from 1998 to 2005 (more than 3.5 million bank-firm relationships). Banks branches are heterogeneous in size and distance to their clients, with larger bank branches serving firms located further away and in more submarkets. Moreover, bank branches characteristics matter for the matching with new firms: I document the existence of positive assortative matching between bank branches and firms. Finally, corporate credit markets exhibit a high level of price dispersion that suggests a key role for search frictions. By adding structure to the search process, I generate a number of theoretical 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) the dynamic of firm-bank matching. In particular, the model predicts a gravity structure for credit flows between cities. When search costs decrease, the gravity equation is distorted: firms meet with more potential lenders and eventually borrow credit at a lower rate.

The second contribution of the paper is to propose a novel instrumental variable strategy for the timing of Broadband Internet diffusion. Broadband Internet operators relied on already existing infrastructures - local copper loops networks used for phone calls transmission and larges optic fiber cables used by railways and highways companies - to gradually deploy the ADSL technology, with one objective: to connect the maximum number of customers while minimising the cost of investment. Accordingly, I solve a simple maximisation 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, it is the interaction between pre-existing infrastructures with the density of population, controlling for firm density, which predicts the timing of the arrival of Broadband Internet. In other terms, 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 banks' branching strategy, investment and credit supply to small businesses (or, conversely, to small firms' demand for credit). It provides natural ground for an event-study identifying how information and communication technologies affect credit markets, through a reduction in search frictions.

The third contribution of the paper is to structurally estimate the model. Using my instrument variable for the staggered diffusion of Broadband Internet in France, from 1999 to 2005, as a shock that reduced search frictions, I can estimate parameters $\theta$ and $\gamma$ – 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. In practice, I first verify that credit flows between cities follow a gravity equation that is distorted by the staggered roll-out 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. On the contrary, the effect is negative when two neighbouring cities, already very closely tied economically, are connected by internet. For robustness, I provide simulation evidence of the performance of the Poisson pseudo-maximum likelihood (PPML) estimator for the estimation of gravity equations in a dynamic setting and with many zeros, extending the work of Santos Silva and Tenreyro (2006, 2011). I confirm that PPML estimator with fixed-effects is well behaved and consistently estimates the time-varying treatment effect, even with more than 90% of zeros. I also show that my baseline results are barely affected by the addition city pair fixed-effects and of the lag dependent variable as covariate.

My main results rely on regressions carried out at the urban unit level. This geographical level simply matches the level of the treatment and allows me to deal with 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 do not affect my estimates. This empirical approach provides a causal interpretation for a number of recent developments observed in the banking market: Kroszner and Strahan (1999) shows how the large-scale adoption of information technology reduced the dependence on geographical proximity between customers and banks, and Petersen and Rajan (2002) documents the erosion of the local nature of small business lending, with increasing distance between small firms and their lenders in the United States, and also new communication habits.[4]

Finally, I assess the implications of my findings on the cost of debt for small firms, using the structure of my model. In practice, I plug my empirical estimates for $\theta$ and $\gamma$, 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 medium-sized cities than in the largest French cities. Firms initially located far from bank branches, or which did not have a wide variety of potential banking contacts, benefited more from the reduction of search frictions, since it allowed them to match with new or better banking partners. In this respect, the spread of Broadband Internet reduces spatial inequalities in access to credit and foster local bank competition in rural and isolated submarkets. My results also have a natural interpretation in terms of misallocation. As search frictions generate misallocation of credit entirely orthogonal to risk (contrary to information asymmetries that induce credit rationing on observables), a reduction in search frictions reduce the cost of debt for every firm.

**Related literature**. This paper contributes to the literature on corporate finance and search and matching.

There is a vast theoretical and empirical literature in corporate finance and microeconomics 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).[56]

[4]The large diffusion of information and communication technologies represented a profound change for the banking industry and Broadband Internet was the catalyst for this numerical transformation. As digitization proceeded apace, transaction costs decreased and the rise of online financial services allowed firms to search for the best banking partner in a faster and more efficient way, leading to structural changes in banking markets. Those developments are prominent in France: inter-regional credit flows have grown by 15% and the average firm-bank distance has increased by 10% between 1998 and 2005. My empirical approach provides a causal interpretation for those facts, suggesting that innovations in information technology – namely, Broadband Internet diffusion – have reduced the role of both transaction and search costs in shaping credit outcomes, allowing firms to search for credit further afield and leading to structural changes in local credit markets.

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

[6]This paper is also directly related to the literature that explores the determinants of bank-firm matching. Firm and bank size appear as key characteristics (Stein, 2002; Hubbard et al., 2002; Cole et al., 2004; Berger et al., 2005), along with geographic proximity (Petersen and Rajan, 1995, 2002), export country specialization (Paravisini et al., 2015), monitoring capacity (Jing, 2014) and bank capitalization (Schwert, 2018). However, little evidence exists on the importance of branch characteristics, despite the fact that bank branches and loan officers are the main contact point for SME searching for the

This literature focuses on asymmetric information and has long highlighted the role of bank-firm relationships in alleviating agency frictions that shape credit supply (for surveys see, e.g., Boot, 2000; Degryse et al., 2009; Udell, 2015). Yet, the role of search frictions, in particular in the formation of bank-firm relationships, has been overlooked despite its potentially equal importance.[7] I propose 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 emphasize 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 agency frictions that I structurally take to the data.[8] As far as I know, this paper is the first one to account for search frictions in the formation of bank-firm relationships from a micro perspective and provide a causal evidence of how a reduction in search frictions affect the allocation of credit.[9] 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).

This study also belongs to a literature that studies ICT diffusion and bank competition. Hauswald and Marquez (2003) find that the *dissemination of information*, i.e. improved access to information, erodes informational advantages and increases the intensity of bank competition, ultimately reducing borrowers' expected loan rate. Hauswald and Marquez (2006) extend that model by allowing endogenous investment by banks in information processing technology and assuming that the bank–borrower distance to have a negative effect on the precision of banks' information. Similar to Vivès and Ye (2021), they find that more intense bank competition reduces banks' incentive to invest in information processing and that borrowers pay lower rates when

right banking partner (Berger et al., 1997). I contribute to this literature by leveraging bank branch-level data. I show that branch characteristics (size, distance to client, growth, specialization) matter for the matching with small firms. I use these empirical evidence as a motivation for my theoretical framework, where bank branches are heterogeneous in size and distance to clients.

[7]Macro-finance papers have studied the aggregate effect of search frictions in credit market, starting with den Haan et al. (2003) and Wasmer and Weil (2004). Recently, Argyle et al. (2019) and Allen et al. (2019) have shown how search frictions affect mortgage and consumer credit markets.

[8]I model loan prices are a function of agency frictions that depend on productivity and distance and that capture screening and monitoring costs (Degryse and Ongena, 2005; Hauswald and Marquez, 2006).

[9]Recently, Boualam (2018) builds a general equilibrium theory of bank lending relationships in an economy subject to search frictions and limited enforceability, in a related macro approach. The interaction between search and agency frictions generates a slow accumulation of lending relationship capital, distorts the optimal allocation of credit and leads to slow recoveries. While its approach features a directed search process, I model search as random. Not only this is for tractability purpose, but it is also consistent with the fact that entrepreneurs – who can't devote much time and ressources to this costly search – don't observe bank branches and loan officers characteristics before meeting with them.

they are located farther from the bank that screens them. The mechanism herein departs from this previous research: the ICT diffusion affects bank competition and credit allocation through a reduction of search frictions faced by firms. However, my model also predicts that ICT diffusion could also have an effect on competition and loan prices through a reduction in screening and monitoring costs, but this effect is absorbed by the city $\times$ time fixed-effects in the empirical analysis.

My model delivers a gravity equation for inter-regional credit flows which echoes a 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). Recently, Brei and von Peter (2018) estimated gravity equations on international banking flows using a PPML approach. They find a substantial role of distance for banking, even with immaterial transport cost, pointing to the role of information frictions. My paper complement and extend these findings by estimating a gravity equation for within-country bank credit flows with two types of informational asymmetries. It confirm the prominent role for distance, even for local credit markets, but also underlines the relevance of search frictions.

This paper also relates to the vast literature in labor (see for a survey, e.g., Rogerson et al., 2005) and trade (Rauch, 2001; 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 herein is in the spirit of trade models (Eaton et al., 2018; Lenoir et al., 2018) where firms undergo a random search process, it incorporates agency frictions that replace traditional iceberg transport costs. Finally, some recent papers explore how the diffusion of information and communication technologies – including Broadband Internet – affect such frictions (Allen, 2014; Lendle et al., 2016; Steinwender, 2018; Akerman et al., 2018; Malgouyres et al., 2021; Bhuller et al., 2019). 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. I 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 and the change in competition, eroding banks' rents. Second, I develop a novel IV for the timing of Broadband Internet diffusion. As the observed pace of 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.

The rest of the paper is organized as follows. Section 2 present the data. Section 3 presents the model and the main predictions that guide the empirical analysis. Section 4 presents the empirical context and details the instrument variable strategy. Section 5 describes my empirical methodology. Section 6 details the results. Finally, Section 7 studies the implications for the cost of debt through the lens of my model. Section 8 concludes.

## 2 Data

In this section, I provide a brief description of the data I use to study local credit markets in France and how they were affected by Broadband Internet diffusion over the period 1998-2005. I combine multiple proprietary data sets from the Banque de France about firm-bank relationships, branch-level credit exposure, interest rates of new loans, with unique data on Broadband Internet expansion at the city-level.

### 2.1 The Credit Register

The French Credit Register is a large data set of bank-firm linkages available at the Bank of France over the period 1998-2005. This credit register aims to collect data on bank exposures to residents on a monthly basis to monitor and control systemic risk. The monthly data comes 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, the bank's commitments on credit lines and guarantee as well as medium and long-term leasing, factoring and securitized loans at the branch level. Recipients are single businesses, corporations, sole-proprietorship engaged in professional activities. They 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 data set excluding individual entrepreneurs covers information on more than 1.9 millions of bank-firm relationships, corresponding to more than 1.4 million of unique firms or corporations (SIREN), 341 unique banks and 15,925 bank branches serving firms. The smallest banks in my sample own only one branch. In contrast, the largest one (i.e. Crédit Lyonnais) owns more than 2,800 branches all across the country. This data allows me to document novel facts on bank-branches heterogeneity that guide my theoretical assumptions . In particular, I show that bank branches differ markedly from each others with respect to (i) their total credit exposure and number of clients (see Figure 2), (ii) their average distance to clients (iii) the number of markets in which they operate and (iv) their portfolio specialization (see Appendix B.1).

## 2.2 Loan Rates (M-Contran)

The M-Contran database gathers quarterly survey about all new loans granted to French firms from 2006 Q1 to 2016 Q4. This information is collected by the Banque de France in order 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 institution. 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 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 survey provides rich information on a wide range of relevant loans characteristics, such as the size of the loan, the loan's precise purpose (investment, treasury, leasing etc.), its maturity at issuance, whether it is fixed-rate or adjustable rate, and whether it is secured or not.

Table 1: Explaining Price Dispersion

| $R^2$ | (1)<br>Time + Bank FE | (2)<br>(1) + Dep. FE | (3)<br>(2) + Sector FE | (4)<br>(3) + Branch FE | (5)<br>(4) + Loan | (6)<br>(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**: $R^2$ 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$_u$ $+ \epsilon_{ijtu}$ Dependant variable is the bank interest rate. Loan$_{ijtu}$ is a vector of loan characteristics (term in months, amount, type of rate: fixed or variable), Firm$_{ijtu}$ is a vector of firm characteristics (age, size, debt, investment grade, turnover). Column (1) include time and bank fixed effects. In columns (2) to (6), I sequentially add a county fixed effect (i.e. French "Départments"), a sector (NACE Rev. 2 French classification) fixed-effect and a quarter fixed-effect.

I use this data to show that the law of one price does not hold in French credit submarkets for SMEs. Formally, I estimate a price equation (see details in Appendix B.4) that aims to explain the observed variation in loan prices. Table 1 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 which suggests the 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 Firm Location, Creation and Geography

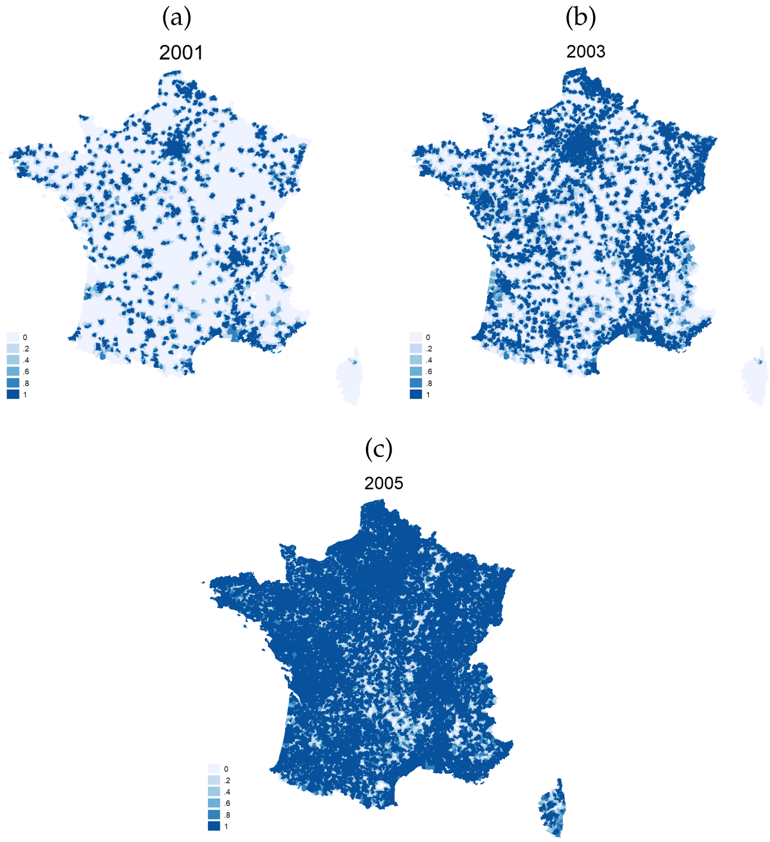
**Firm Location and Creation.** In order to conduct different empirical exercises along this paper, I gather a rich set of (i) administrative data on the creation date and the location French firms' establishments (from the exhaustive SIRENE database). Firms' establishment locations allow me to identify mono-establishment or mono-city firms (which is a key feature of my identification strategy), while the year of creation is useful to study firm entry and first banking partner choice in Appendix B.2.

**Geographical data.** I gather geographical data for mainland France about more than 36.000 municipalities, 2,000 urban units and 762 urban areas. An urban unit (UU) is a commune or group of communes that includes on 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 concerned has more than half of its population in this built-up area. The largest geographical unit that we consider in this paper is the urban area (UA). An urban area 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 resident population with a job works in the pole or in municipalities attracted by it. Finally, municipalities or cities are the finest unit of measurement that I use for distance computation, firm and branch locations. In the city-level part of my empirical analysis, I aggregate bank branches credit exposure, distance to clients, etc., at the urban unit level.

## 2.4 Broadband Internet Data

I use the unique data collected by Malgouyres et al. (2021) about the date of upgrade to ADSL for each Local Exchange (LE)'s in mainland France. The historical operator (France Télécom) had to make this data available to other operators as well as websites allowing consumers to gauge the quality of their line for regulatory reasons.

Figure 1: Broadband internet roll-out in France




**Notes:** This figure shows the roll-out of Broadband Internet for all city 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 city with no internet connection ($Z_{ut}$ equal to 0). This unique data was collected by Malgouyres et al. (2021) and contain the date of upgrade to ADSL for each Local Exchange (LE)'s in mainland France. The historical operator (France Télécom) had to make this data available to other operators as well as websites allowing consumers to gauge the quality of their line for regulatory reasons. Additionally, the authors gather data from the regulatory agency (ARCEP) regarding the geographical coverage of each LE. Combining both datasets,they construct a continuous measure $Z_{ut}$ described in Section 2.

Additionally, the authors gather data from the regulatory agency (ARCEP) regarding the geographical coverage of each LE. Combining both datasets, Malgouyres et al. (2021) construct a continuous measure of broadband access of city $i$ at year $t$, denoted $Z_{it}^{\text{city-level}}$, which is a time-weighted percentage of area covered in city $i$.[10] I construct

[10] Formally, the measure used by Malgouyres et al. (2021) writes:

a additional continuous measure of connectivity at the urban-unit level, $Z_{ut}^{\text{urban-unit}}$ which is simply the (weighted) sum of $Z_{it}^{\text{city-level}}$ for each $i \in u$:

$$Z_{ut}^{\text{urban-unit}} = \sum_{i \in u} w_i \cdot Z_{it}^{\text{city-level}} \tag{2}$$

where $i \in u$ denotes the municipalities included in urban unit $u$ and $w_i$ is the weight of city $i$ in the total population of urban unit $u$. Equation 2 implies that $Z_{ut}^{\text{urban-unit}}$ is continuous between 0 and 1. Namely, 0 implies that no firm located in $u$ enjoy an ADSL connection during year $t$, while 1 indicates that all firms benefit from it all the year long. Figure 1 shows the roll-out of Broadband Internet for all urban-units in mainland France, from 1998 to 2005. The dark areas represent a large degree of coverage (a high $Z_{ut}$). In 2000, those 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.

# 3 Model

Motivated by empirical evidence I document on search frictions in French credit markets (Appendix B.2), 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; First, informational asymmetries affect banks ability to screen projects. Second, search frictions hinder firms ability to locate and match with the right financing partner, as in Eaton et al. (2018) and Lenoir et al. (2018).[11] The model captures the key empirical evidence presented in section B. In the following sections, I start by summarizing the main assumptions; Then, I derive analytical predictions on aggregate credit flows and firm-bank matching.

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

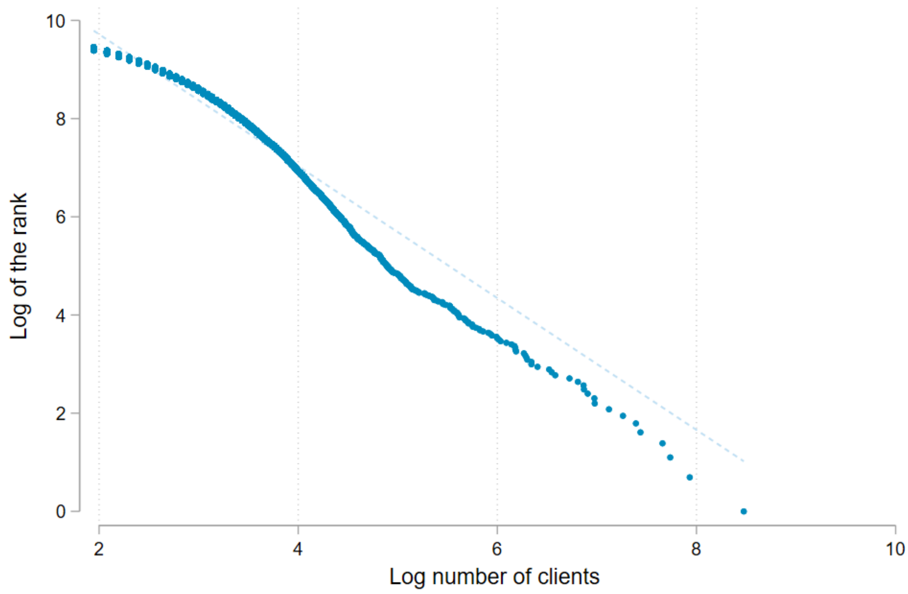
$\widetilde{Z}_{it}$ will be equal to one if all of its areas have had access for the entire year. It will be equal to 1/2 if the entire city has had access to broadband over half the year $t$.

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

## 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.

Figure 2: 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 5 clients. Formally, it shows 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: #1 being the largest branch, #2 the second largest, and so on.

**Supply side.** There is a continuum of bank branches in each submarket $u$, of measure $N_u = S_u \cdot z_{min}^{-\theta}$, with $S_u$ indicates the size of the submarket. Bank branches produce and provide credit with a single factor constant return-to-scale production function[12]. For the sake of simplicity, I make no distinction between short-term and long-term loans and I do not model more complex credit types as leasing or factoring. Bank

[12]Constant return-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 return-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 return-to-scale does not affect the predictions of the model.

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 a bank branch $b_u$ located in submarket $u$ is independently drawn from a Pareto distribution of parameter $\theta$ and support $[z_{min}, +\infty[$:

$$z_{b_u} \sim \text{Pareto } (z_{min}, \theta) \tag{3}$$

This Pareto assumption is data driven: Figure 2 shows that the relationship between branch log size – measured as its total numer of clients – and log rank is close to a straight line with a slope close to 1.[13] Thus, the number of bank branches located in submarket $u$ that can provide credit with efficiency above $z$ writes $N_u(z) = S_u \cdot z^{-\theta}$.

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, bank 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 $\gamma$ and support $[z^f_{min}, +\infty[$:

$$z_{e_u} \sim \text{Pareto } (z^f_{min}, \gamma) \tag{4}$$

such as the number of firms in a submarket writes $F_u = S_u \cdot {z^f_{min}}^{-\gamma}$.[14] In order 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 consequence, an entrepreneur has to undergo costly search process.

**Search and matching.** I build on Eaton et al. (2018) where matching between buyers and sellers is random. Each entrepreneur meets with a discrete number of bank branches, some located in their own submarket, some located remotely. This ran-

[13]Here branch size is used as a proxy of branch productivity, which can't be measured accurately.
[14]I remain agnostic about the entrepreneur production function that can be either Cobb-Douglas or CES, without altering the model predictions.

dom search process is a reduce form for the active search for banking partners: entrepreneurs need to gather information about bank branches characteristics, contact loan officers and, finally, 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 $\Theta_{e_u}$, which is the sum of potential banking partners met in each of the N submarkets. $\Theta_{e_u}$ 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 a credit from their *optimal* banking partner (i.e. the first-best match)[15]. 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. (2018) 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 entrepreneurs heterogeneity matters and 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 $z_{e_v} \in [0,1]$). $\kappa_{uv}$ (also $\in [0,1]$) can be interpreted as a pair-specific $u$-$v$ inverse measure of the strength of 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:

$$\text{Card}(\Theta_{e_v}(u)) = z_{e_v}\kappa_{uv} \cdot S_u z_{min}^{-\theta} \tag{5}$$

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 modelling 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 $\Theta_{e_v}$ while entrepreneurs with a low productivity may end up with only a few bad quotes. Second, heterogeneity in $\kappa_{uv}$ across submarkets

[15]Note that, without search frictions, all entrepreneurs located in the same submarket may end up applying for credit from the exact same bank branch, if only branch productivity matters (Cerqueiro et al. (2011) shows that branch heterogeneity goes far beyond productivity). This will lead to positive assortative matching and directed search.

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 branches 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 the fact 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) 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 its investment project and get 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 depends on 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 the entrepreneur $e_v$ writes $r_{b_u,e_v} = \frac{c_u d_{uv}}{z_{e_v} z_{b_u}}$[16]. First, the interest rate increases with $d_{uv}$, a function of branch-firm distance, and the unit cost of production. Second, loan rates are negatively correlated to both branch and entrepreneur productivity. This negative relationship between prices and branch productivity summarizes and aggregates all the costs linked to information asymmetries faced by the loan officer: screening and monitoring, that play a key role in most of bank-firm matching models (Hauswald and Marquez, 2003; Vivès and Ye, 2021) where monitoring and screening is more costly for a bank if there is more distance between its expertise and the entrepreneur's project characteristics.

The assumption of bank branches pricing at their marginal cost is strong, in particular in a context of credit markets subject to informational asymmetries. As pointed out in Lenoir et al. (2018), an alternative is to assume that bank branches compete à la Bertrand: the branch that offers the best contract terms doesn't price at its marginal cost, but 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

[16]The bank branch $b_u$ program writes: $\Pi_u = z_{match} Kr - Kc_u d_{uv}$ 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.

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 $\Theta_{e_v}$, the entrepreneur $e_v$ decides to match with the one offering the best contract terms (i.e. lowest interest rate). The interest rate paid by $e_v$ writes:

$$r_{e_v} = \underset{b_u \in \Theta_{e_v}}{\operatorname{argmin}} \{ \frac{c_u d_{uv}}{z_{b_u} z_{e_v}} \} \tag{6}$$

The poisson search process combined with the Pareto distribution of brank branches size allows an analytical formula for $r_{e_v}$. Eaton et al. (2018) demonstrates that the assumption of Poissons draws into a Pareto distribution delivers a Weibull distribution for the minimum interest rate introduced in equation 6. Formally:

$$\mathbb{P}\big(r_{e_v} \leq r\big) = W_{e_v}(r) = 1 - \exp\left( - r^{\theta} z_{e_v}^{\theta+1} \sum_{u=1}^{N} S_u \cdot (c_u d_{uv})^{-\theta} \kappa_{uv} \right) \tag{7}$$

Conditional on $r$ fixed, 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 $\kappa_{uv}$) the better the contract terms will be, on average. Finally, the entrepreneur's productivity directly impacts on the likelihood to be matched with a branch that offers a low interest rate because more productive entrepreneurs draw a larger set $\Theta_{e_v}$ of potential lenders that somehow compensates the adverse effect of search frictions. Note that a larger $\theta$ also alleviates the negative impact of search frictions, but mostly in favor of less productive entrepreneurs: indeed, less heterogeneity between bank branches advantages the entrepreneurs 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 branch's productivity. I then investigate how a shock on search frictions modifies those predictions.

### 3.2.1 Aggregate Credit Flows

In Appendices B.1 and B.3, I document the existence of inter submarket lending: while most of the credit is distributed locally, with borrowers and lenders located in the same urban unit, some branches operate in remote submarkets. Let $\Pi_{uv}$ 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 one if the entrepreneur is matched with a branch in $u$, zero otherwise:

$$\Pi_{uv} = \frac{\sum_{e_v=1}^{F_v} \mathbb{I}\{\mathrm{M}(e_v) = u\}}{\sum_{e_v=1}^{F_v} 1} \tag{8}$$

where $\mathrm{M}(e_v) = u$ indicates that entrepreneur $e_v$ decides to match with a branch from $u$. Lenoir et al. (2018) shows that using the law of large numbers, $\Pi_{uv}$ 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 a 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's heterogeneity, I show that this condition holds as the likelihood to ultimately match with a submarket $u$ does not depend on $z_{e_v}$. Thus, equation 8 writes:

$$\Pi_{uv} = \mathbb{E}_{e_v}\Big[\mathbb{I}\{\mathrm{M}(e_v) = u\}\Big] = \mathbb{P}\Big[\{\mathrm{M}(e_v) = u\}\Big] \tag{9}$$

I consider a level of price r (and a level of productivity $z_{e_v}$) fixed. Let $D_{u,e_v}(r)$, be the number of branches from $u$ met by entrepreneur $e_v$ that propose 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:

$$\mathbb{P}\Big[\{\ \mathrm{M}(e_v) = u\}|r\Big] = \frac{\int_{D_{e_v}(r)>0} \mathbb{P}\Big[\{\ \mathrm{M}(e_v) = u\}|r, D_{e_v}(r)\Big]\, \mathrm{dF}(D_{e_v}(r))}{\int_{D_{e_v}(r)>0} 1\, \mathrm{dF}(D_{e_v}(r))} \tag{10}$$

In equation 10, 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:

$$\begin{aligned}
&\mathbb{P}\Big[\{\ \mathrm{M}(e_v) = u\}|r\Big] = \\
&\qquad \sum_{n=1}^{+\infty}\sum_{n_u=0}^{n}\Bigg[\mathbb{P}\Big[\{\ \mathrm{M}(e_v) = u\}|r, D_{u,e_v}(r) = n_u, D_{k\neq u,e_v}(r) = n - n_u\Big] \\
&\qquad \times \mathbb{P}\Big[D_{u,e_v}(r) = n_u\Big] \times \mathbb{P}\Big[D_{k\neq u,e_v}(r) = n - n_u\Big]\Bigg] \\
&\qquad \times \mathbb{P}\Big[D_{e_v}(r) > 0\Big]^{-1}
\end{aligned} \tag{11}$$

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

$$\begin{aligned}
\mathbb{P}\Big[\{\ \mathrm{M}(e_v) = u\}|r\Big] &= \frac{z_{e_v}^{\theta+1}\kappa_{uv}S_u(c_u d_{uv})^{-\theta}}{z_{e_v}^{\theta+1}\sum_{k=1}^{N}\kappa_{kv}S_k(c_k d_{kv})^{-\theta}} \\
&\times 1 - \exp\left(-r^\theta z_{e_v}^{\theta+1}\sum_{u=1}^{N} S_u \cdot (c_u d_{uv})^{-\theta}\kappa_{uv}\right) \\
&\times \mathbb{P}\Big[D_{e_v}(r) > 0\Big]^{-1}
\end{aligned} \tag{12}$$

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 7, leading to the following simplification for equation 11 :

$$\mathbb{P}\Big[\{\ \mathrm{M}(e_v) = u\}|r\Big] = \frac{z_{e_v}^{\theta+1}\kappa_{uv}S_u(c_u d_{uv})^{-\theta}}{\sum_{k=1}^{N} z_{e_v}^{\theta+1}\kappa_{kv}S_k(c_k d_{kv})^{-\theta}} \tag{13}$$

First, equation 13 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 $e_v$ in $u$ mechanically increases but so do the number of branches drawn in others 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 branches productivity share the same shape parameter $\theta$ 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:

$$\Pi_{uv} = \mathbb{E}_{ev}\Big[\{\mathbb{I}\ \mathrm{M}(e_v) = u\}\Big] = \frac{\kappa_{uv}S_u(c_u d_{uv})^{-\theta}}{\sum_{k=1}^{N}\kappa_{kv}S_k(c_k d_{kv})^{-\theta}} \tag{14}$$

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 15, I derive two predictions about aggregate credit flows and the impact of a shock on search frictions $\kappa_{uv}$.

*Prediction 1: Gravity Equation for Bank Credit.* As shown in Lenoir et al. (2018), a log-linearization of equation 15 delivers a gravity equation for the share of credit distributed in $v$ by branches located in $u$:

$$\log \Pi_{uv} = \log \kappa_{uv} - \theta \cdot \log d_{uv} + \mathrm{FE}_v + \mathrm{FE}_u \tag{15}$$

where $\mathrm{FE}_u$ stands for $\log S_u c_u^{-\theta}$ and $\mathrm{FE}_v$ 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 ans bank holdings (Coeurdacier and Martin, 2009). Recently, Okawa and van Wincoop (2012) proposed 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. First-order condition of equation 15 with respect to $\kappa_{uv}$ leads to:

$$\frac{\partial \ln \Pi_{uv}}{\partial \kappa_{uv}} = \underbrace{\frac{\partial \ln \kappa_{uv}}{\partial \kappa_{uv}}}_{(a)} + \Big[ - \underbrace{\frac{\partial \ln \sum_{k=1}^{N} \kappa_{kv} S_k (c_k d_{kv})^{-\theta}}{\partial \kappa_{uv}}}_{(b)} \Big] \tag{16}$$

Two opposite mechanisms are at stake. First, a reduction of 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 in average more bank branches from $u$ in $\Theta_{e_v}$. 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) rewrites $\frac{S_u (c_u d_{uv})^{-\theta}}{\sum_{k=1}^{N} \kappa_{kv} S_k (c_k d_{kv})^{-\theta}} \geq 0$, such that the *competition effect* is negative and may compensate the direct effect of *connectivity*. Equation 16 simplifies in:

$$\frac{\partial \ln \Pi_{uv}}{\partial \kappa_{uv}} = \frac{1}{\kappa_{uv}} - \frac{S_u(c_u d_{uv})^{-\theta}}{\sum_{k=1}^{N} \kappa_{kv} S_k (c_k d_{kv})^{-\theta}} \geq 0 \tag{17}$$

Note that the effect of a reduction of search frictions $\kappa_{uv}$ in heterogeneous across submarkets $u$. The larger or closer a submarket, the smaller the total impact. This captures the fact that large and nearby submarket already benefits 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 of search frictions.

### 3.2.2 Branch-Entrepreneur Matching

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 that ultimately match with a particular bank branch and (ii) the impact of a reduction of search frictions on the matching equilibrium. Both predictions can be confronted to firm-branch relationship data from the Credit Register and to motivating empirical work presented in Appendix B.

*Prediction 3: Positive Assortative Matching*. I consider a bank branch located in $u$ and its likelihood $F_{b_u}(e_v)$ to lend to the 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 then, the likelihood for $b_u$ to be the lowest cost supplier. Formally:

$$\begin{aligned} F_{b_u}(e_v) &= \mathbb{P}\Big(b_u \in \Theta_{e_v}\Big) \times \mathbb{P}\Big(\underset{\Theta_{e_v}}{\text{argmin}}\{\frac{c_u d_{uv}}{z_{b_u} z_{e_v}}\} = b_u\Big) \\ &= \mathbb{P}\Big(b_u \in \Theta_{e_v}\Big) \times \Big(1 - \mathbb{P}(r_{e_v} < r_{bu})\Big) \end{aligned} \tag{18}$$

From equation 7 for minimum price distribution, I have an analytical formula for $\mathbb{P}(r_{e_v} \leq r_{bu})$[17]. 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 $e_v$ ultimately decide to borrow from $b_u$, namely $F_{b_u}(e_v)$:

$$F_{b_u}(e_v) = z_{e_v}\kappa_{uv} \times \exp\Bigg( - (c_u d_{uv})^{\theta} z_{b_u}^{-\theta} z_{e_v} \sum_{u=1}^{N} S_u \cdot (c_u d_{uv})^{-\theta} \kappa_{uv}\Bigg) \tag{19}$$

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

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 19 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 offices are likely to match only with unproductive entrepreneurs that do not enjoy a large set of potential partners. This is consistent with empirical evidence from the Probit regression presented in section B: the product of branch and firm size is positively correlated with branch-firm matching, as the branch growth rate is.

*Prediction 4: Firm-Branch Matching and Search Frictions*. The first-order condition of equation 19 with respect to $\kappa_{uv}$ indicates how the matching process is affected by a shock on bilateral search frictions between any two submarkets $u$ and $v$.

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

Similar to *Prediction 2*, the impact of a reduction of 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. 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 20 offers a reduced-form expression that summarizes both effects:

$$\frac{\partial \ln F_{b_u}(e_v)}{\partial \kappa_{uv}} = \frac{1}{\kappa_{uv}} - \frac{S_u z_{e_v}}{z_{b_u}^{\theta}} \tag{20}$$

The impact of search frictions varies along the distribution of branch productivity. High-productivity branches (high $z_{b_u}$) benefit more form the reduction of 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 offer 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 are relatively passive, contrary to what is usually assumed in the banking literature where search frictions are absent (Matutes and Vives, 1996; Cordella and Levy Yeyati, 2002; Hauswald and Marquez, 2003, 2006). In particular, loan officers price credit loans at their marginal price and do not actively search or compete for customers. They meet with potential borrowers, screen their investment project and offer a price. I argue that those simplifying assumptions not only allow to make the model very tractable, but also make sense in the context of local credit markets for SMEs financing. The firms I consider here are (very) small ones that are not likely to be spontaneously approached by bankers: surveys actually show that small firms are often the one bearing the search cost (FED, 2014). Second, a part of the competition between banks may occur at a national level (advertising, marketing, etc.) rather than at the local level where bank branches are meeting points between firm demand and bank supply of credit. 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 doesn't price at its marginal cost, but equals the marginal cost of the branch with the *second best* offer, but, under the assumption of inelastic demand, competition à la Bertrand would not affect the model predictions about firm-branch matching.

Another key aspect of my theoretical approach is that bank branching strategy is not modelled: the number and the location of bank branches is exogenous. While bank branches management is out of the scope of this paper, I take this into account in my empirical analysis, with 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 2005 as a large scale quasi-natural experiment to study the impact of a reduction of search frictions on credit markets. First, I present the technological and institutional context of Broadband Internet diffusion in France. Second, I document the adoption of ICTs by French banks and How high-speed Internet is likely to reduce search frictions for SMEs. Third, I discuss the identification strategy and propose a new IV for Broadband Internet deployment.

## 4.1 The Rise of Online Banking

In the early 2000's, the large diffusion of ICTs represented 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 with personal-loan applications submitted with a few swipes on a mobile phone, and time to cash can now be as short as a few minutes (McKinsey, 2018). Not only transaction costs decreased: the rise of online price comparison services and brokers allows individuals to search for the best banking partner in a faster and more efficient way[18]. The combination of both transaction and search costs reduction resulted in a severe disruption of search frictions for individuals.

Regarding corporate credit and SME lending, the situation is mixed. The loan officer remains the most relevant interlocutor and the ultimate decision-maker in SME lending, as He is ideally suited to understand client's 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: recent survey on UK SMEs (Ernst and Young, 2018) shows that financial services used by SMEs are mainly online banking (85%) and second, branch-based banking (43%), emphasizing the growing importance of digitization. Among other examples, emails allow firms to easily contact a loan officer, online data and document sharing speeds up the meeting process and reduces transaction costs. Bank websites are showcases designed to attract new customers and

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

provide information about financial products. Finally, customer online areas facilitate communication.

In order to document the fact that French banks had started their digital transition process at the beginning of the 2000's, I gather new data on large French bank adoption of ICTs; First, I check the existence of the bank website with a firm customer area before 2000 (using the waybackmachine.com website). Second, I 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 2000's, with a sophisticated website, while their domain name were registered in average in the mid-90's, showing an early preoccupation for online visibility. In particular, the average website demonstrate the willingness of banks to improve the accessibility of basic but crucial information to 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 or to ask for information about financial products and services. For each bank branch, the phone number, the contact email and the physical address were immediately available. This represents a sharp reduction of search costs born by the entrepreneur.

## 4.2 The diffusion of Broadband Internet in France

### 4.2.1 The ADSL technology

The ADSL (Asymmetric Digital Subscriber Line) 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 the ADSL technology, bandwidth and bit rate are said to be asymmetric, meaning greater towards 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 Local Exchange (LEs), since the intensity and the quality of the analog signal decreases as it is routed over the copper lines. Local Exchanges are the telephone exchanges owned by the incumbent operator France Télécom (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 the telephone network, 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) towards the world wide web. A key feature of the 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 the installation of an equipment inside the LE called a DSLAM (Digital subscriber line access multiplexer) that is required in order to recover the data transmitted via ADSL on the local copper loop and adapt it so 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 I will use in my empirical analysis.

### 4.2.2 The ADSL roll-out in France

The ADSL technology became popular during the 1990s, as many OECD countries were planning the expansion of services related to information and communications technology. In the early 2000s in France, the deployment of the technology beyond France's largest cities was slow. The causes for this staggered deployment are multiple. First, France Telecom (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, at the same time France Telecom had to invest massively in upgrading its LEs to ADSL, it went through a debt crisis which 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 France Télécom. Under the impulse of the government – which increased its stake in the firm during the 2002 bailout of the firm – France Telecom pledged in 2003 to cover 90% of the French (metropolitan) population by the end of 2005, i.e. all local exchanges (LEs) with more than 1000 lines, for a total investment of 600 M euros (750 M 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 France Telecom to undergo the investment until 2002. After 2002, with the strong impulse of the government, France Telecom 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 2 more years than anticipated in 2003.

### 4.3 Identification and Instrumental Variable

My identification strategy is based on the gradual diffusion of the new technology in different LEs over space and time. Note that the question of what were the criteria for deciding to *treat* one LE before another has been studied in 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 slow down by the sunk cost of upgrading the infrastructure, consistent with statistical analysis[19].

In this paper, I propose an instrument variable strategy based on a theoretical optimal investment plan for infrastructure upgrading. The ADSL technology combines local copper loops and a large optic fiber network. Thus, when France Télécom decided to connect 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 a LE was proportional to the number of inhabitants newly connected. 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 a LE upgrading and I use the distance (in km) to the closest optic fiber as a proxy of the connection costs.

[FIGURE 12 ABOUT HERE]

A key feature of this instrumental variable strategy lies in the exogeneity of the distance between a 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. France Télécom leased the existing infrastructure to those company in order to faster Broadband expansion. Figure 12 displays the location of around 13,000 Local Exchanges, highways and railroads already existing

[19] 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 is carried out in their paper. Different to their paper, my treatment is continuous and I can't rely on a dynamic event-study approach and check for the pre-trends. For this reason, I propose an IV strategy for the timing of Broadband expansion.

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.

[FIGURE 13 ABOUT HERE]

Formally, I predict the *optimal connection rank* $\hat{R}_i$ for the Local Exchange $i$, 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 preexisting and time-unvarying city characteristics:

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

$$\hat{R}_i = \hat{\alpha} + \hat{\beta_1} \cdot \text{Density}_{i,1998} + \hat{\beta_2} \cdot \text{Shortest Distance}_{i,1998}$$

Figure 13 shows the rank correlation between optimal versus observed connection ranks. The combination of exogenous connection *gains* and *costs* have a strong predictive power, with a R-square 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 urban units as $\mathbb{C}_{uvt} = Z_{vt} \times Z_{ut}$. $\mathbb{C}_{uvt}$ 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 instrument variable strategy delivers a similar measure of between urban unit connectivity 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 roll-out affects inter-regional flows through a reduction in search frictions and finally (iii) the Broadband Internet roll-out 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 urban-unit level for all single-city SMEs in my sample, in order to compute bilateral credit shares $\Pi_{uvt}$. $\Pi_{uvt}$ 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 $v$. $\Pi_{uvt}$ lies in $[0,1]$, 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, $\Pi_{uvt}$ equals 1 implies that firms in $v$ are fully financed by branches in $u$. In an economy without inter-regional exchanges, all $\Pi_{uvt}$ would equal 0, except $\Pi_{uut}$ 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, 2011) show that the Pseudo-Poisson Maximum Likelihood (PPML) estimator, introduced by Gourieroux et al. (1984), is a promising workhorse for the estimation of gravity equations, in particular in the presence of many zeros. It is perfectly suited for the estimation of multiplicative models, without log-linearization of the dependent variable.[20] I adopt this standard approach and rely on the new estimator for pseudo-poisson regression models with multiple high-dimensional fixed-effects developed by Correia et al. (2019). Formally, I estimate equation 15 in its multiplicative form:

$$Y_{uvt} = \exp\left[\ln S_{vt} + \ln M_{ut} + \beta_1 \ln dist_{uv} + \beta_2 \ln X_{uvt}\right] + \epsilon_{uvt} \tag{21}$$

where $Y_{uvt}$ is the bilateral dependent variable (credit flows, shares, number of clients served, etc.) aggregated at the urban-unit level. $u$, $v$ and $t$ are indices for origin (the urban unit from which the bank branches operate), destination (the urban unit in which the borrowing firms are located) and time. $S_{vt}$ and $M_{ut}$ are the origin urban unit $\times$ year and the destination urban unit $\times$ year fixed effects; fixed-effects ensure 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 urban units $u$ 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 effect as my goal is to identify the coefficient $\beta_1$ associated with physical distance $dist_{uv}$ between $u$ and $v$. Here, the physical distance captures both the monitoring costs

[20]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 dependant variable to be Poisson distributed (Fally, 2015)

and the search frictions: according to my model and consistent to gravity estimation in international trade and finance, $\beta_1$ should be negative.

**Impact of Broadband Internet Diffusion.** I study the impact of a reduction of search frictions on aggregate credit flows using the staggered diffusion of Broadband Internet, i.e. prediction 16 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 investigate the impact of time-varying trade policies as trade agreements, I consistently identify the effect of time-varying connectivity between urban units using a dynamic PPML estimator with fixed-effect 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 of search frictions):

$$Y_{uvt} = \exp\left[\ln S_{ut} + \ln M_{vt} + \beta_1 \ln dist_{uv} + \beta_2 \ln X_{uvt} + \beta_3\,\mathbb{C}_{uvt}\right] + \epsilon_{uvt} \qquad (22)$$

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* $\times$ *year* and *destination* $\times$ *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[21]. Another widely used specification includes dyadic fixed-effects, namely *origin-destination* FE, in order 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 22 including pair-specific fixed-effect: a direct consequence is that I cannot estimate the coefficient $\beta_1$ relating to the physical distance $dist_{uv}$ in this specification.

**PPML: Difference-in-Difference with Many Zeros?** 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) shows 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 this paper, I extend and adapt their simulation exercise to the exact case of my empirical setting. Not only my estimating sample contains a vast majority of urban unit pairs that do not exchange credit over the entire period 1998-2005 (more than 95% of the credit shares equal zero) but also I have to deal with panel data and a difference-in-difference approach. In Appendix C I present simulation evidence on the performance of the

[21] Broadband Internet not only affects search frictions, but might also impact urban unit sizes, exogenous production costs or branch and firms productivity – in this setting, that does not bias the $\beta_3$ point estimate because of the urban units-time fixed-effects.

PPML estimator when the panel data is 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. My results confirm and extend the findings of Santos Silva and Tenreyro (2006, 2011), showing that both the PPML and the GPML estimators are well behaved in the two cases considered. In particular, the coefficient of interest $\beta_3$ is consistently estimated with two-way unit $\times$ time fixed effects. These findings are an additional reason that justify my empirical approach and the validity of my estimation procedure.

## 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 in order 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-from-treatment.

The level of observation being a branch $b$ located in city $u$, I'm able to discretize the treatment status by setting treatment status to 1 after the city experienced its highest increase in the predicted treatment variable $\hat{Z}_{it}^{\text{city-level}}$. Formally, I define the year of treatment as $t_{i0} = \text{argmax}_t\{\Delta\hat{Z}_{it}^{\text{city-level}}\}$ and discretized treatment status as $\mathbb{1}\{t \geq t_{i0}\}$. The year of treatment for each city is denoted $t_{i0}$. 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 $d \in \{-6, -5, ..., +4, +5\}$. The main estimating equation is as follows:

$$Y_{b(u)t} = \sum_{\substack{d=-5 \\ d\neq -1}}^{d=5} \beta_d \times \mathbb{1}\{t = d + t_{0u}\} + \mathbf{x}'_{ut}\delta + \alpha_{b(u)} + \psi_{r(u),t} + \varepsilon_{ut} \tag{23}$$

where $\alpha_{b(u)}$ and $\psi_{r(u),t}$ are fixed effects for the branch b located in city i 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 variables for $d = -5$ and $d = -1$. That restriction is necessary to avoid multi-colinearity 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 results with alternative normalizations in the robustness section.

The specification presented in equation (23) includes leads and lags. The inclusion of leads allows us to assess the presence of pre-trends. We also estimate a simpler "semi-dynamic" specification where only the lags of the treatment are included, as presented in equation (24):

$$Y_{b(u)t} = \sum_{d=0}^{d=5} \beta_d \times \mathbb{1}\{t = d + t_{0u}\} + \mathbf{x}'_{ut}\delta + \alpha_v + \psi_{r(u),t} + \varepsilon_{ut} \tag{24}$$

The event-study coefficients $\widehat{\beta}_d$ estimated from equation (23) can be interpreted causally under the identifying assumption that, conditional on receiving broadband over the period considered and conditional on bank branch and city fixed-effects, the *timing* of broadband roll-out 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 the inclusion of a large set of controls and fixed effects meant to account for city and well as 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 urban unit), and iii) the share of credit granted remotely (i.e. to remote clients).

# 6 Results

This section presents the results. I first show that inter-regional credit flows follow a gravity equation by estimating equation (21), using a pseudo-Poisson maximum likelihood approach. Then, I estimate equation (22) and document an increase in inter-regional credit flows triggered by the staggered roll-out of Broadband Internet, associated with a reduction in search frictions. Finally, I estimate equation (23) 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 2 shows the results for the estimation of equation (21), using a pseudo-Poisson maximum likelihood approach described in section 5.1. The dependant variable is the credit shares $\Pi_{uvt}$, 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) display the results with no control, origin urban unit × year and destination urban unit × 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 (*département*) and the lagged log of trade flows between counties. The coefficient of interest $\beta_1$, associated to the log of the distance between urban unit $u$ and $v$, is negative and close to -2. This magnitude implies that firms borrow credit from banks located in a close-by urban unit 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.

Table 2: GRAVITY EQUATION FOR INTER-REGIONAL CREDIT FLOWS

| | **Share of credit in v borrowed from u:** $\Pi_{uvt}$ | | | | |
|---|---|---|---|---|---|
| | (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**: PPML estimation of equation (21). $dist_{uv}$ = 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. The sample consists of all origin-destination-year combinations where at least one firm is located with positive credit. Robust standard errors in in parentheses.

It is not evident to compare this magnitude with the literature: as far as I know, 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. Recently, 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 indicates that distance-related frictions are likely to be more important at a very local level than for international credit flows. This confirm 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 keeping only pairs of urban units with positive credit flows. Table 4 shows the results. The coefficient of interest is still negative and significant, but of a smaller magnitude: $\beta_1$ is now close to -1 and remains stable to the addition of 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 (16) of my model by running a PPML estimation of the augmented gravity equation (22). The main results of my paper are presented in Table 3, where the first column reproduces column (4) of Table 2 for comparison purposes. In column (2), I include the continuous variable $\mathbb{C}_{uvt}$ for observed internet interconnection between two urban units. Similar to Table 2, in order to estimate an effect on distance, origin and destination $\times$ year fixed effects are included but not pair specific fixed effects. The estimate is positive and significant, in line with the model prediction. It suggests that a technology-induced reduction in search frictions distort the gravity equation for credit flows. The amount of credit exchanged between firms and banks located in connected urban units increases relative to other not-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 urban unit $u$ 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 an heterogeneous effect of brandband 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 for a reduction of search frictions. I formally test this prediction by including the interaction variable for the bilateral distance (expressed as 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 neighbouring cities, already economically very closely tied, are interconnected by internet.

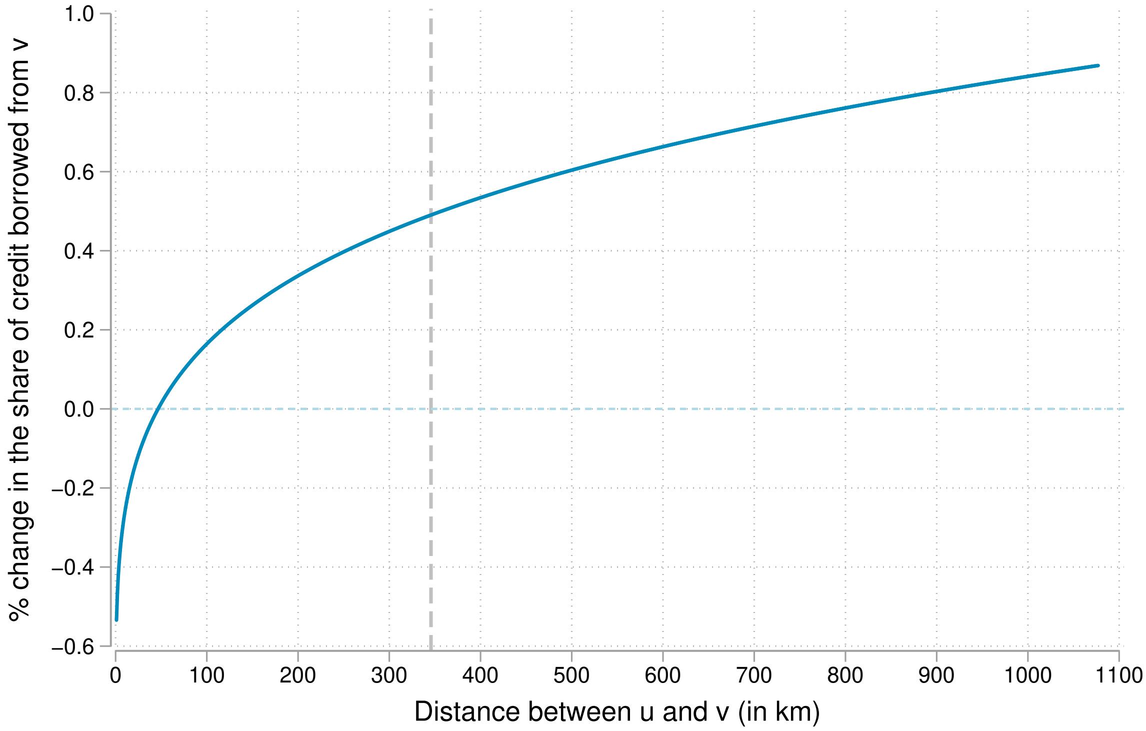
Table 3: TECHNOLOGY-INDUCED REDUCTION IN SEARCH FRICTIONS

| | **Share of credit in v borrowed from u:** $\Pi_{uvt}$ | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Log distance $dist_{uv}$ | -1.765*** | -1.760*** | -1.844*** | -1.764*** | -1.865*** |
| | (0.004) | (0.004) | (0.006) | (0.004) | (0.006) |
| $\mathbb{C}_{uvt}$ | | 0.370*** | 0.800*** | | |
| | | (0.058) | (0.059) | | |
| Log distance $dist_{uv}$ x $\mathbb{C}_{uvt}$ | | | 0.230*** | | |
| | | | (0.008) | | |
| $\hat{\mathbb{C}}_{uvt}$ | | | | 0.058 | 0.399*** |
| | | | | (0.059) | (0.062) |
| Log distance $dist_{uv}$ x $\hat{\mathbb{C}}_{uvt}$ | | | | | 0.199*** |
| | | | | | (0.008) |
| log (Trade Flows) | 0.473*** | 0.475*** | 0.470*** | 0.474*** | 0.468*** |
| | (0.004) | (0.004) | (0.004) | (0.004) | (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**: PPML estimation of equation (21). $dist_{uv}$ = 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 urban units. $\mathbb{C}_{uvt}$ 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 instrument variable strategy delivers a similar measure of between urban unit connectivity 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. The sample consists of all origin-destination-year combinations where at least one firm is located with positive credit. Robust standard errors in in parentheses.

An important concern with those results is that potential endogeneity of internet take-up biases these estimates. In columns (4) and (5), I estimate equation (22) using the instrument variable 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 urban unit $u$ now increases by only 6% when $u$ and $v$ are connected.

Figure 3: Heterogeneity of the effect with respect to distance




**Notes:** This figure plots the marginal effect of broadband internet with respect to distance between to interconnected cities $u$ and $v$, as estimated in equation (22). 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 %. The vertical grey dashed line represent the average distance between two cities in my sample, and correspond to the overall effect estimated in column (4) of Table 3.

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 instrumental variable. In particular, the interaction term is positive and significant. In economic terms, theses results mean that an increase in internet availability of 10 percentage points increases credit flows for an urban unit at the 25th distance percentile by 24% less than for an urban unit at the 75th distance percentile. Figure 3 illustrates this heterogeneity. While the effect of being interconnected is negative when cities are geographically nearby (competition effect dominates), it increases sharply and because positive after the 50th kilometer. After the 100th kilometer the slope of the curve is then much flatter. Figure A3 in Appendix 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 dimensions. I first estimate the augmented gravity equation (16) by adding bilateral pair fixed-effects that control for any unobserved characteristics of the urban unit pair that are constant over time (see Head and Mayer, 2014). Table 5 shows the results in columns (2) and (4). 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. Second, in order to take into account the dynamic nature of credit flows, I replicate the baseline results by adding the lag dependant variable to the regressors in Table 6. This implies that credit relationships existing in the previous year provides a basis for the credit flows observed in the current year.[22] The main findings are confirmed by this alternative specification. Finally, I estimate models (21) and (22) on the extensive margin of credit flows. I use the share of relationships between bank located in $u$ and firms located in $v$ divided by the number of banking relationships on $v$, rather than the share of credit used above. Results are presented in Tables 7 and 8 and 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.

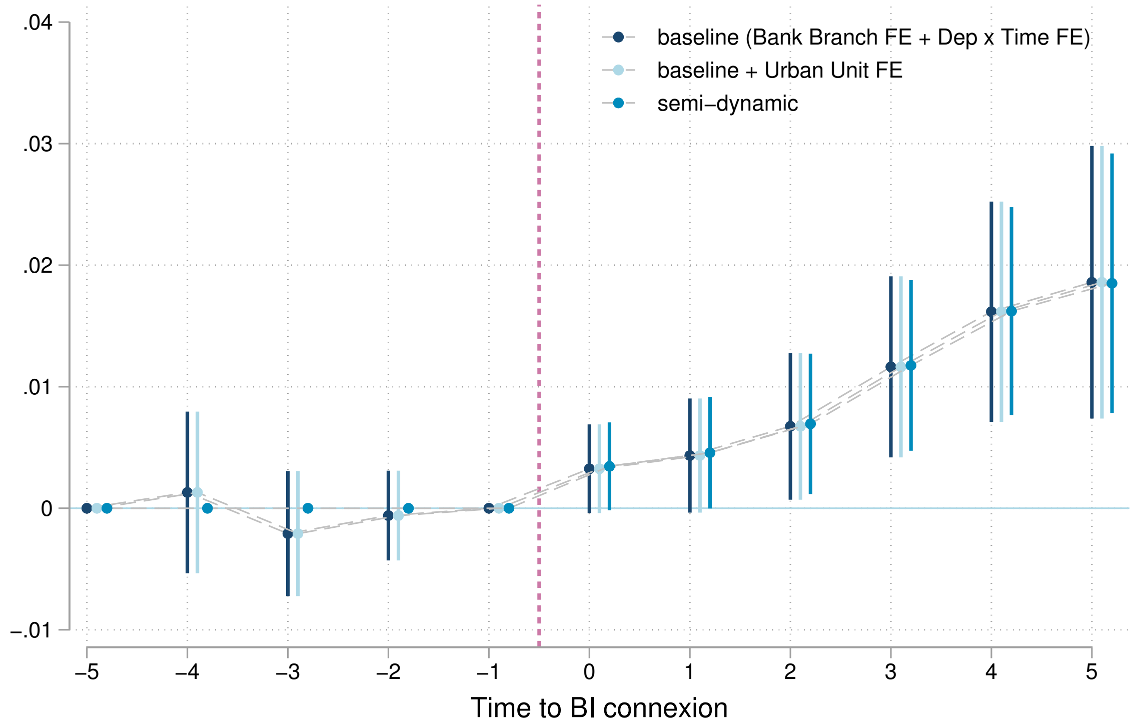
**Firm-bank matching.** The model predicts that broadband internet diffusion affects the firm-branch matching process: the reduction in search frictions allows very distant firms and bank to match and, in turn, the share of remote firms (located outside $u$) financed by a bank located in $u$ increases. I formally test prediction (20) 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 lent 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, within cities, and the share of remote credit is close to zero. On the contrary, if search is free, firms can meet with and borrow from banks located far away and the share of credit lent remotely is large. I show the results for different specifications and assess the robustness of the results. I then turn to the extensive margin: the share of firms located remotely financed by a

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

bank in a given city.

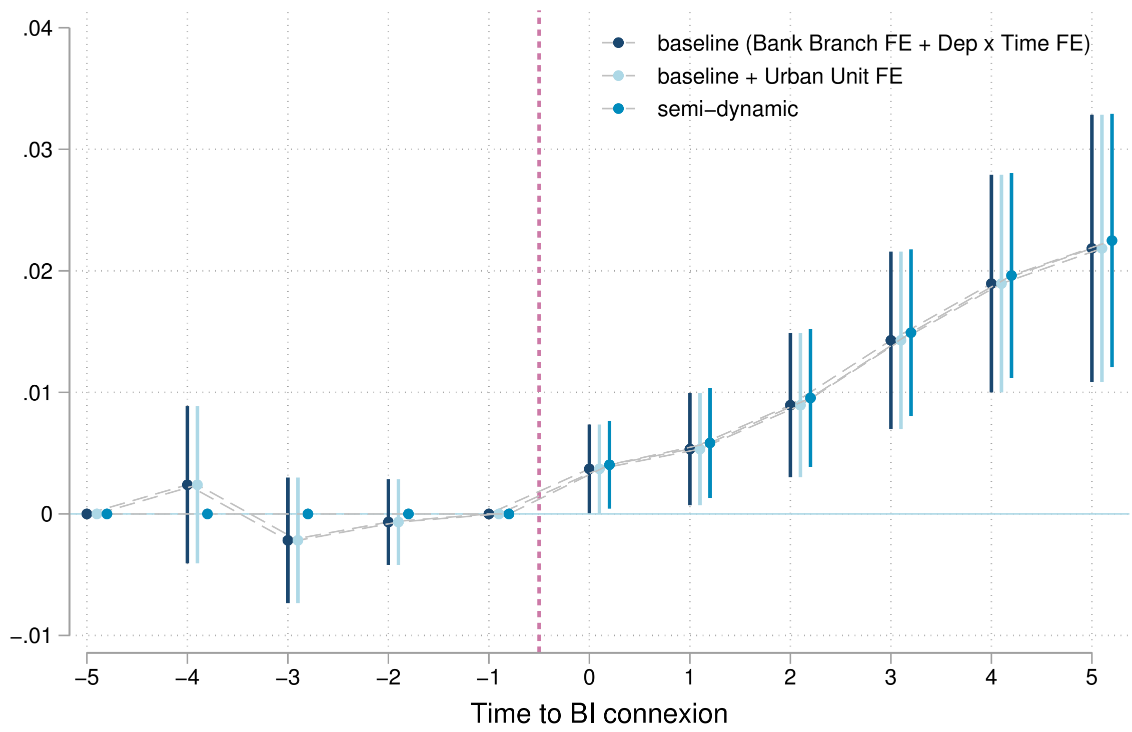
Figure 4: SHARE OF CREDIT TO REMOTE FIRMS




**Notes:** This figure plots estimates for specification in equation (23) – fully dynamic – and equation (24) – semi-dynamic. The dependant variable is the share of credit granted to firms located outside the bank's urban unit. The sample include all bank branches with a positive credit exposure. The baseline specification (navy blue) includes bank branch fixed effects and year × county fixed-effects. Urban unit fixed-effects are included in the alternative fully-dynamic specification (light blue). 95% confidence interval are presented. Standard errors clustered at the city level.

Figure 4 displays the results, plotting estimated coefficients from equation (23). The dark blue dots report results from the baseline specification including the bank branch and county-year fixed-effects. 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 that. The coefficient for d = 5 in that specification is 0.018 suggesting that the expansion of access to broadband internet increased the bank branch-level share of credit lent to distant firms by about 10%, 5 years after the period of largest expansion. Not only this is in line with the model predictions but also it is consistent with the aggregate impact on credit flows documented above. Our second specification adds urban unit fixed-effects to the regression that aim to control for time invariant city characteristics. The light blue dot show the estimates. Here again, I find no sign of a pre-trend prior to broadband expansion contrasting with a steady growth afterwards. The estimated effect after five years is virtually the same than 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 24). The regression should in theory 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.

Figure 5: Share of remote clients




**Notes:** This figure plots estimates for specification in equation (23) – fully dynamic – and equation (24) – semi-dynamic. The dependant variable is the share of clients financed outside the bank's urban unit. The sample include all bank branches with a positive credit exposure. The baseline specification (navy blue) includes bank branch fixed effects and year × county fixed-effects. Urban unit fixed-effects are included in the alternative fully-dynamic specification (light blue). 95% confidence interval are presented. Standard errors clustered at the city level.

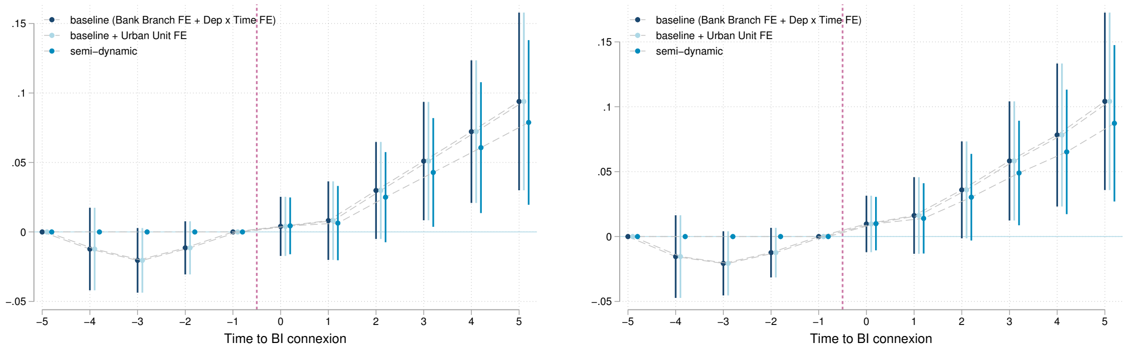
I then explore the possible mechanisms behind the bank-level increase in share of credit lent to remote clients documented above. I study how the extensive margin of credit, i.e, the number of banking relationships, is affected by the technology-induced reduction in search frictions (as opposed to the average credit, the intensive margin). For this purpose, the dependant variable I consider is the share of remote firms financed by bank branches located in a given city, which is defined as the number of clients located outside the bank's urban unit divided by the total number of clients. Figure 5 displays 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 increased the bank branch-level share of remote clients by about 12%, 5 years after the period of 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 cities credit flows is mainly driven by the creation of new relationships (new matches between banks and firms located in different cities), and (ii) that those new credit relationships are in average smaller than the existing ones.

A direct consequence of these results is that firm-bank distance should increase following Broadband Internet expansion, because 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 dependant variable. Figure 6 displays the results.

Figure 6: Firm-bank distance

(a) Unweighted

(b) Weighted by credit exposure



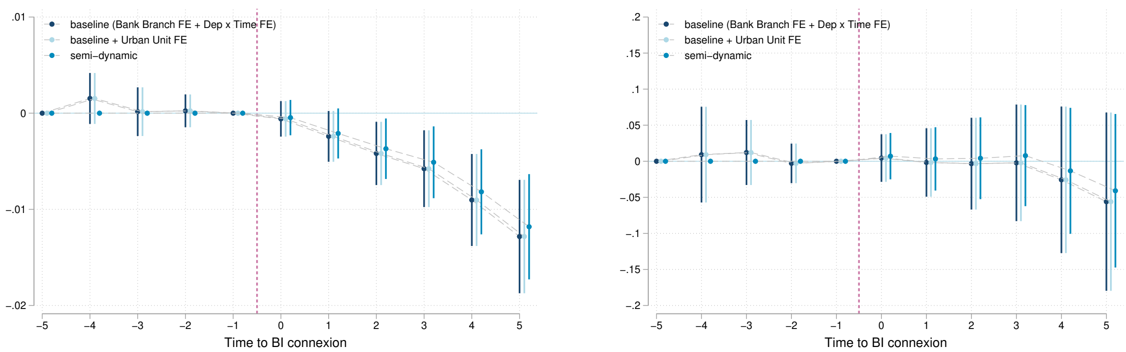

**Notes:** These figures plot estimates for specification in equation (23) – fully dynamic – and equation (24) – semi-dynamic. In the left panel, the dependant variable is the (log) average firm-bank distance measured at the branch level. In the right panel, the (log) distance is weighted by credit exposure. The sample include all bank branches with a positive credit exposure. The baseline specification (navy blue) includes bank branch fixed effects and year × county fixed-effects. Urban unit fixed-effects are included in the alternative fully-dynamic specification (light blue). 95% confidence interval are presented. Standard errors clustered at the city level.

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 in average 10% further, 5 years after the shock. Figure 6(a) shows that the result is not significantly affected if one considers the average distance weighted by credit exposure. These results are in line with Kroszner and Strahan (1999) which points out that innovation in information technology reduced the dependence on geographical proximity between customers and banks in the US, starting in the 70s. This is also consistent to Petersen and Rajan (2002) that documents the erosion of the local nature of small business lending, with increasing distance between small firms and their lenders in the United States but also new communication habits. Similar trends are observed in France: inter-regional credit flows have grown by 15% and the average firm-bank distance has increased by 10% between 1998 and 2005. As far as I know, this paper is the first to provide a causal interpretation for those 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 in local credit markets. This suggests that innovations in information technology 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 in local credit markets. Additionally, Figures A1 and A2 in Appendix A show no impact on bank branch size, measured as total credit or the number of clients, neither on the average loan size (intensive margin). Those findings are consistent with gravity results presented above, that emphasize the extensive margin channel, and confirm that Broadband Internet allowed to establish new firm-bank relationships between submarkets, rather than to increase the size of existing credit relationships.

Finally, I study the consequences for bank branches 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 total number of relationships and the portfolio average investment grade.

Figure 7: Bank risk

(a) Share of risky borrowers (b) Average borrowers risk




**Notes:** These figures plot estimates for specification in equation (23) – fully dynamic – and equation (24) – semi-dynamic. In the left panel, the dependant variable is the the ratio of risky relationships to total number of relationships measured at the branch level. In the right panel, the portfolio average investment grade.. The sample include all bank branches with a positive credit exposure. The baseline specification (navy blue) includes bank branch fixed effects and year $\times$ county fixed-effects. Urban unit fixed-effects are included in the alternative fully-dynamic specification (light blue). 95% confidence interval are presented. Standard errors clustered at the city level.

Results in Figure 7 indicate no risk taking by connected bank 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 5) decreases slightly, by 10% five years after the connection.

# 7 Implications for the cost of debt

In this last 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 (15) 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 $\kappa_{uv}$ for a pair of connected cities formally writes:

$$\Delta\,\mathrm{ln}\kappa_{uv} = \hat{\gamma}\,\Delta\,\mathbb{C}_{uv} = -0.058 \tag{25}$$

The distribution of the minimum loan price (Equation 7) obtained by an entrepreneur located in $v$ rewrites as follow:

$$\begin{aligned}\mathbb{P}\big(r_{e_v} \leq r\big) = W_{e_v}(r) &= 1 - \exp\Big(-r^{\theta} z_{e_v}^{\theta+1}\sum_{u=1}^{N} S_u \cdot (c_u d_{uv})^{-\theta}\kappa_{uv}\Big)\\ &= 1 - \exp - \left(\frac{r}{z_{e_v}^{-\frac{\theta+1}{\theta}}\big(\sum_{u=1}^{N} S_u \cdot (c_u d_{uv})^{-\theta}\kappa_{uv}\big)^{-\frac{1}{\theta}}}\right)^{\theta}\end{aligned}$$

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

$$P_{e_v} = \mathbb{E}[r|e_v] = z_{e_v}^{-\frac{\theta+1}{\theta}}\Big(\sum_{u=1}^{N} S_u \cdot (c_u d_{uv})^{-\theta}\kappa_{uv}\Big)^{-\frac{1}{\theta}} \times \Gamma\big[1 + \frac{1}{\theta}\big] \tag{26}$$

where $\Gamma$ stands for the Gamma function. By taking the log of Equation 26, I can isolate the effect on $P_{e_v}$ of the technology-induced reduction in search frictions caused by the diffusion of Broadband Internet:

$$\Delta lnP_{e_v} = ln\,\frac{P_{e_v}(1)}{P_{e_v}(0)} = -\frac{1}{\theta}\,ln\left(\frac{\sum_{u=1}^{N}\hat{\kappa}_{uv}(1)\cdot S_u(0)\cdot(c_u d_{uv})^{-\theta}(0)}{\sum_{u=1}^{N}\hat{\kappa}_{uv}(0)\cdot S_u(0)\cdot(c_u d_{uv})^{-\theta}(0)}\right) \tag{27}$$

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 (27) simplifies as follow:

$$\Delta ln\hat{P}_{e_v} = -\frac{1}{\theta}\,ln\left(\frac{\sum_{u=1}^{N}\hat{\kappa}_{uv}(1)\cdot F\hat{E}_u(0)\cdot dist_{uv}^{-\hat{\nu}\hat{\theta}}}{\sum_{u=1}^{N}\hat{\kappa}_{uv}(0)\cdot F\hat{E}_u(0)\cdot dist_{uv}^{-\hat{\nu}\hat{\theta}}}\right) \tag{28}$$

While Broadband Internet may affect all variables in this equation as the firm productivity $z_{e_v}$, the cost $c_u$ and the $\theta$ paramater of the bank branches' size distribution, I keep them constant to conduct counterfactual exercises. I finally 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 ($\theta$, $dist_{uv}$) in order to compute $\Delta ln\hat{P_{e_v}}$ the change in the lower expected cost of debt triggered by the BI-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 results echoes the conclusions of Hauswald and Marquez (2003) that shows 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 firm search and I do not model the process by which banks search for customers (and information about those customers), but the intuition of an easier *dissemination of information* is similar. In a recent theoretical approach, Vivès and Ye (2021) also studies how the diffusion of information technology affects competition in the lending market. Here again, they do not model search frictions (in their model, ICT reduces 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.[23]

Figure 14 illustrates the spatial dimension of the decline in firms' cost of debt. Dark blue cities are the one in which this decline is the strongest (larger than the $75^{th}$ percentile) while the light grey ones indicates a reduction lower than the median. Figure 15 gives a similar point of view with a focus on the Paris region. 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 results is in line with the model intuition that a reduction in search frictions precipitates an increase in competition due to the ability for firm 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 were not highly constrained by the search frictions and the decline is low. On the contrary, in isolated and rural submarkets where firms had to make costly efforts to multiply meetings with different bankers and eventually match

[23] In this paper, 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 project (through productivity $z_b$ or $d_{uv}$). This effect is absorbed by the urban unit $\times$ time fixed-effects in my empirical analysis.

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 D1. I show in particular 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. This suggests that the age structure of the workforce is a proxy for the probability of technology adoption, in line with evidence from Meyer (2011).

# 8 Conclusion

I develop a new theory of firm-bank matching subject to search frictions. I provide a causal evidence on how such frictions affect firm-bank matching and the allocation of bank credit, using the staggered roll-out 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 by 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. On the contrary, the effect the effect is almost null when two neighbouring cities, already economically very closely tied, 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 highlight the role of transactions and search cost in shaping firm's access to credit. Credit markets with high search frictions make financing by bank credit both difficult, time-consuming and onerous, especially for small businesses. This conclusion calls for a variety of economic policies aiming at to make the process of searching and applying for credit more fluid, efficient and less burdensome, in particular 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.

**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**, "L'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.

**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.

_ **, Nathan H. Miller, Mitchell Petersen, Raghuram Rajan, and Jeremy Stein**, "Does function follow organizational form? Evidence from the lending practices of large and small banks," *Journal of Financial Economics*, 2005, *76* (2), 237–269.

**Bhuller, Manudeep, Andreas Kostøl, and Trond Christian Vigtel**, "How Broadband Internet Affects Labor Market Matching," Memorandum 10/2019, Oslo University, Department of Economics 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.

**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.

**Cole, Rebel, Lawrence G. Goldberg, and Lawrence White**, "Cookie Cutter vs. Character: The Micro Structure of Small Business Lending by Large and Small Banks," *Journal of Financial and Quantitative Analysis*, 2004, *39* (02), 227–251.

**Cordella, Tito and Eduardo Levy Yeyati**, "Financial opening, deposit insurance, and risk in a model of banking competition," *European Economic Review*, 2002, *46* (3), 471–485.

**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.

**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.

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

**Eaton, Jonatahn, Samuel Kortum, and Francis Kramarz**, "Firm-to-Firm Trade: Imports, exports, and the labor market," Discussion papers 16048, Research Institute of Economy, Trade and Industry (RIETI) March 2018.

**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.

**FED**, "Small Business Credit Survey," 2014.

**Gabaix, Xavier**, "Power Laws in Economics: An Introduction," *Journal of Economic Perspectives*, 2016, *30* (1), 185–206.

**Gourieroux, Christian, Alain Monfort, and Alain Trognon**, "Pseudo Maximum Likelihood Methods: Theory," *Econometrica*, 1984, *52* (3), 681–700.

**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.

**Hubbard, Robert, Kenneth Kuttner, and Darius N Palia**, "Are There Bank Effects in Borrowers' Costs of Funds? Evidence from a Matched Sample of Borrowers and Banks," *The Journal of Business*, 2002, *75* (4), 559–81.

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

**Ioannidou, Vasso and Steven Ongena**, ""Time for a Change": Loan Conditions and Bank Behavior when Firms Switch Banks," *Journal of Finance*, 2010, *65* (5), 1847–1877.

**Jing, Harvey**, "Financial strength and monitoring capital: Evidence from firm-bank matching," *Working paper, Xiamen University.*, 2014.

**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," Technical Report 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).

**Matutes, Carmen and Xavier Vives**, "Competition for Deposits, Fragility, and Insurance," *Journal of Financial Intermediation*, 1996, *5* (2), 184–216.

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

**McKinsey**, "The lending revolution: How digital credit is changing banks from the inside," *McKinsey Report*, 2018.

**Meyer, Jenny**, "Workforce age and technology adoption in small and medium-sized service firms," *Small Business Economics*, October 2011, *37* (3), 305–324.

**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.

**OECD**, "Enhancing SME access to diversified financing instruments," 2018.

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

**Paravisini, Daniel, Veronica Rappoport, and Philipp Schnabl**, "Specialization in Bank Lending: Evidence from Exporting Firms," Working Paper 21800, National Bureau of Economic Research December 2015.

**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.

**Rauch, James E.**, "Business and Social Networks in International Trade," *Journal of Economic Literature*, December 2001, *39* (4), 1177–1203.

**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.

**Schwert, Michael**, "Bank Capital and Lending Relationships," *Journal of Finance*, 2018, *73* (2), 787–830.

**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.

**Stein, Jeremy C.**, "Information Production and Capital Allocation: Decentralized versus Hierarchical Firms," *The Journal of Finance*, 2002, *57* (5), 1891–1921.

**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.

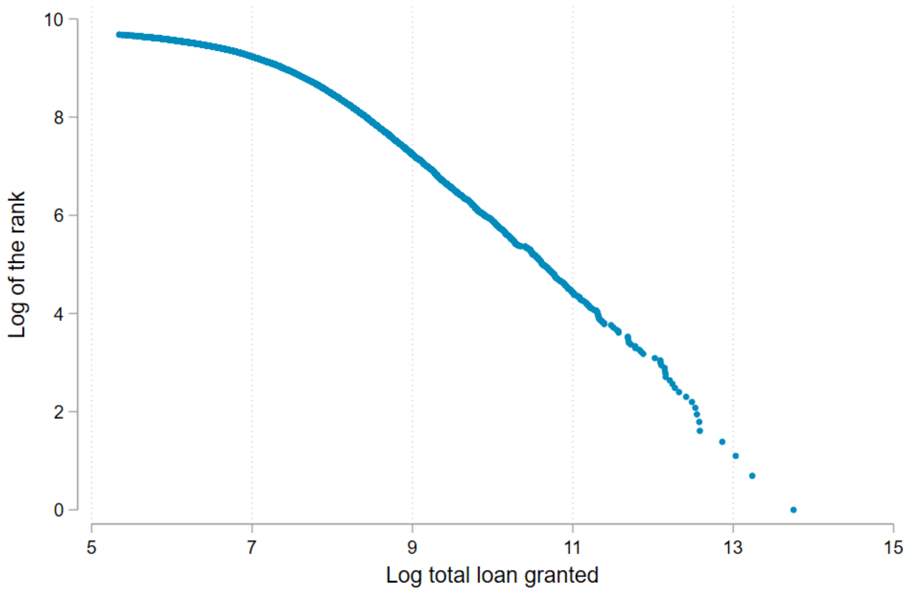
**Vivès, Xavier and Zhiqiang Ye**, "Information Technology and Bank Competition," *CEPR Working Paper*, 2021.

**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.

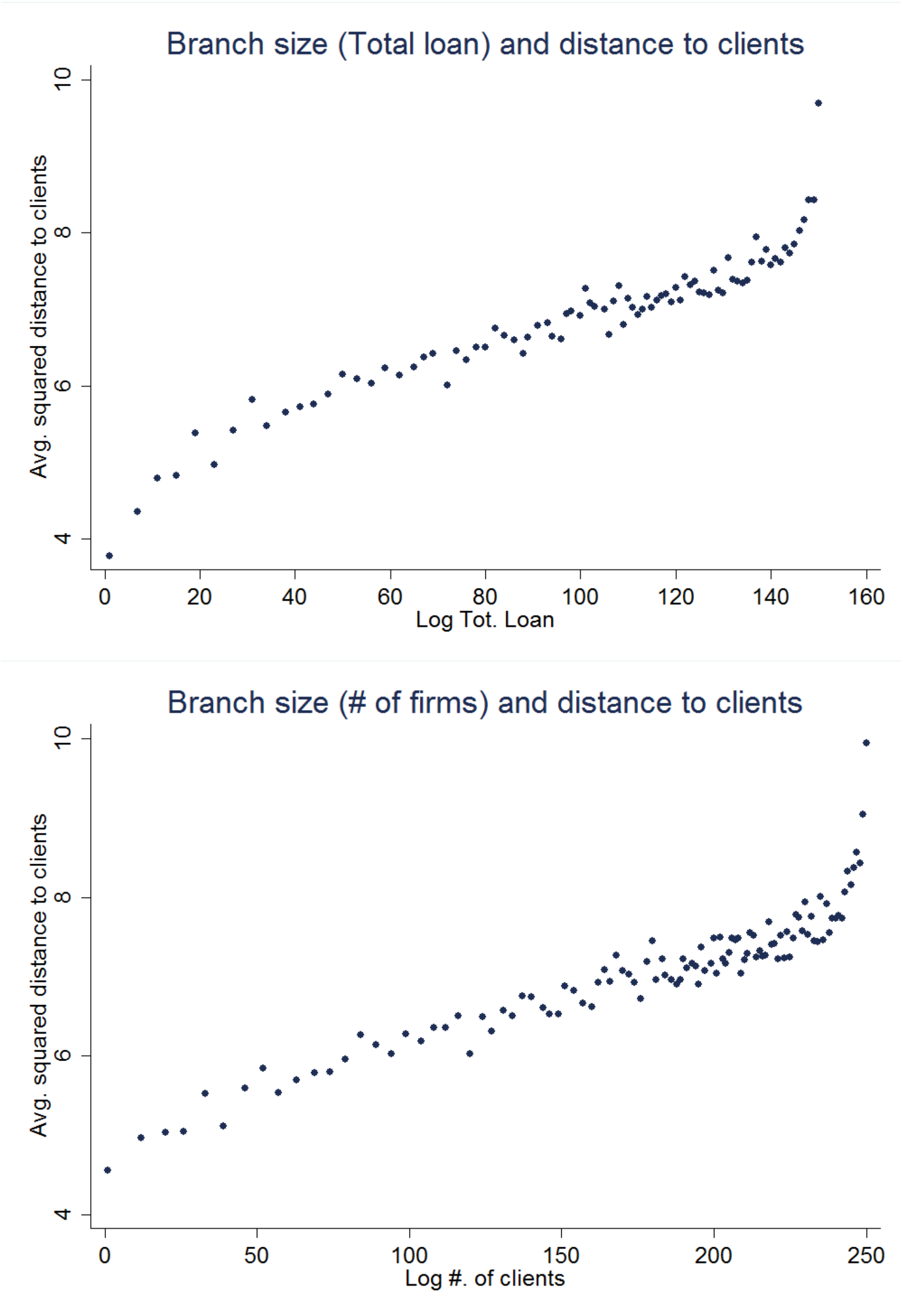
# 9 Figures

Figure 8: Branch Rank versus Size (Total credit)



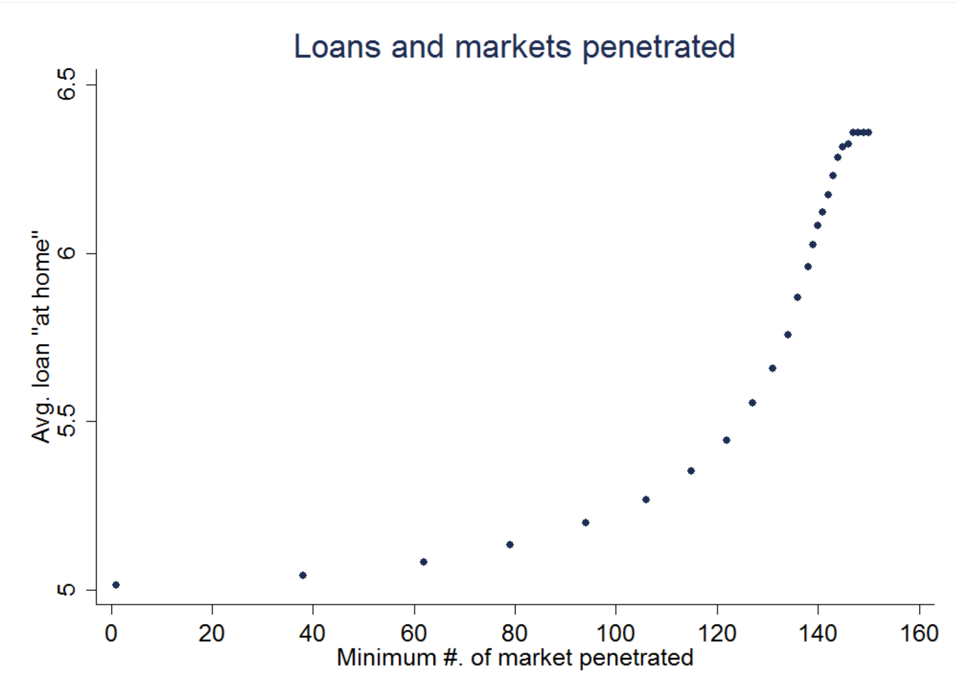

**Notes:** This figure displays the distribution of bank branch size for the last quarter of 2005, for all banks with at least 5 clients. Formally, it shows scatter plot of the log (Size) against the log (Rank). I compute the size of a branch as its number of clients. Bank branches are ranked by size: #1 being the largest branch, #2 the second largest, and so on.

Figure 9: Branch Size and Average Distance to Clients



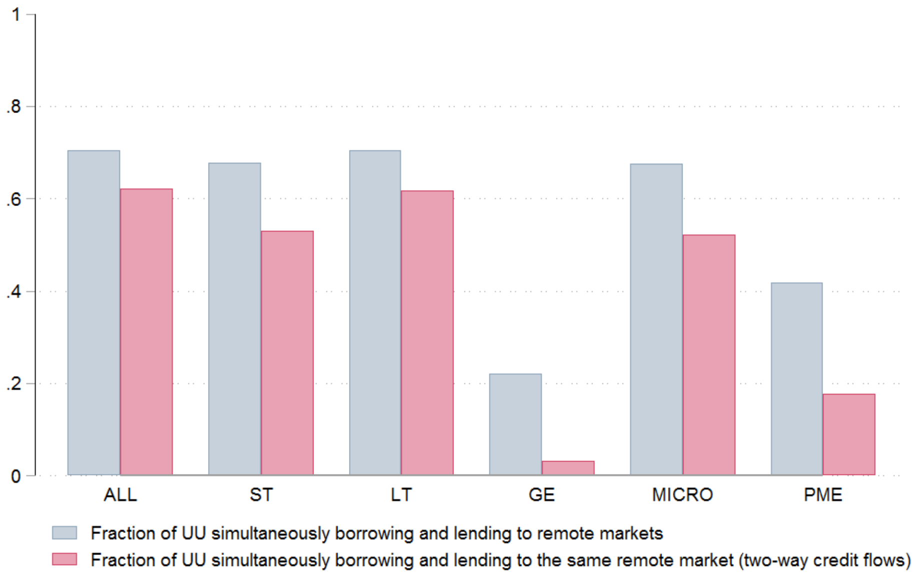

**Notes:** This figure displays shows the positive correlation between branch size measured as total credit exposure (top panel) and, alternatively, as the number of clients (bottom panel) and average square geographic distance between the branch and its clients, in kilometers, for the last quarter of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

Figure 10: Average Loan Locally and Number of Distant Markets Penetrated



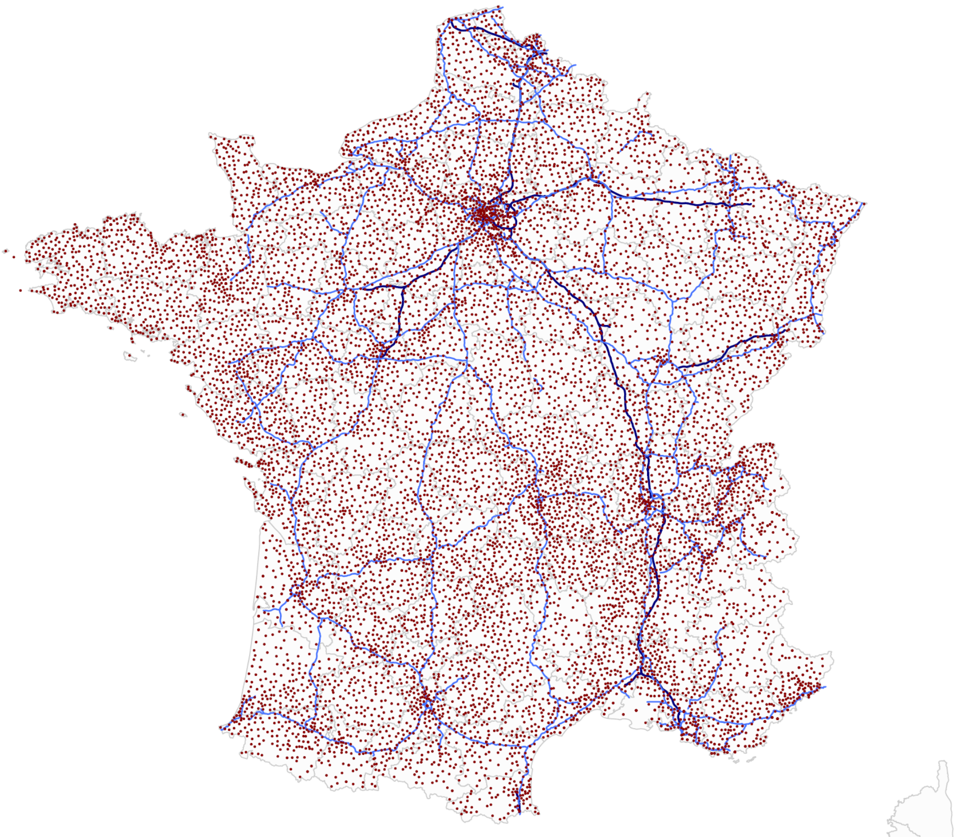

**Notes:** This figure displays the log average credit size of a branch in its local submarket ("at home") for the group of branches that operate at least in $k$ remote submarkets, with $k$ on the $x$ axis. Bank branches are ranked and grouped based on the number of remote submarkets (i.e. urban units)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

Figure 11: Inter-Submarket and Two-Way Lending



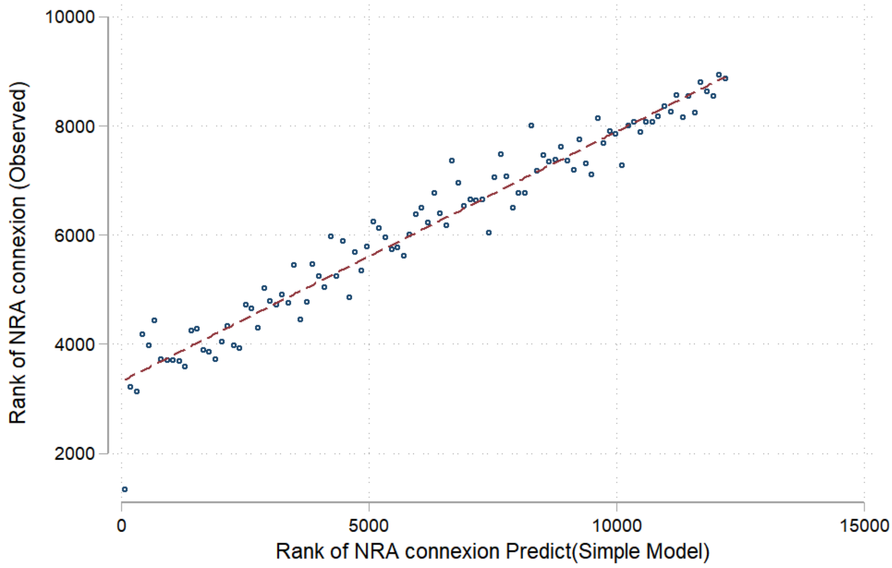

**Notes:** This figure shows the share of urban units that *borrow* and *lend* simultaneously to remote submarkets (in blue). ALL indicates that all types of credit and all type of clients are included in the computation, while ST stands for short-term credit, LT for long-term. GE indicates that urban units *borrow* and *lend* simultaneously to large firms, MICRO to very small and PME to medium size firms. The red bars indicate the share of urban units simultaneously *borrow* and *lend* to the same distant submarket, with similar sub-categories.

Figure 12: Local Exchanges, Highways and Railroads before 1999



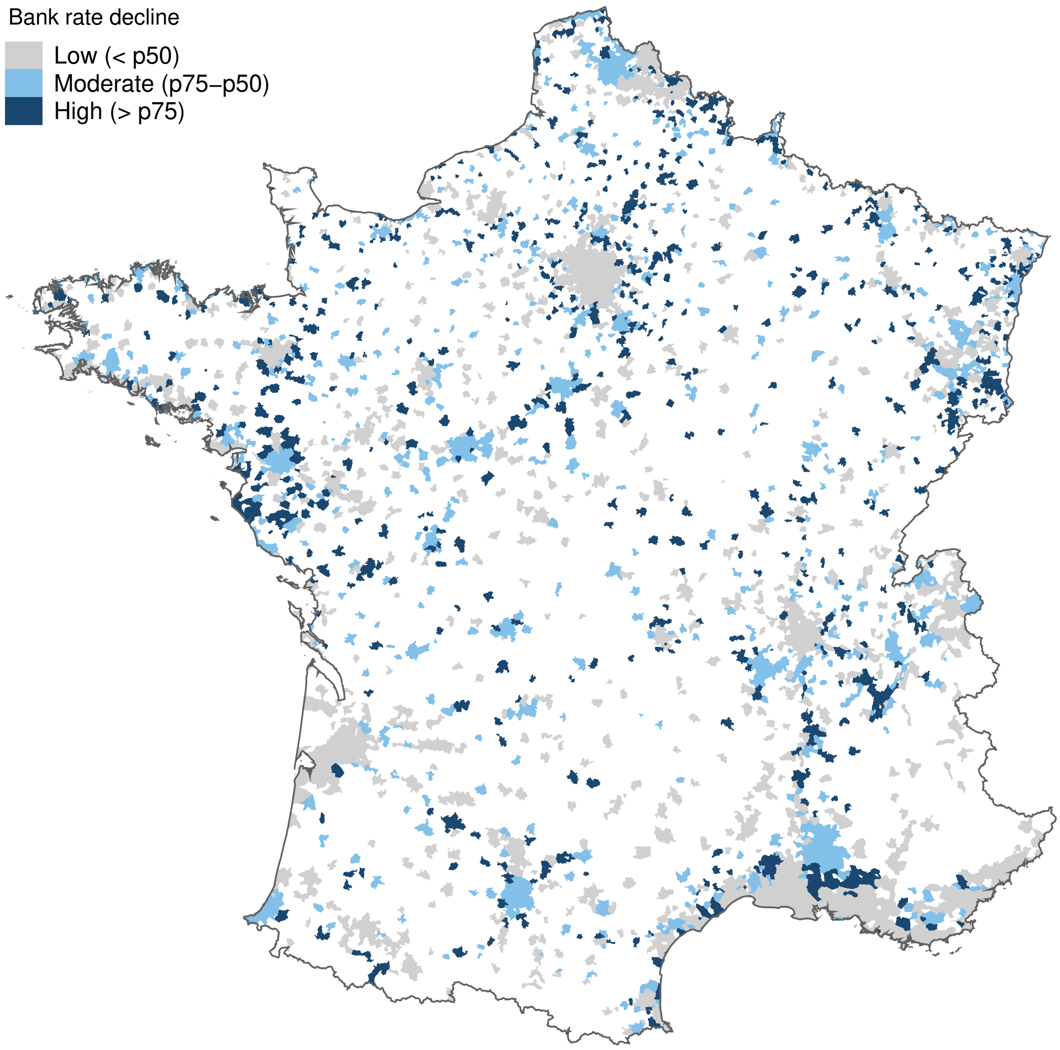
**Notes:** This figure displays the location of around 13,000 Local Exchanges (red dots), highways (light blue lines) and railroads (dark blue lines) already existing before the beginning of Broadband Internet expansion in France.

Figure 13: Optimal connection rank predicted vs. observed connection rank



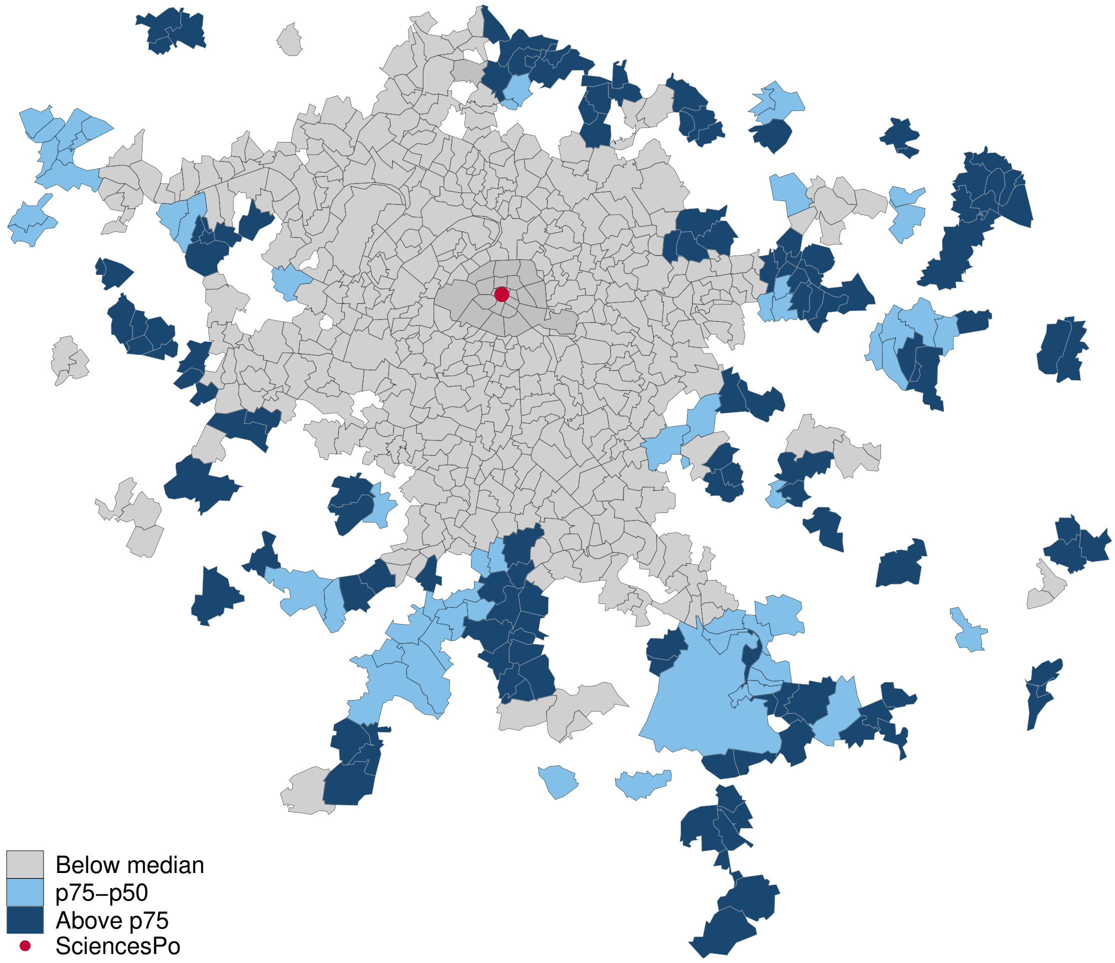

**Notes:** This figure shows the rank correlation between optimal versus observed connection ranks. The combination of exogenous connection *gains* and *costs* have a strong predictive power, with a R-square close to 0.70. The *optimal connection rank* $\hat{R}_i$ is predicted for each Local Exchange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

Figure 14: 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 one in which this decline is the strongest, larger than the $75^{th}$ percentile. The light blue areas undergo a decline of the cost of debt higher than the median (but lower than the $75^{th}$ percentile). Finally the light grey ones indicates a reduction lower than the median.

Figure 15: Reduction of the cost of debt triggered by a reduction in search frictions: zoom in Paris region




**Notes:** This map illustrates the spatial dimension of the decline in firms' cost of debt with a focus on the Paris region. Dark blue cities are the one in which this decline is the strongest, larger than the $75^{th}$ percentile. The light blue areas undergo a decline of the cost of debt higher than the median (but lower than the $75^{th}$ percentile). Finally the light grey ones indicates a reduction lower than the median. The red dot indicates the localisation of SciencesPo, i.e. the center of Paris.

# 10 Tables

Table 4: 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**: PPML estimation of equation (21). $dist_{uv}$ = 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. The sample consists of origin-destination-year combinations with positive credit flows, where at least one firm is located with positive credit.

Table 5: TECHNOLOGY-INDUCED REDUCTION IN SEARCH FRICTIONS: PAIR FIXED-EFFECTS

| | **Share of credit in v borrowed from u:** $\Pi_{uvt}$ | | | |
|---|---|---|---|---|
| | (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) $\times$ Year FE | Yes | Yes | Yes | Yes |
| Destination (v) $\times$ 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**: PPML estimation of equation (22). $dist_{uv}$ = 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 urban units. $\mathbb{C}_{uvt}$ 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 instrument variable strategy delivers a similar measure of between urban unit connectivity 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 $\times$ years and destination $\times$ years, columns (2) and (4) add pair fixed-effects. The sample period is 1997-2005. The sample consists of all origin-destination-year combinations where at least one firm is located with positive credit.

Table 6: Technology-Induced reduction in search frictions with lags

| | Share of credit in v borrowed from u $\Pi_{uvt}$ | | | | | |
|---|---|---|---|---|---|---|
| | (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) $\times$ Year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Destination (v) $\times$ 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**: PPML estimation of equation (22) with lag dependant 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 urban units. $\mathbb{C}_{uvt}$ 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 instrument variable strategy delivers a similar measure of between urban unit connectivity 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 $\times$ years and destination $\times$ years. The sample period is 1997-2005. The sample consists of all origin-destination-year combinations where at least one firm is located with positive credit.

Table 7: Technology-Induced reduction in search frictions: extensive margin

| | Share of firms in v borrowing from u: $\Pi_{uvt}$ | | | | |
|---|---|---|---|---|---|
| | (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) $\times$ Year FE | Yes | Yes | Yes | Yes | Yes |
| Destination (v) $\times$ 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**: PPML estimation of equation (21). $dist_{uv}$ = 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 urban units. $\mathbb{C}_{uvt}$ 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 instrument variable strategy delivers a similar measure of between urban unit connectivity 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 $\times$ years and destination $\times$ years. The sample period is 1997-2005. The sample consists of all origin-destination-year combinations where at least one firm is located with positive credit.

Table 8: TECHNOLOGY-INDUCED REDUCTION IN SEARCH FRICTIONS: EXTENSIVE MARGIN

| | **Share of firms in v borrowing from u $\Pi_{uvt}$** | | | | |
|---|---|---|---|---|---|
| | (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) $\times$ Year FE | Yes | Yes | Yes | Yes | Yes |
| Destination (v) $\times$ Year FE | Yes | Yes | Yes | Yes | Yes |
| Origin-Destination FE | No | No | No | Yes | Yes |
| $R^2$ | 0.65 | 0.66 | 0.68 | 0.69 | 0.69 |
| Observations | 24545811 | 24545640 | 24545640 | 24545640 | 250780 |

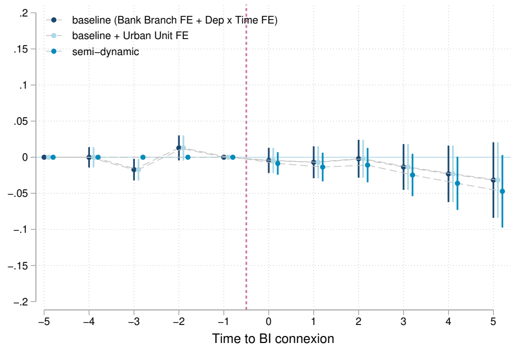
**Notes**: PPML estimation of equation (21). $dist_{uv}$ = bilateral distance. $\mathbb{C}_{uvt} = Z_{vt} \times Z_{ut}$, is a continuous variable that indicates the degree of broadband internet inter-connectivity between two urban units. $\mathbb{C}_{uvt}$ belongs to $[0, 1]$. This measure captures the ability for firms located in $u$ to locate and communicate with bank branches located in $v$, using the world wide web. My instrument variable strategy delivers a similar measure of between urban unit connectivity 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 $\times$ years and destination $\times$ years. The sample period is 1997-2005. The sample consists of all origin-destination-year combinations where at least one firm is located with positive credit.

# Appendix

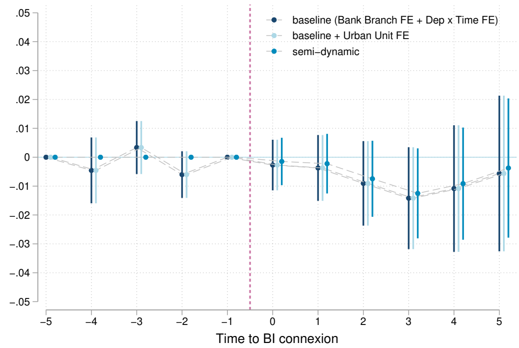
## A Additional results

Figure A1: Bank branch total and average credit

(a) Total credit



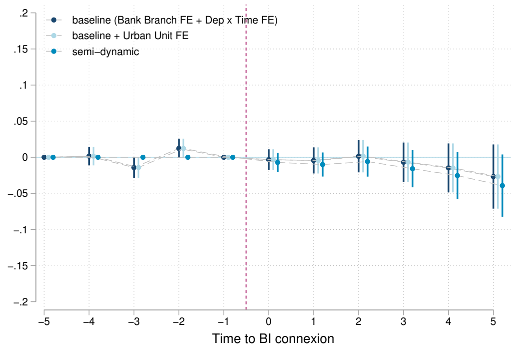

(b) Average credit



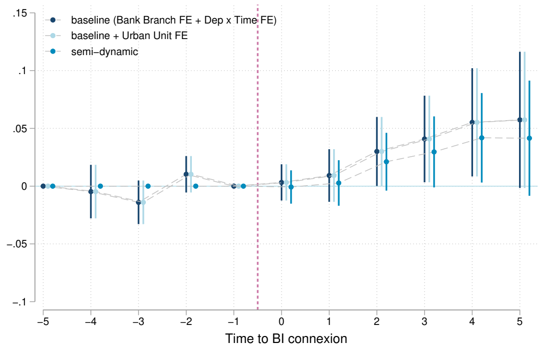

**Notes:** These figures plot estimates for specification in equation (23) – fully dynamic – and equation (24) – semi-dynamic. In the left panel, the dependant variable is the (log) total credit measured at the branch level. In the right panel, the (log) average credit per loan. The sample include all bank branches with a positive credit exposure. The baseline specification (navy blue) includes bank branch fixed effects and year × county fixed-effects. Urban unit fixed-effects are included in the alternative fully-dynamic specification (light blue). 95% confidence interval are presented. Standard errors clustered at the city level.

Figure A2: Bank branch total and remote clients

(a) Number of clients

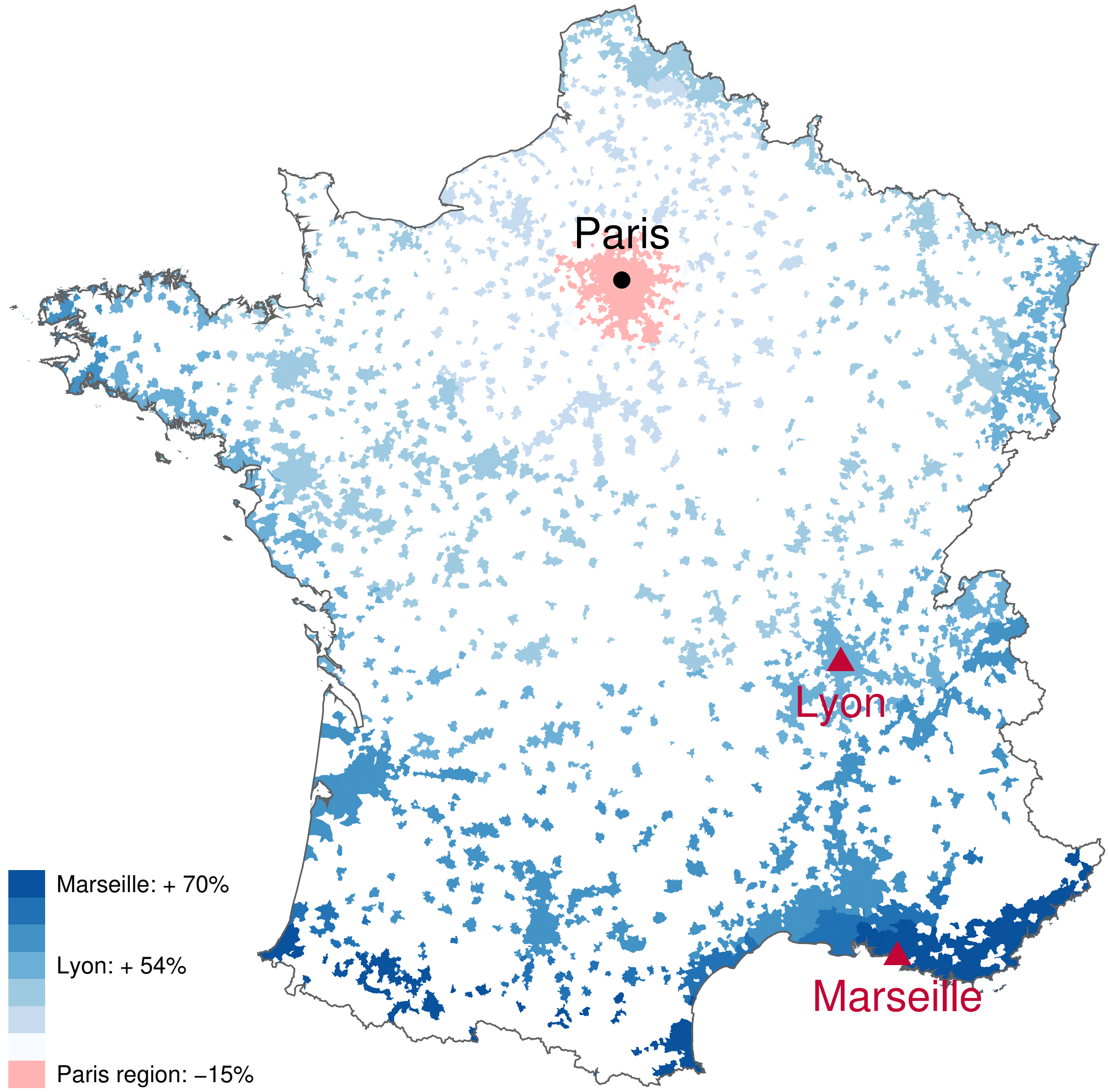



(b) Number of remote clients




**Notes:** These figures plot estimates for specification in equation (23) – fully dynamic – and equation (24) – semi-dynamic. In the left panel, the dependant variable is the (log) number of clients measured at the branch level. In the right panel, the (log) number of remote clients. The sample include all bank branches with a positive credit exposure. The baseline specification (navy blue) includes bank branch fixed effects and year × county fixed-effects. Urban unit fixed-effects are included in the alternative fully-dynamic specification (light blue). 95% confidence interval are presented. Standard errors clustered at the city level.

Figure A3: Heterogeneous effect of BI 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 shows Marseille and Lyon, respectively the second and the third biggest french cities. Dark blue indicates an effect in the the 90th percentile while light red indicates the negative effect of being connected to Paris for cities located close to Paris.

Table A1: PPML WITH MANY ZEROS IN A DYNAMIC SETTING: SIMULATION RESULTS

**Data Generating Process I** Obs. = 200 x 200 x 10, a = 50 b = 0

$\mathbb{P}(Y_{uvt} = 0) = 0.97$

| *Estimator* | $\beta_1$ | | $\beta_2$ | | $\beta_3$ | |
|---|---|---|---|---|---|---|
| | 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** Obs. = 200 x 200 x 10, a = 1 b = 5

$\mathbb{P}(Y_{uvt} = 0) = 0.88$

| *Estimator* | $\beta_1$ | | $\beta_2$ | | $\beta_3$ | |
|---|---|---|---|---|---|---|
| | 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 in Section 5. 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.

# B Search Frictions in Credit Markets

In this section, I present a 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; then, I describe the geography of credit flows and I provide evidence of substantial price dispersion; Finally, I mirror those empirical facts with new survey evidence.

## B.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 offices characteristics appear as critical factors to firms, especially SMEs, when choosing their financial services providers (Berger et al., 1997). While prior literature shows that the matching of firms and banks is endogeneous and depends on firm and bank size (Stein, 2002; Hubbard et al., 2002; Cole et al., 2004; Berger et al., 2005), geographic proximity (Petersen and Rajan, 1995, 2002), export country specialization (Paravisini et al., 2015), monitoring capacity (Jing, 2014) and bank capitalization (Schwert, 2018), little evidence exists on the importance of branch characteristics. In this section, I focus on bank branches – rather than bank – heterogeneity. I document four important facts. Branches differ markedly from each others 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.** Figure 2 and 8 display the distribution of branch size for the last quarter of 2005. I compute the size of a branch as its total credit exposure (2) and, alternatively, its number of clients (8). Then, I rank branches by size[24]: #1 being the largest branch, #2 the second largest, and so on; finally, I plot the log (Size) versus the log (Rank).

[FIGURE 2 ABOUT HERE]

[24] Regressing the log rank on log size, I find the following:

$$\log(\text{Rank}) = 12.29 \underset{[0.002]}{\mathbf{-1.04}} \cdot \log(\text{Size})$$

The relationship is close to a straight line (R2=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. A slope of approximately 1 has been found repeatedly using data on city and firm size, stock markets returns, etc. (Gabaix, 2016)

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 branch size follows a power law (i.e. Pareto distribution): a few number of very 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 9 shows the positive correlation between branch size (measured as total credit exposure and, alternatively, as the number of clients) and average square geographic distance between the branch and its clients, in kilometers, for the last quarter of 2005.

[FIGURE 9 ABOUT HERE]

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 is. Bank branches differ in their role and their ability to finance remote clients. Smaller branches focus on their very local market and concentrate on proximity lending: example (branches with less than 10 clients are in average 15.5 km from them). On the contrary, largest branches grant credit to firms located in other cities or regions (branches with more than 100 clients (top 1%) are, in average, 80 km from them). This heterogeneity may be the result of banks internal organization and simply reflects the hierarchy between local branch offices and large business centers. This 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 article.

**Branch Size and Market Entry.** How does a bank branch ability to finance distant clients relate to its local activity? I rank and group bank branches based on the number of remote submarkets (i.e. urban units) where they operate: all my branches lend at least to one submarket while none are active in all. Figure 10 depicts average credit size in branch local submarket ("at home") for the group of branches that operate at least in $k$ remote submarkets, with $k$ on the $x$ axis.

[FIGURE 10 ABOUT HERE]

The relationship between the average credit exposure in the branch local submarket and the number of remote submarkets penetrated is strongly positive and corroborates the findings of 9. The larger a branch is locally, the higher the number of distant submarkets served. This striking regularity reveal 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 than Duquerroy et al. (2019), I show that, within a bank and a urban unit, branches specialize in several dimensions: industry, type of loans and type of businesses. In particular, some bank branches appear to finance heavily SMEs in comparison with other branches located in a same submarket and belonging to a 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 urban unit.

## B.2 Endogenous Firm-Branch Matching

In section B.1, I documented the fact that bank branches markedly differ from each others in several dimensions. Thus, search frictions may arise if firms pay attention to those branch characteristics and devote time and ressources to locate and match with the *ideal* interlocutor. Does this observed heterogeneity affect how firms search for the right banking partner ? Does it lead to endogenous matching ? To address those questions, I investigate the importance of branch characteristics in new firms matching decisions.

**Data Sources.** I use a new dataset of firm creations from 2002 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 ranges from 9,958 in 2002 to 14,178 in 2005. After their establishment, firms match with a single bank branch: the period of time running from firm entry to firm first banking match is two years on average. However, more than 50% of those new firms find their banking partner within their first year of existence. For the sake of simplicity, I keep in my final sample firms that match with their first banking partner in less than 5 years, which represent 90% of the observations: I end up with 45,685 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.

**Regression Specification.** 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 specialisation 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 the 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 :

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

where $\Phi$ is the c.d.f. of the standard normal distribution; Spec. $\text{Sector}_{j,t-1}$ (respectively Spec. $\text{Size}_{j,t-1}$) takes the value 1 if branch $j$ is specialized in lending to firms from the same sector (respectively same size) as firm $i$, the year before firm $i$ entry. $\text{Size}_{j,t-1}$ is the log size of branch $j$, measured as total credit exposure and, alternatively, as the number of clients and Size $\text{Growth}_{j,t-1}$ is the growth rate of branch $j$, both measured the year before firm $i$ entry. $\text{Distance}_{i,j}$ is simply the log geodesic distance in kilometers between branch $j$ and firm $i$. All explanatory variables are lagged so the branch $j$ characteristics are measured before firm $i$ entry and are not contaminated by the realized firm-branch matching. The dependent variable is an indicator function that takes the value 1 if firm $i$ is borrowing from branch $j$ at t.

**Coefficients Interpretation.** The coefficient $\alpha$ 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 (see Agarwal and Hauswald, 2010) as being close to clients eases the acquisition and the use of private information in informationally opaque credit markets.
Coefficients $\beta_1$ and $\beta_2$ control for heterogeneity in size and dynamics of bank branches documented in section B.1. $\beta_1 > 0$ and $\beta_2 > 0$ would 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 $\beta_3$ and $\beta_4$ control for the direct impact of branch specialization (industry or size category) on firm's choice of banking partner: this relates to Paravisini et al. (2015), documenting that banks specialize in one export market and that specialization affects a firm's choice of new

Table B1: Endogeneous 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**: Probit estimation of different specifications of Equation (29). Dependent variable is an indicator function that takes the value 1 if firm $i$ is borrowing from branch $j$ at 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 $i$ 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 $i$ entry. Distance$_{i,j}$ is simply the log distance between branch $j$ and firm $i$. All explanatory variables are lagged. Standard errors clustered at the urban unit level.

lenders and how to finance exports. $\beta_3 > 0$ (respectively, $\beta_4 > 0$) would mean 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' $i$ industry (alternatively, the firm' $i$ size category).

**Results.** Table B1 shows the marginal effects from the Probit estimation of different specifications of equation (29). Standard errors are clustered at the urban unit level. In every specification, all coefficients are statistically significant (at the 1 percent confidence level), and of the expected signs. Both physical proximity ($\alpha$), branch size ($\beta_1$) and branch growth ($\beta_2$) increase the likelihood that a firm match 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, I find reduced-form evidence that bank branch heterogeneity does matter 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, which could suggest the presence of search costs if firms devote time and ressources to gather information about potential banking partners before applying for a loan.

## B.3 The Geography of Bank Credit

I describe the geography of credit flows and document new facts consistent with the presence of search and matching frictions. In a frictionless world in which all bank branches are similar, meeting with many bankers is not costly and firms always benefit from borrowing from neighbouring banks,[25] if possible, and inter-urban unit lending should be marginal. In particular, if a branch and a firm located in the same urban unit are simultaneously offering and applying for credit, a match is likely to occur locally. Therefore, we shouldn't observe a branch and a firm, while being located in the same urban unit, simultaneously lending and borrowing from another distant urban unit.

**Simultaneous inter-submarket lending.** In order to verify this assumption, I construct credit flows between each pair of urban units, using data from the French Credit Register between 1998 and 2005. Credit flows from $u$ to $v$ are defined as the total credit granted by all bank branches located in $u$ to remote firms located in $v$: urban units may be lenders, borrowers or both. Table 11 presents the results.

[FIGURE 11 ABOUT HERE]

The first bar (in blue) indicates that more than 70 % of urban units *borrow* and *lend* simultaneously to remote submarkets. One could think of it as being the result of regional specialization. A firm located in a urban unit where branches are not able to offer a very specific type of credit, or to lend to a specific industry / firm size, may be forced to search for creditors elsewhere. Yet, the other blue bars indicates that a vast majority of urban units *borrow* and *lend* simultaneously the same type of credit to remote submarket: SMEs in urban unit $v$ borrow long-term credit (LT) to branches in urban unit $u$, while branches in urban unit $u$ simultaneously lend LT credit to SMEs in urban unit $z$.

**Simultaneous two-way lending.** More interestingly, I show that *two-way lending* occurs. Red bars on figure 11 indicate that more than 60 % of urban units simultaneously *borrow* and *lend* to the same distant submarket. Some bank branches located in Paris finance firms in Bordeaux, while some bank branches in Bordeaux finance firms in

[25] A large body of work has highlighted the importance of geographical distance on firm-bank relationship, especially for small firms (Petersen and Rajan (1995), Hauswald and Marquez (2006)). Degryse and Ongena (2005) show that lenders located in the vicinity of small firms face significantly lower transportation and monitoring costs: *Banks derive market power ex ante from their relative physical proximity to the borrowing firms or ex post from private information they obtain about firms during the course of the lending relationship.* Consequently, most of the lending activity should concentrate within a city or an urban unit, between a firm and a branch in close proximity to each other.

Paris. This fact is robust when I control for the type of credit and the type of firm (red bars 2 to 6).

## B.4 Price Dispersion

Finally, inspired by the labor market literature, I document the fact 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 show 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). Note that controlling for loans and borrowers characteristics does not affect my result. Formally, I estimate equation (30) that aims to explain the observed variation in loan prices:

$$\begin{aligned} \text{Interest rate}_{ijtu} = {} & \text{Loan}_{ijtu} \cdot \rho_1 + \text{Firm}_{ijtu} \cdot \rho_2 \\ & + \text{FE}_{s(u)} + \text{FE}_j + \text{FE}_t + \text{FE}_u + \epsilon_{ijtu} \end{aligned} \tag{30}$$

where $i$ stands for the borrower, $j$ for the bank branch, $t$ for the quarter and $u$ indicates the urban area in which both the firm and the bank branch operate. $\text{Loan}_{ijtu}$ is a vector of loan characteristics (term in months, amount, type of rate: fixed or variable), $\text{Firm}_{ijtu}$ 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 "Departments"), a sector (NACE Rev. 2 French classification) fixed-effect and a quarter fixed-effect. Table 1 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.

As I am mostly interested in the explanatory power of these different groups of observable variables and fixed-effects, I only report the $R^2$ of each regression. 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. Note that if my reported $R^2$ is somehow larger than Cerqueiro et al. (2011) findings, this is due to the very strict inclusion of fixed-effects, notably bank branch FE, in my empirical model.

## B.5 Survey Evidence

In this section, I document a series of evidence of search frictions from recent surveys that are consistent with empirical findings presented above. First, industry reports show that locating the right banking partner is not straightforward and that firms commonly multiply loan applications: the FED (2014) survey indicates that 3 applications are submitted on average (2.7 institutions contacted). Second, surveys highlight the fact that the application process for a loan is time-consuming: 33 hours are spent applying for credit on average according to the FED (2014) report, consistent to Infosys (2018) survey: *SMEs spend over 25 hours – simply on their loan request paperwork – and have to approach numerous banks with their application*. Among the SMEs, 26% (29%) deplore a difficult application process with large (small) banks.

Additionally, surveys highlight important transaction delays. Industry reports show delays in the range of 45-90 days between the application and the closing date. The Infosys (2018) survey indicates that 24% (29%) of SMEs report a long wait for the credit decision or funding with large banks (small banks), namely *high underwriting, transaction and search costs*. The OECD (2018) report "Enhancing SME access to diversified financing instruments" corroborates this conclusion: transaction costs are particularly high in relative terms for micro-entreprises, start-ups, young SMEs. This costly search, among other factors, may result in firm resignation: 37% of businesses appear to give up their search for finance and cancel their spending plans after their first rejection (BIS/BMG Research, 2018).

# C PPML: Difference-in-Difference with Many Zeros?

In this section, I test the performance of the PPML estimator in the exact case of my empirical setting. Not only my estimating sample contains a vast majority of urban unit pairs that do not exchange credit over the entire period 1998-2005 (more than 95% of the credit shares equal zero) but also I have to deal with panel data and a difference-in-difference approach. I present simulation evidence on the performance of the PPML estimator when the panel data is 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($\beta\, X'_{uvt}$). A the best of my knowledge, this is the first simulation evidence of the performance of the PPML estimator in this particular setting, that 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 bank branches located in $v$ that serve firms in $u$, and $z_{juvt}$ is the amount of credit that each those banks lent their clients located in $u$. Because $m_{uvt}$ and $z_{juvt}$ are independant, $\mathbb{E}(Y_{uvt}|X_{uvt}) = \mathbb{E}(m_{uvt}|X_{uvt}) \times \mathbb{E}(z_{juvt}|X_{uvt})$. 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 conditionally 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 $x_{1u}$ and $x_{1v}$. $x_{2uvt}$ is a time-varying variable equal to 1 with 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 $\epsilon_{uvt}$ is drawn from a standard normal, such that $x_{2uvt} = \mathbb{I}(w_{2uvt} > 0.6)$. $x_{3uvt}$ is a city pair specific treatment dummy variable that equals one after the time of treatment $t$ and zero 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 $\beta_0$ = 0, $\beta_1$ = -1, $\beta_2$ = 1 and $\beta_3$ = 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) considering the quadratic specification $\mathbb{V}ar(m_{uvt}) = a\mathbb{E}(m_{uvt}|X_{uvt}) + b\mathbb{E}(m_{uvt}|X_{uvt})^2$ so that:

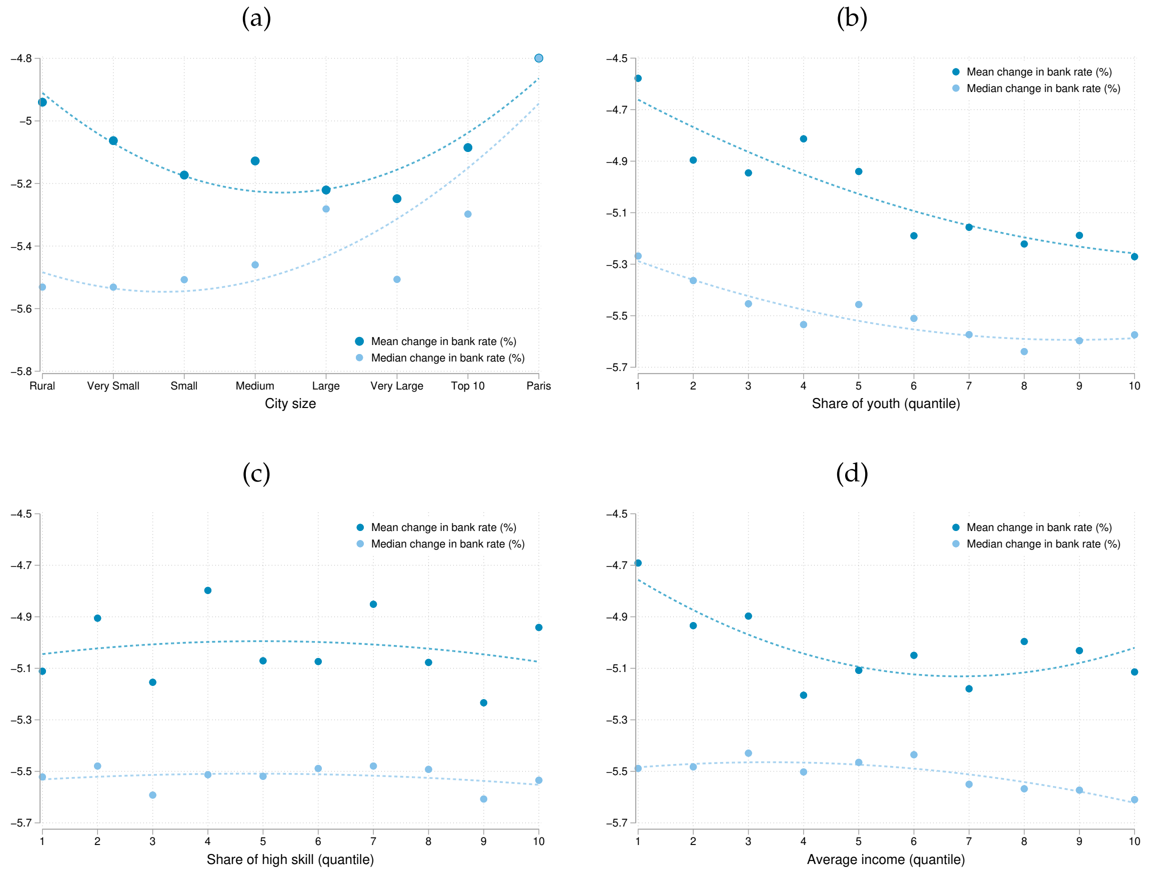
$$\mathbb{V}ar(Y_{uvt}|X_{uvt}) = (1 + 2a)\mathbb{E}(m_{uvt}|X_{uvt}) + 2b\mathbb{E}(m_{uvt}|X_{uvt})^2$$

Picking the value of $a$ and $b$ allows to generate a high probability of zeros and different heteroskedasticity patterns. Table A1 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 $\times$ 200 $\times$ 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[26] with and without origin $\times$ year and destination $\times$ year fixed-effects. These results confirm and extend the findings of Santos Silva and Tenreyro (2006, 2011), showing that both the PPML and the GPML estimators are well behaved in the two cases considered. In particular, the coefficient of interest $\beta_3$ is consistently estimated with two-way unit $\times$ fixed effects.These findings are an additional reason that justify my empirical approach and the validity of my estimation procedure.

[26]GPML stands for Gamma Pseudo-Maximum Likelihood is a PPML-like estimation procedure in which the dependant variable is a share instead of the value in level.

# D Heterogeneity in the decline of the cost of debt

Figure D1: 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 figure (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.