

# Geographical-Proximity Bias in P2B Crowdlending Strategies\*

Carole GRESSE<sup>†</sup>      Hugo MARIN<sup>‡</sup>

This version: June 9<sup>th</sup>, 2021

## Abstract

Using data from a peer-to-business crowdlending platform that exploits an auction-driven system to fund corporate loans, we show that non-professional investors are subject to a geographical-proximity bias. They are more likely to win the auctions of borrowers located close to their place of residence notwithstanding that they are not better informed about their creditworthiness. Unexpectedly, this behavioral bias distorts the loan rate discovery process by increasing the cost of funding for borrowers. This adverse effect results from the greater ability of local investors to submit winning bids at an early stage. This ability is gained from their experience in previous auctions of geographically close borrowers. This suggests that the familiarity feeling stemming from geographical closeness strengthens investor attention, and thereby improves lenders' knowledge about the dynamics of the order flow in local borrowers' auctions.

*JEL codes:* G11, G14, G23, G43.

*Keywords:* geographical-proximity bias; peer-to-business crowdlending; crowdfunding; behavioral finance; loan performance; price discovery process.

---

\*We would like to thank Sabrina Buti, Jérôme Dugast, Zoran Filipović, Jean-Paul Laurent, Gaëlle Le Fol, Jacopo Magnani, Tamara Nefedova, and Evgenia Passari as well as participants at the 14<sup>th</sup> financial risks international forum, at the 37<sup>th</sup> Spring AFFI conference, and at the finance doctoral seminar of Paris Dauphine-PSL, for helpful comments. We thank Unilend for providing us with the data. We gratefully acknowledge financial support from the Fintech Chair of Université Paris Dauphine - PSL. The views expressed in this paper are of our own.

<sup>†</sup>Carole Gresse is at Université Paris Dauphine - PSL, DRM, CNRS, France.

<sup>‡</sup>Hugo Marin is at Université Paris Dauphine - PSL, DRM, CNRS, France.      Email: hugo.marin@dauphine.psl.eu

# 1. Introduction

While small and medium businesses (SMBs) are important to any economy, accessing to funds is usually a major issue for them.<sup>1</sup> For instance, SMBs accounted for 99.9% of businesses and 47.5% of the total number of employees in the U.S. in 2018.<sup>2</sup> Yet their access to financial markets is fairly limited. They therefore depend on banks to finance their operating needs and investments. Peer-to-business crowdfunding<sup>3</sup> help alleviate this difficulty by making investors finance SMBs directly, without intermediation, on electronic platforms, either via equity (equity crowdfunding), or via debt (crowdlending), in the hope of obtaining financial gains.<sup>4</sup> Since crowdfunding platforms offer numerous projects to fund, investors must screen and select firms that they deem likely to yield positive returns. In the case of crowdlending, positive returns are ensured by the full repayment of loans so that borrowers' creditworthiness should be at the heart of investors' screening strategies. However, given that the average financial literacy of non-professional investors participating in crowdlending is by nature weaker than that of professional lenders, we fear behavioral biases to be strong on those platforms and to consequently cause overinvestment or underinvestment issues.

In this study, we investigate the decision-making process of non-professional investors to check whether this process is influenced by a behavioral bias that we designate as the *geographical-proximity bias*. This bias, which resembles the home bias at the international level, is however quite different in essence. The home bias, i.e., the tendency to overinvest in domestic assets relative to foreign assets, can be caused by exogenous factors such as transaction costs, foreign exchange risk, fiscal barriers, a lack of information about foreign

---

<sup>1</sup>Abraham and Schmukler (2017) address this issue and suggest some solutions to overcome it.

<sup>2</sup>More details about the data on American SMBs available at <https://www.sba.gov/sites/default/files/advocacy/2018-Small-Business-Profiles-US.pdf>.

<sup>3</sup>Crowdfunding can be defined as a way of raising funds from a large number of investors, out of the banking system, generally through an internet platform. In peer-to-business crowdfunding, fund raisers are corporations while, in peer-to-peer crowdfunding, fund raisers are individual borrowers.

<sup>4</sup>Peer-to-business crowdfunding should not only be seen as a way to raise funds for firms. Belleflamme et al. (2013) show that some entrepreneurs use crowdfunding as a marketing means. Chemla and Tinn (2020) elaborate the idea that firms use crowdfunding to analyze consumers' preferences and to gauge demand for their products.

assets, etc. On the contrary, the geographical-proximity bias does not relate to such external factors. It pertains to a broader notion, known as the familiarity bias, which consists in overinvesting on assets we are more familiar with. With the geographical-proximity bias, familiarity stems from geographical closeness. In other words, if investors are subject to this bias, they would overinvest in firms located near their place of residence because geographical closeness would provide them with a feeling of familiarity.

Using data from a French peer-to-business crowdlending platform that uses an auction system to fund loans, we examine whether geographical closeness between lenders and borrowers creates biases in the bidding strategies of lenders. Our identification strategy relies on the implementation of a bidding algorithm that submits bids on behalf of investors. We interpret the implementation of the bidding algorithm as an exogenous decrease in the geographical-proximity bias, provided the bias exists. We find evidence that investors have a preference for firms that are located in the same administrative division as them, but only during the period before the implementation of the bidding algorithm. Investors located in the same administrative division as the borrower are more likely than other investors to participate in the auction, and they are also more likely to win it. We then investigate whether this preference to lend money to geographically close borrowers is information-driven or purely behavioral. By analyzing loan performance, we find that investors who lend money to borrowers located in their administrative division are as likely as other investors to experience a default on those loans. This proves that investors do not own private information on the creditworthiness of closely located firms but it supports the idea that they develop a familiarity feeling towards them. Last, we investigate whether the geographical-proximity bias deteriorates the efficiency of the loan rate discovery process during the auction. We find that the presence of geographically close investors in the auctions is associated with an increase in the borrowers' cost of funding. We explain this result by the fact that closely located investors time their bids differently from other investors. They participate earlier in the auction and then become passive in the final moments, when the competition in rate

takes place. This decreases the level of competition between bidders in the pricing mechanism, and thereby increases the borrower's cost of funding. We also find that the positive link between borrowing rates and the weight of local investors in auctions is driven by the weight of local investors who have already participated in previous local auctions. This suggests that those investors have inferred information on the order flow of local auctions from their previous experience. This knowledge, probably resulting from increased attention due to a familiarity feeling, allows them to submit winning bids at higher rates in further local auctions, which in turn tends to increase the borrowing rates of those auctions.

The remainder of the paper is structured as follows. Section 2 reviews the relevant literature and presents our hypotheses. Section 3 describes the data and the strategy used to identify the geographical-proximity bias. Section 4 tests the existence of the geographical-proximity bias and section 5 analyzes its impact on the efficiency of the price discovery in the auction system. We provide robustness checks in Section 6. Section 7 concludes.

## 2. Literature review and hypotheses development

A substantial extent of the literature on crowdfunding<sup>5</sup> has focused on investors' behavioral biases. Most of the past studies on this matter have addressed the case of peer-to-peer crowdlending, notably because of the ease of accessing the data of Prosper.com, one of the leading U.S. peer-to-peer crowdlending platform. [Ceyhan et al. \(2011\)](#) study the dynamics of investors' bidding behavior in an auction-driven crowdlending platform. They show that investors have a herding behavior at the beginning and at the end of each auction. [Dorflleitner et al. \(2019\)](#) show that investors stop investing on a peer-to-business crowdlending platform after experiencing a default, which leads to portfolio under-diversification. [Hervé and Schwienbacher \(2018\)](#) show that investors are more likely to use round-numbers when levels of uncertainty rise on an equity crowdfunding platform. More broadly, [Adhami et al. \(2019\)](#) investigate the return of investors of a peer-to-business crowdlending platform and

---

<sup>5</sup>See [Moritz and Block \(2016\)](#) for a detailed literature review on crowdfunding.

show that returns are inversely related to loan risk. We contribute to this growing literature by examining a behavioral bias that has been scarcely investigated in this context: the geographical-proximity bias.

## 2.1. The geographical bias and related literature

The geographical-proximity bias is a specific form of a broader concept: the familiarity bias, which can be defined as a tendency to prefer what is familiar. For instance, [Heath and Tversky \(1991\)](#) show that individuals prefer betting on their own judgment when they are familiar with the matter. Such behaviors exist in stock markets. For example, [Frieder and Subrahmanyam \(2005\)](#) show that individual investors hold portfolios that are overweighted towards stocks with high visibility whereas institutionals' holdings are more related to firm size and beta. The familiarity bias can also distort investors' expectations as [Benartzi \(2001\)](#) shows that investors that are "very familiar" with a specific stock are more likely to extrapolate past performance to future performance. This type of familiarity bias can be designated as a geographical-proximity bias when familiarity results from geographical closeness. Such a bias has already been identified in financial markets. [Huberman \(2001\)](#) shows that investors tend to invest a relatively large portion of their portfolios in their local phone company. [Seasholes and Zhu \(2010\)](#) show that investors hold portfolios that are overly weighted towards local stocks while they do not earn abnormally higher returns on those stocks. This suggests that investors' preferences for local stocks is a behavioral bias. [Wolf \(2000\)](#) finds evidence of a home bias in commodity trades in the U.S. at the state level, which suggests that international trade barriers do not fully explain the higher propensity of agents to trade locally.

The familiarity bias in general, and the geographical-proximity bias in particular, are also present in online markets. [Hortaçsu et al. \(2009\)](#) use data from two auction sites, eBay and Mercadolibre, and find that transactions are more likely to occur between counterparts located in the same city. [Agrawal et al. \(2011\)](#) highlight the existence of a familiarity bias

on a crowdfunding platform that connects artists and investors. They show that this bias is mainly driven by the family and friends of the artists. In a study close to our research, [Lin and Viswanathan \(2016\)](#) exploit the moves of borrowers across states to show the existence of a home-state bias on a U.S. peer-to-peer crowdlending platform. They also find evidence that this bias is mainly behavioral. We directly contribute to this strand of the literature by showing that investors that are more familiar with some firms because of their geographical proximity are more likely to lend to them.

The geographical-proximity bias resembles the widely documented home bias<sup>6</sup> in the sense that both biases result in overinvesting in home assets, home being the home country in one case and a more restricted local area in the other one. However, their determinants radically differ. Although some authors as [Lewis \(1999\)](#) fail to explain the home bias in U.S. equity markets with rational factors, it is generally acknowledged that the home bias is caused by exogenous factors such as trading costs, fiscal barriers, or lack of information. For instance, [Ahearne et al. \(2004\)](#) explain the home bias in equity holdings by the presence of information asymmetries. On the contrary, the geographical-proximity bias is fully behavioral and not related to information. For that reason, finding investors' preferences for local borrowers on crowdlending platforms does not suffice to prove the existence of that bias. We also need to check that those preferences are not driven by superior information held on closely located firms. This point is important for two reasons. First, there exists examples in the literature of apparent familiarity biases that were proved to be information-based choices. For example, [Massa and Simonov \(2006\)](#) show that Swedish investors heavily invest in stocks related to their non-financial income. They however find that this bias is not behavioral but rather information-driven. Second, several papers show that rational creditworthiness screening or rational information inference may explain patterns observed on crowdfunding platforms. [Zhang and Liu \(2012\)](#) find that investors do not passively follow

---

<sup>6</sup>For instance, [Disdier and Head \(2008\)](#) conduct a broad study to show that the home bias is persistent through different markets, methodologies, and time periods. Similarly, [Hau and Rey \(2008\)](#) show that mutual funds from various countries are affected by the home bias.

the actions of other investors but instead actively try to infer the creditworthiness of borrowers by observing the lending decisions of other investors. They interpret such behavior as rational herding. Vallée and Zeng (2019) study the adverse selection issue arising from heterogeneous levels of sophistication between investors on a crowdlending platform. They show that sophisticated investors lose part of their out-performance relative to unsophisticated investors after a substantial decrease in the borrowers' public information reported by the platform. Comparably, Mohammadi and Shafi (2017) show that institutional investors have better performance than individual lenders because of a higher ability in screening on a peer-to-business crowdlending platform.

## 2.2. Hypotheses development

This paper aims to show whether the decision-making process of non-professional investors is affected by a geographical-proximity bias. This bias leads investors to invest more heavily in projects located near their location of residence. We therefore expect investors located close to a borrower to be more likely to participate in and to win its auction. To determine whether an investor is located close to a borrower, we use the geographic administrative divisions of the French national territory. France has three levels of geographic administrative divisions, of which the highest is the *région* or region.<sup>7</sup> *Régions* divide into *départements*, in turn composed of cities, named *communes* in French. On average, the surface area of a *département* is four times that of a county in the U.S. and *communes* are equivalent to U.S. municipalities. We consider that a lender is geographically close to a borrower if the lender lives in the same administrative division as the borrower, the administrative division level under consideration being the *département*,<sup>8</sup> and we construct the two following testable hypotheses:

*H1: Investors located in the same administrative division as the borrower are more likely*

---

<sup>7</sup>Until 2015, there were 25 *régions* in France. In 2015, their number was reduced to 18. In our analysis, we exclusively consider the pre-2015 administrative division.

<sup>8</sup>Unfortunately, our data does not allow to pursue the analysis at a finer level, namely at the city-level.

*to participate in the borrower's auction.*

*H2: Investors located in the same administrative division as the borrower are more likely to win the borrower's auction.*

Finding evidence in support of H1 and H2 would not be enough to prove the existence of a geographical-proximity bias because the higher likelihood of investors to lend to borrowers located close to them could be explained by an informational advantage about the borrowers' creditworthiness over the rest of the investors. As shown by Iyer et al. (2016) with data from Prosper.com, peer lenders use both hard and soft information to predict the rate of default of borrowers, and in our case, local investors may well have a privileged access to soft information about the borrowers of their administrative division. We therefore compare two theoretical hypotheses: (1) the view that investing more heavily on nearby borrowers is the outcome of a behavioral bias, referred to as the behavioral hypothesis; and (2) the view that those heavier investments are due to informational advantages, referred to as the rational hypothesis. According to the rational hypothesis, lenders located close to the borrower should have a greater ability than other lenders to avoid underperforming loans. Under the behavioral explanation, investors close to the borrower do not have more information about the creditworthiness of the borrower. They should therefore have either an equal or a higher probability to experience a default on the loan than the rest of the investors. To distinguish between the two hypotheses, we test H3.

*H3: Investors located in the same administrative division as the borrower are less likely to experience a default than the rest of the investors.*

Not rejecting H3 would validate the rational hypothesis while rejecting H3 would validate the behavioral hypothesis. More precisely, rejecting H3 after finding evidence in support of H1 and H2 would prove the existence of a geographical-proximity bias.

Finally, we aim to determine whether the geographical-proximity bias may affect the efficiency of the price discovery process in the auction system. Under the rational hypothesis, the presence of local investors, i.e., investors of the same administrative division of the



borrower, should improve the efficiency of the auction's price discovery process because their bids would be more informative. On the contrary, under the behavioral hypothesis, the presence of local investors either at best has no impact on, or deteriorates the efficiency of the price discovery process of the auction system by overinvesting on local firms' loans. To investigate this issue, we model theoretical fundamental loan rates as a function of common loan-rate determinants in a first stage, and in a second stage, we examine how loan rate deviations from those fundamental values relate to the presence of local investors in auctions. This translates into the following hypothesis.

*H4: Loan rate deviations from their fundamental values increase with the presence of local investors in auctions.*

### **3. Data and identification strategy**

The data comes from Unilend, a French peer-to-business crowdlending platform that provides French small and medium businesses with unsecured fixed-rate amortized loans. Those data are particularly appropriate for our research for three reasons. First, Unilend was a major actor of the crowdlending market over our observation period. Unilend was created in 2013 and was, at that time, the first peer-to-business crowdlending platform in France. At the end of 2016, it had become the third biggest crowdlending platform in France, with a market share slightly exceeding 15% of total lent money.<sup>9</sup> Second, Unilend only targets retail investors as potential lenders whereas many other crowdlending platforms, such as Prosper.com in the U.S. or Lendix in France, allow both retail investors and institutionals to fund loans. Third, Unilend is one of the seldom platforms on which loan rates are set by investors in Dutch auctions. All other platforms with sizeable market shares<sup>10</sup> use fixed-rate procedures in which investors bid at a rate previously determined by the platform according to the loan

---

<sup>9</sup>As of today, Unilend is still an important player. It has the typical business model of a marketplace which requires a critical mass of lenders and a steady flow of borrowers to be profitable. Due to a lack of borrowers, it went bankrupt in October 2018, and was acquired by another crowdlending platform, named Pretup, in December 2018.

<sup>10</sup>In France, the PrêtStory platform also operates Dutch auctions but its market share is tiny.

maturity and credit risk.

### **3.1. The Unilend auction system and bidding algorithm**

On Unilend, loans are funded through a sealed-bid auction system where investors place bids with the amount of money they are willing to lend and the interest rate they are willing to receive. The amount of the loan, its term, and the maximum duration of the auction are announced on the platform before the auction starts. Once the auction starts, investors bid in competition with each other at rates bounded between 4 and 10%. At the end of the auction, bids are sorted in ascending order of interest rates and bids' amounts are cumulated from the lowest interest rate up to the first rate at which the loan amount can be cleared. Each winning investor lends the amount of her winning bid(s) and receives the monthly service of the loan at the rate she bid. The maximum winning interest rate in an auction, i.e., the clearing rate, is referred to as the equilibrium rate of the loan. The funding rate of the borrower is computed as the average of all winning bids' rates weighted by the corresponding bid amounts. It is referred to as the average rate in the remainder of the paper. In the case that bids are inadequate to clear the loan amount, the loan is not funded.

Auctions' durations can be decomposed into two periods: the coverage period during which the cumulated bid amounts do not clear the loan amount and the competition period that starts as soon as an additional bid allows to clear the loan amount. During the coverage period, every bidder is a winner and there is no competition between bidders. During the competition period, bidders compete in rates with each other, and only bids with the lowest rates submitted potentially win the auction. We further consider the ratio of the competition period duration divided by the total duration of the auction as a measure of the competition intensity in the auction.

A specific feature of our data is the implementation of a bidding algorithm by the platform during the sample period. This algorithm, named Autolend, bids on behalf of investors according to parameters they preset on the platform. After the implementation

of this algorithm, bids can be either manual or automated. Investors can freely activate or deactivate Autolend. When activating Autolend for the first time, investors must declare their reservation rate, i.e., the minimum interest rate at which they accept to lend, for all possible loan maturity/credit rating combinations. Once those parameters are set, the algorithm automatically bids on their behalf at the maximum rate accepted in an auction any time this maximum rate is higher than or equal to their reservation rate. When a bid is outbid, the algorithm bids again at a new rate equal to the previous bid rate minus one tick, provided that this new rate is compatible with the investor's reservation rate.

### **3.2. Available information about borrowers**

Before the start of each auction, the platform provides information about the borrower and its project. More specifically, Unilend publishes on its platform the income statements and balance sheets<sup>11</sup> of the borrower up to two fiscal years before the year of the auction.<sup>12</sup> They also describes the borrower's business and the reason why it is seeking funds. In addition, Unilend credit analysts assign each loan a credit rating by using both hard information from the balance sheets and income statements and soft information from interviews with the firm's management, generally the CEO. Unilend uses a rating grid ranging from zero to five stars with an increment of 0.5 stars, the number of stars decreasing with default risk. Only projects with a credit rating above or equal to three stars are offered to investors. Since we expect investors to rely on those credit ratings to assess the riskiness of the borrowers, we check, as a preliminary test, to what extent those ratings are actually related to financial variables known as meaningful about creditworthiness. We run an ordered logistic regression in which the borrower's credit rating is the dependent variable, and firm size, cash holdings, leverage, and profitability measures are the regressors. Our estimates show that Unilend ratings significantly relate with all tested variables in the expected way. Firm size, cash scaled by total assets, and the operating margin are all positively and statistically correlated

---

<sup>11</sup>The income statements and balance sheets communicated to investors are not always audited.

<sup>12</sup>Only the year before the auction starts is generally provided.

with the borrower's credit rating, while the gearing ratio has a negative and statistically significant impact on that rating. Those findings confirm that the ratings advertised to investors on the platform are reliable synthetic measures of credit risk.

### 3.3. Sample and descriptive statistics

The initial sample provided by Unilend comprises 318 auctions run on their platform from November 2013 to October 2016 and accounting for a total of 1.6 million bids. Nine of those auctions were not completed because of insufficient demand from investors. They were thus removed from the sample. The Autolend algorithm was first used in April 2016. 247 loans were completed<sup>13</sup> before its implementation, and 62 loans were completed after. We also removed the auctions for which we had no data either on the size of total assets or the sales of the borrowing firm one fiscal year before the auction. This leaves us with a final sample of 290 completed auctions. Those 290 auctions account for 1,470,365 bids submitted by 9,059 unique investors. 11.8% of bidders are non-physical investors, i.e., companies. 89% of non-professional investors are men and non-professional investors.

We can identify in the data the *département* of residence of borrowers and investors, which is instrumental to study geographical proximity. Borrowers in our sample are located in 63 different *départements*. Active investors, i.e., investors that have submitted at least one bid in one of the 290 auctions composing our final sample, are located in 100 different *départements*. Bids submitted by investors located in the same *département* as the borrower represent 3.42% of the total number of bids.

Table 1 reports the descriptive statistics on the characteristics of the loans, the borrowers, the auctions, and the investors for our 290 sampled auctions. Loan amounts range from €10,000 to €400,000, with an average close to €75,000. Loan maturities range from 12 to 60 months, with an average of 42 months. Borrowers have on average a total asset size of 1.72 million euros and sales equal to 1.88 million euros in the fiscal year preceding the auction.

---

<sup>13</sup>An auction is completed if it has received enough bids to reach the loan's amount.

Investors are on average 43.6 years old at the time they submit their bids.

**Table 1 about here**

In Table 2, we compare auctions completed before the implementation of Autolend with auctions completed after Autolend. We test the difference in mean and median between the two groups of auctions for the main variables used in this paper. While the economic characteristics of borrowers and funded projects are relatively similar over the two sub-periods, the characteristics of the bids in the auctions are completely different.

**Table 2 about here**

Borrowers are similar in terms of total assets, sales, leverage,<sup>14</sup> and Unilend credit rating. The differences in loan size and loan maturity are statistically significant, the average loan amount and the average loan term being slightly smaller in the post-Autolend period, but those differences remain small in economic terms.

In contrast with the stability of loan and borrower characteristics, the bidding frequency is much higher in the auctions completed after the implementation of the bidding algorithm. Both the average number of distinct bidders and the average number of bids per auction substantially increase, with 719 distinct investors and 1,461 bids per auction on average in the period before the implementation of Autolend versus 1,467 investors and 20,845 bids after. This bidding frequency increase is a direct consequence of bid automation. Further, the algorithm monitors all auctions with equal attention whereas investors may not realize that an auction has opened or may forget that the auctions they were interested in are still open.<sup>15</sup> The constant monitoring by the Autolend algorithm leads to a higher level of competition between investors during the auctions. To provide evidence on this increased competition, we compute the relative proportion of the auction duration in which bidders

---

<sup>14</sup>We define debt leverage as the sum of suppliers debts, financial debts and other debts divided by total assets.

<sup>15</sup>Investors receive an e-mail when a new auction opens and also when their last bid in a live auction has been outbid.

are in competition with each other, i.e., the ratio competition period duration over auction total duration, referred to as the competition degree in the rest of the paper. For auctions completed before the implementation of Autolend, the competition period represents 45% of the total auction duration, while the competition period represents 88% of the total auction duration after the introduction of Autolend.

Because the algorithm constantly monitors auctions, it allows loans to be funded rapidly, sometimes in less than one second, which in turn increases the relative share of the competition period in the total auction duration. Consequently, the duration of auctions completed after the implementation of Autolend is much shorter, with an average of three days against 12 days for auctions completed before Autolend.

As a consequence of this increased competition, both the average and the equilibrium rates are lower for the auctions completed after the implementation of the algorithm. The average funding rate decreased, on average, from 8.23% for the pre-Autolend subsample to 6.42% for the post-Autolend subsample. We compare those relatively high interest rates with the rates at which banks lent money to French SMBs for similar loan term and loan amount between 2013 and 2016.<sup>16</sup> On average, banks lent money to SMBs at 2.22% during the period before Autolend, and at 1.71% during the period after Autolend. The large difference suggests that crowdlending is essentially used by SMBs that cannot access bank loans because they are considered too risky. An alternative explanation may be that SMBs prefer to borrow money through crowdlending platforms rather than banks because the borrowing process is lighter and faster. For instance, income statements and balance sheets do not need to be audited in order to borrow money through a crowdlending platform and firms receive the loan's principal on their bank account a couple of days after the end of the auction. However, the difference in funding rates makes this second explanation hardly credible, and we believe that crowdlending is complementary to banks rather than being a

---

<sup>16</sup>This analysis is based on the data from *Banque de France*, available at <https://www.banque-france.fr/statistiques/credit/credit/taux-des-credits-aux-entreprises>. Each completed loan is matched with the aggregated data on loan rates for SMBs from *Banque de France*.

substitute. Crowdlending platforms mainly finance SMBs that banks are not willing to lend to because they are deemed too risky. This is consistent with the findings of [Tang \(2019\)](#) who shows that peer-to-peer crowdlending is a complement to bank lending for small loans. It also confirms the predictions of the model of [de Roure et al. \(2019\)](#) that loans originated through crowdlending are, on average, riskier than bank loans.

### **3.4. Identification strategy**

As noted by [Lin and Viswanathan \(2016\)](#), the identification of the home-state bias is challenging. So is the identification of the geographical-proximity bias. Our identification strategy relies on the implementation of Autolend. This algorithm was implemented by the Unilend crowdlending platform in order to improve bidding mechanisms and accelerate their auction procedures. With Autolend, investors set in advance the parameters of their bids, i.e., at their reservation rate for a given loan maturity and credit rating. Their bids are then automatically submitted according to those preset parameters, so that they do not need to monitor the platform. Because geographical location is not part of the Autolend parameters, bids submitted by the algorithm on behalf of investors are free of any geographical-proximity bias. For auctions that were completed after the implementation of Autolend, automated bids represent on average 94.11% of the total number of bids in each auction. Automated bids represent 51.72% of winning bids, and they fund 47.74% of the loan amount on average. Thus, the impact of bias-free automated bids is large. As a consequence, the geographical-proximity bias should be much more present in auctions completed before the introduction of the bidding algorithm than in auctions completed after. Our identification strategy is therefore based on the comparison between auctions completed before the implementation of the bidding algorithm, for which we should observe the geographical-proximity bias if it exists, with the auctions completed after the implementation of Autolend, which are supposed to be less impacted by this bias.

A concern about our identification strategy is that active investors may differ between

the two periods. Indeed, if investors were different between the two periods, then the difference in the intensity of the geographical-proximity bias in both periods would not necessarily be caused by the implementation of the bidding algorithm, but could rather result from the difference in the profile of active investors. Fortunately, most of the investors that were active before the implementation of the algorithm were still active after. Before the implementation of Autolend, there were 7,908 active investors against 5,707 active<sup>17</sup> investors in the post-Autolend period. Out of the 5,707 investors that were active after the implementation of the bidding algorithm, 4,548 investors (79.69%) were already active in the period before. In addition, as reported in Table 2, the two populations of investors are similar in terms of age.

## 4. Testing the existence of the geographical-proximity bias

We begin this section by showing that investors prefer investing in closely located companies, i.e., companies located in their *département* of residence, further designated as local companies. We then investigate whether this preference is information-based or behavioral.

### 4.1. Testing the existence of investor preferences for local companies

To investigate whether crowdlending investors are subject to a geographical-proximity bias, we examine their likelihood to participate in an auction in the following logistic regression:

$$\begin{aligned}
 Participation_{i,j} = & \alpha_1 \cdot Geographical\ Closeness_{i,j} \\
 & + \alpha_2 \cdot Auction\ Duration_j + \alpha_3 \cdot Age_i + \alpha_4 \cdot X_j + \epsilon_{i,j}
 \end{aligned}
 \tag{1}$$

---

<sup>17</sup>An investor is considered to be active if she has at least submitted one bid in the auctions of the period considered.



where  $Participation_{i,j}$  is a dummy equal to one if investor  $i$  bids at least once in auction  $j$ , zero otherwise. Our main variable of interest is dummy  $Geographical\ Closeness_{i,j}$  which equals one if investor  $i$  participating in auction  $j$  is located in the same administrative division as the borrower (i.e., the same *département*), and zero otherwise. Investors for whom this dummy equals one will further be designated as local investors, term "local" being defined with respect to the borrower's place of residence. We control for the pre-announced duration of the auction, taken in logarithm and denoted  $Auction\ Duration_j$ , because investors have a higher likelihood to participate in auctions that last longer. We use  $Age_i$ , the age of investor  $i$  taken in logarithm, as a control for investor  $i$ 's wealth and investing experience. We expect it to positively impact the probability of participation.  $X_j$  is a vector of control variables at the auction level. This vector includes the loan amount and the loan maturity, both taken in logarithm, plus three dummies standing for the credit rating of the loan. Each of those dummies, further denoted  $Rating\ n$ , equals one if the credit rating is  $n$  stars, zero otherwise,  $n$  being alternatively 3.5, 4, and 5, and the base case being a three-star rating. The rationale for the controls composing  $X_j$  is as follows. Due to the finite number of investors registered on the platform, large-size loans take longer to be cleared in an auction. For that reason, the larger the loan amount, the easier to participate in the auction. All things equal, the loan maturity is expected to negatively impact the decision of participating in the auction, as a longer maturity means both a longer time before recovering the funds invested and greater default risk. Last, rating dummies provide a synthetic assessment of the riskiness of the investment. They should thus have an impact on the decision to participate in the auction. We also introduce fixed effects by administrative divisions to take into account the differences in economic development of those divisions. As an illustration, the Paris area as well as the east part of the French territory have much denser industrial networks than the rest of the country. The number of auctions in those more economically-active geographical areas is thus greater, which makes the likelihood of participating in auctions of those areas statistically greater, all things being equal. From there, finding a significantly positive  $\alpha_1$

coefficient in Regression (1) would be in support of the hypothesis that investors participate more in the auctions of local companies (H1).

The population of investors considered for the regression are those that have already signed in on the platform at the time of auction  $j$ . The full sample of auctions is divided into two groups: the first group is composed of the auctions completed before the implementation of the bidding algorithm and the second group is composed of the auctions completed after. As discussed in the previous section, our identification strategy relies on the comparison of those two groups, since the first group of auctions should be more impacted by that bias than the second one. As illustrated by the statistics of Table 2, the number of bids and the number of unique bidders per auction are much higher after the implementation of Autolend. Indeed, the automation of bid submissions completely changes the way auctions unfold. For those reasons, we run logistic regression (1) on both sub-samples, and we expect  $\alpha_1$  to be significantly positive for pre-Autolend auctions but statistically insignificant for the other group of auctions.

The results reported in Table 3 are in line with expectations. Local investors are statistically more likely than others to participate in auctions before the implementation of the bidding algorithm, but this effect disappears after the implementation of Autolend. These findings are in support of H1. Regarding control variables, an investor is more likely to participate in the auction if she is older. Age can be seen as a proxy for personal wealth here. The wealthier the investor, the higher her likelihood to participate. As expected, investors' participation likelihood is positively correlated with the duration of the auction. The correlation between the likelihood to participate in the auction and the size of the loan is unclear as the coefficients are of opposite signs over the two sub-periods.

### **Table 3 about here**

As expressed in H2, we also expect investors of the same administrative division as the borrower to have a higher chance to win the auction. We test this by conducting the

following logistic regression at the bid level:

$$\begin{aligned} \text{Winning Bid}_{i,j,t} = & \beta_1 \cdot \text{Geographical Closeness}_{i,j} + \beta_2 \cdot \text{Age}_i + \beta_3 \cdot X_j \\ & + \beta_4 \cdot X_{i,j,t} + \epsilon_{i,j,t} \end{aligned} \quad (2)$$

where  $\text{Winning Bid}_{i,j,t}$  is a dummy equal to one if the bid of investor  $i$  submitted at time  $t$  in auction  $j$  is served at equilibrium and ultimately funds the loan, zero otherwise. If H2 holds,  $\beta_1$  should be significantly positive for the group of projects funded before the implementation of Autolend and statistically insignificant for the other group of projects.  $X_j$  is the vector of controls already used in Equation (1), which comprises the loan amount, the loan term, and rating dummies.  $X_{i,j,t}$  is a vector of controls at the bid level. It includes the interest rate and the timing of the bid. The interest rate obviously determines the probability of winning the auction, as in a Dutch auction, the bids with the lowest rates have priority. The bid timing also impacts the likelihood to win the auction because the visibility on the potential clearing rate increases at the end of the auction. Submitting late thus allows investors to fine-tune their bid rates with the most up-to-date information on the actual demand for the loan, and thereby to get a higher chance to win the auction. The timing of a bid submitted at time  $t$  by investor  $i$  in auction  $j$  is computed as a score in which the numerator is the difference between the submission time of the bid under consideration and the time of the first bid submitted in the auction, and the denominator is the time elapsed between the first bid and the last bid submitted in the auction. This score, which ranges from zero to one, decreases with the earliness of the bid.

Results are reported in Table 4. For the group of auctions completed before the implementation of Autolend, bids submitted by local investors are significantly more likely to win the auction ( $\beta_1$  significantly positive), but  $\beta_1$  is statistically insignificant for the auctions completed after the introduction of the bidding algorithm. Those results are in support of H2. Estimated coefficients of controls are as expected. The lower the interest rate of a bid and the later its submission, the more likely the bid to win the auction. The likelihood to

win an auction decreases with the credit rating of the loan. This can be explained by the fact that riskier loans attract less investors. Most risk-averse investors may consider that, over a certain level of risk, projects are not worth investing at less than 10%, the maximum interest rate permitted on the platform. As a consequence, a lower rating leads to lower competition between investors which, in turn, increases the bid's likelihood to win the auction. The same kind of explanation can be brought up to explain the positive correlation found between the term of the loan and the bid's likelihood of winning the auction, as longer terms are associated with higher default risk. Last, the likelihood of winning an auction increases with the loan amount due to the limited elasticity of the funds offer on the platform.

**Table 4 about here**

## **4.2. Preferences for local companies: informed investing or behavioral bias?**

We have shown, in the previous section, that crowdlenders invested more heavily in the loans of local companies in the pre-Autolend period. We now investigate whether this is rational investing based on superior information held on nearby companies (rational hypothesis) or whether this is a familiarity bias related to geographical closeness (behavioral hypothesis). To this aim, we follow the methodology of [Lin and Viswanathan \(2016\)](#) and we analyze the performance of pre-Autolend loans after issue.

The data on loan performance is available for 204 loans out of the 236 loans auctioned before the implementation of Autolend. 164 loans were fully repaid at maturity; five loans are considered delinquent, i.e., the payment of the principal is late by more than 90 days but less than 9 months; and 35 loans are in default. A loan is proved to be in default in the following four cases: (1) the payment of the principal is late by more than 9 months; (2) the loan is under a recovery procedure; (3) the borrowing firm went bankrupt; or (4) the

borrowing firm is going into receivership. In total, 40 loans out of the 204 (19.60%) sampled loans are delinquent or in default.<sup>18</sup> Finally, out of these 40 distressed loans, 28 had a credit rating of 3 stars, 8 were rated 3.5 stars and 4 loans had a rating of 4 stars.

We investigate the likelihood of auction winners to invest in delinquent or defaulting loans by running the following logistic regression:

$$Default_{i,j} = \gamma_1 \cdot Geographical\ Closeness_{i,j} + \gamma_2 \cdot Age_i + \gamma_3 \cdot X_j + \epsilon_{i,j} \quad (3)$$

where subscript  $i$  identifies an investor who has successfully participated in auction  $j$ .  $Default_{i,j}$  is a dummy equal to one if the borrower of auction  $j$  has subsequently defaulted on its loan, zero otherwise. We also study the case where dummy  $Default_{i,j}$  includes delinquent loans. We control for investor age ( $Age_i$ ) and for the same loan characteristics as in previous regressions (vector  $X_j$ ), namely loan term, loan rating, and loan amount. We include fixed effects per administrative division. Under hypothesis H3, we expect  $\gamma_1$  to be significantly negative.

Results are reported in Table 5. Regarding control variables, a lender is less likely to experience a default as the loan's rating increases. The probability to invest in a defaulting loan increases with the loan amount. As expected, loans with longer terms are more likely to become delinquent or to be in default. In columns (2) and (4) of Table 5, we add a variable that measures the degree of competition in the auction by the ratio of the competition period duration over the total auction duration.<sup>19</sup> We add this covariate to take into account a potential herding behavior that led investors to blindly follow the crowd and fund the borrower because of a high demand coming from other investors. A higher degree of competition during the auction is associated with a lower probability of default, which suggests that the

---

<sup>18</sup>As a means of comparison, the ratio of non-performing bank loans subscribed by French SMBs was 4.05% in 2011 and 3.90% in 2015 according to the OECD. This confirms that the loans originated through crowdlending platforms are riskier than bank loans.

<sup>19</sup>The competition period corresponds to the auction time period when the loan amount can already be cleared with submitted bids and bidders compete in rate.

crowd is informed and that investors are subject to rational herding behavior.<sup>20</sup>

### Table 5 about here

Regarding our main variable of interest,  $\gamma_1$  coefficients are not found to be significantly negative. Investors that funded a borrower located close to their place of residence are statistically as likely as other lenders to have invested in a subsequently defaulting or delinquent loan, which leads us to reject H3. This means that local investors do not own privileged information about the creditworthiness of the borrower. As a consequence, the higher probability of investors to win the auctions of borrowers located close to their place of residence is not information-driven (rejection of the rational hypothesis), and it should be considered as a behavioral bias (behavioral hypothesis). This result also rules out the alternative hypothesis that the geographical-proximity bias could be a proxy for a local network effect based on personal relationships between lenders and managers of the same geographical area. If lenders geographically close to the borrower personally knew its managers, they should have private information about the creditworthiness of the borrower, and they should experience a significantly lower rate of default when investing on local companies. Therefore, rejecting H3 rules out both the rational hypothesis and the network-effect alternative hypothesis. On the contrary, it supports the hypothesis that the geographical-proximity bias is mainly behavioral.

Now that we have shown that the tendency of investors to over-lend to geographically close borrowers is not driven by superior information but is rather behavioral, we investigate whether this bias impacts the efficiency of the price discovery process in the auction system.

---

<sup>20</sup>According to [Chemla and Timm \(2019\)](#), in light of the recent literature, crowdfunding harnesses the wisdom of crowds when investors' valuations are independent from one another. We can consider that this is the case in Unilend auctions, where the characteristics of each investor's bids are not visible from other participants.

## 5. Impact of the geographical-proximity bias on the price discovery process

In this section, we examine how the geographical-proximity bias impacts the efficiency of the price discovery process in the auction system. Given that geographically close lenders are subject to a behavioral bias, their presence in an auction may draw the clearing rate away from its fundamental value. We investigate this issue with a two-stage methodology: we first estimate the deviations of observed loan rates from their fundamental values; we then analyse how those deviations relate to the geographical-proximity bias.

### 5.1. Test design

We start our analysis by comparing observed loan rates with their fundamental values. Since fundamental rates are unobservable, we model them by regressing the observed average borrowing rate and the observed equilibrium rate of each auction on several characteristics of the loan and the borrower deemed as essential in the formation of interest rates. Those rate determinants include the loan amount, the loan maturity, the borrowing firm size measured by total assets in logarithm, credit rating dummies, plus four dummies identifying the purpose of the loan, i.e., the type of project for which the firm borrows money. Dummy *Commercial development* equals one for activity-growth projects, zero otherwise; dummy *Intangible investment* equals one for loans funding R&D or any other type of intangible assets, zero otherwise; dummy *Tangible investment* equals one for loans funding tangible assets, zero otherwise; dummy *Financial operations* equals one for loans financing liabilities repayment or liabilities restructuring, zero otherwise; and the base case in the regression corresponds to firms borrowing money to finance their working capital. In addition, we add year fixed effects to take into account general trends in interest rates across time as well as potential changes in the demand for crowdlending loans.

The estimates of this predictive model are reported in Table 6, column (1) displaying the estimates for average borrowing rates and column (2) displaying the estimates for equilibrium rates. Average borrowing rates and equilibrium rates are positively correlated with the loan amount, positively correlated with the term of the loan, and negatively correlated with the borrower's credit rating. The size of the borrowing firm does not appear to have any impact on the rates. We find that loans funding commercial development obtain a higher borrowing rate and a higher equilibrium rate, on average, than loans funding other types of project. Borrowing money to invest in intangible assets (a new website for instance), tangible assets (e.g., new production machine), or to carry out financial operations such as debt restructuring has no effect on the borrowing rate or the equilibrium rate of the auction. Error terms extracted from this model are then considered as pricing errors, or, in other words, as deviations from fundamental values.

#### Table 6 about here

In a second stage, we test whether the geographical-proximity bias has an impact on the efficiency of the price discovery process (H4) by running the following quantile regression:

$$Loan\ Rate\ Deviation_j = \theta_1.Closeness\ Degree_j + \theta_2.Rating_j + \epsilon_j \quad (4)$$

where *LoanRateDeviation<sub>j</sub>* is the difference between the equilibrium (average) rate resulting from auction *j* and the estimated fundamental value of that equilibrium (average) rate. *Closeness Degree<sub>j</sub>* is a measure of the relative weight of local investors in auction *j*. Under the generic notation *Closeness Degree<sub>j</sub>*, we use three different metrics of the degree of presence of local investors. The first metric, named *Local funding*, is the total amount of funds coming from local investors divided by the loan amount. The second metric, named *Local bid amount*, is the cumulative amount of all bids submitted by local investors divided by the cumulative amount of all bids submitted in the auction. The third metric, named



*Local bid number*, is the number of bids submitted by local investors divided by the total number of bids submitted in the auction. While the first variable only takes winning bids into account, the other two variables include both winning and losing bids. By definition, those three variables are positively correlated with the intensity of the geographical-proximity bias in an auction. If bids submitted by investors who are subject to the geographical-proximity bias deteriorate the efficiency of the price discovery process (H4), we expect  $\theta_1$  to be significantly different from zero. We also control for the loan's credit rating as the deviation from efficiency may depend on risk-aversion.

We estimate the coefficients of Regression (4) by using a quantile model because we expect the relation between the outcome variable and the degree of presence of geographically close investors in the auction to be non-linear. First, the outcome variable of the regression and the variables gauging the degree of presence of geographically close investors in the auction are not normally distributed. Second, we expect the relation between the dependent variable and the presence of local investors to be characterized by threshold effects, in the sense that the impact of the geographical-proximity bias on the rate deviation could become significant only once the weight of geographically close investors in the auction is high enough, and that the impact might be a jump rather than a proportional response.

## 5.2. Results

Whereas the predictive model of loan rates' fundamental values is estimated over the full sample, the second-stage regressions of loan rate deviations (Regressions 4) are only conducted on the subsample of auctions completed before the implementation of Autolend. Results are reported in Table 7. Columns (1), (3) and (5) analyze the deviations of equilibrium rates while columns (2), (4) and (6) analyze the deviations of average borrowing rates, those deviations being defined as the observed value minus the estimated fundamental value. All coefficients of the variables gauging the presence of local investors in the auction are positive and statistically significant, which is in support of H4. The intensity of the

geographical-proximity bias in pre-Autolend auctions thus deteriorates the efficiency of the price discovery process by increasing the borrower's rate of funding.

**Table 7 about here**

### 5.3. Interpretation

The positive relation between the geographical-proximity bias and loans' funding rates is puzzling, especially when we put it in perspective with the finding that local investors have a higher likelihood of winning an auction as shown in Table 4. On the one hand, the positive link between the weight of local investors and funding rates would suggest that local investors bid at higher rates than others. On the other hand, local investors' higher likelihood of winning the auction would suggest that they bid at lower rates. To solve this puzzle, we analyze the bidding strategies of geographically close investors by looking at three characteristics of those strategies: the bid interest rate, the bid timing, and the number of bids submitted in each auction.

We start by analyzing the interest rate of winning bids with the following quantile regression:

$$\begin{aligned}
 \text{Bid Rate Deviation}_{i,j,t} = & \psi_1 \cdot \text{Geographical Closeness}_{i,j} \\
 & + \psi_2 \cdot \text{Bid Timing}_{i,j,t} + \psi_3 \cdot \text{Bid Timing}_{i,j,t}^2 + \psi_4 \cdot \text{Age}_i + \psi_5 \cdot \text{Rating}_j + \epsilon_{i,j,t}
 \end{aligned} \tag{5}$$

where  $\text{Bid Rate Deviation}_{i,j,t}$  is the difference between the interest rate of the winning bid of investor  $i$  submitted at time  $t$  in auction  $j$  and the estimated fundamental value of the average rate of auction  $j$ . We control for the timing of the bid. The later the bid is submitted, the more competitive in rate it has to be to win against competitors. We also add the square of the bid's timing as a covariate because the relationship between the timing of the bid and the interest rate of the bid is not linear. We control for age considering that it proxies for

investing experience. We expect older investors to be less prone to discount bid rates in order to win auctions. We also control for the credit rating of the borrower by including vector *Rating* which comprises the three rating dummies already used in previous regressions. The coefficients of Regression (5) are estimated by using a quantile regression as the outcome variable is not normally distributed.

Results are reported in Table 8. Local investors obtain statistically higher interest rates than other winning investors. The aforementioned results only hold for the group of auctions completed before the introduction of Autolend as the coefficient of dummy *Geographical Closeness* is not statistically significant for the group of auctions completed after the implementation of the bidding algorithm. This finding is consistent with the results reported in Table 7, according to which the presence of geographically close investors increases the cost of funding of borrowers. Regarding the control variables, as expected, late submissions are associated with lower interest rates, and the interest rates of winning bids increase with the age of the investor and default risk.

### **Table 8 about here**

The results reported in Table 8 can still not explain the higher likelihood of local investors' bids to win the auction. This leads us to study their timing and their frequency. For the group of auctions completed before the implementation of Autolend, we compare the average timing of bids submitted by geographically close investors with the average timing of bids submitted by other investors. Results are reported in panel A of Table 9. It appears that there is no difference in timing between the two groups of investors when considering both accepted and rejected bids. However, when considering winning bids only, we find that geographically close investors submit their bids earlier than the rest of the crowd. More precisely, local investors submit their last bid in the auction earlier than others. In panel B of Table 9, we compare the average number of bids submitted per auction for the two groups of investors. Although both groups of investors submit a similar number of bids during

each auction, we find that local investors participate more during the coverage period of the auction and less during the competition period. The fact that local investors participate more during the coverage period than other investors explains why they earn significantly higher interest rates as there is no competition in rate during that period. Moreover, by participating less during the competition period, they reduce the level of competition at a crucial time for the equilibrium rate formation. This explains the increase in the borrowers' cost of funding associated with their presence in the auction.

### **Table 9 about here**

The findings displayed in Tables 8 and 9 mean that local crowdlenders are able to win the auctions of local borrowers with early bids at higher rates than other winners. To investigate this phenomenon, we split the investor population in two groups: (1) those who have already participated in at least two auctions of local borrowers, referred to as experienced local investors, and (2) all the others, i.e., those who have participated in one auction at most of a local borrower. We then run the regressions of Table 7 over each sub-population. We find that the positive relation between auction rates and the weight of local investors is totally driven by the weight of the experienced ones. We therefore conclude that experienced local investors have extracted information about the dynamics of the order flow in previous auctions of local borrowers. They are then able to submit winning bids at higher rates than other winners, at an early stage, in further local borrowers' auctions. This suggests that the familiarity feeling for geographically close businesses strengthens investor attention and learning abilities, and thereby improves investor knowledge about the dynamics of local auctions' order flow.

## 6. Robustness checks

Our identification strategy lies on the comparison between two groups of auctions that did not take place during the same time period. The statistical significance of the geographical-proximity bias in the period before the implementation of Autolend could therefore be driven by time-specific unobserved variables. To address this concern, we run the regressions of Table 3, Table 4, and Table 8 by reducing the sample to auctions completed in 2016, the year when the Autolend algorithm was launched. In that year, 38 auctions were completed before the implementation of Autolend, and 54 after the algorithm's implementation. Notwithstanding the small size of the sub-sample, the results are totally similar to those obtained with the full sample.

Our identification strategy also relies on the fact that the samples before and after the implementation of Autolend are comparable. However, the sample size for the period after Autolend is significantly smaller than the sample size for the period before Autolend in Tables 3 and 8. One could therefore suspect the absence of statistical significance for the coefficients identifying the geographical-proximity bias in the post-Autolend regressions of Tables 3 and 8 to be caused by the smaller size of the post-Autolend sub-sample. We show that it is not the case in two ways. First, as previously mentioned, our findings are robust to reducing the observation period to the 2016 year for which the pre-Autolend sub-sample and the post-Autolend sub-sample are comparable in size. Second, we use the bootstrapping technique to make the pre-Autolend and the post-Autolend sub-samples similar in size. More precisely, we run bootstrapped regressions in the period before the implementation of Autolend, i.e., for column (1) in Tables 3 and 8, using the same sample size as for the regressions in the period after the implementation of Autolend. The coefficient of dummy *Geographical Closeness* remains positive and statistically significant at the 1% level in both tables before the implementation of Autolend. The difference in statistical significance of the coefficient of dummy *Geographical Closeness* between the two periods in Tables 3 and

8 is therefore not explained by differences in sample sizes.

We also test whether the geographical-proximity bias exists at larger distances by considering the next-level administrative division, i.e., the French *régions* instead of the *départements*.<sup>21</sup> On average, the surface area of a *région* is seven times larger than that of a *département*. When running all regressions at the *région*-level, we find that the geographical-proximity bias also exists at that level but with almost no impact on the price discovery process.

## 7. Conclusion

This paper investigates whether individual investors are subject to a geographical-proximity bias by testing whether the fund lenders of a peer-to-business crowdlending platform overlend to firms located close to their place of residence. Using the implementation of a bidding algorithm that submits bids on behalf of investors as an exogenous shock on behavioral biases, we provide evidence that the geographical-proximity bias exists.

Investors located in the same administrative division as borrowers are more likely to participate and win their auctions, and thereby to lend them funds. We then show that creditors located in the same administrative division as borrowers are as likely as creditors located further away to experience a default on their claims. This invalidates the hypothesis that investors would lend more to closely located borrowers because they would own more information about their creditworthiness. On the contrary, it suggests that the preference to lend money to geographically close borrowers is behavioral to such extent that it might hamper the auction's price discovery process. We investigate whether this is the case and we find that the geographical-proximity bias indeed has a significant impact on the quality of equilibrium loan rates. The presence of geographically-biased investors in the auction is associated with an increase in the borrower's cost of funding. This result is likely to be explained by the way such investors time their bids. Geographically close investors submit

---

<sup>21</sup>For a definition of French *régions*, cf. Section 2.2 page 7.

their bids early, and they become passive towards the end of the auction. By not participating at the end of the auction, when the competition between bidders is at its peak, geographically close investors decrease the level of competition in the auction and therefore increase the borrower's cost of funding.

Furthermore, we find that the positive link between auction rates and the weight of local investors is driven by experienced local investors, i.e., local investors who have already participated in at least two auctions of local borrowers. We infer from this finding that investor attention and learning capacities are increased when participating in the auctions of geographically close borrowers. This allows crowdlenders to acquire some specific knowledge about the dynamics of the order flow of local auctions. They are then able to submit winning bids at higher rates in further local auctions, which in turn tends to increase the borrowing rates of those auctions.

From a regulatory perspective, we recommend the use of bidding algorithms, such as Autolend, to bid on behalf of non-professional investors in auctions, as they are free from behavioral biases. Another solution to avoid the adverse effect of the geographical-proximity bias of non-professional investors would be to anonymize borrowers so that they cannot be recognized by investors.

Last, this study highlights that familiarity feelings, among which familiarity stemming from geographical proximity, distort investment decisions with potential harmful effects on price efficiency. We believe that this finding based on crowdlending data has external validity for any financial activity involving fund offering by non-professional investors, and we expect individual investors to overinvest in assets that are familiar to them in many contexts including bond, stock, or real estate investments.

## References

- Abraham, Facundo, and Sergio L. Schmukler, 2017, Addressing the SME Finance Problem, *World Bank Policy Research Working Paper* .
- Adhami, Saman, Gianfranco Gianfrate, and Sofia A. Johan, 2019, Risks and Returns in Crowdlending, *SSRN Electronic Journal* .
- Agrawal, Ajay, Christian Catalini, and Avi Goldfarb, 2011, The Geography of Crowdfunding, *SSRN Electronic Journal* .
- Ahearne, Alan G., William L. Grier, and Francis E. Warnock, 2004, Information Costs and Home Bias: An Analysis of US Holdings of Foreign Equities, *Journal of International Economics* 62, 313–336.
- Belleflamme, Paul, Thomas Lambert, and Armin Schwiabacher, 2013, Individual Crowdfunding Practices, *Venture Capital* 15, 313–333.
- Benartzi, Shlomo, 2001, Excessive Extrapolation and the Allocation of 401 (k) Accounts to Company Stock, *The Journal of Finance* 56, 1747–1764.
- Ceyhan, Simla, Xiaolin Shi, and Jure Leskovec, 2011, Dynamics of Bidding in a P2P Lending Service: Effects of Herding and Predicting Loan Success .
- Chemla, Gilles, and Katrin Tinn, 2019, How Wise Are Crowds on Crowdfunding Platforms?, *SSRN Electronic Journal* .
- Chemla, Gilles, and Katrin Tinn, 2020, Learning Through Crowdfunding, *Management Science* 66, 1783–1801.
- de Roure, Calebe, Lorian Pelizzon, and Anjan V. Thakor, 2019, P2P Lenders versus Banks: Cream Skimming or Bottom Fishing?, *SSRN Electronic Journal* .



- Disdier, Anne-Célia, and Keith Head, 2008, The Puzzling Persistence of the Distance Effect on Bilateral Trade, *The Review of Economics and Statistics* 90, 37–48.
- Dorflleitner, Gregor, Lars Hornuf, and Martina Weber, 2019, Paralyzed by Shock: The Portfolio Formation Behavior of Peer-to-Business Lending Investors, *Max Planck Institute for Innovation & Competition Research Paper* .
- Frieder, Laura, and Avanidhar Subrahmanyam, 2005, Brand Perceptions and the Market for Common Stock, *Journal of Financial and Quantitative Analysis* 40, 57–85.
- Hau, Harald, and Helene Rey, 2008, Home Bias at the Fund Level, *American Economic Review* 98, 333–38.
- Heath, Chip, and Amos Tversky, 1991, Preference and Belief: Ambiguity and Competence in Choice Under Uncertainty, *Journal of Risk and Uncertainty* 4, 5–28.
- Hervé, Fabrice, and Armin Schwienbacher, 2018, Round-Number Bias in Investment: Evidence from Equity Crowdfunding, *Finance* 39, 71–105.
- Hortaçsu, Ali, F. Martínez-Jerez, and Jason Douglas, 2009, The Geography of Trade in Online Transactions: Evidence from eBay and MercadoLibre, *American Economic Journal: Microeconomics* 1, 53–74.
- Huberman, Gur, 2001, Familiarity Breeds Investment, *The Review of Financial Studies* 14, 659–680.
- Iyer, Rajkamal, Asim Ijaz Khwaja, Erzo FP Luttmer, and Kelly Shue, 2016, Screening Peers Softly: Inferring the Quality of Small Borrowers, *Management Science* 62, 1554–1577.
- Lewis, Karen K., 1999, Trying to Explain Home Bias in Equities and Consumption, *Journal of Economic Literature* 37, 571–608.
- Lin, Mingfeng, and Siva Viswanathan, 2016, Home Bias in Online Investments: An Empirical Study of an Online Crowdfunding Market, *Management Science* 62, 1393–1414.

- Massa, Massimo, and Andrei Simonov, 2006, Hedging, Familiarity and Portfolio Choice, *The Review of Financial Studies* 19, 633–685.
- Mohammadi, Ali, and Kouros Shafi, 2017, How Wise Are Crowds? A Comparative Study of Crowds and Institutions in Peer-to-Business Online Lending Markets, *Swedish House of Finance Research Paper* .
- Moritz, Alexandra, and Joern H. Block, 2016, Crowdfunding: A Literature Review and Research Directions, in *Crowdfunding in Europe*, 25–53 (Springer).
- Seasholes, Mark S., and Ning Zhu, 2010, Individual Investors and Local Bias, *The Journal of Finance* 65, 1987–2010.
- Tang, Huan, 2019, Peer-to-Peer Lenders Versus Banks: Substitutes or Complements?, *The Review of Financial Studies* 32, 1900–1938.
- Vallée, Boris, and Yao Zeng, 2019, Marketplace Lending: A New Banking Paradigm?, *The Review of Financial Studies* 32, 1939–1982.
- Wolf, Holger C., 2000, Intranational Home Bias in Trade, *Review of Economics and Statistics* 82, 555–563.
- Zhang, Juanjuan, and Peng Liu, 2012, Rational Herding in Microloan Markets, *Management Science* 58, 892–912.

Table 1: Summary statistics

This table reports the mean, the median, the standard deviation, the minimum value, and the maximum value of several characteristics of the 290 loans of our sample, their auctions, and their borrowers. Among those characteristics, accounting items are those of the fiscal year preceding the auction. Leverage designates the ratio of total debts over total assets. The loan rating is the credit notch assigned to the loan by Unilend. For auctioned loans, those notches range from 3 to 5, with a 0.5 increment, 5 being the best creditworthiness. The equilibrium rate is the clearing rate of the auction. The average borrowing rate is the weighted average rate of the bids served in the auction, i.e., the rate paid by the borrower on its loan. The auction duration is that announced on the platform before the start of the auction. The competition degree is the duration of the competition period - that is the phase of the auction during which the cumulative amount of bids already clears the loan and investors compete in rate - divided by the total duration of the auction.

	N	Mean	Median	St. Dev.	Min	Max
Loan amount	290	74,367.240	55,000.000	55,103.330	10,000.000	400,000.000
Loan term (in months)		41.979	36.000	13.458	6.000	60.000
Total assets (in millions)		1.722	0.591	4.205	0.000	46.694
Sales (in millions)		1.880	0.754	5.531	0.000	83.532
Leverage		0.550	0.563	0.183	0.008	0.906
Loan rating		3.259	3.000	0.349	3.000	4.000
Number of bids		5,070.224	1,531.500	8,206.224	217.000	39,690.000
Number of distinct bidders		858.266	754.500	519.120	159.000	3,240.000
Investor age (in years)		43.599	43.559	0.871	40.316	46.105
Average borrowing rate (in %)		7.890	8.211	1.573	4.000	9.903
Equilibrium rate (in %)		8.137	8.450	1.642	4.000	10.000
Auction duration (in hours)		247.000	195.000	189.000	0.070	1,083.517
Competition degree		0.527	0.570	0.342	0.000	1.000

Table 2: Comparing loans before and after the implementation of Autolend

This table compares the means of several characteristics of the 290 loans of our sample, their auctions, and their borrowers, before and after the implementation of the Autolend algorithm. Among those characteristics, accounting items are those of the fiscal year preceding the auction. Leverage designates the ratio of total debts over total assets. The loan rating is the credit notch assigned to the loan by Unilend. For auctioned loans, those notches range from 3 to 5, with a 0.5 increment, 5 being the best creditworthiness. The equilibrium rate is the clearing rate of the auction. The average borrowing rate is the weighted average rate of the bids served in the auction, i.e., the rate paid by the borrower on its loan. The auction duration is that announced on the platform before the start of the auction. The competition degree is the duration of the competition period - that is the phase of the auction during which the cumulative amount of bids already clears the loan and investors compete in rate - divided by the total duration of the auction. We test the difference in means between the two subsamples with a *t*-test and the difference in medians with a Wilcoxon test, and we report the associated *p*-value.

	Pre-Autolend period	Post-Autolend period	<i>t</i> -test	Wilcoxon test
N	236.000	54.000		
Loan amount	76,854.450	63,497.220	0.089	0.023
Loan term (in months)	43.182	36.722	0.004	0.003
Total assets (in millions)	1.792	1.416	0.387	0.613
Sales (in millions)	1.978	1.451	0.268	0.743
Leverage	0.552	0.541	0.666	0.676
Loan rating	3.250	3.296	0.404	0.413
Number of bids	1,460.801	20,844.741	0.000	0.000
Number of distinct bidders	718.873	1,467.463	0.000	0.000
Investor age	43.643	43.408	0.040	0.081
Average borrowing rate (in %)	8.226	6.425	0.000	0.000
Equilibrium rate (in %)	8.499	6.552	0.000	0.000
Auction duration (in hours)	285.947	78.267	0.000	0.000
Competition degree	0.445	0.884	0.000	0.000

Table 3: Investor likelihood of participating in an auction

This table reports the results of logistic regressions in which the dependent variable is a dummy equal to 1 if, for a given investor and a given auction, the investor has participated in the auction, 0 otherwise. Column 'Pre-Autolend period' and column 'Post-Autolend period' present the results for the pre-Autolend and the post-Autolend subsamples respectively. *Geographical closeness* equals 1 for local investors, 0 for others. Rating  $n$  equals 1 if the loan's rating is  $n$ , 0 otherwise. *Investor age*, *Auction duration*, *Loan amount* and *Loan term* are expressed in years, hours, euros, and months respectively, and taken in logarithm. Standard errors are reported in parentheses. \*\*\*, \*\* and \* denote the statistical significance of the parameter estimates at the 1%, 5%, and 10% level respectively.

	Pre-Autolend period	Post-Autolend period
	(1)	(2)
Geographical closeness	0.064*** (0.014)	0.033 (0.023)
Investor age	0.551*** (0.009)	0.475*** (0.013)
Auction duration	0.554*** (0.011)	0.007** (0.004)
Loan amount	-0.053*** (0.011)	0.228*** (0.013)
Rating 3.5	0.098*** (0.007)	0.361*** (0.018)
Rating 4	0.162*** (0.013)	0.444*** (0.022)
Rating 4.5	-0.018 (0.030)	
Loan term	-0.109*** (0.008)	-0.071*** (0.017)
Intercept	-5.841*** (0.093)	-5.781*** (0.131)
Fixed effects	Administrative division	Administrative division
Number of observations	951,364	465,398

Table 4: Bid likelihood of winning an auction

This table reports the results of logistic regressions in which the dependent variable is a dummy equal to 1 if the bid of a given investor in a given auction submitted at a given time is served in the auction. Column 'Pre-Autolend period' and column 'Post-Autolend period' present the results for the pre-Autolend and the post-Autolend subsamples respectively. *Geographical closeness* equals 1 for local investors, 0 for others. *Rating n* equals 1 if the loan's rating is *n*, 0 otherwise. *Interest rate* is the interest rate of the bid submitted. It ranges from 4% to 10% in the coverage period, i.e., as long as the cumulated bid amount does not clear the loan. *Bid timing* measures the earliness of the bid in the auction by a ratio ranging from 0 to 1, 0 corresponding to the timing of the first submission and 1 corresponding to the last one. *Investor age*, *Loan amount*, and *Loan term*, measured in years, euros, and months respectively, are taken in logarithm. Standard errors are reported in parentheses. \*\*\*, \*\* and \* denote the statistical significance of the parameter estimates at the 1%, 5%, and 10% level respectively.

	Pre-Autolend period	Post-Autolend period
	(1)	(2)
Geographical closeness	0.224*** (0.022)	0.009 (0.037)
Interest rate	-0.680*** (0.005)	-1.556*** (0.009)
Bid timing	1.226*** (0.013)	4.242*** (0.029)
Investor age	-0.220*** (0.014)	0.200*** (0.019)
Rating 3.5	-0.709*** (0.010)	-0.649*** (0.018)
Rating 4	-1.024*** (0.016)	-1.412*** (0.026)
Rating 4.5	-1.821*** (0.049)	
Loan term	1.976*** (0.013)	1.476*** (0.017)
Loan amount	1.924*** (0.010)	1.489*** (0.016)
Intercept	-23.212*** (0.120)	-16.458*** (0.161)
Number of observations	338,492	1,110,209

Table 5: Likelihood to invest in a defaulting or delinquent loan

This table reports the results of logistic regressions in which the dependent variable is set to 1 for a given auction winner and a given loan if the loan is in default (columns 1 and 2) or the loan is either in default or considered delinquent (columns 3 and 4). A loan is considered delinquent if the repayment of its principal is late by more than 60 days but less than 9 months. A loan is considered in default if: (1) the repayment of its principal is late by more than 9 months; (2) the loan is under a recovery procedure; (3) the borrower went bankrupt; or (4) the borrowing firm is going into receivership. We only consider loans completed before the implementation of Autoland and for which we have information about their performance. *Geographical closeness* equals 1 for local investors, 0 for others. Rating  $n$  equals 1 if the loan's rating is  $n$ , 0 otherwise. *Competition degree* is the duration of the competition period - that is the phase of the auction during which the cumulative amount of bids already clears the loan and investors compete in rate - divided by the total duration of the auction. *Investor age*, *Loan amount*, and *Loan term*, measured in years, euros, and months respectively, are taken in logarithm. Standard errors are reported in parentheses. \*\*\*, \*\* and \* denote the statistical significance of the parameter estimates at the 1%, 5%, and 10% level respectively.

	Default		Default or delinquent	
	(1)	(2)	(3)	(4)
Geographical closeness	0.008 (0.055)	0.007 (0.056)	-0.009 (0.054)	-0.010 (0.054)
Rating 3-5	-0.386*** (0.037)	0.049 (0.041)	-0.914*** (0.036)	-0.608*** (0.038)
Rating 4	0.390*** (0.051)	0.718*** (0.054)	0.267*** (0.051)	0.524*** (0.053)
Rating 4.5	-21.995 (1,179.405)	-21.457 (1,179.398)	-22.175 (1,179.404)	-21.601 (1,179.400)
Investor age	0.008 (0.043)	0.034 (0.043)	0.007 (0.042)	0.028 (0.042)
Loan amount	0.164*** (0.028)	-0.727*** (0.042)	0.397*** (0.026)	-0.304*** (0.037)
Loan term	1.331*** (0.051)	0.714*** (0.056)	1.789*** (0.052)	1.250*** (0.056)
Competition degree		-2.094*** (0.076)		-1.811*** (0.072)
Intercept	-8.119*** (0.425)	4.515*** (0.615)	-12.327*** (0.402)	-2.133*** (0.560)
Fixed effects	Administrative division	Administrative division	Administrative division	Administrative division
Number of observations	80,778	80,778	80,778	80,778

Table 6: Modelling fundamental interest rates

This table reports the estimates of OLS regressions run to model the fundamental values of loan rates. The dependent variable is alternatively the observed average borrowing rate (Column 1) or the observed auction equilibrium rate (Column 2). Rating  $n$  equals 1 if the loan's rating is  $n$ , 0 otherwise. *Commercial development*, *Intangible investment*, *Tangible investment*, and *Financial operations* are dummies identifying the type of project for which the firm borrows money. *Loan amount*, *Loan term*, and *Total assets* are expressed in euros, months, and million euros respectively, and are taken in logarithm. Robust standard errors are reported in parentheses. \*\*\*, \*\* and \* denote significance of the parameter estimates at the 1%, 5% and 10% level, respectively.

	Average rate	Equilibrium rate
	(1)	(2)
Loan amount	0.927*** (0.096)	1.032*** (0.100)
Loan term	1.360*** (0.149)	1.394*** (0.154)
Rating 3.5	-0.584*** (0.111)	-0.606*** (0.116)
Rating 4	-0.749*** (0.183)	-0.753*** (0.196)
Rating 4.5	-1.690*** (0.574)	-1.494** (0.600)
Total assets	-0.009 (0.025)	-0.014 (0.027)
Commercial development	0.297** (0.151)	0.308* (0.158)
Intangible investment	0.284 (0.201)	0.319 (0.210)
Tangible investment	0.043 (0.148)	0.059 (0.155)
Financial operations	0.676 (0.422)	0.674 (0.441)
Intercept	-6.611*** (1.157)	-6.985*** (1.226)
Fixed effects	Year	Year
Number of observations	290	290
Adjusted R <sup>2</sup>	0.745	0.744



**Table 7: Impact of the geographical-proximity bias on the rate discovery process**

This table reports the estimates of quantile regressions of loan rate deviations from their fundamental values on measures of the geographical-proximity bias intensity. Our dependent variable is the difference between the observed rate and the rate predicted by the model of Table 6. For each loan, the interest rate considered is alternatively the average borrowing rate (Columns 2, 4, and 6) and the auction equilibrium rate (Columns 1, 3, and 5). Loans considered are those completed before the implementation of Autolend. *Local funding* is the share of the loan amount funded by local investors. *Local bid amount* is the share of the auction's total cumulative bid amount submitted by local investors. *Local bid number* is the number of bids submitted by local investors divided by the total number of bids submitted in the auction. Rating  $n$  equals 1 if the loan's rating is  $n$ , 0 otherwise. Standard errors are reported in parentheses. \*\*\*, \*\*, \* and \* denote the statistical significance of the parameter estimates at the 1%, 5%, and 10% level respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	Equilibrium rate	Average rate	Equilibrium rate	Average rate	Equilibrium rate	Average rate
Local funding	1.909*** (0.685)	2.174* (1.260)				
Local bid amount			1.871*** (0.454)	2.039*** (0.555)		
Local bid number					2.538** (0.987)	2.756*** (1.034)
Rating 3.5	0.043 (0.101)	0.038 (0.129)	0.049 (0.107)	0.028 (0.113)	0.015 (0.116)	-0.006 (0.116)
Rating 4	0.128 (0.357)	0.232 (0.248)	0.180 (0.350)	0.234 (0.215)	0.195 (0.354)	0.198 (0.219)
Rating 4.5	0.051 (0.524)	-0.040 (0.338)	0.112 (0.592)	0.031 (0.274)	0.262 (0.795)	0.185 (0.480)
Intercept	0.046 (0.077)	-0.039 (0.087)	0.039 (0.071)	-0.039 (0.074)	0.053 (0.079)	-0.013 (0.082)
Number of observations	236	236	236	236	236	236

Table 8: Impact of geographical closeness on accepted bids' rates

This table reports the estimates of quantile regressions in which the dependent variable is the difference between the interest rate of each winning bid in a given auction and the fundamental value of the average borrowing rate as predicted by the model of Table 6 for that auction. Column (1) and Column (2) present the results for the pre-Autolend and the post-Autolend periods respectively. *Geographical closeness* equals 1 for local investors, 0 for others. *Bid timing* measures the earliness of the bid in the auction by a ratio ranging from 0 to 1, 0 corresponding to the timing of the first submission and 1 corresponding to the last one. *Investor age* is taken in logarithm. Rating  $n$  equals 1 if the loan's rating is  $n$ , 0 otherwise. Standard errors are reported in parentheses. \*\*\*, \*\* and \* denote the statistical significance of the parameter estimates at the 1%, 5%, and 10% level respectively.

	Pre-Autolend period	Post-Autolend period
	(1)	(2)
Geographical closeness	0.152*** (0.019)	0.001 (0.009)
Bid timing	-0.142*** (0.040)	-1.379*** (0.054)
Bid timing <sup>2</sup>	0.113*** (0.037)	1.783*** (0.042)
Investor age	0.058*** (0.010)	-0.0001 (0.005)
Rating 3.5	0.040*** (0.007)	0.155*** (0.004)
Rating 4	0.188*** (0.009)	-0.238*** (0.003)
Rating 4.5	0.071*** (0.007)	
Intercept	-0.141*** (0.038)	-0.560*** (0.023)
Number of observations	111,602	42,311

Table 9: Comparing the bidding strategies of local and non-local investors

This table compares the bidding strategies of local and non-local investors in auctions completed before the implementation of Autolend. Panel A reports, for local and non-local investors, the average timing of their bids, their accepted bids, their first bid, and their last bid submitted in an auction. The timing of a bid is the earliness of that bid in the auction, measured by a score between 0 and 1, 0 corresponding to the timing of the first submission and 1 corresponding to the last one. Panel B reports, for local and non-local investors, the average number of bids submitted per auction as well as the average of three ratios, the number of bids submitted in the coverage period divided by the total number of bids, the number of bids submitted in the competition period divided by the total number of bids, and the number of bids submitted in the coverage period divided by the number of bids submitted in the competition period. The coverage period is the phase of the auction during which cumulated bid amounts do not clear the loan amount and each bidder is a winner. The competition period starts as soon as the cumulated bid amount clears the loan. In that phase, only bidders with bid rates lower than or equal to the clearing rate are winners. The differences in means between local and non-local investors are tested with *t*-tests which *p*-values are reported in the last column.

	Local investors	Non-local investors	<i>t</i> -test
Panel A : Average bid timing			
All bids	0.6012	0.5996	0.595
Accepted bids	0.6891	0.7172	0.000
First bid in the auction	0.4441	0.4450	0.837
Last bid in the auction	0.6655	0.6804	0.001
Panel B : Average number of bids			
Total # of bids	1.9890	2.0338	0.135
# bids coverage period / Total # of bids	0.4802	0.4372	0.000
# bids competition period / Total # of bids	0.5198	0.5628	0.000
# bids coverage period / # bids competition period	0.2990	0.2840	0.078