Do banks and microfinance institutions compete? Microevidence from Madagascar*

Pierrick Baraton — *Agence Française de Développement*

Florian Léon — *University of Luxembourg, CREA*

April 18, 2019

Abstract

This paper examines whether the loan strategy of a microfinance institution is shaped by the entry of a bank. Specifically, we investigate whether the distance between a borrower of a microfinance institution and the closest bank influences loan conditions provided by the microfinance institution. We use an original panel dataset of 32,374 loans granted to 14,834 borrowers provided by one of the largest microfinance institutions in Madagascar between 2008 and 2014. We find that the closer a bank is located to a given MFI borrower, the larger the loan obtained and the less collateral required. We also find that the effect is stronger for clients that could be more easily caught by banks (i.e., large firms and clients without a previous relationship with the MFI).

Keywords: Microfinance; Banks; Competition; Loan conditions

JEL classification: G21; O16

*We would like to thank Mialy Ranaivoson and Nathanielle Razaniajatovo for their substantial efforts in localizing commercial bank branches, and Olivier Santoni and Joanne Wirtz for their helpful support in the use of the mapping software. We are also grateful to the Central Bank of Madagascar and our partner MFI for the data they provided. We thank two anonymous referees, Mathilde Bauwin, Laurent Bergé, Luisito Bertinelli, Arnaud Bourgain, Mélanie Chaudey, Lisa Chauvet, Timothée Demont, Nènè Oumou Diallo, Sylvain Marsat, Julia Meyer, Andrea Presbitero, Grégoire Rota-Gaziosi, Alabert Winkler and conference participants at the IDE conference (Clermont-Fd, 2016), the DIAL Conference (Paris, 2017), the MBF conference (Nanterre, 2017) and the 8th international workshop on microfinance (Oslo, 2018) for their helpful comments. Any errors are our own. The views expressed herein are those of the authors and should not be attributed to the AFD. First version: December 2015. Authors emails: pierrick.baraton@gmail.com; florian.leon@uni.lu (corresponding author). Version accepted in *Economic Development and Cultural Change*

1 Introduction

Limited access to formal credit is a major barrier to growth for developing economies, especially for small firms (Beck and Demirguc-Kunt, 2006) and microenterprises (De Mel et al., 2008; McKenzie and Woodruff, 2008). Since the 1970s, microfinance has emerged as a powerful tool to reach borrowers who are excluded from the formal financial system (Armendáriz and Morduch, 2010). Microfinance can be viewed as a response to market failures in capital markets, filling the gap between moneylenders who charge usurious interest rates and commercial banks[1] who are unwilling to provide financing to people in poverty. This view holds that microfinance institutions (MFIs) and commercial banks operate in two segmented markets. MFIs target low-income people and entrepreneurs excluded from bank financing due to a lack of collateral or insufficiently sized financing needs.

However, this commonly held view of dually segmented financial markets has recently been challenged by new strategies developed by both MFIs and commercial banks. On the one hand, there is a process of commercialization of the microfinance industry, implying the entry of for-profit MFIs and an increase of competition (Armendáriz and Szafarz, 2011; De Quidt et al., 2018). Several microfinance institutions have changed their legal status from NGO to shareholder-owned financial entity, and in extreme cases, some MFIs have evolved into commercial banks (e.g., Prodem in Bolivia, Bandhan in India and Microcred in Madagascar). A number of MFIs, without changing their legal status, have begun to develop their range of services to match the growing financial needs of small businesses by offering larger loans with longer maturities. On the other hand, a less discussed phenomenon is that some commercial banks have begun to target smaller firms by developing special products or acquiring microfinance institutions. This ”downscaling” process began in Latin America in the 1990s and has since experienced significant growth in other areas of the world (Ferrari and Jaffrin, 2006).

Our aim is to examine whether banks and MFIs continue to operate in two segmented markets or have begun to compete for similar clients. If both financial intermediaries

[1]For the reminder of this article, the terms ”commercial bank” and ”bank” are used interchangeably to refer to formal lenders (i.e., registered financial entities) which have not historically offered financial products or used lending techniques designed specifically to target poor populations.

focus on two different types of borrowers, we could expect that changes in the banking market have no impact on MFIs' operations, and vice versa. However, if both markets overlap, changes in either industry would likely affect the other. Consequently, policy-driven shocks (such as changes in regulation) or economic-driven shocks (such as the entry of new actors) in one industry could have unexpected consequences for the other.

In this paper, we test whether MFI loan conditions are affected by the entry of a new bank in the vicinity of the MFI borrower. We implement this using a new dataset from one of the largest MFI in Madagascar. We expect that MFIs and banks compete if MFI credit officers are sensitive to the entry of a new bank in their loan decision. Madagascar is a perfect testing ground for our question because some banks have recently adopted "downscaling" strategies and because our partner MFI initiated its upscaling strategy a few years ago. Three commercial banks in Madagascar, including the largest bank, have developed specific products for microfirms and small firms. Meanwhile, our partner MFI provides individual loans with an upper limit of \$60,000 (200 million ariary).[2] However, in reality, it continues to offer mainly microloans below \$1,000. Finally, while some countries have developed credit bureau or credit registry including all loans provided by banks and MFIs (for example, in Rwanda and Bolivia), in Madagascar, the credit registry for microfinance institutions was consolidated with the registry for banks only in 2017. Therefore, there is a scope for exploiting private information and offering better credit conditions to better clients if an outside lender enters the market. Investigating whether our partner's business strategy has been shifted by bank presence gives us initial insight into how MFIs react to the development of the banking sector in a low-income country.

We use a large dataset containing information on 32,374 loans (14,834 borrowers) that were distributed from 2008 to 2014. Loans are always granted to entrepreneurs (individuals), for a productive purpose, never for individual consumption. We argue that MFIs and banks compete if loan conditions offered by our partner MFI are affected by bank proximity to MFI clients. We consider two measures of loan conditions, namely, loan amount and collateral requirements.[3] Loan amount refers to quantity, while collateral

[2] For all figures in ariary and their value in dollar, we employ the following parity: \$1=3,150 ariary, insofar as parity varies between 3,000 and 3,300 ariary per dollar.

[3] Alternative terms of contracts such as maturity and interest rates could theoretically be considered. However, loan maturity and interest rate are not at the discretion of credit officers. As explained in the following, maturity has a low variability insofar as 90% of loans have a maturity within one month. In

requirement is a form of price in a fixed price environment as ours (Fisman et al., 2017). In line with recent literature on the role of distance in lending (Petersen and Rajan, 2002; Degryse and Ongena, 2005), we assume that the distance between an MFI client and the closest commercial bank branch is a good measure of competitive pressure induced by the bank's presence. Indeed, the probability of an MFI client being wooed by a commercial bank increases as the distance between the client and commercial bank decreases due to transportation and informational costs.

We study whether credit conditions offered by our partner's MFI to its clients are shaped by the presence of a bank in the vicinity. A major issue of identification occurs because borrowers and banks do not randomly locate. Our main identification strategy is based on the inclusion of borrower fixed effects. In doing so, we investigate whether the credit condition dynamic is shaped by a change in distance between the borrower and the closest bank rather than comparing credit conditions obtained by different borrowers at varying distance. This approach allows us to control for all time-invariant unobserved characteristics affecting location, firms' performance and credit terms. We find that bank proximity improves loan conditions offered by our partner MFI. Firms in the vicinity of a bank can secure larger loans (higher quantity) with less collateral (at lower price). The economic impact of distance is far from negligible. For instance, a one standard deviation decrease in distance raises the loan amount by \$110 (approximately 10% of its mean value) and reduces the collateral ratio by more than 15 percentage points (approximately 5 percent of its mean).

To improve identification, we propose two additional tests. Indeed, our baseline framework does not allow us to control for time-variant unobserved factors that could affect bank branch location, firms' performance and credit conditions. We run two indirect tests to address this issue. First, we employ the night light intensity at the district level (lowest administrative division) to control for local economic growth. Night light intensity is correlated with the level of activity and, contrary to surveys, is available for all districts in different years (except in 2014). Second, we include municipality-period dummies that capture all time-variant unobservable factors occurring at the municipality

addition, nominal interest rates were set at 18% before 2011 and 21% after 2011 for (almost) all loans. The credit officers have no discretionary power in the setting of interest rate.

level (the second smallest administrative division in Madagascar). Both tests are in line with our main results. We also consider alternative robustness checks considering remaining econometric issues such as heterogeneity between treated and untreated borrowers, omitted variables, attrition or sample selection (we obtain location of borrowers for half of clients only).

Finally, we show that the effect of distance is stronger for large firms, consistent with the idea that MFIs and banks compete only for the transparent borrowers. However, old firms do not seem to benefit more from competition. In addition, we document that firms with a previous lending relationship with our partner MFIs tend to suffer from competition, contrary to new borrowers. This result indicates that our partner MFI tries to use credit terms mainly to retain new clients, at the detriment of captured ones. In other words, the effect of competition seems stronger for clients that could be more easily caught by banks (i.e., large firms and clients without a previous relationship with the MFI).

This paper is at the crossroads of two strands of papers. Our article is directly linked to the literature on the consequences of competition in microfinance, especially on loan conditions. Systematic evidence shows an intensification of competition in MFIs due to the entry of new actors, often attracted by the success of incumbents and seeking opportunity of profit (Assefa et al., 2013; Baquero et al., 2018; De Quidt et al., 2018). These trends have raised some concerns about the risk of overindebtedness and the viability of traditional nonprofit MFIs (McIntosh and Wydick, 2005; McIntosh et al., 2005) and drift away in the mission of microfinance from poverty alleviation to profit maximization (Cull et al., 2007; Mersland and Strøm, 2010). In this paper, we dedicated special attention to another topic by investigating whether an intensification of competition improves or deteriorates loan conditions. The implication of competition on loan condition is theoretically ambiguous, as documented in the banking literature (Léon, 2015). The *market power view*, builds on standard economic theory, argues that competitive pressure will improve loan conditions. The alternative view notes that, in the presence of information asymmetries, competition can actually be harmful. The quality of screening and lenders' incentives to invest in information acquisition technologies are higher in less competitive markets. Some papers have investigated this issue in the case of microfinance. De Quidt et al. (2018) investigate how recent change in market structure (i.e., increased competition and a shift towards for-profit MFIs) explains the shift in lending patters from joint liability contract to individual loans. This model is in line with the observation made by Navajas et al. (2003) on the change in lending condition of the major incumbent in Bolivia after the entry of an important competitor. Closest paper to ours is Baquero et al. (2018) that investigate how competition between MFIs shapes the credit terms. Specifically, they find that interest rates charged by nonprofit MFIs are insensitive to concentration, while profit-oriented MFIs charge lower interest rates in less concentrated markets. They also document that nonprofit MFIs are sensitive to the presence of commercially oriented MFIs. The loan rates of nonprofit MFIs increase when there is a higher proportion of profit-oriented MFIs in the market. The explanation advanced by the authors states that nonprofit MFIs are forced to develop niche markets to stay in the business. We complement this literature in two ways. First, we focus on nonpricing elements of loan contracts. Indeed, as explained below, credit officers of our partner MFI have some discretionary power on amount and collateral requirements but not on rate and maturity. Second, we investigate the competitive pressure induced by the entry of a bank and not by the impact of competition among MFIs. As Baquero et al. (2018), we show that there is a spillover between the different segments of the market. While they focus on the distinction between commercially-oriented MFIs and nonprofit MFIs, we regard the interaction between on (for-profit) MFIs and banks.

We also complement the literature on the interactions between banks and MFIs. This literature groups together macroeconomic papers and microeconomic evidence. Preliminary works have exploited cross-country analysis to scrutinize whether and how banking sector development affects MFIs' financial (Ahlin et al., 2011; Hermes et al., 2011) and social performances (Cull et al., 2014; Vanroose and D'Espallier, 2013) with mixed results.[4] Ahlin et al. (2011) and Cull et al. (2014) conducted similar work, wherein they

[4]For instance, Cull et al. (2014) document that the development of commercial banks gives MFIs, especially commercially oriented ones, incentives to explore new market niches (e.g., smaller loans, lending targeted to women). Vanroose and D'Espallier (2013) provide more conflicting conclusions. While MFIs offer small loans in countries where the formal banking sector is more developed, MFIs reach less clients in these countries.

investigate the impact of banking development on interest rates charged by MFIs. Both studies show that MFIs offer lower interest rates in more financially developed countries. Cull et al. (2014) also find that the impact of banking development is stronger for commercially oriented MFIs. We extend this macroeconomic literature in two ways. First, our microeconomic approach allows us to shed light on the mechanisms through which banking development affects MFI business operations. The impact of banks on MFIs is often explained by the competitive pressure induced by banks on MFIs (Ahlin et al., 2011; Cull et al., 2014). According to this view, it is because banks increase competition that MFIs offer lower interest rates. However, a challenging explanation ("complementary effect") is that MFIs use financial services in their day-to-day operations. An expansion of the banking sector therefore increases access to banking services for MFIs at lower costs (deposits services, refinancing, and so on). Consequently, MFIs can offer the same services at lower costs and/or offer new financial services and/or grant loans to new customers.[5] Our work offers a way to disentangle both explanations. The complementary effect occurs at the MFI level (headquarters for more important financial services) or at the branch level. We therefore assume that all credit officers in the same office benefit from the same financial services. However, and it is our main argument here, competitive pressure occurs at the borrower level. Some borrowers are more able to be caught by banks because they are closer to bank office and/or because they are more bankable (larger, older). In this work, we control for complementary effect by adding credit officer dummies and exploiting variation within borrower history. A second contribution is that we analyze aspects of loan contracts other than interest rates. The existing literature often focuses on the *extensive margin* effect (number and characteristics of borrowers)[6]

[5]For instance, consider the main result reported by Ahlin et al. (2011) that interest rates are reduced in more financially developed countries. This reduction is mainly explained by better loan repayment and lower operational costs. They explain this by the competitive pressure induced by banks. But, the complementary effect can also explain this result. If banks open new branches in remote areas, this facilitates the business of MFIs operating in these areas. For instance, managers are not forced to come in the main city to deposit funds collected that reduce operational costs. In addition, this may induce a pressure on borrowers to repay if they expect a loan from the bank in the future (and know that MFI's officers and bank's officers communicate). It should be noted that even observing that commercial-oriented MFIs benefit more than others from banking development (Cull et al., 2014) is not sufficient to prove the existence of competition between banks and MFIs because commercially oriented MFIs rely more on banking services than other MFIs.

[6]Ideally, we would complement our study by an analysis of extensive margin. Unfortunately, a consistent investigation would require that we have access to a survey of borrowers and nonborrowers as in Brown et al. (2016). These data are not available in Madagascar rendering identification challenging. In

and neglect the impact of bank development on *intensive margin*. When the *intensive margin* is considered, as in Ahlin et al. (2011) and Cull et al. (2014), the analysis focuses on interest rates only (due to the lack of data). Our study reveals that other loan contract terms (namely, loan amount and collateral requirements here) play a role in discriminating between borrowers.

As explained above, another and recent trend in the literature of the interaction between MFIs and banks relies on rich microeconomic datasets for specific countries. Using a rich survey, Brown et al. (2016) document that the creation of an MFI branch increases the percentage of households with a bank account in South-Eastern Europe. Recently, Agarwal et al. (2018) exploit a rich dataset reporting all loans granted to households in Rwanda to investigate whether the expansion of credit cooperative networks spurs credit access not only by credit cooperatives but also by banks. They document that approximately 10% of new borrowers switch from credit cooperatives to banks. We extend this second strand of works in two directions. First, while they focus on the impact of MFI expansion on banks, we do the opposite. We investigate whether MFIs react to the entry of a bank. Second, contrary to these papers, our analysis is based on entrepreneurs/microfirms and not on households.

The remainder of the paper is organized as follows. Section 2 and Section 3 expose the context of our study and the conceptual framework, respectively. Section 4 discusses the econometric methodology. Section 5 presents the data and variables. Section 6 displays the main econometric results and Section 7 presents robustness checks. Section 8 discusses the implications of our findings and concludes.

2 Context

To assess the relevance of our investigation, it is important to have basic information on the credit market in Madagascar and on our partner MFI. The country has approximately 25 million inhabitants. Recent history has been marked by political turmoil, inducing irregular and limited growth. Therefore, Madagascar continues to belong to the least

a companion investigation, we studied whether bank distance affects the characteristics of clients (age and size) at the community level. However, our results are sensitive to specification and do not allow us to draw definitive conclusions.

developed country group according to the World Bank classification, with a GDP per capita below $500.

The Malgasy financial system is not well developed, and credit access by firms and households is challenging. The ratio of credit to private sector is 13 percent of GDP in 2017. The financial sector is dominated by banks, despite many microfinance institutions operating in the country. Banks account for more than 90 percent of assets, credit and deposits. There were twelve banks at the end of 2014.[7] Despite its limited development, bank penetration has expanded over time, as indicated in Table A1 in the Appendix. According to the Central Bank data, 94 bank branches opened between 2008 and 2014, representing 40% of bank branches in 2014 (227 bank branches). While the majority of branches are located in the capital city (40% of bank branches), all provinces have witnessed an increase in the number of branches over time.

While the majority of banks are commercial banks focusing mainly on medium and large firms and rich households, three banks deserve special attention for our analysis: *Bank of Africa*, *AccesBanque* and *Microcred*.

Bank of Africa (BOA), which belongs to the eponym pan-African banking group, is the largest commercial bank operating in Madagascar. BOA has developed the most widespread network over the country, with 79 branches in 2014 (92 in 2018). Its core business consists of offering financial services to formal firms and households. Contrary to other commercial banks, BOA has a product dedicated to micro and small firms, even informal ones ("mesofinance credit"). It grants loans up to 50 million ariary ($\simeq$ $15,000) and lasting up to five years. This product is quite similar to the three year 200 million ariary ($\simeq$ $60,000) loans from the microfinance institution we study, except for the interest rate and compliance and accounting requirements asked by the BOA.

The two other banks (*AccesBanque* and *Microcred*) differ in their nature to other commercial banks (including BOA). *AccesBanque* is a full bank, but a part of its activity focuses on small firms and households. *Microcred* was a former MFI that became a bank (and it is no longer an MFI for authorities). In this way, their lending technologies are close to those developed by MFIs. However, both have to respect bank regulation.

[7]In 2018, there were only 11 banks because *BICM* is in liquidation. It should be noted that this liquidation does not alter our findings, insofar as *BICM* have a limited number of offices (all in the capital city) in 2014 and does not account in the measurement of distance.

Alongside banks, there are 25 MFIs operating in Madagascar. MFIs in Madagascar are classified into three categories (1, 2, and 3). Categories 1 and 2 account for the largest number of MFIs, but category 3 is composed of the three largest MFIs in Madagascar. Our partner MFI belongs to the third category.[8] Our partner is specialized in the financing of microenterprises and entrepreneurs in urban and semi-urban areas. Its headquarters are in Antananarivo but there are credit offices across all the country.

Our partner MFI offers a wide range of saving, money transfer, and credit products. Concerning credit, the MFI focuses on both enterprises (micro, small, and medium enterprises) and individual entrepreneurs. Clients can be prospected by a credit officer on its work place or apply for a loan directly at branches. Our partner MFI grants two main types of individual loans. The very small enterprise loans amounts to a maximum of 20 million ariary (around $6,000) and a minimum of 200,000 ariary (approximately $60) for a maximum of 18 months (min. 3 months) maturity. For the small and medium enterprises, the maximum amount rises up to 200 M ariary (approximately $60,000) for a 36 month maturity. The amount granted (and collateral requirements required) depends on the entrepreneurs' reimbursement capacity, which is estimated by the credit officer based on basic accounting documents (simplified profit and loss, and balance sheet, which are often recomputed from scratch because fewer entrepreneurs keep accurate accounts) and soft information obtained during *in situ* visits by the credit officer. Although credit officers may propose large loans, in reality, our partner continues to offer mainly microloans. In 2014, half of the loans it granted were below $500 and less than 2.5% of loans exceeded $5,000. The majority of loans have a maturity of one year. The commercial sector amounts to more than 60% of its portfolio, and the services sector amounts to approximately 25%.

3 Hypotheses

Our aim in this paper is to investigate whether commercial banks and MFIs, operating theoretically on two different markets, compete or not. In doing so, we employ data on

[8] Due to confidentiality, we are not allowed to divulge our partner's name.

credit contract terms granted by one of the largest MFI in Madagascar. We are particularly interested in determining whether competition induced by bank presence influences loan terms. A critical step consists of defining a good measure of competitive pressure induced by banks on MFIs.[9] We turn to the recent literature investigating the role of distance in banking (Petersen and Rajan, 2002; Degryse and Ongena, 2005). Banks can extract rent from their relative proximity to borrowing firms not only due to transportation costs but also to informational advantages. For the lender, higher distance results in higher monitoring costs (Sussman and Zeira, 1995) and more difficulty in assessing the borrower's trustworthiness (Hauswald and Marquez, 2006). For the borrower, higher distance results in higher prospecting costs as it decreases their awareness of the availability and conditions of the loans offered (especially in the absence of advertising as may be the case, particularly in developing countries) and increases the cost of information (as it takes more time to reach the nearest branch). Empirical investigations (Degryse and Ongena, 2005; Bellucci et al., 2013) confirmed that bank-borrower proximity matters to explain credit conditions.

In this paper, we focus on the role of distance between an MFI's borrower and the closest commercial bank. Indeed, banking literature not only documents that borrower-lender distance matters but also that distance between a borrower and her closest competitor explains variation in credit conditions (Degryse and Ongena, 2005; Bellucci et al., 2013). Switching costs are reduced for borrowers located in the vicinity of a competing bank due to lower transportation and informational costs. In line with these arguments, we assume that the probability of an MFI's client being wooed by a commercial bank is higher for those located in the vicinity of a commercial bank, but this probability decreases with distance between the borrower and the closest competing bank.[10]

Assuming that borrowers located in the vicinity of a bank are more likely to obtain bank loans, we study whether the distance between the borrower and the closest bank

[9]There are a large number of indices of competition in the banking literature. However, these measures imply strong data requirements and are not well-adapted to consider competition between different types of lenders (see Léon, 2014).

[10]These statements are especially true considering that regular banks do not use wandering credit officers to prospect clients in a large area but are rather directly solicited by customers. In addition, it is worth noting that although mobile banking is currently developing in Madagascar, it does not enable people to obtain credit and its reach remain small for the moment. Therefore, we do not believe that mobile banking could influence our results.

affects loan conditions offered by our partner MFI. In the absence of competition (banks and MFI operate on two different markets), the MFI will be insensitive to the entry of a bank and does not adapt its lending policy to this change. However, if MFIs and banks compete, the MFI will react to the entry of a bank by improving its offers to retain its current clients and avoid a flight to bank. We therefore propose the following mutually exclusive hypotheses:

Hypothesis 1

The closer the MFI borrower to the new bank branch, the better the credit conditions offered by the MFI.

Hypothesis 2

The distance between an MFI's borrower and the closest bank does not affect loan conditions.

4 Empirical strategy

We study whether the presence of a bank affects credit conditions faced by borrowers of our partner MFI. We follow the literature (Degryse et al., 2009; Behr et al., 2011) and employ a linear specification as follows:

$$y_{istj} = \beta d_{it} + \Delta \mathbf{X}_{it} + \mu_i + \nu_t + \eta_s + \tau_j + \epsilon_{it} \tag{1}$$

where y_{istj} is the dependent variable capturing credit terms for borrower i in sector s in period t granted by credit officer j; d_{it} is the distance between the borrower i and the closest bank in period t; $\mathbf{X}_{it}$ is a matrix of variables controlling for firm's characteristics and lending relationship; μ_i, ν_t, and η_s are borrower i, period t, and sector s fixed effects, respectively.[11] Finally, supply-side factors are considered through the inclusion of credit officer dummies that correct for all unobserved time-invariant characteristics of the lender (τ_j).

[11] To be precise, we employ a within estimator at the borrower level that is equivalent to include borrower fixed effects but allows us to preserve degrees of freedom. In addition, we see that a non-negligible share of borrowers (approximatively one third) switch from one sector to another during the period of observations.

Our coefficient of interest is β. Hypothesis 1 implies that distance between the borrower and the closest bank influences credit conditions (by improving credit conditions).

Adding fixed effects allows us to focus on within variation and to avoid all bias induced by time-invariant unobserved characteristics that affect location and credit terms. In other words, we investigate whether the credit condition dynamic is shaped by a change in distance between the borrower and the closest bank rather than comparing credit conditions obtained by different borrowers at different distances. It should be noted that it is crucial to control for the intensity of the lender-borrower relationship. If not included, we fail to distinguish between the impact of distance and the effect of lending relationship that plays a central role in microfinance lending technologies.[12] The inclusion of borrower fixed effects has an additional advantage. If MFI's branches are located in the vicinity of bank branches, the distance between a bank and a client merely reflects the distance between the lender (here, the credit officer) and the borrower. Recent papers have shown that distance between the borrower and the lender affects information asymmetry in microfinance (Pedrosa and Do, 2011; Presbitero and Rabellotti, 2014)[13], and therefore credit conditions (Degryse and Ongena, 2005). Adding individual dummies allows us to control for this aspect because the distance between borrowers and their credit office is time-invariant in our study.[14]

In spite of its improvements, the inclusion of borrower fixed effects is not immune to criticism because we control only for unobserved time-invariant factors. However, time-varying factors (such as economic growth at the local level) may impact both bank's willingness to open a branch and credit conditions offered by MFIs to borrowers. In Section 6.3, we offer two alternatives specifications to control for this issue. First, we compute for each year and for the smallest administrative division (district, also called *fokontany*), the nightlight luminosity, which is a proxy of economic activity and/or pop-

[12] Our econometric findings are robust to the exclusion of lending relationship variable.

[13] Pedrosa and Do (2011) show that the intensity of screening increases with distance and Presbitero and Rabellotti (2014) show that distance increases information asymmetry and moral hazard.

[14] Even if clients sometimes change of credit officers, it is worth noting that the distance between MFI clients and their credit offices is time-invariant because clients change for a credit officer in the same branch.

ulation density (Chen and Nordhaus, 2011; Henderson et al., 2011). Second, we add municipality-period dummies that allows us to control for all unobserved shocks (beyond those captured by night light luminosity) that could affect a municipality for each period considered.

5 Data and variables

5.1 Data

5.1.1 Client file

The unique dataset we analyze consists of all loans granted over the period from January 1, 2008 to December 31, 2014 by one of the largest MFIs in Madagascar. Loans are always granted to entrepreneurs (individuals), for a productive purpose, never for individual consumption. For a first loan, the loan officer always visit the business of the entrepreneur to check whether the loan amount is consistent with the activity. Our partner shares with us its customer file. For each loan granted, we obtain data on the loan terms, as well as information on the borrower's business and the lender-borrower relationship. In addition, we have data on the precise location (latitude and longitude) of half of clients.

The initial database comprised 74,599 loans made to 35,472 borrowers. However, before selecting the final dataset used in the regressions, we applied some filters. We first removed double-counting and observations for those loans where at least one variable is lacking. We then trimmed the top and bottom 1% for each outcome and independent variable to avoid the presence of outliers.[15] Finally, we excluded observations with missing information on geographical location. By the end of 2014, our partner collected the location of 46% of its clients.[16] The final sample includes 14,834 borrowers representing

[15]Some exceptions are made for the age of the firm and the number of employees where the bottom 1% is zero and concerns a large number of observations (and is not an outlier).

[16]Since 2010 our partner has collected the precise location (latitude and longitude) of its clients. To date, 16,636 clients out of 35,472 clients are geolocated when we consider the whole sample (46.9%). From an empirical perspective, our estimates are subject to a sample selection bias insofar as our partner only provided GPS information for half of the clients in our study. There is no explicit rule to determine which clients are chosen and which are not. However, the choice to select some clients and exclude others is certainly a nonrandom decision, as indicated in Table A2). In Appendix B, we run a three-step model initially developed by Wooldridge (1995) to deal with sample selection issue in fixed effect model. We show that this issue is not crucial in our study.

32,374 observations.

5.1.2 Bank location

We complement our client database by identifying the location of every bank branch operating in Madagascar. As of December 31, 2014, we identified 227 bank branches operated by 12 commercial banks. We refer to the register of the Malagasy National Bank[17] to identify all of the commercial banks operating in the country. We focus exclusively on brick-and-mortar branches because we are concerned with credit activity. In other words, we do not consider ATMs or small banking service points (in hotels for instance) as branches. These infrastructures are used to provide basic financial services (deposits, withdrawals, money exchange) but are not used to grant loans. We hand collect the postal address of each branch on their website. Using addresses and Google Maps©, we obtain the precise location of all branches (latitude and longitude). It is worth noting that only half of the branches had a postal address accurate enough to be geolocated thanks only to the internet. We complement our database with *in situ* visits to obtain the precise location of unlocated branches. Finally, to obtain a time-varying measure, we complement data on branch locations by collecting the list of active branches by year from 2008 to 2014. To do so, we employ the annual list of branches provided by the Central Bank. We collect the list of branches operating in 2008 and identify new branches in each subsequent year.

The Appendix (Table A1) shows the number of bank branches by region and by year. We see that the number of bank branches increased by 40% from 2008 (134 branches) to 2014 (227 branches) and the expansion occurred in all provinces. We also observe two important waves in bank branch opening in 2010 and in 2012. Finally, the capital region accounts for half of the total bank branches, and its share increased over time (40% of all branches in 2008).

5.1.3 Limitations of our dataset

Our dataset has several shortcomings. First, the customer file used here reports information on loans after the acceptance phase. As a result, we have no information on

[17] http://www.banque-centrale.mg/index.php?id=m8_5_1

the amount asked by borrowers and on rejected borrowers. In addition, we do not know whether the borrower was able to repay.

Second, we lack information about the credit experience of borrowers with other intermediaries (before, during and after the loan acceptance). Borrowers are formally not allowed to ask for a loan from another institution. However, due to the lack of credible threat, this might occur.[18] We are therefore unable to investigate overindebtedness, repayment behavior or switching after the entry of a bank.

In addition, if a borrower disappears from our dataset, we are unable to know the reasons behind (borrowers that defaulted and did not obtain subsequent loans; borrowers without a need for credit; or borrowers that obtained a loan from a bank or another MFI). Nonetheless, we do not believe that this shortcoming invalidates our analysis because this bias is certainly an attenuation bias. There is no rational reason to expect that our partner MFI offers better credit conditions to borrowers in difficulty retaining them. However, we may expect that borrowers switching from our partner to a bank would certainly obtain better loan conditions if they would have decided to stay in relationship with the MFI. Therefore, our analysis is unable to estimate this effect and we therefore underestimate the true effect.

Regarding the bank branch location, we retrieve data provided by the Central Bank in the registry of bank branches for each year. We did not observe bank closure but only bank branch opening. We are unable to know if this is due to the absence of closure or if it is because the Central Bank does not remove inactive banks in its registry. We assume that all bank branches reported in the registry are active.

Finally, while we obtained the precise location of banks, we do not have such data for MFIs. The best we obtained is a dataset reporting location of MFIs at the municipality level for the last year. In other words, we are unable to track the opening of new MFI branches over time and changes in competition between MFIs. We expect that the inclusion of municipality-period dummies allows us to control for all shocks occurring at the municipality level, including increasing competition between MFIs.

[18] We surveyed 243 borrowers of our partner MFI in 2014, and only three of them declare to have a loan with another MFI, one from a money lender, and 18 from a bank.

5.2 Variables

5.2.1 Loan contracts

Data on credit loan terms are used to compute our outcome variables. Our partner MFI provides us with four different loan conditions: loan amount, interest rate, maturity, and collateral requirements. Loan amount and interest rates are deflated using the consumer price index. For collateral requirements, we compute the ratio of collateral pledged to total loans. Maturity is expressed in days. The descriptive statistics, reported in Table 1, document that loan amount represents $1,129 on average. The real interest rate is 12.6%, and the average loan has a maturity of one year. Guaranteed collateral represents 2.8 times the total value of the loans.

In the baseline analysis, we focus exclusively on loan amount and collateral requirements. Loan amount and collateral requirements capture two different aspects. Loan amount may proxy availability of credit in a context where borrowers cannot access to total amount of funding they ask for. Collateral requirements are more related to price in a context of fixed interest rates that is our situation, as explained in the following (Fisman et al., 2017).[19]

We do not consider maturity and (real) interest rate in the baseline due to the lack of variability, as documented in the last column of Table 1. Indeed, the majority of the loans have a maturity of one year (90% of the loans have a maturity between 365 days and 395 days), and interest rates vary between two values in nominal terms (18% or 21%). The determination of interest rate is not at the discretion of the credit officer: Until 2011, the nominal interest rate is 18% for all loans raising to 21% afterwards. In our sample only 1.6% of loans have an interest rate different from these two numbers. Nonetheless, in an extension, we consider (real) interest rates and maturities as well as an alternative definition of collateral requirements based on its composition (rather than its level).

[Insert here Table 1]

[19] There may be other ways for the MFI to compete, such as with the quality of its services, commercial advice dispensed by credit officers, application costs and time etc. Unfortunately, our database does not allow us to consider these aspects.

5.2.2 Distance

A crucial step consists of building a measure of competitive pressure from banks. To compute the distance between a given borrower and the closest bank, we use the precise location of borrowers provided by our partner MFI and the location of banks that we have hand-collected. Using QGIS©, an open source geographic information system, we computed the Euclidian distance between a given borrower and the closest bank.

Our measure differs from the literature (e.g., Degryse and Ongena, 2005) in two aspects. First, for the sake of simplicity and accuracy, we assessed the shortest distance in kilometers and did not use the shortest traveling time.[20] Second, we compute a time-variant measure of distance. It is common in the literature on developed countries to use time-invariant distance because bank networks do not significantly change over time. However, the network of bank branches in Madagascar has dramatically expanded from 2008 to 2014 in Madagascar. More than 40% of branches in 2014 were not active in 2008 (93 to 227), as shown in Table A1. As a result, using distance in 2014 to proxy distance in 2008 can be misleading. Obtaining a time-variant distance variable has the additional advantage of allowing us to consider within variation and therefore provide a better identification.

Our primary measure of distance is the distance from the borrower (whose location is time-invariant) and the closest bank (which can change over if a new bank opens in the vicinity) for each year. However, the same reduction of distance (e.g., 500 meters) does not have the same implication for a client located at one kilometer to the closest bank and for a client located at five kilometers. We may expect that the effect is stronger for the former. To consider this point, we also compute the logarithm of distance.

In addition to continuous measures of distance (linear or logarithm), we also consider two discrete measures. First, for each client, we create a treatment dummy equal to one in period t if the distance observed in t is lower than that observed in the initial period. For instance, let us assume that the entry of a new bank in 2011 reduces the distance between client A and her closest bank. The dummy will take a value of 0 from 2008 to 2011 and then a value of 1. If this client did not experience a reduction of distance, the

[20]Due to the lack of information on the road network in Madagascar, computing the shortest distance for each borrower would have been very difficult and inaccurate.

dummy value is zero from 2008 to 2014. We expect an opposite sign to be observed for distance and log of distance. This dummy can be assimilated to a treatment dummy and it captures how loan dynamics are shaped by a reduction in distance. As a result, this dummy is easy to interpret. However, this measure comes with a major shortcoming; it assumes that all changes in distance are similar.

In addition to the dummy variable, we create categorical dummies. The first dummy takes a value of 1 if a borrower is located at less than 500 meters of the closest bank branch and 0 otherwise. The second dummy takes a value of 1 if the distance between the client and the closest bank branch is between 500 meters and one kilometer, and so on until 2000 meters. We omit a dummy for distance above 2000 meters. Coefficients associated with each category reflect the comparison of situation between a situation where the closest bank is located at more than 2000 meters and a new situation where a bank is located in the new category. For instance, coefficient associated with the second category captures the change of credit condition for a borrower those closest bank where remote (more than 2 kilometers) and a new situation where closest bank is comprised between 500 meters and 1 kilometer.

5.2.3 Control variables

The list of control variables includes information on business characteristics and on borrower-lender relationship intensity. The literature shows that business characteristics are important determinants of loan contract terms in banking (Degryse et al., 2009) and in microfinance (Behr et al., 2011). In particular, opaque firms obtain less advantageous credit conditions. Opacity is often assessed by size (Berger et al., 2001). We therefore add two measures of size, namely total sales (in current USD) and number of employees. In addition, we control for business activity by using dummies for the business sector.

In addition, the banking literature underlines the importance of controlling for the lending relationship (Degryse et al., 2009). This aspect is particularly important in this work for two reasons. First, our identification strategy (see Section 4) implies that we control for the lending relationship. Second, microfinance lending technologies are based on dynamic incentives (Armendáriz and Morduch, 2010). We therefore expect a loan amount to increase with the length of the lending relationship. Regarding collateral, Behr et al. (2011) document that collateral requirements are relaxed over the course of the lending relationship. Following Behr et al. (2011), we proxy the lending relationship using the number of loans obtained by the borrowers.[21] As shown in Table 1, the average firm financed by our partner is eight years old and has two employees and total monthly sales of approximately $1,762. On average, borrowers have had a relationship with our partner MFI for two years.

6 Baseline result

We test whether MFIs and banks compete by scrutinizing whether our partner MFI is sensitive to the entry of a new bank at the proximity of its clients. We consider two different loan characteristics: loan amount and collateral requirements.

6.1 Loan amount

We first report results regarding the effect of distance on the size of loans in Table 2. The hypothesis of competition between MFIs and banks is validated if a reduction in distance increases the total amount offered by MFIs. Put differently, we expect that the coefficients associated with continuous measures of distance are negative (columns [1-2]) and positive for discrete measures of distance (columns [3-4]).

The coefficient associated with distance in meters (column [1]) has the expected sign and is statistically significant at the usual thresholds, in line with Hypothesis 1. Using the log of distance instead of distance alters statistical significance (significant at the 15% level), but the sign is as expected. Employing discrete measures of distance (columns [3-4]) provides a similar conclusion as a linear measure of distance. Borrowers experiencing a reduction of distance obtain better credit conditions, namely, here larger loans.

Distance is not only statistically significant but also economically significant: a one standard deviation increase in distance induces a reduction of the loan amount by more

[21] An alternative measure is the duration of the relationship (Petersen and Rajan, 1994; Berger and Udell, 1995) in years, which we utilize without altering our conclusions. The two measures of the lender-borrower relationship are closely related with a correlation coefficient exceeding 0.9

than \$110 (approximatively 10% of the mean value of loan amount).[22] Using alternative measures of distance provides similar findings. The fact to experience a decrease in distance raises the loan amount by more than \$70 (column [3]). The results reported in column [4] provide evidence that the effect of distance is not homogenous. They indicate that only clients located in a circle of less than one kilometer of a bank obtain larger loans. The observed impact is far from anecdotal: the fact to be located in a circle of less than 500 meters increases the loan amount by more than \$250 (comparatively to a previous situation with a distance above 2000 meters, the reference).

Regarding control variables, we note that larger firms obtain access to larger loans. In addition, borrowers benefit from longer and more intense relationships (loan number) with the lender through larger loans. Dynamic incentives are used in microfinance to reduce moral hazard (Armendáriz and Morduch, 2010), and it is therefore normal to observe a positive correlation between the number of loans and the loan amount.

[Insert here Table 2]

6.2 Collateral requirements

We then study the determinants of collateral requirements by considering the ratio of collateral value to total loan value in columns [5-8]. Collateral ratio is a form of price in the context of fixed interest rates as ours (Fisman et al., 2017). According to Hypothesis 1, we expect that the collateral-to-loan ratio is lower for firms in the vicinity of a bank. Put differently, we expect that $\beta > 0$ for continuous measures of distance (columns [5-6]) and $\beta < 0$ for discrete measures of distance (columns [7-8]). The coefficient associated with linear measure of distance (column [5]) is positive as expected but not statistically significant. However, the coefficient is significant at the 5% level when we consider the natural logarithm of distance (column [6]). This finding is confirmed when we focus on alternative measures of distance: borrowers located in the vicinity of a bank pledge less collateral. Coefficients are statistically significant and have the expected negative sign

[22] Table 2 reports standardized coefficients. The standardized coefficients are interpreted as the standard deviation change in the dependent variable when the independent variable is changed by one standard deviation (Bring, 1994). Put differently, the impact of one standard deviation of variable X is computed as follows: $\sigma_y * \hat{\beta}_x$ where σ_y is the standard deviation of Y and $\hat{\beta}_x$ the estimated standardized coefficient of X. Here the impact of distance is obtained as $\sigma_y * \hat{\beta}_d = 2207 * 0.051 = 112$.

in models with treatment dummy (columns [7]) and distance categories (columns [8]). These results tend to confirm Hypothesis 1.

The economic impact of distance on collateral seems worthy of attention, albeit not huge. A one standard deviation increase in (log of) distance induces an increase of the collateral-to-loan ratio by almost 15%. The amplitude given by discrete measures are rather similar. The fact to experience a reduction in distance reduces this ratio by 12% (column [7]). The results from column [8] indicate that clients located in a circle of 500 meters of a bank have a collateral ratio that is reduced by 37.5 points (comparatively to those located at 2000 meters and more) and by more than 25 points for those with a distance between 500 meters and one kilometer. These differences are far from anecdotal: the average value of collateral ratio equals 280% (standard deviation equals 120%).

Regarding control variables, our model provides interesting results. Our finding indicates that the lending relationship relaxes collateral requirements, confirming findings obtained by Behr et al. (2011) in Mozambique. The coefficients associated with the number of loans granted by our partner are negative and statistically significant at the 1% level. However, larger firms do not seem to pledge less collateral.

6.3 Identification issue

Our identification strategy is based on the inclusion of borrower fixed effects that allow us to control for all time-invariant unobservable characteristics. However, one might raise concerns about time-variant unobserved factors affecting both bank branch location, firms' performance and credit conditions. To address this issue, we should employ control for economic information at the local level. Unfortunately, time-varying economic, demographic, and social indicators are, at the best, available at a regional level in Madagascar and often only at the national level. Nonetheless, we offer two alternative specifications in the following.

6.3.1 Inclusion of night light intensity

First, we employ the night light intensity at the district level. In the footsteps of Chen and Nordhaus (2011) and Henderson et al. (2011), luminosity is often used to proxy economic activity and/or population density at a subnational level, especially in low-income countries. We follow this approach and compute night light intensity for the smallest administrative division in Madagascar, namely, the district (also called *fokontany*). An advantage is that luminosity data are available for each year and vary over time, which allows us to keep the borrower fixed effect. As a result, luminosity captures changes in local economic conditions over time.

Before presenting the results, we detail some methodological points. The data are made available by the National Geophysical Data Center of the National Oceanic and Atmospheric Administration of the U.S. and originate from images taken by satellites of the Defense meteorological Satellite Program of the U.S. Department of Defense.[23] Night light intensity data are available at the pixel level, each pixel corresponding to 30x30 arc second grids (i.e., an area of 0.86 square kilometer on the equator). Data produced report average visible, stable nighttime lights and cloud-free coverage, where ephemeral events (such as fires) as well as background noise are removed and only light from sites with persistent lightning is included. Values of night light intensity range from 0 (no luminosity) to 63 (maximum luminosity). Since April 2012, monthly data are available, and annual data have not been updated since 2013. Contrary to annual data, monthly data are raw data unfiltered for clouds, moon light or other confounding factors (such as fires). Monthly and annual data are not directly comparable. Consequently, we rely exclusively on annual data that cover 5 years out of 6 (instead of 2 for monthly data) and have the advantage of being filtered to obtain stable night light intensity.

Annual data are available from 1992 to 2013, which implies two consequences for us. First, we drop all observations in 2014 when we consider night light intensity. Second, insofar as our period of reference is the semester (and not the year), we consider annual data for both semesters in each year.

For our empirical analysis, luminosity data are aggregated from the pixel level to the respective observational unit, namely, the district that is the lowest administrative division in Madagascar.[24] There are 17,544 districts in Madagascar, but borrowers are

[23]We use the latest version of nighttime lights (Version 4 composites: F16 (2008-09) and F18 (2010-2013)). Data are extracted from `https://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html` (last accessed on July, 2018).

[24]Madagascar is divided into five main subdivisions: 6 provinces (called "*faritany*"), 22 regions ("*faritany*"), 144 counties ("*fivondronana*"), 1,395 municipalities ("*kaominina*") and 17,544 districts ("*fokon-*

located in 1,249 districts. Considering the district allows us to have very precise information on local economic growth. We compute for each district the average value of night light intensity. A major drawback is that night light is available until 2013 (see below). As a result, our sample is reduced by one-quarter when we include this variable (24,075 observations from 10,779 borrowers instead of 32,374 observations from 14,834 borrowers).

[Insert here Table 3]

Table 3 displays the results when we include night light intensity as a control variable. In Panel A, we report results for loan amount and in Panel B those for collateral-to-loan ratio. Because we exclude the last year (2014) due to the lack of data, we present results for observations from 2008 to 2013 without luminosity variables and then include this control variable. In summary, we observe that the inclusion of night light intensity does not dramatically change our findings. Specifically, when we compare columns with and without luminosity variable, coefficients and statistical significance associated with distance are unchanged. In addition, night light intensity is never statistically significant.[25] In other words, we believe that we will obtain very similar results from those obtained in Table 2 if we get access to luminosity data for 2014.

Our main conclusions regarding the impact of distance on loan amount and collateral ratio are largely not altered when we exclude the last year. While we observe that the impact of distance on loan amount is reduced in statistical terms, this is due to a reduction in the sample. Nonetheless, distance seems to continue to play a role as documented with linear measure and dummy variable. For collateral ratio, our findings are closely similar in statistical and economical terms when we exclude 2014 (without or with luminosity variable).

tany").

[25] One might raise concerns about the fact that all district experienced a similar trend in luminosity. To test this possible issue, we run night light intensity on year dummies for all district (within model). Inclusion of year dummies explains less than 40% of (within) variation. In other words, more than three fifths of (within) variability is unexplained by time dummies.

6.3.2 Inclusion of municipality-period dummies

We consider a second test for identification by adding municipality-semester dummies. The best way to control for all unobserved time-variant factors consists of adding dummies that control for all shocks occurring at the local level. Ideally, we should add dummies interacting district (the smallest administrative division) with semester (our period of reference). However, this approach is unfeasible due to the huge number of parameters to estimate (more than 10,000 dummies). As a result, we consider the second lowest division, namely, the municipality.

Before adding municipality-semester dummies, it is interesting to give a glimpse of municipality characteristics to gauge whether municipality-semester dummies capture local environment. Municipality covers, on average, an area of 37.7 square kilometers. However, the mean is driven by outliers in rural areas (large municipalities), and the median municipality has an area equal to 20 square kilometers. Municipalities had on average 13,000 inhabitants in 2009.[26] Less than one quarter of municipalities have more than 15,000 inhabitants (median equals 10,000 and first quartile 6,500 inhabitants). Borrowers are located in 236 municipalities across Madagascar. These municipalities are similar to other municipalities of Madagascar in terms of areas but they differ in terms of the number of inhabitants. On average, there were 24,200 inhabitants in 2009, with a median of 14,415 inhabitants (first quartile equals 9,271 and the last quartile 23,385). This finding is not surprising insofar as our partner MFI operates mainly in urban areas. On average, there are 18 borrowers per municipality-period with a minimum of 1 and a maximum of 311. Based on these raw statistics, we assume that municipality is a correct level to capture local environment.

The municipality-semester dummies enable us to control for any shock that occurs at the municipality level over time. We believe that this approach allows us to control for a large range of time-invariant unobserved factors and therefore provide unbiased estimations. A limitation of this approach (explaining why we do not employ it in our baseline) is the number of parameters to estimate (more than 2,000 parameters because borrowers are located in 236 different municipalities).

[26] Unfortunately, the last relevant survey on population by communes had been provided in 2009.

[Insert here Table 4]

Results on models including municipality-semester dummies are displayed in Table 4. Coefficients associated with distance, with the exception of log of distance, have the expected sign and are statistically significant when we study the determinants of loan amount (Panel A). In Panel B, we investigate the determinants of the collateral ratio. As in the baseline model, coefficients associated with distance have the expected sign but are not always statistically significant at the usual threshold, especially when we consider continuous measures. In a nutshell, after controlling for time-variant unobserved factors, our findings remain largely robust (and are even reinforced for loan amount).

7 Sensitivity tests

In this section, we scrutinize whether our findings are robust to additional tests. We also investigate the conditional impact of distance according to the borrower's characteristics (size, age and the existence of a previous relation with our partner).

7.1 Robustness checks

All tables for robustness checks are displayed in online Appendix.

7.1.1 Restricting the common support

Our econometric intuition is based on the idea that borrowers experienced a reduction in distance will obtain favorable loan conditions. In the baseline model, we consider both borrowers who experienced a reduction in distance (called treated in the following) and those who do not (untreated). In the spirit of the impact evaluation model, borrowers without change in distance can be considered as a control group and other borrowers as a treatment group (this analogy is particularly suited when we consider the treated dummy). However, as indicated in Table A2, we see that treated borrowers differ from untreated borrowers in their (initial) characteristics.[27] We see that treated borrowers initially have better loan conditions (higher amount, lower collateral requirements); they

[27] We thank an anonymous referee inviting us to consider this point.

are larger, older and have a longer relationship with our partner MFI. However, treated firms are more distant from bank office in the initial period and operate in less dynamic areas (according to night light luminosity). In our baseline analysis, we control for this initial difference through the inclusion of borrower fixed effects. However, one might raise concerns that both groups differ not only in their initial characteristics but also in their evolution over time (a form of rejection of parallel trend in difference-in-difference). Unfortunately, because "treatment" (reduction in distance) occurs at different moments of the observation periods, we cannot test this hypothesis.

However, we present additional regressions that restrict the common support by matching treated borrowers with untreated borrowers sharing the same characteristics. In doing so, we employ coarsened exact matching. The coarsened exact matching is an exact matching because it simply matches a treated unit to all control units with the same covariate value. However, it is coarsened insofar as we transformed continuous variables into categorical ones before applying the matching procedure.[28] We match two borrowers if and only if they are in the same category for all variables (sales, employees, age, previous loan, distance to a bank, and industry). In a robustness check, we add as a matching variable the level of luminosity at the fokontany-level (Match 2) or the credit officer dummy (Match 3). We consider the initial characteristics of borrowers to match treated to untreated (initial is defined as the first loan over the period 2008-2014). After matching, the common support is sharply reduced because some untreated borrowers are not similar to treated borrowers and because for some treated borrowers, we are unable to find "identical" untreated borrowers.[29]

In addition to the matching approach, we restrict our common support exclusively on borrowers that experienced a reduction in distance over time. Exploiting only within variation among the treated borrowers allows us to control for unobserved differences

[28]We match on a variety of variables, including total sales (divided into seven categories), number of employees (5), firm age (4), a dummy whether borrower obtained a loan from our partner before 2008 (2), the distance to the bank (7), and industry (15).

[29]In our baseline analysis we have 14,834 borrowers (32,374 observations) among them 2,152 had experienced a reduction in distance. When we consider our first matching procedure (Match 1), our sample is reduced to 5,186 borrowers (13,615 observations) among them 1,423 treated and 3,763 untreated. In our second matching strategy (Match 2) including the level of luminosity, we keep only 2,014 borrowers (5,990 obs.) those 772 are treated. In our third matching (Match 3) that include credit officer, we get 2,159 borrowers (6,198 obs.) with 802 treated borrowers.

between treated and untreated borrowers. This approach is similar in spirit of the event study methodology used in finance (MacKinlay, 1997) or in program impact evaluation (McIntosh et al., 2011). This approach relies exclusively on changes before and after treatment and avoids including borrowers with different characteristics.

The results of our estimation based on coarsened exact matching (Match 1, Match 2, Match 3) and on the sample of treated borrowers exclusively (Treated) are displayed in the Appendix (Table A3). For loan amount, our main findings are unchanged when we consider continuous measures of distance, but models with discrete measures of distance are more sensitive to the change of common support. Although coefficients are in line with our main findings (positive), their economic and statistical magnitude are sometimes reduced when we consider matching. For the models explaining the collateral ratio, findings are closely in line with our main findings reported in Table 2. Our results are even reinforced when we restrict the sample of borrowers considered. In other words, these additional tests tend to confirm our main results and give support to our baseline findings.

7.1.2 Keeping borrowers with more than three (five) loans over the period

As indicated in Table 1, the number of loans per borrower ranges from 1 to 20, with an average of 2.9. This issue raises concerns because our identification rests on variation within a borrower's loan history.[30] As a result, all borrowers with only one loan do not help in their identification. In addition, there is a risk that the correlation between the number of loans and the probability of experiencing a reduction in distance to a bank biases our econometric results. Indeed, we see in our data this correlation: the probability of experiencing a reduction in distance is only 12% for borrowers with 2 loans but exceeds 50% for those with 5 loans and more.

To address this issue, we rerun our baseline model in Table A4, but we keep only borrowers with at least three loans in Panel A and with at least five loans in Panel B.[31] Despite a reduction in the number of observations (from 34,834 to 18,816 in Panel A and to 9,811 in Panel B), our main findings are not altered. The econometric results provided

[30]We thank an anonymous reviewer to point this issue

[31]We also run a model that remove all borrowers with one loan (7,215 borrowers), and results are very close to those reported in Table A4 (Panel A).

in Table A4 are very similar to those obtained in our baseline model (Table 2).

7.1.3 Attrition

Another potential econometric issue is the attrition issue. For almost half of borrowers (6,646 among 14,834), we do not have data in the last year (2014). However, the attrition seems orthogonal to the treatment (reduction in distance). In an unreported analysis, we scrutinize whether the probability of experiencing a reduction in distance is correlated with attrition. In doing so, we run a cross-sectional model (14,834 observations) in which we explain a dummy taking value one for attrition (if no observation is available in 2014) and zero in the absence of attrition. The results, available upon request, indicate that the probability of experiencing a reduction in distance is uncorrelated with the attrition. Nonetheless, to account for this possible issue, we adopt a simple framework. We consider only individuals for which we have data in the last year (2014) and rerun the baseline model on this subsample. The results displayed in Table A5 show that this issue does not seem to bias our findings.

7.1.4 Inclusion of additional control variables

Brick and Palia (2007) note that loan terms are jointly determined. We further employ one extension of Eq. 1 by adding other characteristics of loan terms as control variables to test the sensitivity of our baseline results. Specifically, we include interest rates, maturity and collateral ratio when we consider the determinants of loan amount and interest rates, maturity and loan amount when we consider the determinants of collateral ratio. However, including other loan terms might induce an endogeneity problem due to reverse causation and unobserved third factors. As a result, these results should be treated with caution.

The results, reported in Table A6 in the Appendix, confirm our main conclusion. The statistical effect of distance on loan amount (columns [1-4]) is robust to the inclusion of other credit conditions, but its economic effect is reduced in different specifications. This may reflect the fact that distance alters other terms of loan contracts (such as collateral ratio). We also note that the interest rate tends to be weakly correlated with the size of the loan, contrary to maturity (positive correlation) and collateral requirements (negative correlation). Turning to the collateral-to-loan ratio in columns [5-8], our findings regarding distance are largely unaltered. In addition, we note that only the loan amount is negatively correlated with the collateral ratio.

In spatial models, lending conditions may depend on three parameters (see Degryse and Ongena, 2005): (i) the distance between the borrower and the lender, (ii) the distance between the borrower and the closest competing lender, and (iii) the number of competitors. In our baseline model, we control for the distance between the borrower and the closest competing bank (interest variable) and for the distance between the borrower and the lender (by adding borrower fixed effects). However, we cannot control for the number of competitors due to a lack of data. While we do not believe that this could affect our findings, one might argue that the distance between the lender and the closest competing bank may capture changes in the number of competitors. In the following, we try to test whether our findings are sensitive to that point. Unfortunately, we have limited information on the total number of lenders, especially other MFIs. Nonetheless, we add the distance between the lender (our partner's agency) and the closest competing bank. This distance is an imperfect proxy of the competitive pressure induced by the presence of alternative lender in the vicinity, and we expect that a reduction signal an increase in the number of competitors. The results, displayed in Appendix (Table A7), show that our findings are insensitive to the inclusion of this new control variable. In addition, the distance between the lender and the closest competing bank is not significantly correlated with credit contract terms.

7.1.5 Excluding Access Banque and Microcred

The banking system in Madagascar groups together traditional commercial banks, commercial banks with specific products dedicated to small firms (such as *BOA*) and former MFIs transform into banks (such as *Access Banque* and *Microcred*), as documented in Section 2. One might argue that our results are only driven by former actors. To control for this point, we compute alternative measures of distance by excluding *Access Banque* and *Microcred*. The results, displayed in Table A8, provide interesting findings. In columns [1-4], we focus on determinants of loan amount. Coefficients associated with continuous measures of distance have the expected sign, but are not always statistically significant at the usual thresholds. However, we obtain a statistically significant impact of distance on loan amount for discrete measure, in line with our main results. When we concentrate on the collateral ratio, the impact of distance remains unchanged. In other words, our findings are not driven by the inclusion of former MFIs (Access Banque and Microcred) but remain valid when we focus exclusively on other commercial banks, especially *BOA*. In economic terms, we can say that commercial banks compete with our partner MFI.

7.1.6 Alternative dependent variables

To date, we focus on two main characteristics of loan contracts, namely, loan amount and collateral ratio. In the following, we consider three alternative dependent variables: (i) the structure of collateral; (ii) maturity; and (iii) real interest rate.

First, we consider the quality of collateral instead of its level. The presence of a bank may not only affect the quantity of collateral but also the quality of collateral. We therefore also focus on the composition of collateral. Different forms of collateral are required to obtain a loan. To simplify, we can distinguish between personal guarantees and material guarantees. Personal guarantees involve a third party who agrees to reimburse the loan in case of default. Material guarantees (security) are all assets that the lender can seize in the event of default. Because material guarantees directly affect them, borrowers may prefer to limit the amount of material assets that they guarantee for the total loan amount. Better loan conditions therefore imply not only a limited colateral-to-loan ratio but also a limited percentage of material guarantee to collateral. We compute this ratio as our second measure of collateral requirements. Econometric results using the composition of collateral are reported in Table A9 (columns [1-4]). The findings are very similar to those obtained using the amount of collateral and can be summarized as follows. Distance is not related to the level of material collateral in all specifications considering continuous measures. However, findings are in line with Hypothesis 1 when we employ discrete variables (dummy and categories). As previously discussed, for the collateral ratio, we show that coefficients associated with dummy variables for treated and for two first categories are statistically significant. In economic terms, however, the impact of distance is rather limited. For instance, the fact to witness a decrease in distance reduces the ratio of material guarantees by 1.5 percentage points (mean equals 54.5% and standard deviation equals 16.2%).

Second, we consider the two other loan characteristics that we ignore in the baseline: maturity and real interest rates. Both are excluded due to the lack of variability. For the sake of transparency, we present econometric results on models explaining both variables. In columns [5-8], we display the relationship between distance and maturity. Coefficients associated with distance have the expected sign but are not statistically significant. This finding is expected due to the lack of variation in maturities. We next investigate the impact of distance on real interest rates. As explained above, the nominal interest rate takes two values in the large majority of loans: 18% (in 40% of loans) and 21% (in 57% of loans). The lack of variability certainly explained the lack of relationship between distance and real interest rate. It should be noted that employing nominal interest rate or dummies for low rate provide similar findings (i.e., an absence of relationship).

7.2 Heterogenous impact of distance

We now scrutinize whether the relationship between distance and loan contract terms is sensitive to the borrower's characteristics. We consider two main categories: the degree of transparency and the existence of a previous relationship with our partner MFI.

7.2.1 Transparent vs. opaque borrowers: The role of size (and age)

Our story is based on the intuition that the presence of a bank in the vicinity will change the behavior of our partner MFI in terms of credit conditions offered. In the final paragraph, we investigate whether the impact of distance differs according to the type of borrowers. First, we scrutinize whether the most transparent firms are more likely to be captured by formal banks than opaque firms. In line with existing works, we focus on two proxies of opaqueness, namely, firm size (Berger et al., 2001) and firm age (Hyytinen and Pajarinen, 2008). Indeed, banks are certainly more able to attract larger firms and older firms. On the supply-side, lending technologies developed by banks are more effective for borrowers with credit history and able to pledge assets. On the demand side, small and young could be less interested by banking products. A firm is classified as a large firm if its size (assessed by total sales or the number of employees) is above the median. Similarly, a firm is considered as a young firm if its age is below the median in 2008. We consider for each firm the size and the age during the first loan in 2008 to avoid a switch in the classification for the same borrower. Based on this classification, we interact our different measurement of distance with a dummy for transparent (large and old) firms.[32] We expect that the impact of distance is larger for transparent firms.

The results are displayed in Table 5. In Panel A, we report results for loan amount and in Panel B those for collateral ratio. Each block presents the results for one estimation. We first consider a dummy for old firms in the first column, and in the two subsequent columns we interact distance measures with a dummy for large firms. Our findings are partially in line with our expectations. As indicated in the first column, there is no relevant difference between young and old firms. We see two possible explanations for the lack of statistical significance. First, even young firms could be transparent and attracted by banks. Second, old firms are also firms with longer relationships with our partner MFI, and the lending relationship may attenuate the impact of external competition (see below). Findings for large firms are more in line with our expectations, especially for loan amount. The impact of distance on loan amount is stronger for larger firms, irrespective of measure of size considered (number of employees or sales). Nonetheless, we fail to provide a similar conclusion for the collateral ratio. In other words, large firms are more impacted by the entry of a new bank but only for quantity. For price (collateral ratio), we observe no difference between large and small firms.

[Insert here Tables 5]

7.2.2 Captured vs. new borrowers: The role of lending relationship

We then consider the impact of the lending relationship. A large body of literature has investigated the complex relationship between competition and the borrower-lender relationship in banking. Since Petersen and Rajan (1995), many works have investigated whether competition hinders long-term relationships, without providing a clear answer

[32] As a robustness check, we run models on subsamples instead of employing interactions. Results, available upon request, are largely similar. We also consider quartile instead of median without altering our findings.

(see Degryse and Ongena, 2007; Presbitero and Zazzaro, 2011, among others). However, long-term relationships may also mitigate the impact of competition on borrowers. Informational rents obtained by lenders during the lending relationship lead to borrower capture to the extent that such information cannot be communicated credibly to outsiders (Dell'Ariccia and Marquez, 2004; Degryse and Ongena, 2005). In the following section, we test whether borrowers with a previous relationship with our partner are less impacted by the entry of a new bank. We create a dummy variable equal to one if a borrower obtained a loan from our partner before 2008, and 0 otherwise. The results, displayed in the last column of Table 5 indicate that the impact of competition on the collateral ratio is shaped by relationship lending (but not the impact on loan amount). While we observe that our partner MFI tends to reduce its collateral requirements for new borrowers when a bank opens in the vicinity, it tends to increase the ratio of collateral for captured clients. If collateral requirement is a form of price in a context of fixed interest rates (Fisman et al., 2017), this result is in line with findings in banking literature (Degryse and Ongena, 2005).

8 Conclusion

This paper investigates whether the respective strategies of upscaling and downscaling initiated by MFIs and commercial banks have resulted in competition between them. Our econometric analysis provides one main result: the competitive pressure from banks induces better loan conditions for MFI's borrowers (at least for clients without a strong relationship with the lender). We therefore derive from this finding that MFIs and regular banks do not operate in strictly segmented markets, as often believed; instead, they tend to compete.

Specifically, we study whether firms located in the vicinity of a bank obtain better loan conditions from an MFI than they would otherwise. Our intuition is based on the idea that an MFI will offer better loan conditions to keep its clients, only if MFIs and banks are in competition. In the absence of competition, bank proximity should not affect loan conditions offered by MFIs. We employ an original panel data set of 32,374 loans to 14,834 borrowers granted by one of the major MFIs in Madagascar over the period 2008-2014. We find that the proximity of a bank to an MFI client increased the size of the MFI loan obtained and decreased collateral requirements. In addition, we document that the effect of competition seems stronger for clients that could be more easily caught by banks (large firms and clients without a previous relationship with the MFI). We employ different identification strategies to reveal the true effect of competition from banks on MFIs. Nonetheless, future research should examine whether the effects documented by our analysis hold in different contexts (in Africa or elsewhere) and using alternative identification strategies.

An unresolved question is whether competition between MFIs and banks is a good or a bad new for access to finance in low-income countries and for the business of microfinance. On the one hand, our econometric results can be read optimistically. MFIs are able to follow their clients in their growth and offer them better credit conditions. In addition, this strategy may enable MFIs to become more profitable (due to scale economies) and therefore to improve financial inclusion for the poorest entrepreneurs through a cross-subsidization strategy (profitable loans subsidize less profitable ones, i.e., smaller loans due to scale economies). However, our findings can also be interpreted in more nuanced terms. Existing papers have investigated how competition and commercialization affect the microfinance industry and borrowers. Two main issues have been underlined. First, competition induces a risk of overborrowing that could hurt the viability of MFIs (McIntosh and Wydick, 2005; McIntosh et al., 2005). Second, MFIs could drift away to their social mission to favor financial performance (Mersland and Strøm, 2010), for instance by relying more on individual loans than group lending (De Quidt et al., 2018).

In addition to these well-known potential risks, observing that MFIs and banks compete asks the question of the transition from informal to formal lending. Two opposite views emerge. On the one hand, there is a risk of hold-up for borrowers, already identified in the banking literature (Dell'Ariccia and Marquez, 2004). Private information obtained by MFIs may lead to borrower capture to the extent that such information cannot be communicated credibly to banks. This risk is exacerbated because MFI borrowers cannot easily produce hard information and often rely on soft information. Our findings tend to indicate that MFIs have an informational advantage insofar as borrowers with long-term relationships seem captured by MFIs. This could have detrimental effects on economic development directly, by hindering investment, or indirectly, by providing small firms fewer incentives to formalize. On the other hand, obtaining a loan from an MFI allows previously unbanked borrowers to create credit history and favor their switching towards banks. Agarwal et al. (2018) provide a very interesting case study on how loans obtained by credit cooperatives may serve to build a reputation and help unbanked borrowers to have access to bank loans. However, their study takes place in a specific context, namely, Rwanda, where cooperatives and banks exchange information on borrowers through a credit bureau. Additional research is required to shed light on which view is the most accurate and the effect of the environment (such as the design of credit information sharing mechanisms).

From a policy perspective, increasing interactions between MFIs and formal banks implies considering the unexpected spillover effect from one sector to the other. In particular, regulation in one sector may affect the businesses in other sectors. Policymakers should investigate the indirect effect of bank regulation on MFIs (and vice versa). In a recent paper, Tantri (2018) confirms this view by showing that a change in regulation in microfinance had significant spillover effects on banks in India.

References

Agarwal, S., Kigabo, T., Minoui, C., Presibtero, A., and Silva, A. F. (2018). Financial access under the microscope. *mimeo, available at https://ssrn.com/abstract=3247796*.

Ahlin, C., Lin, J., and Maio, M. (2011). Where does microfinance flourish? Microfinance institution performance in macroeconomic context. *Journal of Development Economics*, 95(2):105–120.

Armendáriz, B. and Morduch, J. (2010). *The Economics of Microfinance*. MIT Press, Cambridge, MA.

Armendáriz, B. and Szafarz, A. (2011). On mission drift in microfinance institutions. In *Handbook of Microfinance*, pages 341–366.

Assefa, E., Hermes, N., and Meesters, A. (2013). Competition and the performance of microfinance institutions. *Applied Financial Economics*, 23(9):767–782.

Baquero, G., Hamadi, M., and Heinen, A. (2018). Competition, loan rates and information dispersion in nonprofit and for-profit microcredit markets. *Journal of Money, Credit and Banking*, 50(5):893–937.

Beck, T. and Demirguc-Kunt, A. (2006). Small and medium-size enterprises: Access to finance as a growth constraint. *Journal of Banking & Finance*, 30(11):2931–2943.

Behr, P., Entzian, A., and Guettler, A. (2011). How do lending relationships affect access to credit and loan conditions in microlending? *Journal of Banking & Finance*, 35(8):2169–2178.

Bellucci, A., Borisov, A., and Zazzaro, A. (2013). Do banks price discriminate spatially? Evidence from small business lending in local credit markets. *Journal of Banking & Finance*, 37(11):4183–4197.

Berger, A. N., Klapper, L. F., and Udell, G. F. (2001). The ability of banks to lend to informationally opaque small businesses. *Journal of Banking & Finance*, 25(12):2127–2167.

Berger, A. N. and Udell, G. F. (1995). Relationship lending and lines of credit in small firm finance. *Journal of Business*, pages 351–381.

Brick, I. E. and Palia, D. (2007). Evidence of jointness in the terms of relationship lending. *Journal of Financial Intermediation*, 16(3):452–476.

Bring, J. (1994). How to standardize regression coefficients. *American Statistician*, 48(3):209–213.

Brown, M., Guin, B., and Kirschenmann, K. (2016). Microfinance banks and financial inclusion. *Review of Finance*, 20(3):907–946.

Chen, X. and Nordhaus, W. D. (2011). Using luminosity data as a proxy for economic statistics. *Proceedings of the National Academy of Sciences*, 108(21):8589–8594.

Cull, R., Demirgüç-Kunt, A., and Morduch, J. (2007). Financial performance and outreach: A global analysis of leading microbanks. *Economic Journal*, 117(517):F107–F133.

Cull, R., Demirgüç-Kunt, A., and Morduch, J. (2014). Banks and microbanks. *Journal of Financial Services Research*, 46(1):1–53.

De Mel, S., McKenzie, D., and Woodruff, C. (2008). Returns to capital in microenterprises: evidence from a field experiment. *Quarterly Journal of Economics*, CXXIII(4):1329–1372.

De Quidt, J., Fetzer, T., and Ghatak, M. (2018). Commercialization and the decline of joint liability microcredit. *Journal of Development Economics*, 134:209–225.

Degryse, H., Kim, M., and Ongena, S. (2009). *Microeconometrics of banking: Methods, applications, and results*. Oxford University Press, USA.

Degryse, H. and Ongena, S. (2005). Distance, lending relationships, and competition. *Journal of Finance*, 60(1):231–266.

Degryse, H. and Ongena, S. (2007). The impact of competition on bank orientation. *Journal of Financial Intermediation*, 16:399–424.

Dell'Ariccia, G. and Marquez, R. (2004). Information and bank credit allocation. *Journal of Financial Economics*, 72(1):185–214.

Ferrari, A. and Jaffrin, G. S. (2006). Commercial bank downscaling - lessons from experience. *Access Finance - The World Bank Group*, 14.

Fisman, R., Paravisini, D., and Vig, V. (2017). Cultural proximity and loan outcomes. *American Economic Review*, 107(2):457–492.

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

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1):153–161.

Henderson, V., Storeygard, A., and Weil, D. N. (2011). A bright idea for measuring economic growth. *American Economic Review*, 101(3):194–99.

Hermes, N., Lensink, R., and Meesters, A. (2011). Outreach and efficiency of microfinance institutions. *World Development*, 39(6):938–948.

Hyytinen, A. and Pajarinen, M. (2008). Opacity of young businesses: Evidence from rating disagreements. *Journal of Banking & Finance*, 32(7):1234–1241.

Léon, F. (2014). Measuring competition in banking: A critical review of methods. *CERDI Etudes et Documents*, 12:1–44.

Léon, F. (2015). Does bank competition alleviate credit constraints in developing countries? *Journal of Banking & Finance*, 57:130–142.

MacKinlay, A. C. (1997). Event studies in economics and finance. *Journal of Economic Literature*, 35(1):13–39.

McIntosh, C., Janvry, A., and Sadoulet, E. (2005). How rising competition among microfinance institutions affects incumbent lenders. *Economic Journal*, 115(506):987–1004.

McIntosh, C., Villaran, G., and Wydick, B. (2011). Microfinance and home improvement: using retrospective panel data to measure program effects on fundamental events. *World Development*, 39(6):922–937.

McIntosh, C. and Wydick, B. (2005). Competition and microfinance. *Journal of Development Economics*, 78(2):271–298.

McKenzie, D. and Woodruff, C. (2008). Experimental evidence on returns to capital and access to finance in mexico. *World Bank Economic Review*, 22(3):457–482.

Mersland, R. and Strøm, R. Ø. (2010). Microfinance mission drift? *World Development*, 38(1):28–36.

Navajas, S., Conning, J., and Gonzalez-Vega, C. (2003). Lending technologies, competition and consolidation in the market for microfinance in bolivia. *Journal of International Development*, 15(6):747–770.

Pedrosa, J. and Do, Q.-T. (2011). Geographic distance and credit market access in Niger. *African Development Review*, 23(3):289–299.

Petersen, M. A. and Rajan, R. G. (1994). The benefits of lending relationships: Evidence from small business data. *Journal of Finance*, 49(1):3–37.

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

Petersen, M. A. and Rajan, R. G. (2002). Does distance still matter? the information revolution in small business lending. *Journal of Finance*, 57(6):2533–2570.

Presbitero, A. F. and Rabellotti, R. (2014). Geographical distance and moral hazard in microcredit: evidence from colombia. *Journal of International Development*, 26(1):91–108.

Presbitero, A. F. and Zazzaro, A. (2011). Competition and relationship lending: Friends or foes? *Journal of Financial Intermediation*, 20(3):387–413.

Sussman, O. and Zeira, J. (1995). Banking and development. *CEPR Discussion Paper*, 1127.

Tantri, P. L. (2018). Contagious effects of a political intervention in debt contracts: Evidence using loan-level data. *Review of Financial Studies*, forthcoming.

Vanroose, A. and D'Espallier, B. (2013). Do microfinance institutions accomplish their mission? Evidence from the relationship between traditional financial sector development and microfinance institutions' outreach and performance. *Applied Economics*, 45(15):1965–1982.

Wooldridge, J. M. (1995). Selection corrections for panel data models under conditional mean independence assumptions. *Journal of Econometrics*, 68(1):115–132.

Tables

Table 1: Descriptive Statistics

Variable	Obs.	Mean	Std. Dev.	Min	Max	CV[†]
Loan contract						
Amount[‡]	32,374	1,129	2,207	22.9	40,076	1.95
Interest rate (real)	32,374	0.126	0.027	0.060	0.220	0.21
Maturity	32,374	389.0	49.2	88	1,095	0.13
Collateral/Amount	32,374	2.851	1.272	0.088	10.84	0.45
Securities/Collateral	32,374	0.545	0.162	0	1	0.30
Distance						
- in meters	32,374	2,403	5,529	5.53	88,604	2.30
- in log	32,374	7.027	1.122	1.710	11.39	0.16
- Dummy	32,374	0.156	0.361	0	1	2.31
- Dist<500m	32,374	0.202	0.401	0	1	2.31
- 500m<dist<1000m	32,374	0.258	0.438	0	1	1.70
- 1000m<dist<1500m	32,374	0.173	0.378	0	1	2.19
- 1500m<dist<2000m	32,374	0.103	0.304	0	1	2.19
- dist>2000m	32,374	0.264	0.441	0	1	1.67
Control variables						
Sales[‡]	32,374	1762.6	2792	0.4	24,555	1.58
Employees	32,374	2.257	2.498	0	32	1.11
Loan Number	32,374	2.901	2.487	1	20	0.86
Luminosity	24,075	27.135	23.016	2	63	0.85
Firm age	32,374	10.465	6.631	0	44	0.64

[†] CV=Std. Dev/Mean; [‡] Data are deflated and in USD

Table 2: Determinants of loan amount and collateral ratio

	Amount				Collateral ratio			
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	-0.051***				0.044			
	(-2.83)				(0.64)			
Log(distance)		-0.039				0.097**		
		(-1.61)				(2.49)		
Dummy			70.05**				-0.120***	
			(2.17)				(-3.83)	
dist<500m				259.4*				-0.375***
				(1.85)				(-3.42)
500m<dist<1000m				172.0*				-0.259***
				(1.67)				(-2.70)
1000m<dist<1500m				-55.04				-0.108
				(-0.62)				(-1.20)
1500m<dist<2000m				-29.31				-0.045
				(-0.47)				(-0.53)
Sales (in current USD)	0.092***	0.092***	0.092***	0.092***	-0.000	-0.000	-0.000	-0.000
	(5.56)	(5.57)	(5.57)	(5.56)	(-0.02)	(-0.01)	(-0.02)	(-0.00)
Employees	0.050***	0.050***	0.050***	0.050***	-0.005	-0.005	-0.004	-0.004
	(2.97)	(2.96)	(2.95)	(2.95)	(-0.34)	(-0.33)	(-0.29)	(-0.30)
Loan number	0.347***	0.348***	0.348***	0.348***	-0.471***	-0.471***	-0.472***	-0.472***
	(11.93)	(11.95)	(11.96)	(11.98)	(-9.58)	(-9.56)	(-9.62)	(-9.58)
Obs.	32,374	32,374	32,374	32,374	32,374	32,374	32,374	32,374
Nb. Borrowers	14,834	14,834	14,834	14,834	14,834	14,834	14,834	14,834
R2 (within)	0.0750	0.0748	0.0747	0.0755	0.0291	0.0297	0.0300	0.0302

The dependent variable is the total loan amount in deflated USD in columns [1-4] and collateral ratio in columns [5-8]. The table shows estimations with four different measures of distance (for categories distance, the omitted category is distance above 2,000 meters). Within estimator (at the borrower level) is used and period, industry and credit officer dummies are included in all specifications. Standard errors are clustered at the borrower level. t-statistics are reported in parentheses. Standardized coefficients are reported for continuous variables (distance, log(distance) and control variables) and usual coefficients for binary variables (dummy and categories). *, **, and *** indicate significance at the 10%, 5% and 1% respectively

Table 3: Inclusion of night light intensity

	Panel A: Determinants of amount of loan							
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	-0.032*	-0.032*						
	(-1.69)	(-1.69)						
Log(distance)			-0.016	-0.016				
			(-0.81)	(-0.80)				
Dummy					43.87*	44.05*		
					(1.66)	(1.66)		
dist<500m							105.2	104.4
							(0.93)	(0.92)
500m<dist<1000m							119.4	118.9
							(1.16)	(1.15)
1000m<dist<1500m							-70.27	-70.63
							(-0.75)	(-0.75)
1500m<dist<2000m							-39.31	-39.46
							(-0.66)	(-0.67)
Night light intensity		0.002		0.002		0.002		0.002
		(1.23)		(1.16)		(1.25)		(1.08)
Obs.	24,075	24,075	24,075	24,075	24,075	24,075	24,075	24,075
Nb. Borrowers	10,779	10,779	10,779	10,779	10,779	10,779	10,779	10,779
R2 (within)	0.0793	0.0793	0.0791	0.0791	0.0792	0.0792	0.0798	0.0798

	Panel B: Determinants of collateral ratio							
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	0.050	0.050						
	(0.80)	(0.80)						
Log(distance)			0.102***	0.103***				
			(2.77)	(2.78)				
Dummy					-0.141***	-0.141***		
					(-3.79)	(-3.78)		
dist<500m							-0.431***	-0.434***
							(-3.84)	(-3.86)
500m<dist<1000m							-0.314***	-0.316***
							(-3.10)	(-3.12)
1000m<dist<1500m							-0.162*	-0.163*
							(-1.73)	(-1.74)
1500m<dist<2000m							-0.037	-0.037
							(-0.39)	(-0.40)
Night light intensity		0.009		0.010		0.009		0.010
		(1.22)		(1.28)		(1.13)		(1.36)
Obs.	24,075	24,075	24,075	24,075	24,075	24,075	24,075	24,075
Nb. Borrowers	10,779	10,779	10,779	10,779	10,779	10,779	10,779	10,779
R2 (within)	0.0290	0.0291	0.0298	0.0299	0.0302	0.0302	0.0306	0.0306

The dependent variable is the total loan amount in deflated USD in Panel A and collateral ratio in Panel B. The table shows estimations with four different measures of distance (for categories distance, the omitted category is distance above 2,000 meters). Within estimator (at the borrower level) is used and period, industry and credit officer dummies, as well as control variables (Sales, number of employees, and loan number), are included in all specifications. Standard errors are clustered at the borrower level. t-statistics are reported in parentheses. Standardized coefficients are reported for continuous variables (distance, log(distance) and control variables) and usual coefficients for binary variables (dummy and categories). *, **, and *** indicate significance at the 10%, 5% and 1% respectively

Table 4: Inclusion of municipality-semester dummies

	Amount				Collateral ratio			
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	-0.120**				0.112			
	(-2.08)				(0.46)			
Log(distance)		-0.032				0.086*		
		(-0.93)				(1.66)		
Dummy			97.58**				-0.016	
			(2.28)				(-1.29)	
dist<500m				331.23**				-0.317**
				(2.01)				(-2.11)
500m<dist<1000m				229.05*				-0.220*
				(1.70)				(-1.67)
1000m<dist<1500m				20.21				-0.121
				(0.19)				(-1.08)
1500m<dist<2000m				32.99				-0.073
				(0.44)				(-0.78)
Obs.	32,374	32,374	32,374	32,374	32,374	32,374	32,374	32,374
Nb. Borrowers	14,834	14,834	14,834	14,834	14,834	14,834	14,834	14,834
R2 (within)	0.170	0.170	0.170	0.170	0.111	0.111	0.111	0.112

The dependent variable is the total loan amount in deflated USD in columns [1-4] and collateral ratio in columns [5-8]. The table shows estimations with four different measures of distance (for categories distance, the omitted category is distance above 2,000 meters). Within estimator (at the borrower level) is used and period, industry and credit officer dummies are included in all specifications. In all specifications, we include control variables (Sales, number of employees, and loan number) and municipality-semester dummies. Standard errors are clustered at the borrower level. t-statistics are reported in parentheses. Standardized coefficients are reported for continuous variables (distance, log(distance) and control variables) and usual coefficients for binary variables (dummy and categories). *, **, and *** indicate significance at the 10%, 5% and 1% respectively

Table 5: Conditional impact on borrower's characteristics

Panel A: Loan amount

Type = 1 if	Age>Median		Nb. Empl>Median		Sales>Median		Had a loan before 2008	
	Coef.	(t-stat)	Coef.	(t-stat)	Coef.	(t-stat)	Coef.	(t-stat)
Distance	-0.060**	(-2.25)	-0.018	(-1.04)	0.015	(1.13)	-0.061***	(-2.63)
Distance*Type	0.005	(0.24)	-0.045***	(-2.74)	-0.070***	(-3.85)	0.009	(0.59)
Log(distance)	-0.035	(-1.54)	0.004	(0.24)	0.031**	(2.15)	-0.033	(-1.06)
Log(distance)*Type	-0.047	(-0.34)	-0.344***	(-2.17)	-0.334***	(-3.06)	-0.044	(-0.37)
Dummy	16.86	(0.98)	-11.50	(-0.29)	-36.96	(-1.46)	46.91	(1.16)
Dummy*Type	83.53*	(1.68)	199.3**	(2.87)	154.4***	(3.50)	53.68	(0.89)
dist<500m	175.5	(1.52)	-24.50	(-0.37)	-149.6**	(-2.43)	240.6	(1.32)
(dist<500m)*Type	127.7	(0.57)	739.2**	(2.42)	596.1***	(3.12)	41.25	(0.16)
500m<dist<1000m	113.5	(1.11)	94.13	(0.87)	-51.21	(-1.07)	201.7	(1.46)
(500m<dist<1000m)*Type	78.64	(0.52)	145.9	(0.73)	348.6**	(2.12)	-57.77	(-0.29)

Panel B: Collateral ratio

Type = 1 if	Age>Median		Nb. Empl>Median		Sales>Median		Had a loan before 2008	
	Coef.	(t-stat)	Coef.	(t-stat)	Coef.	(t-stat)	Coef.	(t-stat)
Distance	0.074	(0.85)	-0.068	(-0.84)	0.087	(1.23)	0.052	(0.95)
Distance*Type	-0.031	(-0.38)	0.142**	(2.15)	-0.037	(-0.43)	-0.002	(-0.03)
Log(distance)	0.0808*	(1.67)	0.072	(1.55)	0.185***	(2.84)	0.173***	(-1.06)
Log(distance)*Type	0.117	(0.52)	0.205	(0.93)	-0.381	(-1.54)	-0.362*	(-1.78)
Dummy	-0.196	(-5.25)	-0.150***	(-4.20)	-0.209***	(-4.55)	-0.236***	(-6.22)
Dummy*Type	0.165**	(1.95)	0.091	(1.53)	0.150	(1.60)	0.268***	(4.76)
dist<500m	-0.367***	(-2.60)	-0.444***	(-3.54)	-0.661***	(-4.00)	-0.703***	(-4.86)
(dist<500m)*Type	-0.001	(-0.04)	0.210	(0.83)	0.457**	(2.14)	0.654***	(3.07)
500m<dist<1000m	-0.136***	(-2.55)	-0.364***	(-3.30)	-0.439***	(-3.20)	-0.403***	(-3.18)
(500m<dist<1000m)*Type	0.121	(0.61)	0.304	(1.39)	0.283	(1.50)	0.267	(1.44)

The dependent variable is the total loan amount in deflated USD in Panel A and collateral ratio in Panel B. The table only reports results for distance measure and interaction between distance and dummy taken value 1 if age is above the median value in the first column (old firms), if the number of employees is above the median in the second column (large firms), if the total sales are above the median in the third column (large firms) and for firms having a previous relationship with our partner MFI in the fourth column (captured borrowers). Each block is the result of one estimation. For all specifications, control variables as well as period, industry and credit officer dummies are included but unreported and within estimator (at the borrower level) is used. Standard errors are clustered at the borrower level. t-statistics are reported in parentheses. Standardized coefficients are reported for continuous variables (distance, log(distance) and control variables) and usual coefficients for binary variables (dummy and categories). *, **, and *** indicate significance at the 10%, 5% and 1% respectively

Do banks and microfinance institutions compete? Microevidence from Madagascar

Appendix

Appendix A Additional Tables

Table A1: Bank branches from 2008 to 2014, by province and region

Year	2008	2009	2010	2011	2012	2013	2014
Antananarivo (Centre)	*61*	*65*	*88*	*96*	*107*	*109*	*115*
Itasy	1	1	1	1	1	1	2
Analamanga	54	58	78	84	89	91	95
Vakinankaratra	5	5	7	9	13	13	13
Bongolova	1	1	2	2	4	4	5
Antsiranana (North)	*19*	*19*	*20*	*20*	*21*	*25*	*27*
Diana	9	9	10	10	11	13	15
Sava	10	10	10	10	10	12	12
Fianarantsoa (East)	*15*	*15*	*15*	*15*	*20*	*20*	*21*
Amoron'i Mania	3	3	3	3	3	3	3
Haute Matsiatra	3	3	3	3	8	8	9
Vatovavy-Fitovinany	5	5	5	5	5	5	5
Ihorombe	1	1	1	1	1	1	1
Atsimo-Atsinanana	3	3	3	3	3	3	3
Mahajanga (North-West)	*10*	*10*	*11*	*12*	*15*	*17*	*17*
Sofia	2	2	2	2	3	4	4
Boeny	7	7	8	9	11	12	12
Betsiboka	0	0	0	0	0	0	0
Melaky	1	1	1	1	1	1	1
Toamasina (North-East)	*20*	*21*	*23*	*25*	*27*	*29*	*31*
Alaotra-Mangoro	7	7	7	8	9	9	10
Atsinanana	8	8	10	11	12	14	15
Analanjirofo	5	6	6	6	6	6	6
Toliara (South-West)	*9*	*11*	*11*	*11*	*14*	*14*	*16*
Menabe	2	3	3	3	3	3	4
Atsimo-Andrefana	4	4	4	4	7	7	7
Androy	0	1	1	1	1	1	1
Anosy	3	3	3	3	3	3	4
TOTAL (Madagascar)	**134**	**141**	**168**	**179**	**204**	**215**	**227**

Table A2: t-test

	Geolocated		t-test	Treated		t-test
	No	Yes	(p-value)	No	Yes	(p-value)
Loan characteristics						
Amount†	889.9	997.9	<0.01	887.8	1183.7	<0.01
Rate	0.096	0.123	<0.01	0.131	0.091	<0.01
Maturity	396.4	389.6	<0.01	388.6	388.9	0.81
Collateral ratio	3.014	3.232	<0.01	3.131	2.939	<0.01
Security to collateral	0.563	0.578	<0.01	0.857	0.563	<0.01
Borrowers characteristics						
Distance to bank	-	-	-	2485.7	3122.8	<0.01
Sales†	1857.4	1829.8	0.66	1380.3	2088.2	<0.01
Employees	2.208	2.142	0.07	1.909	2.560	<0.01
Age (firm)	10.976	9.255	<0.01	8.416	12.305	<0.01
Nb relationship	1.764	1.482	0.07	1.201	2.238	<0.01
Luminosity	31.860	26.186	<0.05	27.972	22.597	<0.01
# borrowers	*18,836*	*16,636*		*12,682*	*2,152*	

† Data in deflated USD
Means are obtained for the first loan over the period 2008-2014 (initial condition)

Table A3: Models using matching procedures and restricting to treated borrowers only

	Amount				Collateral ratio			
	Match 1	Match 2	Match 3	Treated	Match 1	Match 2	Match 3	Treated
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance [A]	-0.054***	-0.054**	-0.064**	-0.046***	0.137***	0.135*	0.135	0.036
	(-2.92)	(-2.45)	(-2.38)	(-3.28)	(2.04)	(1.66)	(1.54)	(0.57)
Log(distance) [B]	-0.018	-0.032	-0.006	-0.036*	0.120***	0.103	0.137**	0.083**
	(-0.86)	(-0.95)	(-0.29)	(-1.76)	(2.74)	(1.52)	(2.04)	(2.31)
Dummy [C]	60.38*	48.88	45.09	70.20**	-0.158***	-0.178***	-0.196***	-0.159***
	(1.71)	(0.79)	(1.34)	(1.98)	(-3.93)	(-3.06)	(-3.57)	(-4.40)
dist<500m [D]	171.7	460.7*	240.3*	269.9*	-0.422***	-0.657***	-0.705***	-0.357***
	(1.05)	(1.81)	(1.78)	(1.78)	(-3.14)	(-3.04)	(-3.32)	(-3.10)
500m<dist<1000m [D]	141.6	452.0*	188.9	176.2*	-0.271**	-0.378*	-0.477**	-0.244**
	(1.12)	(1.85)	(1.30)	(1.66)	(-2.21)	(-1.95)	(-2.48)	(-2.47)
1000m<dist<1500m [D]	-46.88	119.2	-91.98	-48.22	-0.088	-0.325*	-0.308*	-0.104
	(-0.47)	(0.55)	(-0.65)	(-0.51)	(-0.81)	(-1.79)	(-1.78)	(-1.15)
Matching on								
- Sales	X	X	X		X	X	X	
- Employees	X	X	X		X	X	X	
- Previous rel.	X	X	X		X	X	X	
- Distance to bank	X	X	X		X	X	X	
- Industry	X	X	X		X	X	X	
- Luminosity		X				X		
- Credit officer			X				X	
Obs.	13,615	5,990	6,198	9,091	13,615	5,990	6,198	9,091
Nb. Firms	5,186	2,014	2,159	2,152	5,186	2,014	2,159	2,152
Treated	1,423	772	802	2,152	1,423	772	802	2,152
Untreated	3,763	1,242	1,357	0	3,763	1,242	1,357	0

Each cell displays the coefficient and associated t-statistic for our interest variable (distance) for one specific model. The dependent variable is the total loan amount in deflated USD in columns [1-4] and collateral ratio in columns [5-8]. Match 1 refers to our coarsened exact matching procedure when treated borrowers (those experienced a reduction in distance) and untreated borrowers are matched according to (initial value of) sales, number of employees, the existing previous relationship with our partner MFI, distance to the closest bank and the industry. In match 2, we add the luminosity intensity at the fokontany level in the matching procedure and in Match 3 we add credit officer. The columns with label "Treated" refer to models focusing on borrowers that experienced a reduciton in distance (treated borrowers). Each row (from [A] to [D]) represents models with different measure of distance (level [A], log of distance [B], dummy [C] and categories [D]). Models is similar to those used in Table 2. Within estimator (at the borrower level) is used and period, industry and credit officer dummies are included in all specifications. In all specifications, we include control variables (Sales, number of employees, and loan number). Standard errors are clustered at the borrower level. t-statistics are reported in parentheses. Standardized coefficients are reported for continuous variables (distance, log(distance) and control variables) and usual coefficients for binary variables (dummy and categories). *, **, and *** indicate significance at the 10%, 5% and 1% respectively

Table A4: Keeping only borrowers with at least 3/5 loans

Panel A: Keeping borrowers with at least three loans over the period

	Amount				Collateral ratio			
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	-0.044***				0.053			
	(-2.62)				(0.76)			
Log(distance)		-0.045*				0.111***		
		(-1.81)				(2.87)		
Dummy			73.20**				-0.125***	
			(1.86)				(-3.71)	
dist<500m				286.0*				-0.399***
				(1.86)				(-3.53)
500m<dist<1000m				189.1*				-0.279***
				(1.69)				(-2.84)
1000m<dist<1500m				-70.17				-0.130
				(-0.73)				(-1.42)
1500m<dist<2000m				-34.18				-0.034
				(-0.49)				(-0.39)
Obs.	18,816	18,816	18,816	18,816	18,816	18,816	18,816	18,816
Nb. Borrowers	4,433	4,433	4,433	4,433	4,433	4,433	4,433	4,433
R2 (within)	0.083	0.083	0.083	0.085	0.029	0.030	0.030	0.030

Panel B: Keeping borrowers with at least five loans over the period

	Amount				Collateral ratio			
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	-0.055**				0.053			
	(2.37)				(0.76)			
Log(distance)		-0.048				0.106**		
		(-1.54)				(2.14)		
Dummy			51.53*				-0.120***	
			(1.66)				(-2.93)	
dist<500m				218.5				-0.447***
				(1.30)				(-3.16)
500m<dist<1000m				226.6*				-0.421***
				(1.66)				(-3.49)
1000m<dist<1500m				-39.11				-0.142
				(-0.32)				(-1.22)
1500m<dist<2000m				-60.58				-0.070
				(-0.73)				(-0.64)
Obs.	9,811	9,811	9,811	9,811	9,811	9,811	9,811	9,811
Nb. Borrowers	1,755	1,755	1,755	1,755	1,755	1,755	1,755	1,755
R2 (within)	0.112	0.112	0.112	0.113	0.036	0.037	0.037	0.039

The dependent variable is the total loan amount in deflated USD in columns [1-4] and collateral ratio in columns [5-8]. The table shows estimations with four different measures of distance (for categories distance, the omitted category is distance above 2,000 meters). Within estimator (at the borrower level) is used and period, industry and credit officer dummies are included in all specifications. In all specifications, we include control variables (Sales, number of employees, and loan number). Standard errors are clustered at the borrower level. t-statistics are reported in parentheses. Standardized coefficients are reported for continuous variables (distance, log(distance) and control variables) and usual coefficients for binary variables (dummy and categories). *, **, and *** indicate significance at the 10%, 5% and 1% respectively

Table A5: Attrition issue

	Amount				Collateral ratio			
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	-0.078***				-0.028			
	(-2.78)				(-0.27)			
log(distance)		-0.068*				0.065		
		(-1.81)				(1.15)		
Dummy			89.92*				-0.069*	
			(1.94)				(-1.77)	
Dist<500m				483.5**				-0.304*
				(2.11)				(-1.91)
500m<dist<1000m				358.7**				-0.182
				(2.18)				(-1.36)
1000m<dist<1500m				64.76				-0.026
				(0.48)				(-0.20)
1500m<dist<2000m				48.47				0.035
				(0.50)				(0.30)
Observations	18,980	18,980	18,980	18,980	18,980	18,980	18,980	18,980
Nb. Borrowers	8,417	8,417	8,417	8,417	8,417	8,417	8,417	8,417
R2 (within)	0.0777	0.0778	0.0773	0.0791	0.0343	0.0346	0.0346	0.0354

The dependent variable is the total loan amount in deflated USD in columns [1-4] and collateral ratio in columns [5-8]. The table shows estimations with four different measures of distance (for categories distance, the omitted category is distance above 2,000 meters). Within estimator (at the borrower level) is used and period, industry and credit officer dummies are included in all specifications. In all specifications, we include control variables (Sales, number of employees, and loan number). Standard errors are clustered at the borrower level. t-statistics are reported in parentheses. Standardized coefficients are reported for continuous variables (distance, log(distance) and control variables) and usual coefficients for binary variables (dummy and categories). *, **, and *** indicate significance at the 10%, 5% and 1% respectively

Table A6: Adding other contract terms

	Amount				Collateral ratio			
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	-0.045***				0.038			
	(-2.65)				(0.56)			
Log(distance)		-0.025				0.092**		
		(-1.25)				(2.37)		
Dummy			52.11*				-0.113***	
			(1.81)				(-3.70)	
dist<500m				199.1*				-0.355***
				(1.66)				(-3.04)
500m<dist<1000m				83.22				-0.246**
				(0.69)				(-2.30)
1000m<dist<1500m				-54.38				-0.111
				(-0.95)				(-1.01)
1500m<dist<2000m				-19.57				-0.050
				(-0.63)				(-0.35)
Interest rate (real)	-0.187*	-0.187*	-0.187*	-0.186*	0.090	0.091	0.091	0.090
	(-1.86)	(-1.86)	(-1.86)	(-1.85)	(1.36)	(1.39)	(1.38)	(1.36)
Maturity	0.258***	0.257***	0.258***	0.257***	-0.005	-0.004	-0.005	-0.005
	(12.78)	(12.80)	(12.79)	(12.81)	(-0.37)	(-0.34)	(-0.36)	(-0.35)
Collateral ratio	-0.034***	-0.034***	-0.034***	-0.033***				
	(-4.64)	(-4.63)	(-4.62)	(-4.59)				
Amount (in USD)					-0.119***	-0.118***	-0.118***	-0.118***
					(-4.95)	(-4.93)	(-4.92)	(-4.88)
Obs.	32,374	32,374	32,374	32,374	32,374	32,374	32,374	32,374
Nb. Borrowers	14,834	14,834	14,834	14,834	14,834	14,834	14,834	14,834
R2 (within)	0.266	0.266	0.266	0.266	0.0351	0.0357	0.0360	0.0361

The dependent variable is the total loan amount in deflated USD in columns [1-4] and collateral ratio in columns [5-8]. The table shows estimations with four different measures of distance (for categories distance, the omitted category is distance above 2,000 meters). Within estimator (at the borrower level) is used and period, industry and credit officer dummies are included in all specifications. In all specifications, we include control variables (Sales, number of employees, and loan number). Standard errors are clustered at the borrower level. t-statistics are reported in parentheses. Standardized coefficients are reported for continuous variables (distance, log(distance) and control variables) and usual coefficients for binary variables (dummy and categories). *, **, and *** indicate significance at the 10%, 5% and 1% respectively

Table A7: Inclusion of the distance between MFI agency and the closest bank

	Amount				Collateral ratio			
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	-0.051***				0.046			
	(-3.16)				(0.67)			
Log(distance)		-0.040*				0.092**		
		(-1.65)				(2.33)		
Dummy			61.35*				-0.110***	
			(1.86)				(-3.41)	
dist<500m				257.5*				-0.346***
				(1.73)				(-3.05)
500m<dist<1000m				156.6				-0.235**
				(1.44)				(-2.40)
1000m<dist<1500m				-37.96				-0.089
				(-0.43)				(-0.99)
1500m<dist<2000m				-25.94				-0.031
				(-0.40)				(-0.37)
Distance IMF-bank	-0.012	-0.001	-0.006	-0.000	0.104*	0.072	0.074	0.063
	(-0.57)	(-0.03)	(-0.25)	(-0.02)	(1.79)	(1.20)	(1.26)	(1.07)
Obs.	32,374	32,374	32,374	32,374	32,374	32,374	32,374	32,374
Nb. Borrowers	14,834	14,834	14,834	14,834	14,834	14,834	14,834	14,834
R2 (within)	0.0790	0.0788	0.0787	0.0795	0.0290	0.0295	0.0297	0.0300

The dependent variable is the total loan amount in deflated USD in columns [1-4] and collateral ratio in columns [5-8]. The table shows estimations with four different measures of distance (for categories distance, the omitted category is distance above 2,000 meters). Within estimator (at the borrower level) is used and period, industry and credit officer dummies are included in all specifications. In all specifications, we include control variables (Sales, number of employees, and loan number). Standard errors are clustered at the borrower level. t-statistics are reported in parentheses. Standardized coefficients are reported for continuous variables (distance, log(distance) and control variables) and usual coefficients for binary variables (dummy and categories). *, **, and *** indicate significance at the 10%, 5% and 1% respectively

Table A8: Excluding former MFIs (Access Banque and Microcred)

	Amount				Collateral ratio			
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	-0.029				0.055			
	(-1.03)				(0.66)			
log(distance)		-0.038				0.147***		
		(-1.18)				(4.17)		
Dummy			96.30**				-0.159***	
			(2.40)				(-4.48)	
Dist<500m				283.5*				-0.421***
				(1.66)				(-3.87)
500m<dist<1000m				199.5*				-0.239**
				(1.66)				(-2.42)
1000m<dist<1500m				43.96				-0.108
				(0.43)				(-1.08)
1500m<dist<2000m				80.44				0.010
				(1.01)				(0.11)
Observations	32,374	32,374	32,374	32,374	32,374	32,374	32,374	32,374
Nb. Borrowers	14,834	14,834	14,834	14,834	14,834	14,834	14,834	14,834
R2 (within)	0.0776	0.0778	0.0780	0.0781	0.0292	0.0304	0.0306	0.0305

The dependent variable is the total loan amount in deflated USD in columns [1-4] and collateral ratio in columns [5-8]. The table shows estimations with four different measures of distance (for categories distance, the omitted category is distance above 2,000 meters). Within estimator (at the borrower level) is used and period, industry and credit officer dummies are included in all specifications. In all specifications, we include control variables (Sales, number of employees, and loan number). Standard errors are clustered at the borrower level. t-statistics are reported in parentheses. Standardized coefficients are reported for continuous variables (distance, log(distance) and control variables) and usual coefficients for binary variables (dummy and categories). *, **, and *** indicate significance at the 10%, 5% and 1% respectively

Table A9: Change of dependent variable

	Quality of collateral				Maturity (in days)				Real interest rate			
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
Distance	-0.034				-0.016				-0.001			
	(-0.75)				(-0.82)				(-0.46)			
Log(distance)		0.033				-0.040				-0.005		
		(1.01)				(-1.37)				(-0.33)		
Dummy			-0.015***				0.409				0.0000	
			(-3.51)				(0.76)				(0.30)	
dist<500m				-0.026**				3.500				0.0001
				(-2.03)				(1.40)				(0.30)
500m<dist<1000m				-0.028**				5.220*				0.0000
				(-2.44)				(1.75)				(0.02)
1000m<dist<1500m				-0.009				1.286				0.0003
				(-0.83)				(0.55)				(1.59)
1500m<dist<2000m				-0.005				-0.618				0.0001
				(-0.41)				(-0.32)				(0.66)
Dummies												
- Industry	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
- Credit officer	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
- Period	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	32,374	32,374	32,374	32,374	32,374	32,374	32,374	32,374	32,374	32,374	32,374	32,374
Nb. Borrowers	14,834	14,834	14,834	14,834	14,834	14,834	14,834	14,834	14,834	14,834	14,834	14,834
R2 (within)	0.0697	0.0697	0.0705	0.0703	0.0235	0.0237	0.0235	0.0239	0.985	0.985	0.985	0.985

The dependent variable is the ratio of material guarantees to total guarantees in columns [1-4], maturity in days in columns [5-8], and the real interest rates in columns [9-12]. The table shows estimations with four different measures of distance (for categories distance, the omitted category is distance above 2,000 meters). Within estimator (at the borrower level) is used and period, industry and credit officer dummies are included in all specifications (except in columns [9-12] where period dummies are not included). In all specifications, we include control variables (Sales, number of employees, and loan number). Standard errors are clustered at the borrower level. t-statistics are reported in parentheses. Standardized coefficients are reported for continuous variables (distance, log(distance) and control variables) and usual coefficients for binary variables (dummy and categories). *, **, and *** indicate significance at the 10%, 5% and 1% respectively

Appendix B Three-step Wooldridge procedure

Due to localization of only half of clients, our estimates are subject to a potential sample selection issue. In econometric terms, we suspect a sample selection issue that may bias results. Heckman (1979) provides a simple way to test and control for sample selection in cross-sectional data. However, this issue is more complex for panel data with fixed effects. Different parametrical and non-parametrical methods have been developed to control for sample selection in fixed-effect model (for remainder, our identification strategy is based on the inclusion of borrower fixed effect). In this paper we employ the three-step procedure proposed by Wooldridge (1995). In the following, we briefly present this method.

In a first step, for each period (here, semester) we estimate a selection equation using a standard probit as follows:

$$Pr(s_i = 1) = \Phi(\Delta \mathbf{X}_i + \gamma Z_j + \eta_s) \quad (\forall t = 0, \ldots, T) \tag{2}$$

where s_i is a dummy equals 1 if a borrower is geolocated and 0 otherwise, and Z_j is a selection variable and X_i the list of control variables included in the baseline model (without/with other credit terms). The selection variable must be strongly correlated with the selection rule (here, the likelihood to be geolocated) but not with outcome (here, credit conditions faced by agent i). As selection variable (Z_j), we use the share of geolocated clients by credit officer j, defined as follows:[33]

$$Z_j = \frac{\text{Nb of geolocated clients in the pool of agent } j}{\text{Total nb. of clients in the pool of agent } j}$$

We compute the selection variable Z_j for each period. In Equation 2 we include neither borrower fixed effects (μ_i) nor period fixed effect (ν_t) because we estimate the model per period and we have only one observation by borrower for each period. In addition, we exclude credit officer dummies (τ_j) because this variable is strongly correlated with Z_j (even perfectly correlated when we do not exclude borrower i to compute Z_j). In Figure B1, we report the estimated $\hat{\gamma}$ per period as well as confidence interval. We observe that

[33]We exclude borrower i in the computation of this ratio. But this modification does not change our results.

our selection variable is always positive and highly significant in all periods.

In a second step, we compute the inverse of the Mills ratio for each borrower i for each semester t as follows:

$$\hat{\lambda}_i = \frac{\phi(\hat{\Delta}\mathbf{X}_i + \hat{\gamma}Z_j + \hat{\eta}_s)}{\Phi(\hat{\Delta}\mathbf{X}_i + \hat{\gamma}Z_j + \hat{\eta}_s)} \quad (\forall t = 0, \ldots, T) \tag{3}$$

where $\Phi(.)$ is the cumulative normal distribution function and $\phi(.)$ the density normal function.

In a third step, we re-estimate the baseline model (Eq. 1) by adding the estimated inverse Mills ratio as covariates. Insofar as $\hat{\lambda}_i$ is computed for each period by running a probit model by period, we use a time-variant measure of the inverse of Mills ratio ($\hat{\lambda}_{it}$) allowing us to include usual borrower and time fixed-effects as follows:

$$y_{it} = \beta d_{it} + \Delta\mathbf{X}_{it} + \rho\hat{\lambda}_{it} + \mu_i + \nu_t + \eta_s + \tau_j + \epsilon_{it} \tag{4}$$

According to Wooldridge (1995), a simple test to detect sample selection consists of the t-statistics for ρ. Under the null hypothesis (absence of bias) ρ is statistically equal to 0. If $\rho \neq 0$, we need to correct for sample selection bias. In this case, we cannot use standard errors because $\hat{\lambda}_{it}$ is a generated variable. A simple way to get robust standard errors is by applying the bootstrapping method (we apply 500 replications).

Results from the three-step model are displayed in Table B1). In Panel A, we show models explaining the loan amount. Coefficients associated with the inverse of Mills ratio ($\hat{\lambda}$) are not statistically significant, indicating the absence of a sample selection bias. In addition, results regarding distance variables are unchanged in both econometric and economic terms. In a second step we correct for sample selection for model explaining the collateral-to-loan ratio. The inverse of the Mills ratio ($\hat{\lambda}$) is significant at 10%, indicating that the model is potentially subject to sample selection bias. In spite of it, our conclusions are not altered, and even reinforced, when we control for sample selection.

Figure B1: Estimated parameters of selection variable in selection equation

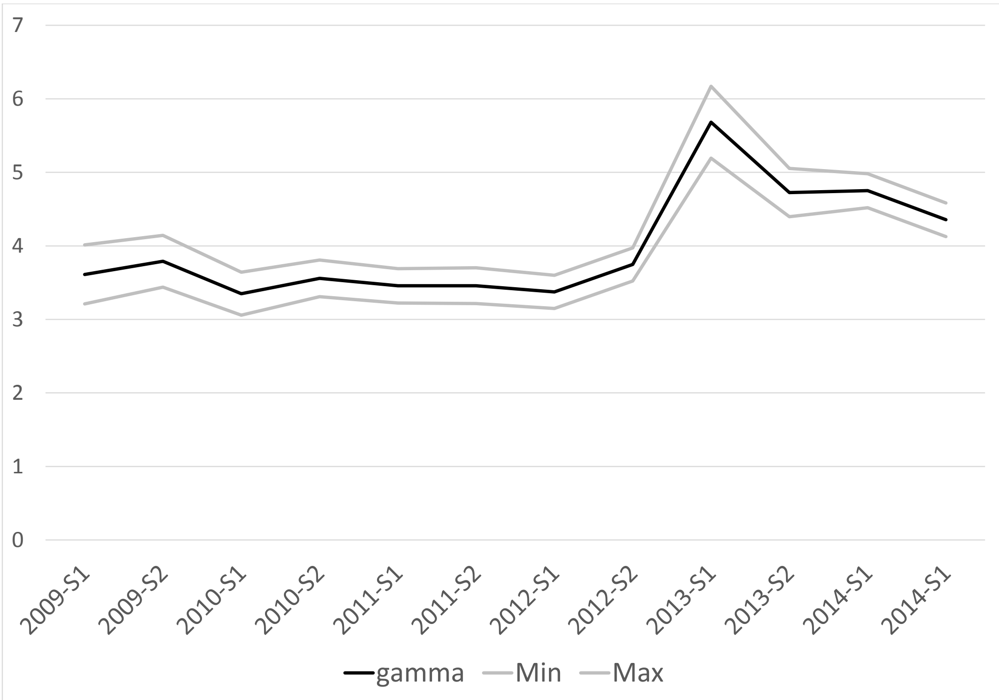

Table B1: Sample selection (Wooldridge procedure)

	Amount				Collateral ratio			
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	-0.054***				0.054			
	(-4.96)				(1.31)			
log(distance)		-0.046**				0.109***		
		(-2.00)				(3.44)		
Dummy			77.65*				-0.128***	
			(1.85)				(-6.87)	
Dist<500m				280.21*				-0.397***
				(1.65)				(-4.79)
500m<dist<1000m				188.02				-0.278***
				(0.77)				(-4.82)
1000m<dist<1500m				41.83				-0.121**
				(0.38)				(-1.99)
1500m<dist<2000m				20.40				-0.0520
				(0.44)				(-0.99)
$\hat{\lambda}$ (p-value)	NS	NS	NS	NS	p<0.10	p<0.10	p<0.10	p<0.10
Observations	32,374	32,374	32,374	32,374	32,374	32,374	32,374	32,374
Nb. Borrowers	14,834	14,834	14,834	14,834	14,834	14,834	14,834	14,834
R2 (within)	0.081	0.081	0.081	0.081	0.070	0.071	0.071	0.071

The dependent variable is the total loan amount in deflated USD in columns [1-4] and collateral ratio in columns [5-8]. The table shows estimations with four different measures of distance (for categories distance, the omitted category is distance above 2,000 meters). Three-step procedure developed by Wooldridge (1995) is employed (see Appendix B for details). Within estimator (at the borrower level) is used and period, industry and credit officer dummies as well as control variables (sales, employees, age, loan number) are included in all specifications. Standard errors are bootstrapped. t-statistics are reported in parentheses. Standardized coefficients are reported for continuous variables (distance, log and control variables) and usual coefficients for binary variables (dummy and categories). *, **, and *** indicate significance at the 10%, 5% and 1% respectively. For $\hat{\lambda}$, we report the p-value (NS: non significant at 10%, p<0.10)