FACULTY OF LAW, ECONOMICS AND FINANCE

CREA
Discussion
Paper
2018-04

Economics

Center for Research in Economics and Management
University of Luxembourg

Do banks and microfinance institutions compete? Microevidence from Madagascar

available online : http://wwwfr.uni.lu/recherche/fdef/crea/publications2/discussion_papers

Pierrick BARATON, CERDI, Université d'Auvergne, France
Florian LEON, CREA, University of Luxembourg

January, 2018

UNIVERSITÉ DU LUXEMBOURG

Do banks and microfinance institutions compete?
Microevidence from Madagascar[*]

Pierrick Baraton[†]

Florian Léon[‡]

Abstract

In recent years, both microfinance institutions (MFIs) and banks across the world have been converging towards the financing of small enterprises with high financing needs. This paper scrutinizes whether banks and MFIs compete each other as a result of recent transformations in both industries. In doing so, we study whether the loan strategy of a microfinance institution is shaped by the local presence of a bank. Specifically, we investigate whether bank proximity influences loan conditions provided by one of the largest microfinance institutions in Madagascar. We employ an original panel dataset of 32,374 loans granted to 14,834 borrowers over the period 2008-2014. We find that the closer a bank is located to a given MFI borrower, the larger the loan obtained and the less collateral required. These results are insensitive to several robustness tests for possible endogeneity of distance, sample selection issue, and alternative specifications. In addition, findings are stronger for larger and more established (older) firms in line with our hypothesis.

[†]CERDI, Université d'Auvergne (pierrick.baraton@gmail.com)

[‡]Corresponding author: CREA, University of Luxembourg (florian.leon@uni.lu)

1 Introduction

Limited access to formal credit is a major growth constraint for developing economies, especially for small firms (Beck and Demirguc-Kunt, 2006) and microenterprises (De Mel et al., 2008; McKenzie and Woodruff, 2008; Banerjee and Duflo, 2014). Since the 1970s, microfinance has emerged as a powerful tool to reach borrowers 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 money-lenders 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 side, a number of MFIs have adopted a "scaling-up" process and begun to develop their range of services to match the growing financial needs of small businesses. This "upscaling" strategy has resulted in MFIs targeting larger firms and more affluent borrowers. Several microfinance institutions have changed their legal status from NGO to shareholder-owned financial entity and, in extreme cases, some MFIs have transformed themselves into commercial banks (e.g., Prodem in Bolivia, Bandhan in India and Microcred in Madagascar). A recent but growing body of literature has investigated whether these changes have induced a drift in the historical mission of MFIs, without providing a clear answer (Armendáriz and Szafarz, 2011). 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 1990's 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

---

[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.

markets or have begun to compete for some of the same clients. If both financial intermediaries focus on two different types of borrowers, we could expect that changes in the banking market to have a limited impact on MFI' operations, and vice versa. However, if both types of lenders clients overlap, changes in either industry would likely to affect the other. As a consequence, 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 order to test our hypothesis, we investigate whether competitive pressure induced by the proximity of a commercial bank to MFI clients affects loans granted to them by one of the largest MFIs in Madagascar.[2] Madagascar is a perfect testing ground for our study question because some banks there have recently adopted "downscaling" strategies and because our partner MFI initiated its upscaling strategy few years ago. At least three commercial banks in Madagascar, including the largest one (*Bank of Africa*) and a former MFI (*Microcred*), have developed specific products for micro- and small firms. Meanwhile, our partner MFI provides individual loans with an upper-limit of $40,000. However, in reality, it continues to offer mainly micro-loans (in 2014, half of the loans it granted were below $500 and less than 2.5% of loans exceeded $5,000). Investigating whether our partner's business strategy has shifted by bank's presence gives us initial insight into how MFIs may react to banking development in a low-income countries.

We use a rich data set containing information on 32,374 loans (14,834 borrowers) from 2008 to 2014. 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] 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.

---

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

[3]Loan officers may discretionarily determine the loan amount and collateral requirements but cannot set interest rates and maturity, as discussed in the following.

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 identification strategy is based on the inclusion of borrower fixed effects.[4] 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 varying distance. We find that bank proximity improves loan conditions for MFI borrowers. All things being equal, firms in the vicinity of a bank can secure larger loans with less collateral from an MFI. The increase in loan amounts granted is significant in statistical and economic terms. For instance, firms located in a circle of less than 500 meters to the nearest bank obtain, on average, $200 more than firms located beyond two kilometers. Results are robust to a battery of sensitivity tests. We draw special attention to improve our identification by running alternative specifications (inclusion of municipality-period dummies that capture all time-variant unobservable factors occurring at the municipality level; focusing only on clients experienced a change in distance). In addition, we show that effect of distance is stronger for well-established (older) and large firms in line with our intuition.

This paper is directly related to a handful of articles that have scrutinized how the development and expansion of the banking and MFI industries are interlinked. These studies often employ cross-country investigation to relate how banking development affects financial performance of MFIs with mixed results (Hermes et al., 2011; Ahlin et al., 2011; Vanroose and D'Espallier, 2013). Recent works have focused on social performance and outreach. 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. Brown et al. (2016) have recently shed new light on the relationship between banks

---

[4]Due to the lack of data, we cannot use instrumental strategy to model bank's decision to locate.

4

and MFIs by adopting a micro-economic approach. They document that the openness of an MFI branch increases the percentage of banked households in South-Eastern Europe.

We extend this literature in several ways. To our knowledge, we are the first to explicitly focus on competition between banks and MFIs. Vanroose and D'Espallier (2013) and Cull et al. (2014) have provided indirect evidence on this topic. However, their results are rather mixed and are subject to alternative interpretations. In particular, observing that banking development favors outreach can be explained by the competition view (MFIs focus on clients not targeted by banks) but also by a complementary effect because MFIs benefit from banking development.[5] Bank expansion may reduce management costs for MFIs (Ahlin et al., 2011) and explain why MFIs are able to serve less affluent clients (Cull et al., 2014). Our analysis allows us to disentangle both effects because it is done at the borrower-level (and complementary effect certainly occurs at the institution or branch level that is controlled for). Our research can also be distinguished from Brown et al. (2016), who point out that financial inclusion goes hand-in-hand with microfinance development (deposit market); however the authors do not explore the presence of competition in the credit market. In this way, our work differs from theirs as it focuses on a different question in a different context (Madagascar instead of Eastern Europe). Our analysis also complements the literature by scrutinizing the *intensive margin* effect, while the existing literature focuses on the *extensive margin* effect. Competition induced by bank presence could result in two effects. On the one hand, the competitive pressure induced by the expansion of commercial banks could motivate MFIs to attract new clients, especially those who are not wooed by banks (*extensive margin*). On the other hand, competition may induce MFIs to offer better loan conditions in order to retain clients (*intensive margin*).[6] Our work indicates that bank competition impacts not only the pools of MFI clients (*extensive margin*) but also conditions faced by incumbent MFI clients (*intensive*

---

[5]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]One exception is Cull et al. (2014) that consider how bank presence impact average interest rates. 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 non-borrowers as in Brown et al. (2016). These data are not available in Madagascar rendering identification challenging. In 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.

margin). Nonetheless, these different contributions provide a similar general conclusion by indicating that banks and microfinance institutions are influenced by development of other intermediaries.

Our article also sheds light on the potential so-called "mission drift" of some MFIs caused by their upscaling strategies. Evidence of mission drift in the microfinance industry as a whole is limited (Cull et al., 2007; Mersland and Strøm, 2010) and previous works cannot really determine if offering larger loans has resulted in crowding out the poorest borrowers (Armendáriz and Szafarz, 2011).[7] Our results show that harsher competition induced by bank proximity motivates MFIs to offer better loan conditions to less opaque clients (i.e., the largest and oldest). Therefore, we could legitimately assume that MFIs are targeting "bankable" clients, contrary to their initial mission, and to the detriment of other more opaque clients. However, focusing on wealthier clients may not necessarily imply a "mission drift" on the part of MFIs if this strategy allows MFIs to extract higher rents in order to continue their initial mission (i.e., serving the poor). In the final section, we provide some indirect statistics indicating that the upscaling strategy implemented by our partner MFI may not lead to crowding out poor people. Specifically, we observe that the new clients served by our partner MFI tend to be more opaque over time. Put differently, our partner not only pursues its best clients but also continues to target small firms.

Finally, our paper also provides marginal contributions to two additional bodies of literature. First, we complement a scant body of literature investigating the determinants of loan terms in microfinance. Some papers have addressed this issue in the context of banking (Brick and Palia, 2007; Degryse et al., 2009). To our knowledge, only Behr et al. (2011) have seriously investigated it in microfinance. They study the implications of the lending relationship on collateral requirements and interest rates and document that the lending relationship alleviates collateral requirements. We confirm this finding and also show that the age of the firm has an unexpected impact: older firms tend to obtain smaller loans at a higher cost (higher level of collateral requirements). This may reflect the fact that older firms invest in riskier endeavors. Second, we also add slightly to a

---

[7]Average loan size used as a proxy of poverty outreach does not allow for distinction between inherent microfinance characteristics (such as progressive lending and cross-subzidation) and mission drift.

6

recent, but growing, body of literature on the role of distance in lending activities. Papers have mainly focused on the effect of distance in banking (e.g. Petersen and Rajan, 2002; Degryse and Ongena, 2005), or even if rarely, in microfinance (Pedrosa and Do, 2011; Presbitero and Rabellotti, 2014). In this paper, we show that distance can be employed to test competition between different intermediaries.

The remainder of the paper is organized as follows. Section 2 exposes the conceptual framework. Section 3 presents the data and variables and Section 4 the econometric methodology. Section 5 and Section 6 display the econometric results and robustness checks, respectively. Section 7 discusses the findings. The final section concludes.

2   Hypothesis tested

MFIs and banks in Madagascar have begun to offer comparable loans and are both converging towards the financing of a new target group: small enterprises with high financing needs. 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 credit contract terms granted by one of the largest MFI in Madagascar. We are particularly interested to know if competition induced by bank presence influences loan terms. A critical step consists in defining a good measure of competitive pressure induced by banks on MFIs.[8] We turn to the recent literature investigating the role of distance in banking (Petersen and Rajan, 2002; Degryse and Ongena, 2005). Banks may extract rent from their relative proximity to the 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

---

[8]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).

Ongena, 2005; Agarwal and Hauswald, 2010; Bellucci et al., 2013) confirmed that credit conditions deteriorate with bank-borrower proximity (even if credit access is improved).

In this paper, we focus on the role of distance between a MFI's borrower and the closest commercial bank. Banking literature also shows that borrowers located at the proximity of a competing bank are more likely to get better credit conditions (Degryse and Ongena, 2005; Bellucci et al., 2013). Indeed, switching costs are reduced by the distance to the closest competitor due to reduction in transportation and informational costs. In line with these arguments, we assume that sandwiching costs are increased with distance between the borrower and the closest alternative lender (here, a commercial bank). In other words, the probability of a MFI's client being wooed by a commercial bank increases with distance between her and the closest commercial bank.[9]

Assuming that borrowers located in the vicinity of a bank are more likely to get bank loans, we study whether the distance between the borrower and the closest bank affects loan conditions offered by our partner MFI. In absence of competition (banks and MFI operate on two different markets), the MFI will be insensitive to the entry of a bank and do 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 in order to retain its current clients and avoid a flight to bank. We therefore make the two mutually exclusive hypotheses:

Hypothesis 1

Banks and MFIs compete if a MFI's borrower located in the vicinity of a commercial bank obtain better credit conditions; In other words, credit conditions are (positively) related to the distance between the borrower and the closest commercial bank

Hypothesis 2

In absence of competition, the distance between a MFI's borrower and the closest bank does not affect loan conditions

---

[9]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.

In the following, we present an empirical framework developed to disentangle between both hypotheses.

3   Data and variables

3.1   Data

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.[10] Our partner shares with us its customer file. For each loan granted, we get 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 a half of clients.

The initial database comprised about 74,599 loans made to 35,472 borrowers. However, before selecting the final data set 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.[11] 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.[12] The final sample includes 14,834 borrowers representing 32,374 observations.

We complement our client database by identifying the location of every bank branch operating in Madagascar. As of December 31, 2014, we identified 154 bank branches operated by 12 commercial banks. We refer to the register of the Malagasy National Bank[13] to identify all of the commercial banks operating. 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

---

[10]MFIs in Madagascar are classified in three categories (1, 2, and 3). Category 3 is made up of the largest MFIs in Madagascar, including our partner.

[11]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).

[12]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%).

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

the internet. We complement our database with *in situ* visits to get 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.

3.2 Variables

3.2.1 Loan contracts

Data on credit loan terms are used to compute our outcome variables. Four different loan conditions are provided by our partner MFI: 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. Our partner provides us with the value of the collateral. 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 amount of material assets that they guarantee for the total loan amount. Better loan conditions therefore imply not only a limited ratio of collateral to loan but also a limited percentage of material guarantee to collateral. We compute this ratio as our second measure of collateral requirements. The descriptive statistics, reported in Table 1, document that loan amount represents $1,129 USD 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.

Our partner gives us information on four loan contract characteristics: loan amount, collateral requirements, interest rate, and maturity.[14] However, in the empirical analysis,

---

[14]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.

we focus exclusively on loan amount and collateral requirements. We do not exploit data on interest rates and maturity due to a lack of variation. 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%).[15] Loan amount and collateral requirements capture two different aspects. Loan amount may proxy availability of credit in a context where borrowers cannot access to complete funds required. Collateral requirements is more related to price in a context of fixed interest rates (Fisman et al., 2017).

Table 1: Descriptive Statistics

Variable	Obs.	Mean	Std. Dev.	Min	Max	CV[†]
Loan contract						
Amount[‡]	32,373	1,129	2,207	22.9	40,076	1.95
Interest rate	32,373	0.126	0.027	0.060	0.220	0.21
Maturity	32,373	389.0	49.2	88	1,095	0.13
Collateral/Amount	32,373	2.851	1.272	0.088	10.84	0.45
Securities/Collateral	32,373	0.545	0.162	0	1	0.30
Individual characteristics						
Sales[‡]	32,373	1762.6	2792	0.4	24,555	1.58
Employees	32,373	2.257	2.498	0	32	1.11
Age (firm)	32,373	8.608	6.371	0	44	0.74
Borrower-lender relationship						
Number	32,373	2.901	2.487	1	20	0.86
Year	32,373	2.254	3.043	0	19	1.35
Distance						
- in meters	32,373	2,403	5,529	5.53	88,604	2.30
- in log	32,373	7.027	1.122	1.710	11.39	0.16
- Dummy	32,373	0.156	0.361	0	1	2.31
- Dist<500m	32,373	0.202	0.401	0	1	2.31
- 500m<dist<1000m	32,373	0.258	0.438	0	1	1.70
- 1000m<dist<1500m	32,373	0.173	0.378	0	1	2.19
- 1500m<dist<2000m	32,373	0.103	0.304	0	1	2.19
- dist>2000m	32,373	0.264	0.441	0	1	1.67

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

3.2.2 Distance

A crucial step consists in building a measure of competition intensity. As expressed above, we employ the distance between a borrower and the closest bank (as the crow flies). The intuition is based on the idea that a client in the vicinity of a bank is more likely to be captured by a bank than is a remote client due to transportation costs and information asymmetry. To compute our interest variable, namely the distance between a given borrower and the closest bank, we used the precise location of borrowers provided by

---

[15]In a previous version, we run models explaining (real) interest rates and maturity. Inclusion of period-dummies capture all variability in interest rates due to deflation. Model explaining maturity fail to provide consistent results for both interest and control variables.

our partner MFI and the location of banks that we had 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 (Degryse and Ongena, 2005; Agarwal and Hauswald, 2010) 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.[16] 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 (63 to 154). As a result, using distance in 2014 to proxy distance in 2008 can be misleading. We therefore compute 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. Obtaining a time-variant distance variable has the additional advantage of allowing us to provide better identification as explained below.

In addition to linear distance, we employ alternative measures of distance. In spite of the advantage of transparency, the linear measure is perhaps not the most fitted. A 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 and almost not existent for the latter. To consider this point, we run regressions by using alternative measures of distance. First, we employ the logarithm of distance. In addition, we consider categorical dummies. The first dummy takes value 1 if a borrower is located at less than 500 meters of the closest bank branch, and 0 otherwise. The second dummy takes value 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 dummy for distance above 2000 meters. Finally, for each client we create a treatment dummy equals to one in period t if distance observed in t is lower than that observed in the initial period.[17] This dummy captures

---

[16]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.

[17]For instance, let us assume that the entry of a new bank in 2011 reduces the distance between client

how loan dynamic is shaped by a reduction in distance.

3.2.3 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 the size (Berger et al., 2001) and age of the firm (Hyytinen and Pajarinen, 2008). We therefore add two measures of size (total sales and number of employees) and the age of the firm. 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 below) implies that we control for lending relationship. Second, microfinance lending technologies are based on dynamic incentives (Armendáriz and Morduch, 2010). We therefore expect 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.[18]. 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 USD. On average, borrowers have had a relationship with our partner MFI for two years.

A and her closest bank. The dummy will take value 0 from 2008 to 2011 and then the value of 1. If this client did not experienced a reduction of distance, the dummy value is zero from 2008 to 2014. We expect an opposite sign that those observed for distance and log of distance.

[18]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

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:

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

where y_{it} is the dependent variable (total amount or collateral requirements) for borrower i in period t; 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. 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).

We further employ one extension of Eq. 1 by adding other characteristics of loan terms as control variables as follows:

$$y_{it} = \beta d_{it} + \Delta \mathbf{X}_{it} + \gamma \tilde{y}_{it} + \mu_i + \nu_t + \eta_s + \tau_j + \epsilon_{it} \tag{2}$$

Specifically, $\tilde{y}_{it}$ includes 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. The rational for this inclusion is that loan terms are jointly determined (Brick and Palia, 2007). However, including other loan terms might induce an endogeneity problem due to reverse causation and unobserved third factors. As a result, our baseline model is run without other loan terms and their inclusion is viewed as a robustness check.

As explained above, we consider only two dependent variables (amount and collateral requirements) due to the lack of variability of interest rate and maturity. We expect that firms close to banks obtain larger loan amount in line with Hypothesis 1 ($\beta < 0$ when we consider linear or log of distance and $\beta > 0$ when we consider treatment dummy and categorical classes). Competition could also influence the collateral pledged by the borrower. Contrary to a common belief, physical collateral is used in individual lending

in microfinance (Behr et al., 2011). Collateral is used by MFIs to handle information asymmetry, either as a screening device or as an incentive for the borrower to behave properly. We believe that lenders may choose to lower the collateral-to-loan ratio in order to retain current borrowers or attract new ones, because this allows clients either to obtain a larger loan or to decrease the consequences in case of default. As a consequence, Hypothesis 1 is validated if distance affects collateral requirements. In other words, we expect that $\beta > 0$ when we consider linear or log of distance and $\beta < 0$ when we consider treatment dummy and categories of distance. Hypothesis 2 is confirmed if coefficients associated with distance are not statistically significant in different specifications (using size of the loan or collateral requirements).

A challenge is to go beyond a simple correlation and provide causal interpretation. One might raise concerns about the fact that credit conditions as well as location (distance) can be affected by unobserved characteristics. For instance, a dynamic entrepreneur may be able to obtain favorable credit conditions and be more likely to be located in a dynamic area (close to bank). As a result, the coefficient associated with distance will only reflect the relationship between unobserved individual characteristics and loan contract terms. We cannot employ instrumental variable strategy (or a natural experiment) due to the lack of data. A relevant instrument should model bank's decision to locate. In doing so, we must employ time-varying variables at sub-regional level.[19] However, such information is not available in Madagascar.

To tackle the identification issue, we include individual fixed effects. 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 distance. 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

---

[19]Madagascar is divided into four main subdivisions: Regions or "faritany" (22), counties or "fivondronana" (114), municipalities or "kaominina" (1,395) and district or "fokontany" (17,544). To be precise, the national statistical institute (INSTAT) provides data on municipalities and district for one year (2009) over our period of reference.

role in microfinance lending technologies.[20]

The inclusion of borrower fixed effects has an additional advantage. If MFI's branches are located in the vicinity of bank branches, 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)[21], 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.[22]

5 Baseline result

We test whether MFIs and banks compete by scrutinizing if 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. As explained in Section 3, we cannot employ interest rate and maturity due to a lack of variability. Hypothesis 1 (existence of competition) will be confirmed if borrowers in the vicinity of a bank should obtain larger loans and pledge less collateral. At the opposite, in absence of competition, loan contract terms offered by MFIs should not be affected by bank presence (Hypothesis 2).

5.1 Loan amount

We first report results regarding the effect of distance on the size of loan in Table 2. The hypothesis of competition between MFIs and banks is validated if the coefficients associated with continuous measures of distance are negative (columns [1-4]) and positive for discrete measures of distance (columns [5-8]).

In all specifications, coefficients associated with distance have the expected sign and

---

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

[21]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.

[22]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 office.

Table 2: Determinants of loan amount

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	-0.053***	-0.051***						
	(-3.29)	(-3.43)						
log(distance)			-0.045*	-0.038*				
			(-1.91)	(-1.85)				
Dummy					78.73**	62.36**		
					(2.45)	(2.17)		
Dist<500m							277.96**	220.27*
							(1.97)	(1.73)
500m<d<1000m							187.15*	110.03
							(1.81)	(1.18)
1000m<d<1500m							-42.76	-51.32
							(-0.48)	(-0.63)
1500m<d<2000m							-21.04	-12.93
							(-0.33)	(-0.23)
Sales	0.091***	0.104***	0.091***	0.104***	0.091***	0.104***	0.091***	0.104***
	(5.48)	(6.91)	(5.48)	(6.91)	(5.48)	(6.91)	(5.48)	(6.91)
Employees	0.053***	0.057***	0.053***	0.057***	0.053***	0.057***	0.053***	0.057***
	(3.13)	(3.92)	(3.12)	(3.92)	(3.11)	(3.91)	(3.10)	(3.90)
Firm age	-0.429***	-0.401***	-0.438***	-0.407***	-0.445***	-0.411***	-0.437***	-0.404***
	(-3.39)	(-3.58)	(-3.45)	(-3.60)	(-3.53)	(-3.68)	(-3.42)	(-3.56)
Loan number	0.369***	0.294***	0.0370***	0.295***	0.370***	0.295***	0.370***	0.295***
	(11.99)	(10.26)	(12.01)	(10.29)	(12.03)	(10.30)	(12.05)	(10.30)
Interest rate		-0.182*		-0.183*		-0.182*		-0.182*
		(-1.92)		(-1.92)		(-1.92)		(-1.91)
Maturity		0.257***		0.257***		0.257***		0.257***
		(13.54)		(13.55)		(13.54)		(13.56)
Collateral/Loan		-0.031***		-0.031***		-0.031***		-0.031***
		(-4.41)		(-4.38)		(-4.38)		(-4.34)
Observations	32374	32374	32374	32374	32374	32374	32374	32374
Nb. Borrowers	14,834	14,834	14,834	14,834	14,834	14,834	14,834	14,834
R2 (within)	0.079	0.270	0.079	0.270	0.079	0.270	0.079	0.270

The dependent variable is the total loan amount (in deflated USD). The table shows estimations with four different measures of distance (for categories distance, the omitted category is distance above 2,000 meters). Within estimator 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 and control variables) and usual coefficients for binary variables (dummy and categories). *, **, and *** indicate significance at the 10%, 5% and 1% respectively

are statistically significant at the usual thresholds, in line with Hypothesis 1. Distance is not only statistical significant but also economically significant: an one standard deviation increase in distance induces a reduction of the loan amount by more than \$110 (mean of loan amount equals \$1,129).[23] Using alternative measure of distance provides similar findings. All things being equal, the fact to experience a decrease in distance raises the loan amount by more than \$60 in the most conservative regression (column [6]). Results reported in columns [7-8] provide evidence that the effect of distance is not homogenous. They indicate that only clients located in a circle of less than 500 meters of a bank obtain larger loans. The observed impact is far from anecdotal: the fact to be located in a circle

---

[23]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: $\sigma_y * \hat{\beta}_x$ 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 $\sigma_y * \hat{\beta}_d = 2207 * 0.051 = 112$

of less than 500 meters increases the loan amount by more than \$200 (comparatively to a previous situation with a distance above 2000 meters, the omitted category).

The statistical effect of distance is robust to the inclusion of other credit conditions but its economic effect is reduced by around one fifth in different specifications. This may reflect the fact that distance alter other terms of loan contracts (as shown below). 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). Regarding control variables, we note that larger firms get access to larger loans, while older firms do not. In addition, borrowers benefit from longer and more intense relationships 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.

5.2 Collateral requirements

We then study the determinants of collateral requirements by considering the ratio of collateral value to total loan value. Results are displayed in Table 3. According to Hypothesis 1, we expect that the ratio of collateral to loan is lower for firms in the vicinity of a bank. Put differently, we expect that $\beta > 0$ for continuous measures of distance (columns [1-4]) and $\beta < 0$ for discrete measures of distance (columns [5-8]). Coefficients associated with linear measure of distance (columns [1-2]) are positive as expected but not statistically significant. However, coefficients are significant at the 1% level when we consider the natural logarithm of distance (columns [3-4]). 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 in models with treatment dummy (columns [5-6]) and distance categories (columns [7-8]). These results tend to confirm Hypothesis 1 and reject Hypothesis 2.

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%. Amplitude given by discrete measures are rather similar.

Table 3: Determinants of collateral ratio

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	0.046 (0.67)	0.046 (067)						
log(distance)			0.105*** (2.74)	0.100*** (2.62)				
Dummy					-0.126*** (-4.00)	-0.121*** (-3.84)		
Dist<500m							-0.386*** (-3.53)	-0.368*** (-3.36)
500m<dist<1000m							-0.268*** (-2.80)	-0.255*** (-2.67)
1000m<dist<1500m							-0.116 (-1.29)	-0.119 (-1.34)
1500m<dist<2000m							-0.050 (-0.60)	-0.052 (-0.62)
Sales	0.001 (0.10)	0.011 (0.89)	0.001 (0.11)	0.011 (0.89)	0.001 (0.11)	0.011 (0.89)	0.002 (0.13)	0.011 (0.90)
Employees	-0.004 (-0.29)	0.001 (0.09)	-0.004 (-0.28)	0.001 (0.10)	-0.003 (-0.24)	0.002 (0.14)	-0.004 (-0.26)	0.002 (0.11)
Firm age	0.436** (2.04)	0.375* (1.75)	0.471** (2.21)	0.410* (1.92)	0.499** (2.33)	0.436** (2.03)	0.477** (2.25)	0.416** (1.95)
Loan number	-0.485*** (9.56)	-0.443*** (-8.60)	0.486*** (-9.56)	-0.443*** (-8.61)	-0.488*** (-9.63)	-0.446*** (-8.67)	-0.487*** (-9.59)	-0.445*** (-8.63)
Interest rate		0.101 (1.53)		0.103 (1.56)		0.102 (1.54)		0.102 (1.54)
Maturity		-0.009 (-0.62)		-0.008 (-0.61)		-0.008 (-0.62)		-0.008 (-0.61)
Amount		-0.112*** (-4.65)		-0.111*** (-4.62)		-0.111*** (-4.62)		-0.110*** (-4.57)
Observations	32374	32374	32374	32374	32374	32374	32374	32374
Nb. Borrowers	14,834	14,834	14,834	14,834	14,834	14,834	14,834	14,834
R2 (within)	0.030	0.035	0.030	0.035	0.030	0.035	0.030	0.035

The dependent variable is the ratio of collateral to loan value. The table shows estimations with four different measures of distance (for categories distance, the omitted category is distance above 2,000 meters). Within estimator 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 and control variables) and usual coefficients for binary variables (dummy and categories). *, **, and *** indicate significance at the 10%, 5% and 1% respectively

The fact to experience a reduction in distance reduces this ratio by 12% (columns [5-6]). Results from columns [7-8] indicate that clients located in a circle of 500 meters of a bank have a collateral ratio that is reduced by 40 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 insofar as the average value of collateral ratio equals 280%.

Regarding control variables, our model provides interesting results. First, we point out that collateral requirements and loan amount are negatively correlated. In addition, we show that coefficients associated with the number of loans granted by our partner is negative and statistically significant at the 1% level. This finding indicates that the lending relationship relaxes collateral requirements, confirming findings obtained by Behr et al. (2011) in Mozambique.

The presence of a bank may not only affect the quantity of collateral, but also the "quality" of collateral. We therefore use an alternative measure of collateral using the composition of collateral. We employ the ratio of material guarantees to total guarantees. Material guarantees are more costly for borrowers than personal guarantees (i.e., third-party). According to our hypothesis, we expect that the ratio of costlier (material) guarantees is positively related with distance, indicating that firms close to banks obtain better credit conditions.

Table 4: Determinants of the composition of collateral

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	-0.031	-0.037						
	(-0.68)	(-0.82)						
log(distance)			0.005	0.0045				
			(1.12)	(0.97)				
Dummy					-0.0149***	-0.0142***		
					(3.71)	(3.56)		
Dist<500m							-0.0276**	-0.0252**
							(-2.13)	(-1.96)
500m<dist<1000m							-0.0296***	-0.0280**
							(-2.69)	(-2.56)
1000m<dist<1500m							-0.0097	-0.0101
							(0.96)	(-1.01)
1500m<dist<2000m							-0.0051	-0.0053
							(-0.56)	(-0.58)
Sales	-0.046***	-0.036***	-0.046***	-0.036***	-0.046***	-0.036***	0.002	0.011
	(-3.40)	(-2.62)	(-3.40)	(-2.62)	(-3.40)	(-2.62)	(0.13)	(0.90)
Employees	0.006	0.012	0.006	0.012	0.006	0.012	-0.004	0.002
	(0.47)	(0.96)	(0.47)	(0.96)	(0.47)	(0.96)	(-0.26)	(0.11)
Firm age	0.407*	0.352*	0.407*	0.352*	0.407*	0.352*	0.477**	0.416**
	(1.91)	(1.66)	(1.91)	(1.66)	(1.91)	(1.66)	(2.25)	(1.95)
Loan number	-0.769***	-0.725***	-0.769***	-0.725***	-0.769***	-0.725***	-0.487***	-0.445***
	(-16.93)	(-16.00)	(-16.93)	(-16.00)	(-16.93)	(-16.00)	(-9.59)	(-8.63)
Interest rate		-0.119***		-0.119***		-0.119***		0.102
		(-6.50)		(-6.50)		(-6.50)		(1.54)
Maturity		0.028		0.028		0.028		-0.008
		(0.48)		(0.48)		(0.48)		(-0.61)
Amount		0.001		0.001		0.001		-0.110***
		(0.10)		(0.10)		(0.10)		(-4.57)
Observations	32374	32374	32374	32374	32374	32374	32374	32374
Nb. Borrowers	14,834	14,834	14,834	14,834	14,834	14,834	14,834	14,834
R2 (within)	0.070	0.075	0.071	0.076	0.071	0.076	0.071	0.075

The dependent variable is the ratio of collateral to loan value. The table shows estimations with four different measures of distance (for categories distance, the omitted category is distance above 2,000 meters). Within estimator 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 and control variables) and usual coefficients for binary variables (dummy and categories). *, **, and *** indicate significance at the 10%, 5% and 1% respectively

Econometric results using the composition of collateral are reported in Table 4. 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 for collateral ratio, we show that coefficients associate to dummy for treated and for two first categories

20

are statistically significant. In economic terms, however, the impact of distance is rather limited. For instance, the fact to experience 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%).

6 Robustness checks

In the baseline model, we present results with different measures of distance and different specifications (without/with other contract terms). In the following, we provide additional robustness checks to be ascertain that our findings are robust to alternative specification.[24]

6.1 Identification issue

A critical issue relies on identification insofar as borrower's and bank's locations are not random. Unfortunately, as explained abvoe, we cannot use an instrumental strategy due to lack of data available to model bank's branch location decision. Nonetheless, to reinforce our identification strategy we present additional tests.

6.1.1 Controlling for time-invariant unobserved factors

We firstly add municipality-semester dummies. Our identification strategy is based on the inclusion of borrower fixed effects that allows 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. In particular, since 2009 Madagascar has been particularly embroiled in political and economic turmoil. This crisis may force some SMEs, that usually resort to banks, to turn

---

[24]We do not include the determinants of the type of collateral because (i) results of robustness checks are in line with specifications explaining the collateral ratio; (ii) the economic impact of distance is somewhat limited for this variable. In addition, we also run additional robustness checks, unreported for sake of brevity (but available upon request): (i) we change the sample by excluding borrowers located in Tananarive; (ii) we run models without credit officer dummies; (iii) we run the model without lending relationship variable insofar as this variable can be endogeneous; and, (iv) we modify method adopted to correct standard errors by running models without correction, with standard errors clustered at the credit officer level, at the municipality-period level, and at the industry-period level. In all cases, our conclusions remain valid.

Table 5: Inclusion of municipality-semester dummies

Panel A: Determinants of amount of loan

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	-0.120** (-2.08)	-0.190** (-2.30)						
Log(distance)			-0.032 (-1.12)	-0.035 (-1.46)				
Dummy					97.58** (2.28)	88.69** (2.25)		
dist<500m							331.23** (2.01)	270.00* (1.94)
500m<dist<1000m							229.05* (1.70)	173.70 (1.37)
1000m<dist<1500m							20.21 (0.19)	32.71 (0.34)
1500m<dist<2000m							32.99 (0.44)	41.69 (0.61)
Obs.	32,374	32,375	32,376	32,377	32,378	32,379	32,380	32,381
Nb. Borrowers	14,834	14,835	14,836	14,837	14,838	14,839	14,840	14,841
R^2 (within)	0.170	0.338	0.170	0.338	0.170	0.338	0.170	0.338

Panel B: Determinants of collateral ratio

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	0.112 (0.46)	0.106 (0.43)						
Log(distance)			0.086* (1.66)	0.083† (1.61)				
Dummy					-0.052 (-1.29)	-0.049 (-1.15)		
dist<500m							-0.317** (-2.11)	-0.294* (-1.95)
500m<dist<1000m							-0.220* (-1.67)	-0.203* (-1.64)
1000m<dist<1500m							-0.121 (-1.08)	-0.120 (-1.07)
1500m<dist<2000m							-0.073 (-0.78)	-0.071 (-0.75)
Obs.	32,374	32,375	32,376	32,377	32,378	32,379	32,380	32,381
Nb. Borrowers	14,834	14,835	14,836	14,837	14,838	14,839	14,840	14,841
R^2 (within)	0.111	0.116	0.111	0.117	0.111	0.116	0.112	0.117

The dependent variables are the total amount of loan (Panel A) and the collateral-to-loan ratio (Panel B). Within estimator is used. The list of control variables includes total sales (in constant US$), the number of employees, the firm age, the number of loan between the firm and our partner IMF). In columns [2], [4], [6] and [8], we also add other contract terms. In all specifications, we include town-semester dummies, credit officer dummies and industry-dummies. Standard errors are clustered at the borrower level and 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 15%, 10%, 5% and 1% respectively.

to MFI in order to meet unexpected expenses or compensate sales decreasing. If bank clients are mainly located next to bank branches, then new clients turning to MFIs may enjoy better loan conditions from MFI because they are certainly less "opaque". In these robustness checks, we therefore add municipality-semester dummies.[25] These dummies enable us to control for any shock that occurs at the municipality level over time, such as the economic crisis that started in 2009 and that may have a different impact depend-

---

[25]Considering fokontany is impossible due to the large number of fokontanies (17,544). In addition, fokontanies differ greatly in terms of size and number of population (and clients); in some of them, we have only one client. Using instrumental variable approach is impossible in our study due to the limited data available.

ing on the municipality.[26] This approach has an additional advantage by controlling for any local shocks, beyond economic deterioration. In particular, as explained above, a change in the number of competitors (MFIs or banks) could affect credit conditions as documented in spatial models (see Degryse and Ongena, 2005). By adding municipality-semester dummies, we also control for this possible bias. 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 with this approach (explaining why we do not employ it in our baseline) is the number of parameters to estimate (more than 2,000 parameters). Results on models including municipality-semester dummies are displayed in Table 5. Coefficients associated with distance, at 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 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).

6.1.2 Focusing on borrowers that experienced a reduction in distance

Our econometric intuition is based on the idea that borrowers experienced a reduction in distance will obtain favourable loan conditions. In the baseline model, we consider borrowers who experienced a reduction in distance and those who do not. To our point of view, it is better to consider both groups of borrowers. In the spirit of impact evaluation model, borrowers without change in distance can be considered as control group and other borrowers as treatment group (this analogy is particularly suited when we consider the treated dummy). Nonetheless, our findings should be robust if we consider only the group of treated. We therefore rerun our baseline models by excluding borrowers that did not experience a decrease in distance. Results reported in Table 6 confirm our main findings, in spite of a sharp decrease in the number of observations.

---

[26]We also consider an alternative way to control for this issue. We keep only old borrowers defined as firms that have been clients to our partner since 2010. As shown in the Appendix (Table A1), the main impact is unchanged, and even reinforced.

Table 6: Keeping only borrowers that experienced a change in distance

Panel A: Determinants of amount of loan

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	-0.044***	-0.043***						
	(-3.28)	(-3.43)						
log(distance)			-0.037*	-0.032*				
			(-1.83)	(-1.86)				
Dummy					98.53***	83.70***		
					(2.73)	(2.58)		
500m<d<1000m							290.88*	240.61*
							(1.91)	(1.74)
1000m<d<1500m							202.34*	126.05
							(1.91)	(1.30)
1500m<d<2000m							-34.43	-35.40
							(-0.36)	(-0.40)
Loan contract	No	Yes	No	Yes	No	Yes	No	Yes
Obs.	9154	9154	9154	9154	9154	9154	9154	9154
Nb. Borrowers	2187	2187	2187	2187	2187	2187	2187	2187
R2 (within)	0.0890	0.288	0.0888	0.287	0.0887	0.287	0.0905	0.288

Panel B: Determinants of collateral ratio

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	0.036	0.031						
	(0.56)	(0.49)						
log(distance)			0.078**	0.074**				
			(2.17)	(2.07)				
Dummy					-0.146***	-0.140***		
					(-4.02)	(-3.88)		
500m<d<1000m							-0.340***	-0.324***
							(-2.95)	(-2.80)
1000m<d<1500m							-0.242**	-0.231**
							(-2.46)	(-2.05)
1500m<d<2000m							-0.099	-0.102
							(-1.09)	(-0.58)
Loan contract	No	Yes	No	Yes	No	Yes	No	Yes
Obs.	9154	9154	9154	9154	9154	9154	9154	9154
Nb. Borrowers	2187	2187	2187	2187	2187	2187	2187	2187
R2 (within)	0.0306	0.0353	0.0317	0.0363	0.0324	0.0370	0.0327	0.0371

The dependent variables are the total amount of loan (Panel A) and the collateral-to-loan ratio (Panel B). Within estimator is used. The list of control variables includes total sales (in constant US$), the number of employees, the firm age, the number of loan between the firm and our partner IMF). In columns [2], [4], [6] and [8], we also add other contract terms. In all specifications, we include semester dummies, credit officer dummies and industry-dummies. Standard errors are clustered at the borrower level and 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 15%, 10%, 5% and 1% respectively.

6.1.3 Transparent vs. opaque borrowers

Finally we test to assess whether our story based on competition between banks and MFIs is validated. We argue that transparent firms are more likely to be captured by formal banks than opaque firms. To test this intuition we run the econometric model on two sub-samples of opaque and transparent firms. To distinguish between groups, we use two usual proxies of transparency, namely the size of the firm (Berger et al., 2001) and the age of the firm (Hyytinen and Pajarinen, 2008). A firm is classified as an opaque one if its size (assessed by total sales or the number of employees) or its age is below the

median. We then regress Equation 1 on different sub-samples.[27] We expect the impact of distance is amplified for older and larger firms. Table 7 displays results regarding loan amount determinants on different sub-samples of old/young firms and large/small firms. Samples are divided according to the median age (Panel A), total sales (Panel B), and the number of employees (Panel C). The first four columns present estimations for the sub-sample of opaque firms (value of age, sales and employees below the median) and the last four columns present those for transparent firms. In a word, while results are not always statistically significant, the economic impact of distance is stronger for transparent firms. In details, we show in Panel A that older firms benefit more from a bank in the vicinity than younger firms. Coefficients associated with continuous measures of distance and treatment dummy are only statistically significant for old firms. Results using firm size to distinguish between opaque and transparent firms are less clear-cut in statistical terms (Panels B and C). Nonetheless, the impact of distance in terms of amplitude is stronger for larger firms. In particular when we consider treatment dummy or categories of distance, we show that large firms benefit more from a proximity with a bank. We extend our analysis by scrutinizing the impact of distance on the ratio of collateral to total loan for both groups of firms. Results by sub-groups of firms, displayed in Table 8, are more clear-cut than those obtained for loan amount and give some support to our intuition: transparent firms benefit more from a proximity of a bank. Coefficients associated with distance variables are often statistically significant only for the group of old or large firms. In addition, the economic impact of distance is stronger for the most transparent firms.

---

[27]Considering quartile instead of median does not alter our findings.

Table 7: Determinants of loan amount, sub-sample analysis

Panel A: Age of the firm

	Young firms (age<Median)				Old firms (age>Median)			
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	-0.025				-0.044**			
	(-0.93)				(-3.04)			
Log(distance)		-0.037				-0.040†		
		(-1.11)				(-1.49)		
Dummy			3.67				116.55**	
			(0.12)				(2.32)	
dist<500m				195.43				266.15
				(1.11)				(1.26)
500m<dist<1000m				51.42				159.22
				(0.40)				(0.98)
1000m<dist<1500m				-68.73				-129.01
				(-0.57)				(-0.91)
1500m<dist<2000m				-4.47				-8.45
				(-0.09)				(-0.09)
Control								
- Borrowers	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
- Loan contract	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dummies								
- Industry	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
- Credit officer	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
- Period	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	16,824	16,824	16,824	16,824	15,543	15,543	15,543	15,543
Nb. Borrowers	9,422	9,422	9,422	9,422	6,809	6,809	6,809	6,809
R2 (within)	0.324	0.324	0.324	0.324	0.266	0.266	0.266	0.266

Panel B: Total sales

	Small firms (Sales<Median)				Large firms (Sales>Median)			
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	-0.015				-0.043***			
	(-0.91)				(-2.99)			
Log(distance)		0.011				-0.034†		
		(0.45)				(-1.59)		
Dummy			44.35**				71.00	
			(2.22)				(1.41)	
dist<500m				-20.03				312.15†
				(-0.38)				(1.49)
500m<dist<1000m				28.78				192.78
				(0.68)				(1.26)
1000m<dist<1500m				26.22				-171.26
				(0.61)				(-1.25)
1500m<dist<2000m				44.12				-57.84
				(1.17)				(-0.62)
Control								
- Borrowers	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
- Loan contract	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dummies								
- Industry	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
- Credit officer	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
- Period	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	16,192	16,192	16,192	16,192	16,175	16,175	16,175	16,175
Nb. Borrowers	9,770	9,770	9,770	9,770	7,283	7,283	7,283	7,283
R^2 (within)	0.231	0.231	0.231	0.231	0.310	0.310	0.310	0.311

Panel C: Number of employees

	Small firms (Nb. Empl.<Median)				Large firms (Nb. Empl.>Median)			
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	-0.055**				-0.050***			
	(-2.14)				(-2.66)			
Log(distance)		-0.024				-0.054		
		(-1.17)				(-1.42)		
Dummy			28.42*				150.20†	
			(1.65)				(1.58)	
dist<500m				68.99				633.17†
				(0.87)				(1.59)
500m<dist<1000m				28.03				51.62
				(1.46)				(0.20)
1000m<dist<1500m				42.02				-173.52
				(0.61)				(-0.72)
1500m<dist<2000m				-10.26				48.07
				(-0.21)				(0.31)
Control								
- Borrowers	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
- Loan contract	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dummies								
- Industry	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
- Credit officer	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
- Period	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	23,477	23,477	23,477	23,477	8,890	8,890	8,890	8,890
Nb. Borrowers	12,488	12,488	12,488	12,488	4,623	4,623	4,623	4,623
R^2 (within)	0.247	0.246	0.246	0.247	0.319	0.319	0.319	0.319

The dependent variables is total amount of loan. The sample is divided according to the age of the firm (panel A), total sales (panel B) and the number of employees (panel C). Within estimator is used and period- and industry-dummies as well as control variables (sales, number of employees, and firm age) are included in all specifications. Other loan terms (maturity, rate and collateral ratio) are included as controls in columns [3], [4], [7] and [8]. Standard errors are clustered at the borrower level. t-statistics are reported in parentheses. Standardized coefficients are reported. †, *, **, and *** indicate significance at the 15%, 10%, 5% and 1% respectively.

Table 8: Determinants of collateral ratio, sub-sample analysis

Panel A: Age of the firm

	Young firms (age<Median)				Old firms (age>Median)			
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	-0.191**				0.059			
	(-2.16)				(0.75)			
Log(distance)		-0.043				0.181***		
		(-0.67)				(3.49)		
Dummy			-0.064				-0.129**	
			(-1.40)				(-2.70)	
dist<500m				-0.032				-0.648***
				(-0.18)				(-4.03)
500m<dist<1000m				-0.059				-0.407***
				(-0.39)				(-2.84)
1000m<dist<1500m				-0.114				-0.241*
				(-0.82)				(-1.81)
1500m<dist<2000m				-0.172				-0.057
				(-1.30)				(-0.46)
Control								
- Borrowers	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
- Loan contract	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dummies								
- Industry	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
- Credit officer	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
- Period	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	16,824	16,824	16,824	16,824	15,543	15,543	15,543	15,543
Nb. Borrowers	9,422	9,422	9,422	9,422	6,809	6,809	6,809	6,809
R2 (within)	0.067	0.066	0.067	0.067	0.031	0.033	0.031	0.034

Panel B: Total sales

	Small firms (Sales<Median)				Large firms (Sales>Median)			
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	0.027				0.058			
	(0.29)				(0.64)			
Log(distance)		0.105				0.104**		
		(1.30)				(2.06)		
Dummy			-0.092*				-0.122***	
			(-1.73)				(-2.84)	
dist<500m				-0.462***				-0.380**
				(-2.63)				(-2.46)
500m<dist<1000m				-0.275*				-0.295**
				(-1.95)				(-2.19)
1000m<dist<1500m				-0.307**				-0.042
				(-2.30)				(-0.33)
1500m<dist<2000m				-0.170				-0.038
				(-1.24)				(-0.33)
Control								
- Borrowers	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
- Loan contract	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dummies								
- Industry	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
- Credit officer	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
- Period	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	16,192	16,192	16,192	16,192	16,175	16,175	16,175	16,175
Nb. Borrowers	9,770	9,770	9,770	9,770	7,283	7,283	7,283	7,283
R2 (within)	0.069	0.070	0.070	0.071	0.031	0.033	0.031	0.034

Panel C: Number of employees

	Small firms (Nb. Empl.<Median)				Large firms (Nb. Empl.>Median)			
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	-0.062				0.295***			
	(-0.87)				(4.93)			
Log(distance)		0.046				0.247***		
		(1.08)				(3.43)		
Dummy			-0.071*				-0.202***	
			(-1.91)				(-3.10)	
dist<500m				-0.287**				-0.580**
				(-2.45)				(-2.17)
500m<dist<1000m				-0.246**				-0.260
				(-2.41)				(-0.94)
1000m<dist<1500m				-0.121				-0.122
				(-1.24)				(-0.46)
1500m<dist<2000m				-0.049				0.012
				(-0.41)				(0.16)
Control								
- Borrowers	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
- Loan contract	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dummies								
- Industry	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
- Credit officer	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
- Period	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	23,477	23,477	23,477	23,477	8,890	8,890	8,890	8,890
Nb. Borrowers	12,488	12,488	12,488	12,488	4,623	4,623	4,623	4,623
R2 (within)	0.052	0.052	0.053	0.053	0.050	0.049	0.046	0.046

The dependent variables is total amount of loan. The sample is divided according to the age of the firm (panel A), total sales (panel B) and the number of employees (panel C). Within estimator is used and period- and industry-dummies as well as control variables (sales, number of employees, and firm age) are included in all specifications. Other loan terms (maturity, rate and collateral ratio) are included as controls in columns [3], [4], [7] and [8]. Standard errors are clustered at the borrower level. t-statistics are reported in parentheses. Standardized coefficients are reported. *, **, and *** indicate significance at the 10%, 5% and 1% respectively.

6.2 Additional econometric issue

In addition to identification issue, our econometric model is subject to alternative potential problems, including sample selection, omitted variable, attrition, or sample dependence. We present some tests that tackle these different issues.

6.2.1 Sample selection issue

From an empirical perspective, our estimates are subject to a sample selection bias. Indeed 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 non-random decision. For instance, in the Appendix (Table A2) we show that older clients and borrowers with larger loans are more likely to be geolocated than others. 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.[28] Different parametrical and non-parametrical methods have been developed. In this paper, we follow the three-step methodology proposed by Wooldridge (1995).[29] Results are displayed in Appendix (Table A3). In Panel A, we show models explaining the loan amount. Coefficients associated with $\hat{\lambda}$ 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.

6.2.2 Inclusion of the distance between the lender and the closest competing bank

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

---

[28]Because our identification strategy is based on the inclusion of fixed effects, we cannot remove them.

[29]Technical details regarding the methodology are provided in the Appendix B.

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.[30] 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 signals an increase in the number of competitors. Results, displayed in Appendix (Table A4), 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 no significantly correlated with credit contract terms.

6.2.3 Excluding Access Banque and Microcred

Banking system in Madagscar groups together traditional commercial banks, commercial banks with specific products dedicated to small firms (such as *Bank of Africa*) and former MFIs transform in banks (such as *Access Banque* and *Microcred*). One might argue that our results are only driven by former actors only. To control for this point, we compute alternative measures of distance by excluding *Access Banque* and *Microcred*. Results, displayed in Table A5, provide interesting findings. In Panel A, we focus on determinants of loan amount. Coefficients associated with distance have the expected sign, even not always statistically significant at the usual thresholds. When we concentrate on collateral ratio, the impact of distance remains significant, even if its economic impact is reduced. In other words, results in Appendix (Table A5) indicate that only a part of the effect is driven by Access Banque and Microcred. Put differently, other (traditional) banks also compete with our partner MFI.

---

[30]We already provide a way to treat this issue when we add municipality-semester dummies.

6.2.4 Attrition and sample dependence

Our result can be driven by attrition and specific sub-samples. We tackle these two issues in the following. To be ascertain that our findings are not driven by attrition 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 sub-sample. Results displayed in Appendix (Table A6) show that this issue does not seem to bias our findings. Finally, the average number of observations per borrower is 2.18 and for some of them we have only one observation. We rerun the model by excluding these borrowers and our findings are unchanged as shown in Table A7 in Appendix.

7 Discussion

Our findings suggest that MFIs and regular banks do not operate in strictly segmented markets as often believed but rather tend to compete (at least for the most transparent borrowers). An unresolved question is whether the upscaling strategy developed by MFIs is good or bad new for borrowers. On the one hand, this strategy may enable MFIs to become more profitable and therefore to improve financial inclusion for the poorest entrepreneurs (cross-subsidization strategy). In addition, competition could actually lead to better loan conditions for small business owners. On the other hand, offering better conditions for the most transparent MSEs can have negative implications for the most opaque firms due to mission drift.

A large body of literature has raised concerns about the commercialization of microfinance and the risk that MFIs will neglect poor people, as wealthier clients cost less[31] (Morduch, 2000). Mission drift is a process along which MFIs increase the average size of the loan granted, not for progressive lending or cross-subsidization purposes[32], but only to maximize their financial profit (Armendáriz and Szafarz, 2011). In other words,

---

[31]Insofar as intermediation activity induces sunk costs, it is more profitable to finance one large loan rather than multiple small loans. It is worth noting that generally MFIs fail to reach the poorest people (Navajas et al., 2000; Lønborg and Rasmussen, 2014).

[32]Progressive lending refers to the idea that existing clients can reach out higher credit ceilings after observing a clean repayment record at the end of each credit cycle. Cross-subsidization, which consists of reaching out to unbanked wealthier clients in order to finance a larger number of poor clients whose average loan size is relatively small.

profit is the goal in itself, not the means to reduce poverty among "unbankable" people. Empirical proof is scarce, mainly because using the average loan size as a proxy for the poverty level of MFI clients does not allow researchers to distinguish cross-subsidization from "mission drift" (Armendáriz and Szafarz, 2011). In this work, we document that our partner MFI tends to follow the growth of its largest clients, especially when these clients can be grabbed by a bank. One might ask whether this trend reflects a form of mission drift and negatively affects the most opaque firms. To provide an indirect answer, we compute the average characteristics of our partner MFI's new clients by semester between 2008 and 2014. According to the mission drift hypothesis, we would expect our partner to serve less opaque firms over time (due to competitive pressure induced by bank network expansion over time). However, results, reported in the Appendix (Table A8)[33], do not support this hypothesis. Our partner has not stopped offering loans to opaque firms. On the contrary, it tended to serve more opaque clients (smaller firms) in 2014 than in 2008. Obviously, these simple descriptive statistics are not a definitive answer but are in line with recent evidence indicating the absence of mission drift (Cull et al., 2007; Mersland and Strøm, 2010). These figures show that a MFI may offer better loan conditions to its better clients without limiting its outreach.

Beyond mission drift that affects on credit access by the lower tier of borrowers, the upscaling strategy employed by MFIs might also negatively affect the upper tier of borrowers due to a hold-up effect that could prevent micro and small enterprises (MSEs) from obtaining more attractive bank financing (that is often cheaper and has more attractive features, such as longer maturities). In fact, by granting large loan amounts with less administrative requirements than banks (such as proof of sound accounting practices), MFIs may unintentionally discourage MSEs from applying for formal loans. This could have detrimental effects on economic development directly, by hindering investment, or indirectly, for example by by providing MSEs fewer incentives to formalize. However, this view assumes that successful MSEs will apply for a formal loan. Due to the lack of data, we cannot provide evidence confirming or denying the hold-up effect. Additional

---

[33]We consider all clients (not only those used in the paper). For each semester, we compute the average size (total sales and employees), the average age, and the average amount of loans granted for new clients (those without a previous relationship with our partner MFI). Depending upon the median, the first quartile or the first decile provides a similar conclusion.

investigation is needed.

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. 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 absence of competition bank proximity should not affect loan conditions offered by MFIs.

To test whether bank proximity affect loan conditions, we employ an original panel data set of 32,374 loans granted to 14,834 borrowers granted by one of the major MFIs in Madagascar over the period 2008-2014. A challenge is to go beyond a simple correlation and provide causal interpretation. Due to the lack of data, we cannot rely on instrumental variable strategy. Our identification strategy is based on the inclusion of borrower fixed effects that allows us to control for all time-invariant variables that affect both borrower's location and borrower's performance and credit conditions. In other words, we exploit only within variation (a change in distance for same borrower over time) and not between variation (differences in distance between borrowers). In robustness checks, we adopt alternative strategies. We firstly add municipality-semester dummies. This strategy allows us to control for time-varying unobserved factors that could affect banks' decision to locate and borrowers' performance occurring at the municipality level. We also focus only on borrowers that experienced reduction in distance to be ascertain that we capture within variation. Finally, we apply a test based on economic intuition. We investigate whether the impact of distance is stronger for transparent firms. The impact of bank presence must be stronger for these borrowers. All tests confirm our baseline results and give us much confidence in our causal interpretation.

We find that the proximity of a bank to an MFI client increased the size of the MFI loan obtained and decreased collateral requirements. These results are statistically and economically significant. For instance, firms located in a circle of less than 500 meters to

the nearest bank obtain, on average, $200 more than firms located beyond two kilometers. Results are robust to a battery of sensitivity tests considering alternative identification strategy, sample selection issue (because only a half of borrowers is geolocated), and alternative specifications. In addition, findings are stronger for transparent firms (well-established and larger firms) in line with our intuition.

We interpret these effects in terms of competition. Our results suggest that these two financial intermediaries actually do compete and that their markets overlap, contrary to the common view. Although previous works have studied the impact of bank development on the composition of MFI portfolios, to our knowledge we are the first to focus on an MFI's reaction in terms of lending conditions. Our findings have important policy implications for access to finance in low-income countries. Favoring the upscaling stategy of MFIs and/or the downscaling strategy of banks could improve access to credit for the largest MFI clients. Our results suggest that these firms could obtain larger loans with less collateral. In addition, a shock in one market (microfinance or banking) may have unexpected consequences for the other market.

Our findings also raise several questions that offer interesting ideas for future research. First, our results are valid in the specific context under investigation here, namely Madagascar and our partner MFI. One challenge for future research is to examine whether the effects documented by our analysis hold in different contexts (external validity). Second, in the last section, we raise concerns about the implications of our findings. Although our results can be read optimistically, they can also be interpreted in more nuanced terms. On the one hand, there is a risk of hold-up by MFIs for the most transparent borrowers. On the other hand, less transparent firms could be neglected due to mission drift. Additional investigation is required to test the existence of both effects.

References

Agarwal, S. and Hauswald, R. (2010). Distance and private information in lending. *The Review of Financial Studies*, 23(7):2757–2788.

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 *The Handbook of Microfinance*, pages 341–366.

Banerjee, A. V. and Duflo, E. (2014). Do firms want to borrow more? testing credit constraints using a directed lending program. *The Review of Economic Studies*, 81(2):572–607.

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*, 37:4183–4197.

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. *The 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.

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

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

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

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. *The Journal of Finance*, 60(1):231–266.

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. *The 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: Journal of the econometric society*, pages 153–161.

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.

Lønborg, J. H. and Rasmussen, O. D. (2014). Can microfinance reach the poorest: evidence from a community-managed microfinance intervention. *World Development*, 64:460–472.

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

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

Morduch, J. (2000). The microfinance schism. *World development*, 28(4):617–629.

Navajas, S., Schreiner, M., Meyer, R. L., Gonzalez-Vega, C., and Rodriguez-Meza, J. (2000). Microcredit and the poorest of the poor: Theory and evidence from bolivia. *World development*, 28(2):333–346.

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. *The Journal of Finance*, 49(1):3–37.

Petersen, M. A. and Rajan, R. G. (2002). Does distance still matter? the information revolution in small business lending. *The 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.

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

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.

Do banks and microfinance institutions compete? Microevidence from Madagascar

Appendix

Appendix A Additional Tables

Table A1: Keeping firms having relationship with MFI since 2010 (or before)

Panel A: Determinants of amount of loan

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	-0.048***	-0.046***						
	(-3.30)	(-3.36)						
Log(distance)			-0.044**	-0.034**				
			(-2.23)	(-1.98)				
Dummy					0.020***	0.015**		
					(2.86)	(2.43)		
dist<500m							0.042**	0.033*
							(2.02)	(1.71)
500m<dist<1000m							0.035*	0.019
							(1.80)	(1.07)
1000m<dist<1500m							-0.006	-0.007
							(-0.40)	(-0.49)
1500m<dist<2000m							-0.004	-0.003
							(-0.48)	(-0.39)
Loan contract	No	Yes	No	Yes	No	Yes	No	Yes
Obs.	14,061	14,061	14,061	14,061	14,061	14,061	14,061	14,061
Nb. Borrowers	3,069	3,069	3,069	3,069	3,069	3,069	3,069	3,069
R2 (within)	0.083	0.277	0.083	0.277	0.083	0.277	0.083	0.277

Panel B: Determinants of collateral ratio

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	0.053	0.047						
	(0.75)	(0.68)						
Log(distance)			0.114***	0.110***				
			(2.86)	(2.75)				
Dummy					-0.053***	-0.051***		
					(-4.03)	(-3.85)		
dist<500m							-0.132***	-0.128***
							(-3.57)	(-3.44)
500m<dist<1000m							-0.108***	-0.104***
							(-3.04)	(-2.93)
1000m<dist<1500m							-0.040	-0.042
							(-1.35)	(-1.41)
1500m<dist<2000m							-0.008	-0.009
							(-0.37)	(-0.39)
Loan contract	No	Yes	No	Yes	No	Yes	No	Yes
Obs.	14,061	14,061	14,061	14,061	14,061	14,061	14,061	14,061
Nb. Borrowers	3,069	3,069	3,069	3,069	3,069	3,069	3,069	3,069
R^2 (within)	0.029	0.036	0.031	0.037	0.031	0.037	0.031	0.037

The dependent variables are the total amount of loan (Panel A) and the collateral-to-loan ratio (Panel B). Within estimator is used. The list of control variables includes total sales (in constant US$), the number of employees, the firm age, the number of loan between the firm and our partner IMF). In columns [2], [4], [6] and [8], we also add other contract terms. In all specifications, we include town-period dummies, credit officer dummies and industry-dummies. Standard errors are clustered at the borrower level. t-statistics are reported in parentheses. Standardized coefficients are reported. *, **, and *** indicate significance at the 10%, 5% and 1% respectively.

Table A2: t-test

	Geolocated			t-test
	All	Yes	No	(p-value)
Loan characteristics				
Amount[†]	1057.9	1176.1	926.6	<0.01
Rate	11.46	12.60	10.20	<0.01
Maturity	387.6	389.3	385.6	<0.01
Collateral/Amount	2.80	2.86	2.73	<0.01
Security/Collateral	54.0	54.4	53.6	<0.01
Business characteristics				
Sales[†]	1719.7	1789.6	1641.3	<0.01
Employees	2.19	2.27	2.11	<0.01
Age (firm)	8.36	8.84	7.82	<0.01
Leding relationship				
Number	2.78	2.92	2.63	<0.01
Years	2.13	2.29	1.95	<0.01
# Obs.	61,953	32,374	29,579	

[†] Data in deflated USD

Table A3: Determinants of loan amount, Wooldridge procedure (sample selection)

Panel A: Determinants of amount of loan

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	-0.054*** (-4.96)	-0.051*** (3.27)						
log(distance)			-0.046** (-2.00)	-0.037** (-2.58)				
Dummy					77.65* (1.85)	60.27* (1.79)		
Dist<500m							280.21* (1.65)	215.66* (1.64)
500m<dist<1000m							188.02 (0.77)	107.00 (0.52)
1000m<dist<1500m							41.83 (0.38)	53.17 (0.70)
1500m<dist<2000m							20.40 (0.44)	13.22 (0.35)
$\hat{\lambda}$	NS	NS	NS	NS	NS	NS	NS	NS
Loan contract	No	Yes	No	Yes	No	Yes	No	Yes
Observations	32374	32374	32374	32374	32374	32374	32374	32374
Nb. Borrowers	14,834	14,834	14,834	14,834	14,834	14,834	14,834	14,834
R2 (within)	0.081	0.271	0.081	0.271	0.081	0.271	0.081	0.271

Panel B: Collateral ratio

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	0.054 (1.31)	0.048 (1.18)						
log(distance)			0.109*** (3.44)	0.104*** (3.37)				
Dummy					-0.128*** (-6.87)	-0.122*** (-6.28)		
Dist<500m							-0.397*** (-4.79)	-0.378*** (-5.14)
500m<dist<1000m							-0.278*** (-4.82)	-0.264*** (-4.50)
1000m<dist<1500m							-0.121** (-1.99)	-0.124** (-2.00)
1500m<dist<2000m							-0.0520 (-0.99)	-0.0539 (-1.05)
$\hat{\lambda}$	p<0.10	p<0.10	p<0.10	p<0.10	p<0.10	p<0.10	p<0.10	p<0.10
Loan contract	No	Yes	No	Yes	No	Yes	No	Yes
Observations	32374	32374	32374	32374	32374	32374	32374	32374
Nb. Borrowers	14,834	14,834	14,834	14,834	14,834	14,834	14,834	14,834
R2 (within)	0.070	0.075	0.071	0.076	0.071	0.076	0.071	0.075

The dependent variable is the total loan amount (in deflated USD). 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). Period, industry and credit officer dummies as well as control variables (sales, employees, age, 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 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)

Table A4: Inclusion of the distance between MFI agency and the closest bank

Panel A: Determinants of amount of loan

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	-0.052*** (-3.24)	-0.051*** (-3.37)						
Log(distance)			-0.044* (-1.80)	-0.037* (-1.73)				
Dummy					77.15** (2.29)	60.30** (2.01)		
dist<500m							274.25* (1.84)	212.57 (1.58)
500m<dist<1000m							184.03* (1.67)	103.53 (1.04)
1000m<dist<1500m							-45.51 (-0.49)	-56.84 (-0.68)
1500m<dist<2000m							-22.70 (-0.35)	-16.37 (-0.28)
Distance IMF-bank	-0.018 (-0.87)	-0.018 (-0.88)	-0.006 (-0.23)	-0.008 (-0.37)	-0.009 (-0.41)	-0.012 (-0.56)	-0.005 (-0.21)	-0.011 (-0.47)
Loan contract	No	Yes	No	Yes	No	Yes	No	Yes
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
R^2 (within)	0.079	0.272	0.079	0.272	0.078	0.272	0.080	0.272

Panel B: Determinants of collateral ratio

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	0.048 (0.68)	0.042 (0.61)						
Log(distance)			0.097** (2.44)	0.092** (2.32)				
Dummy					-0.118*** (-3.67)	-0.113*** (-3.52)		
dist<500m							-0.358*** (-3.17)	-0.340*** (-2.99)
500m<dist<1000m							-0.245** (-2.51)	-0.231** (-2.37)
1000m<dist<1500m							-0.100 (-1.06)	-0.099 (-1.11)
1500m<dist<2000m							-0.037 (-0.44)	-0.039 (-0.46)
Distance IMF-bank	0.111* (1.92)	0.110* (1.91)	0.078 (1.30)	0.078 (1.31)	0.079 (1.36)	0.079 (1.37)	0.069 (1.17)	0.070 (1.19)
Loan contract	No	Yes	No	Yes	No	Yes	No	Yes
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
R^2 (within)	0.029	0.035	0.030	0.035	0.030	0.035	0.030	0.035

The dependent variables are the total amount of loan (Panel A) and the collateral-to-loan ratio (Panel B). Within estimator is used. The list of control variables includes total sales (in constant US$), the number of employees, the firm age, the number of loan between the firm and our partner IMF). In columns [2], [4], [6] and [8], we also add other contract terms. In all specifications, we include town-period dummies, credit officer dummies, industry-dummies and control variables (sales, employees, age, number of loans). Standardized coefficients are reported for continuous variables (distance, log and control variables) and usual coefficients for binary variables (dummy and categories). t-statistics are reported in parentheses. *, **, and *** indicate significance at the 10%, 5% and 1% respectively.

Table A5: Exclusion of Access Banque and Microcred

Panel A: Determinants of amount of loan

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	-0.030	-0.035						
	(-1.08)	(-1.54)						
Log(distance)			-0.041	-0.022				
			(-1.26)	(-0.83)				
Dummy					102.19**	67.41*		
					(2.54)	(1.85)		
dist<500m							298.52*	192.14
							(1.64)	(1.21)
500m<dist<1000m							210.40*	117.20
							(1.75)	(1.06)
1000m<dist<1500m							51.96	8.96
							(0.51)	(0.09)
1500m<dist<2000m							88.15	81.46
							(1.10)	(1.09)
Loan contract	No	Yes	No	Yes	No	Yes	No	Yes
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
R^2 (within)	0.078	0.270	0.078	0.270	0.078	0.270	0.078	0.270

Panel B: Determinants of collateral ratio

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	0.057	0.054						
	(0.67)	(0.65)						
Log(distance)			0.150***	0.146***				
			(4.25)	(4.10)				
Dummy					-0.163***	-0.156***		
					(-4.58)	(-4.41)		
dist<500m							-0.430***	-0.410***
							(-3.96)	(-3.75)
500m<dist<1000m							-0.246**	-0.231**
							(-2.48)	(-2.32)
1000m<dist<1500m							-0.113	-0.110
							(-1.13)	(-1.10)
1500m<dist<2000m							0.006	0.011
							(0.06)	(0.12)
Loan contract	No	Yes	No	Yes	No	Yes	No	Yes
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
R^2 (within)	0.030	0.035	0.031	0.036	0.031	0.036	0.031	0.037

The dependent variables are the total amount of loan (Panel A) and the collateral-to-loan ratio (Panel B). Within estimator is used. The list of control variables includes total sales (in constant US$), the number of employees, the firm age, the number of loan between the firm and our partner IMF). In columns [2], [4], [6] and [8], we also add other contract terms. In all specifications, we include town-period dummies, credit officer dummies, industry-dummies and control variables (sales, employees, age, number of loans). Standardized coefficients are reported for continuous variables (distance, log and control variables) and usual coefficients for binary variables (dummy and categories). t-statistics are reported in parentheses. *, **, and *** indicate significance at the 10%, 5% and 1% respectively.

Table A6: Attrition (keeping borrowers with data for 2014)

Panel A: Determinants of amount of loan

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	-0.078***	-0.070***						
	(-3.53)	(-3.40)						
Log(distance)			-0.080**	-0.060**				
			(-2.18)	(-2.09)				
Dummy					0.020**	0.016**		
					(2.50)	(2.39)		
dist<500m							0.097**	0.074**
							(2.40)	(2.18)
500m<dist<1000m							0.079**	0.059**
							(2.41)	(2.11)
1000m<dist<1500m							0.009	0.006
							(0.36)	(0.27)
1500m<dist<2000m							0.005	0.006
							(0.38)	(0.58)
Loan contract	No	Yes	No	Yes	No	Yes	No	Yes
Obs.	18,962	18,962	18,962	18,962	18,962	18,962	18,962	18,962
Nb. Borrowers	8,414	8,414	8,414	8,414	8,414	8,414	8,414	8,414
R2 (within)	0.081	0.267	0.082	0.267	0.081	0.267	0.083	0.268

Panel B: Determinants of collateral ratio

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	-0.028	-0.033						
	(-0.24)	(-0.32)						
Log(distance)			0.070	0.063				
			(1.23)	(1.11)				
Dummy					-0.019*	-0.017*		
					(-1.88)	(-1.74)		
dist<500m							-0.102**	-0.093*
							(-2.02)	(-1.84)
500m<dist<1000m							-0.070	-0.063
							(-1.51)	(-1.34)
1000m<dist<1500m							-0.010	-0.010
							(-0.26)	(-0.25)
1500m<dist<2000m							0.008	0.008
							(0.28)	(0.28)
Loan contract	No	Yes	No	Yes	No	Yes	No	Yes
Obs.	18,962	18,962	18,962	18,962	18,962	18,962	18,962	18,962
Nb. Borrowers	8,414	8,414	8,414	8,414	8,414	8,414	8,414	8,414
R2 (within)	0.036	0.041	0.036	0.042	0.036	0.042	0.037	0.042

The dependent variables are the total amount of loan (Panel A) and the collateral-to-loan ratio (Panel B). Within estimator is used. The list of control variables includes total sales (in constant US$), the number of employees, the firm age, the number of loan between the firm and our partner IMF). In columns [2], [4], [6] and [8], we also add other contract terms. In all specifications, we include town-period dummies, credit officer dummies, industry-dummies and control variables (sales, employees, age, number of loans). Standardized coefficients are reported for continuous variables (distance, log and control variables) and usual coefficients for binary variables (dummy and categories). t-statistics are reported in parentheses. *, **, and *** indicate significance at the 10%, 5% and 1% respectively.

Table A7: Keeping firms having at least two observations

Panel A: Determinants of amount of loan

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	-0.049***	-0.048***						
	(-3.29)	(-3.42)						
log(distance)			-0.042*	-0.036*				
			(-1.91)	(-1.85)				
Dummy					78.73**	62.36**		
					(2.45)	(2.16)		
500m<d<1000m							277.96**	220.27*
							(1.97)	(1.72)
1000m<d<1500m							187.15*	110.03
							(1.81)	(1.18)
1500m<d<2000m							-42.76	-51.32
							(-0.48)	(-0.63)
Loan contract	No	Yes	No	Yes	No	Yes	No	Yes
Obs.	25431	25431	25431	25431	25431	25431	25431	25431
Nb. Borrowers	7891	7891	7891	7891	7891	7891	7891	7891
R2 (within)	0.0787	0.270	0.0786	0.270	0.0786	0.270	0.0795	0.271

Panel B: Determinants of collateral ratio

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Distance	0.050	0.044						
	(0.75)	(0.67)						
log(distance)			0.104***	0.099***				
			(2.74)	(2.62)				
Dummy					-0.126***	-0.121***		
					(-4.00)	(-3.84)		
500m<d<1000m							-0.386***	-0.368***
							(-3.53)	(-3.36)
1000m<d<1500m							-0.268***	-0.255***
							(-2.80)	(-2.66)
1500m<d<2000m							-0.116	-0.119
							(-1.29)	(-1.34)
Loan contract	No	Yes	No	Yes	No	Yes	No	Yes
Obs.	25431	25431	25431	25431	25431	25431	25431	25431
Nb. Borrowers	7891	7891	7891	7891	7891	7891	7891	7891
R2 (within)	0.0787	0.270	0.0786	0.270	0.0786	0.270	0.0795	0.271

The dependent variables are the total amount of loan (Panel A) and the collateral-to-loan ratio (Panel B). Within estimator is used. The list of control variables includes total sales (in constant US$), the number of employees, the firm age, the number of loan between the firm and our partner IMF). In columns [2], [4], [6] and [8], we also add other contract terms. In all specifications, we include town-period dummies, credit officer dummies and industry-dummies. Standard errors are clustered at the borrower level. t-statistics are reported in parentheses. Standardized coefficients are reported. *, **, and *** indicate significance at the 10%, 5% and 1% respectively.

Table A8: Mission Drift, indirect test

Semester	Amount	Sales	Employees	Age	Amount/Sales
2008-S1	812.61	2034.24	2.35	5.69	139.32
2008-S2	808.40	2312.22	2.58	5.31	125.61
2009-S1	734.27	2124.27	2.33	5.61	116.82
2009-S2	674.59	1473.53	2.00	5.18	140.81
2010-S1	847.77	1841.76	2.33	6.09	169.71
2010-S2	952.05	1721.47	2.16	5.41	166.03
2011-S1	804.67	1708.07	2.25	5.85	169.08
2011-S2	723.59	1444.38	2.13	6.56	154.53
2012-S1	859.07	1477.45	1.88	8.02	159.85
2012-S2	847.25	1452.52	1.67	5.60	192.05
2013-S1	963.07	1542.13	1.70	6.55	189.95
2013-S2	973.56	1529.72	1.83	7.11	230.66
2014-S1	782.67	1197.94	1.76	6.92	213.12
2014-S2	800.77	1222.80	1.73	7.48	209.60

The table reports the average value of amount, total sales, number of employees, age and the ratio of amount to sales to new clients (those that never obtain previous loan from our partner) by semester. Financial data are deflated using the CPI and reported in dollar (1 Ariary=3,144 $US)

Appendix B Three-step Wooldridge procedure

Due to localization of only half of clients, our estimates are subject to a potential sample selection issue. 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:

$$Pr(s_i = 1) = \Phi(\Delta\mathbf{X}_i + \gamma Z_j + \eta_s) \quad (\forall t = 0, \ldots, T) \tag{3}$$

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 include 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:[34]

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

We compute the selection variable Z_j for each period. In Equation 3 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 A1, we report the estimated $\hat{\gamma}$ per period as well as confidence interval. We observe that our selection variable is always positive and highly significant in all periods.[35]

---

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

[35]We extend this model by adding alternative contract terms when we consider these variables in our outcome equation. Results are highly similar for selection variable.

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

$$\hat{\lambda}_i = \frac{\phi(\hat{\Delta}\mathbf{X}_i + \hat{\gamma}Z_j + \hat{\eta}_s)}{\Phi(\hat{\Delta}\mathbf{X}_i + \hat{\gamma}Z_j + \hat{\eta}_s)} \quad (\forall t = 0, \ldots, T) \tag{4}$$

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:

$$y_{it} = \beta d_{it} + \Delta\mathbf{X}_{it} + \rho\hat{\lambda}_{it} + \mu_i + \nu_t + \eta_s + \tau_j + \epsilon_{it} \tag{5}$$

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.

Figure A1: Estimated parameters of selection variable in selection equation