IDB | Invest

Development through
the Private Sector Series
April 2020

TN
No. 21

Free Riding in Loan Approvals:

Evidence from SME Lending in Peru

Authors:
Irani Arraiz
Miriam Bruhn
Benjamin N. Roth
Claudia Ruiz-Ortega
Rodolfo Stucchi

IDB | Invest

Free Riding in Loan Approvals:
Evidence from SME Lending in Peru

Copyright © 2020 Inter-American Investment Corporation (IIC). This work is licensed under a Creative Commons IGO 3.0 Attribution-NonCommercial-NoDerivatives (CC-IGO BY-NC-ND 3.0 IGO) license (http://creativecommons.org/licenses/by-nc-nd/3.0/igo/legalcode) and may be reproduced with attribution to the IIC and for any non-commercial purpose. No derivative work is allowed.

Any dispute related to the use of the works of the IIC that cannot be settled amicably shall be submitted to arbitration pursuant to the UNCITRAL rules. The use of the IIC's name for any purpose other than for attribution, and the use of IIC's logo shall be subject to a separate written license agreement between the IIC and the user and is not authorized as part of this CC-IGO license.

Following a peer review process, and with previous written consent by the Inter-American Investment Corporation (IIC), a revised version of this work may also be reproduced in any academic journal, including those indexed by the American Economic Association's EconLit, provided that the IIC is credited and that the author(s) receive no income from the publication. Therefore, the restriction to receive income from such publication shall only extend to the publication's author(s). With regard to such restriction, in case of any inconsistency between the Creative Commons IGO 3.0 Attribution-NonCommercial-NoDerivatives license and these statements, the latter shall prevail.

Note that link provided above includes additional terms and conditions of the license.

The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the Inter-American Development Bank Group, its respective Boards of Directors, or the countries they represent.

Cover page design: Jimena Vazquez

April 2020

Free Riding in Loan Approvals: Evidence from SME Lending in Peru*

Irani Arraiz⁑, Miriam Bruhn§, Benjamin N. Roth‡,

Claudia Ruiz-Ortega¶, Rodolfo Stucchi†

April 2020

Abstract

This paper provides evidence that commercial lenders in Peru free ride off their peers' screening efforts. Leveraging a discontinuity in the loan approval process of a large bank, the study finds that competing lenders responded to additional loan approvals by issuing approvals of their own. Competing lenders captured almost three-quarters of the new loans to previously financially excluded borrowers, greatly diminishing the profits accruing to the initiating bank. Lenders may therefore underinvest in screening new borrowers and expanding financial inclusion, as their competitors reap some of the benefit. The results highlight that information spillovers between lenders may operate outside credit registries.

JEL Classification: D82, G21, O16

Keywords: Asymmetric information, credit market competition, access to credit

---

* We thank the Entrepreneurial Finance Lab and our partner bank for sharing their data and information, as well as Pierre Freundt from Equifax Peru for patiently answering our questions. We thank Sumit Agarwal, Abhijit Banerjee, Vivek Bhattacharya, Ralph de Haas, Xavier Gine, Daniel Green, Dean Karlan, Michael Kremer, Nicola Limodio, Ernest Liu, Soledad Martinez-Peria, David McKenzie, and Jonathan Zinman for helpful discussions and feedback. Financial support for this project was provided by the Strategic Research Program. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the Inter-American Development Bank, IDB Invest, or the World Bank, their Boards of Directors, or the countries they represent.

⁑ IDB Invest

§ World Bank and IZA

‡ Harvard Business School

¶ World Bank

† IDB Invest

1. Introduction

While the returns to credit and other benefits of financial inclusion have been of central interest in academia and policy, relatively less attention has been paid to the limits of private sector incentives to achieve financial inclusion. In deciding whether to issue credit to a new borrower, a profit-maximizing lender must consider the costs of screening the borrower and issuing the loan, how much debt the borrower can reliably service, and how long the borrower will remain a customer. This final consideration is typically framed based on whether a borrower will eventually join a competing lender after the initial lender has incurred the cost of establishing her reliability (e.g. Petersen and Rajan, 1995), but in principle this form of competition could occur even before the first loan is issued. If lenders are more likely to approve borrowers already approved by their competitors, then lenders that incur the cost of evaluating new or underserved borrowers may not reap the resulting benefits. This phenomenon whereby lenders free ride on the screening efforts of their competitors reduces the incentive to expand credit access and financial inclusion and might warrant policy intervention.

We demonstrate that free riding in loan approvals has a large impact on market outcomes. Specifically, we worked with a large Peruvian bank interested in expanding credit access to small and medium enterprises (SMEs). Our partner bank adopted a new screening technology and determined which SMEs to lend to based on a scoring rule with a strict threshold. Borrowers above the threshold were automatically granted a loan, whereas borrowers below the threshold were offered a loan only if a loan officer deemed it appropriate. Borrowers above the threshold also received more attractive loan terms. Exploiting this threshold along with administrative data from our partner bank and credit registry data from Equifax Peru, we document several findings.

While applicants *without* prior credit histories who score above the threshold were more likely to receive a loan than those who score below it, three-quarters of the additional loans were issued by competing financial institutions rather than from our partner bank. Because the only differences between borrowers on either side of the threshold are whether they were approved for a loan from our partner bank and the resulting loan terms, this is evidence of free riding in loan approvals. We document that free riding in loan approvals is concentrated among non-bank financial institutions and is higher in regions of Peru in which our partner bank faces more competition. Furthermore, we find evidence that our partner bank's loan approvals for these marginal borrowers lead to an increase in the profits of competing financial institutions but not our partner bank.

In contrast, we find that among applicants *with* a prior credit history, nearly all additional borrowing for those above the threshold does come from loans issued by our partner bank. That is, we do not find evidence of free riding in loan approvals for these applicants. Importantly, applicants with credit histories who receive loans only because they are just above the threshold are those whose credit histories otherwise establish them to be unsuitable borrowers. So, the stamp of approval from our partner bank may have carried less weight in competitors' approval processes, as they had other signals of credit (un)worthiness on which to rely.

Taken together, these findings paint a stark picture. Though our partner bank incurred the costs of the novel screening technology, the benefits accrued largely to its competitors. The straightforward implication is that banks may underinvest in expanding the credit supply to underserved borrowers, as doing so entails a private cost but produces a public good. This may justify subsidies to private sector efforts to expand financial inclusion.

Several mechanisms may underlie this phenomenon of free riding in loan approvals. On the supply side, lenders may directly respond to the loan approvals of our partner bank by issuing loan approvals of their own. For instance, borrowers may have shared their loan approval documents with competing lenders, who were then updated about the credit worthiness of these borrowers. On the demand side, borrowers who received a loan approval from our partner bank may have updated their beliefs about their own credit worthiness and redoubled their search for credit from competing lenders. This channel is consistent with Karlan and Zinman (2009), who find that South African consumer loan applicants randomly approved for credit update positively about their self-perceived likelihood to be approved for loans from other lenders. In the case of demand-side mechanisms, other lenders benefit from the loan approvals of our partner bank through the indirect channel of receiving more applicants, rather than by directly responding to the loan approvals themselves.

Importantly, we can rule out some potential mechanisms. On the supply side, we can exclude any mechanism that operates through the credit registry, as our findings are based on loan approvals rather than loan issuances, and loan approvals are not recorded in the Peruvian credit registry. On the demand side, we can rule out complementarities in borrowing whereby an initial loan from our partner bank increases demand for credit from other lenders; few of the borrowers in our sample who received loans from competing lenders first borrowed from our partner bank.

Though we cannot pin down a single explanation, we argue that the precise mechanism by which this free riding occurs does not influence our broad conclusions; so long as a loan approval by our partner bank *causes* (directly or indirectly) other lenders to approve the same borrowers, the private returns of identifying credit worthy borrowers and expanding the financial supply will diverge from the corresponding social returns.

Our paper relates to several literatures. First, we complement the literature examining how variation in the credit supply decisions of one lender influences the credit supply of competitors (Agarwal et al. 2018, Karlan and Zinman 2018, Azevedo et al. 2019, Burke et al. 2019). Through variation in loan terms or access to credit, each of these papers finds evidence that an initial loan from one lender causes other lenders to provide access to credit as well. Our paper is distinct in this literature in that we identify information spillovers that occur even before the initial loan is issued, and are therefore more deleterious to private incentives to expand the credit supply. Also of note, much of this focuses on consumer lending (Burke et al. 2019), or microfinance (Agarwal et al. 2018, Karlan and Zinman, Azevedo et al. 2019), while our focus is on SME lending.

Second, our paper relates to the literature on the consequences of information sharing through credit registries on bank competition and credit supply (eg. Jappelli and Pagano 2002, Djankov et al. 2007, Hertzberg et al. 2011, Liberman 2016, Dobbie et al. 2016, Foley et al. 2018, Sutherland 2018). As noted above, our principal contribution to this literature is to document that information spillovers between lenders can also occur through loan approvals rather than loan issuance. Because loan approvals are not encoded in the Peruvian credit registry, our results imply that information spillovers can occur even before information is encoded in the credit registry, and before the initial lender has derived any benefit from screening new borrowers. And our results imply that information spillovers between lenders may occur even in settings without well-functioning credit registries.

The rest of the paper proceeds as follows: Section 2 discusses the context and novel screening technology and our data sources, Section 3 presents our evidence for free riding in loan approvals, and Section 4 concludes with a discussion of possible mechanisms.

2. Study Context and Data

2.1 Study Context

We collaborated with one of the five largest commercial banks in Peru (we refer to this bank as our "partner bank") in an exercise that started in 2012. At the time, our partner bank had only a small SME portfolio. Among commercial banks, our partner bank had only a 5 percent market share for lending to medium enterprises and a less than 1 percent market share for lending to micro or small enterprises. Our partner bank was thus particularly interested in reaching the micro and small and medium enterprise segment. This segment was dominated by four other banks, which accounted for 90 percent of commercial bank micro and small enterprise lending. Non-bank financial institutions (NBFIs) also played an important role in the micro and small enterprise credit market. Forty-four NBFIs had a combined credit volume that was equal to 60 percent of the commercial bank credit volume for small enterprise credit and 2.4 times the commercial bank credit volume for microenterprise credit.[1]

To better reach the SME segment, our partner bank piloted a psychometric tool for screening loan applicants, with support from the IDB. The psychometric tool was developed by the Entrepreneurial Finance Lab (EFL), a fintech company founded in 2010, and relies on a series of questions that measure personality traits, skills, knowledge, and attitudes that aim to predict the applicant's ability and willingness to repay a loan. The IDB and EFL co-financed the implementation of the psychometric tool during the pilot with our partner bank.

SMEs that applied for a working capital loan (up to 18 months in duration with an average loan size of $3,855) between March 2012 and August 2013 were screened by the EFL tool as part of the application process. Applicants who achieved a score on the EFL application higher than a threshold defined by our partner bank were automatically offered a loan, while those below the EFL threshold were offered a loan only if they were approved under the institution's conventional screening method. Loan terms also differed across the threshold.

We note that as our identification strategy relies on comparing borrowers just above and just below an arbitrary EFL score cutoff, the predictive power of the EFL

[1] All statistics on market share and credit volume are from the Peruvian Bank Supervisor (SBS) for February 2012.

questions is not important for our study. Regardless of EFL predictive power, applicants just above and below the cutoff would have nearly identical predictions of creditworthiness. Critically, however, the borrowers and other lenders could not discern, in all cases, whether borrowers were marginally approved due to their EFL score or if they were approved because our partner bank deemed them creditworthy for substantive reasons. Therefore, both competing lenders and the borrowers themselves may rationally update about the borrowers' creditworthiness after receiving a loan approval from our partner bank.

2.2 Data

We obtained data on 1,883 SMEs that applied for a working capital loan with our partner bank, from two sources: administrative data from our partner bank and Equifax Peru. The administrative data include the EFL score and the date when the SME applied for the loan, as well as the applicant's age, gender, business sales, and whether or not our partner bank approved them for a loan. The applicants in our sample were on average 39 years old and 50 percent of them were female. Average annual business revenues were about US$12,000 (see Column 1 in Table 1). In addition, for each loan applicant, the data include the national ID number (DNI) and, if their business is registered under the business name instead of the individual's name, it also includes the business's tax payer number (RUC).[2]

Our second data source is Equifax Peru, the largest credit bureau in the country. For the DNIs and RUCs in the EFL data, we purchased five years of monthly information on borrowing from regulated financial institutions, covering the period from May 2011 to April 2016.[3]

Equifax collects this information from the Peruvian Bank Supervisor's (SBS) credit registry (*Central de Riesgos*). SBS collects data directly from all regulated financial institutions monthly, covering the universe of commercial banks, as well as all regulated

[2] We originally had 1,909 SMEs in the data. All provided their DNI and 1,327 also provided an RUC. However, for 20 SMEs, the DNIs and RUCs are inconsistent with each other, suggesting typos. We drop these observations from the sample to avoid using wrong information from our second data source. We also drop 6 observations where two DNIs reported the same RUC, that is, three SMEs where two co-owners seem to each have applied for a loan. In these cases, it is not possible to cleanly assign an EFL score to the SME as the unit of observation. Thus, we end up with a sample of 1,883 SMEs. The fraction of the sample dropped is not significantly different below and above the EFL threshold.

[3] By law, Equifax is not allowed to provide data that are older than five years.

NBFIs. For each ID number in any given month, we obtained the total amount borrowed (current and delinquent) from each SBS supervised financial institution in Peru, for the three loan types most relevant in our sample: microloans, loans to small firms, loans to medium firms. If a borrower has more than one loan of the same type with the same institution, Equifax reports only the sum of these loans, with no information on how many loans constitute this total amount.

Our primary outcome of interest is whether loan applicants took out a new loan in the six months following their application to our partner bank. We measure loan take-up for each institution and loan type by creating a dummy equal to one if the amount outstanding of either the DNI or RUC associated with a loan applicant increases by any amount. The immediate post-application period allows six months for loans to be processed and disbursed and provides some time for applicants to potentially shop around with other financial institutions for other loan offers. For a placebo test, we define an analogous outcome measuring whether each loan applicant took out a new loan from each financial institution in the six months immediately preceding his or her loan application with our partner bank.

We also purchased Equifax credit scores for the month when the SME applied for the loan from our partner bank. Here, Equifax included a dummy variable indicating whether this score was primarily based on their credit history, i.e. a "thick file," or on demographics and other sources, such as the Peruvian tax authority (SUNAT), i.e. a "thin file".[4] Our sample includes 1,517 thick file borrowers and 366 thin file borrowers.

We utilize two measures of credit market competition: the log of per capita NBFI lending at the district level, and the per capita number of NBFI branches at the district level. We focus on NBFIs as our results indicate that the increase in loans for thin-file applicants comes from NBFIs, not banks. We note that as our partner bank operates in only eight districts, this is a fairly coarse measure of competition. Total NBFI lending for February 2012 was obtained from the Peruvian Bank Supervisor (SBS) and the 2012 population size from the Human Development Index of the UNDP.

Though we do not directly observe the profits that lenders derive from each loan applicant, we compute an imperfect measure of profits from new loans at the applicant-

[4] For individuals with "thin files", Equifax calculates credit scores based on several variables, including whether they have a co-signer, their income level, and recurring payments, such as tuition, rent, and electricity.

institution level. If the applicant does not have a new loan with a given financial institution six months after the loan application, profits are defined to be zero. If they have a new loan within six months of the loan application, we measure the size of this loan as the sum of all loan balance increases in the six months. The size is set to zero for applicants with no new loans. We obtain the interest income by multiplying the size of the loan by its "typical" interest rate. Typical interest rates are reported by the SBS for each financial institution and each loan type. We do not observe interest rates at the loan level, but anecdotal evidence suggests that interest rates are fairly standardized within products for a given lender. To capture losses, we then subtract the amount in default 24 months after the initial loan application, which could be zero in case of no defaults. We picked 24 months to allow six months for the loans to be issued and 18 months for the termination of loans with the maximum duration.[5]

3. Results

We now describe our identification strategy and results. As our partner bank applied a strict threshold rule based on the EFL score, across which both the likelihood of loan approval and loan terms vary discontinuously, we apply a regression discontinuity (RD) design. Specifically, letting s_i be the borrower's EFL score, $\bar{s}$ be the EFL cutoff, $Y_i(1)$ be the outcome of interest for borrower i conditional on being above the EFL cutoff and $Y_i(0)$ be the outcome of interest for borrower i conditional on being below the EFL cutoff, we estimate $E(Y_i(1) - Y_i(0)|s_i = \bar{s})$ by approximating $\lim_{s_i \to \bar{s}^+} E(Y_i(1)|s_i) - \lim_{s_i \to \bar{s}^-} E(Y_i(0)|s_i)$ (Hahn et al., 2001). We estimate the $\lim_{s_i \to \bar{s}^+} E(Y_i(1)|s_i)$ and $\lim_{s_i \to \bar{s}^-} E(Y_i(0)|s_i)$ using a variety of bandwidths around $\bar{s}$ and polynomial functions of various orders.

In our main results we estimate a linear model on either side of the threshold. Normalizing the threshold $\bar{s} = 0$, and letting y_i represent the observed outcome variable for person i, we estimate

$$y_i = \alpha + \beta \mathbb{I}(s_i \geq 0) + \gamma_1 \mathbb{I}(s_i < 0) s_i + \gamma_2 \mathbb{I}(s_i \geq 0) s_i + \delta X_i + \varepsilon_i$$

[5] If the financial institution stops reporting a loan amount for a given borrower before 24 months are over, we use the amount in default in the last month when the amount was observed (after the first six months after the loan application). If the amount in default exceeds the calculated size of the loan (which may be the case if it includes other loans from the same institution), we replace the amount in default with the size of the loan.

where $\mathbb{I}(\cdot)$ is the indicator function and X_i is a vector of control variables representing age, gender, log business revenues, and Equifax score. Estimates from models with higher order polynomials are included in the appendix. The estimation is done in Stata using the command *rdrobust*, described in Calonico et al. (2014a).

At the outset, in Table 1 we note that none of the demographic or business characteristics are discontinuous across the EFL threshold. Columns 3 through 5 present the RD estimates with polynomials of order 0 through 2, and across all characteristics none exhibit a statistically significant jump at the RD cutoff. Online Appendix Tables A1 and A2 present analogous specifications restricting to the samples of borrowers with and without pre-existing credit histories (i.e. *thick* and *thin file applicants*). This offers some reassurance in our research design. Figure A1 plots distribution of EFL scores in our sample and allows for visual inspection that the distribution appears continuous across the threshold. A formal test using the Stata command *rddensity* (Cattaneo et al., 2018) confirms we cannot reject that there is no sorting around the EFL threshold.

Our primary outcomes of interest are the probability of loan approval from our partner bank and whether the applicant ultimately takes a loan from our partner bank, one of its competitors, or neither. Because there is reason to believe that the magnitude of free riding in loan approvals may differ for loan applicants with thick files and those with thin files, we separately estimate the effects of being above the EFL threshold on the full sample, and on each of the two subsamples.

We now establish our first stage; direct evidence that borrowers just above the EFL threshold are more likely to be approved for loans by our partner bank, and indirect evidence that loan terms are discontinuously more attractive above the EFL threshold. The first row of Table 2 presents our estimates of the effect of being above the EFL threshold on the likelihood an applicant is to be approved for a loan offer from our partner bank. Every point estimate in Table 2 corresponds to the coefficient of interest, β, in a separate RD model. The first row of Figure 1 provides a graphical depiction of the corresponding data and fitted models. All the estimates in Table 2 are from linear models and the associated optimal bandwidth. Alternate specifications with different polynomials and alternative bandwidths are presented in Online Appendix Tables A3 and A4.

The first row of Table 2 demonstrates that across all borrowers in our sample, those who are just above the EFL threshold were 23 percentage points (p-val: 0.00) more likely to be approved for a loan by our partner bank, those with thick files were 25 percentage points (p-val 0.00) more likely to be approved for a loan, and those with a thin files were 10 percentage points (p-val: 0.21) more likely to be approved for a loan. These estimates reflect that most applicants who scored below the EFL threshold were approved for a loan as well. However, while we do not directly observe the terms of loan offers, we note that the second row of Table 2 indicates that across the whole sample of applicants, those just above the EFL threshold were more than twice as likely to accept a loan as those just below the EFL cutoff, compared to a mere 23 percentage points increase in the likelihood of loan approval. That additional loan approvals cannot account for most of the additional borrowing above the threshold indicates that loan offers were discontinuously more attractive at the threshold.

Rows 2 through 5 of Table 2 present our main results – applicants just above the EFL threshold are not only more likely to borrow from our partner bank but also from our partner bank's competitors. Because the EFL score was only observed by our partner bank, these estimates reflect the competitor response to the approval and loan terms of our partner bank. The bottom two rows of Figure 1 provide a graphical depiction of the data and fitted models and alternate specifications with higher order polynomials and alternative bandwidths are presented in Online Appendix Tables A3 and A4.

The second row of Table 2 establishes that across all borrowers in our sample, those just above the threshold are 28 percentage points (p-val: 0.00) more likely to receive a loan from our partner bank in the six months following their application. The additional likelihood that they borrow from a competing lender during the same window is small and not statistically significant. The estimates are qualitatively similar for thick file borrowers.

For thin file borrowers the patterns differ markedly. The additional likelihood that a thin file applicant just above the threshold borrows from our partner bank within six months following their application is only 13 percentage points (p-val: 0.34) and not statistically significant. However, thin file borrowers just above the EFL threshold are 33 percentage points (p-val: 0.06) more likely to borrow from competing financial institutions. Taking the point estimates literally, comparing Rows 2 and 3 suggests that nearly three quarters of the new loans from thin file borrowers are issued by competing lenders. Rows 4 and 5 of Table 2 demonstrate that this effect is highly concentrated amongst non-bank financial institutions rather than traditional banks. Therefore, for borrowers without established credit histories, much of the benefit derived from our partner bank applying the EFL screening technology accrued to competing financial institutions who responded to our partner bank's loan approvals with approvals of their own. For borrowers with established credit histories, competing financial institutions do not make significantly more loans as a result of the loan approval of our partner bank, as might be expected since competitors have other information with which to assess these borrowers.

While the conclusions of Table 2 are confirmed in the graphical plots in Figure 1, there is a notable additional pattern across several of the plots. To the right of the EFL cutoff, there is a downward slope in the probability of borrowing from our partner bank (for the full sample and for thick file borrowers), and in the probability of borrowing from competing financial institutions (for thin file borrowers). This pattern may be due to the fact that the EFL score is correlated with the size and profitability of a borrower's business and may therefore also be correlated with her alternatives to borrowing from commercial lenders. Applicants with higher EFL scores may borrow at lower rates because they have superior alternatives, such as trade credit and financing investment through retained earnings. Nevertheless, for the full sample and for thick file borrowers there is a clear discontinuity at the threshold in the likelihood of borrowing from our partner bank, and similarly for thin file borrowers with competing financial institutions, providing a visual confirmation of the estimates in Table 2.

In Table 3 we present results from our primary placebo test: estimates of the same RD model but for the outcome of borrowing any time in the six months *preceding* loan application from our partner bank. As in Table 2, every point estimate in Table 3 corresponds to the coefficient of interest, β, in a separate RD model. As expected, the RD estimates for the full sample, for thick file borrowers, and for thin file borrowers are small and statistically insignificant. Alternate specifications are presented in Online Appendix Tables A5 and A6.

Bolstering the evidence for free riding, in Table 4 we focus on thin file applicants and demonstrate that the free riding effects are stronger in regions where our partner bank faces more competition. Note that in contrast to Tables 2 and 3, each *column* in Table 4 represents estimates from a separate model. We explore heterogeneous treatment effects based on log NBFI credit per capita in each locality in which our partner bank has a branch. To examine interaction effects, this table uses an OLS specification of the linear model and the bandwidth around the cutoff chosen to be consistent with the models in Table 2. The coefficient of interest is on the interaction between the degree of our partner bank branch level competition and the discontinuity around the EFL threshold. The estimates in columns 1 and 2 demonstrate a negative and statistically significant relationship between the degree of competition facing our partner bank and the additional probability of a thin file applicant borrowing from our partner bank if they exceed the threshold. The estimates in column 2 imply that increasing the level of competition by one standard deviation would correspond to an additional 9.1 percentage point increase in the likelihood a thin file applicant took a loan from our partner bank as a result of being marginally over the threshold.

The pattern for borrowing from competing financial institutions is reversed, albeit less precisely estimated. The estimates in columns 3 through 6 demonstrate a positive, though not statistically significant relationship between the degree of competition facing our partner bank and the probability of a thin file applicant borrowing from a competing lender around the threshold. Table A7 presents results from our alternative measure of competition, log NBFI branches per capita. The results are qualitatively similar.

Table A8 presents the effects on our partner bank's and on competing lenders' profits, measured using the Inverse Hyperbolic Sine (IHS) function. Note that profits for our partner bank and its competitors vary across the threshold due to 1) the differential likelihood that borrowers are approved for a loan, and 2) the differential likelihood that they accept loans from each lender. Columns 3 and 4 demonstrate that thick file borrowers just above the cutoff are on average 242% (p-val 0.00) more profitable for our partner bank than those just below the cutoff. On the other hand, estimates of our partner bank's additional profits from thin file borrowers are much smaller and not statistically significant. But for NBFIs, thin file borrowers just above the threshold are 301% (p-val 0.04) more profitable than those just below the threshold. Therefore, most of the additional profits resulting from screening new thin file borrowers accrued to competing financial institutions.

Table A9 presents estimates of the effect of the EFL threshold on the resulting defaulted balances for borrowers at our partner bank and its competitors. We find no evidence that marginally approved applicants that took loans from our partner bank were differentially reliable from those who took loans from competing financial institutions.

4. Discussion

In a pilot subsidized by the IDB, our partner bank in Peru adopted a novel screening technology in an effort to expand its loan portfolio among SME borrowers. While our partner bank succeeded in expanding financial inclusion, our estimates suggest that for borrowers without a pre-existing credit history, many of the new loans originated from competing financial institutions. As our partner bank followed a threshold rule whereby applicants who scored above a cutoff were automatically approved for a loan and those who scored below it were not, we can exploit exogenous variation in loan approvals; because competing financial institutions could not see the underlying scores, their additional likelihood to approve borrowers above the score threshold is a result of the loan approval of our partner bank.

A number of auxiliary tests bolster the credibility of our identification strategy and results. We note that no observable demographic or business characteristics are discontinuous at the threshold. A placebo test examining the likelihood to borrow from any financial institution in the six months prior to applying for a loan from our partner bank shows no discontinuity in the threshold. We demonstrate free-riding effects for borrowers without prior borrowing histories but not for those with prior borrowing histories, as would be expected if competing lenders and the borrowers themselves already have strong signals to infer the creditworthiness of borrowers in the latter group. Finally, we show that the free-riding effects are stronger in areas where our partner bank faces more competition from non-bank financial institutions.

This phenomenon greatly reduced the benefit accruing to our partner bank of their efforts to expand financial inclusion. Indeed, after the pilot and subsidies ended, our partner bank decided not to continue using the EFL screening technology. We argued in the introduction that this phenomenon may justify subsidies for efforts to extend loans to new borrowers, including the adoption of novel screening technologies.

We have labeled this competition effect “free riding in loan approvals,” as competing lenders benefitted from our partner bank’s investment to expand its credit supply. This free riding may operate through several channels. On the supply side, competing lenders may have responded directly to the loan approvals of our partner bank, if borrowers shared our partner bank’s loan approval as evidence of their creditworthiness.

On the demand side, borrowers who were approved for a loan from our partner bank may have updated about their own creditworthiness and engaged in more vigorous search for alternative sources of credit. In that case, competing financial institutions would be responding only indirectly to the loan offers of our partner bank. We view this mechanism as inconsistent with the fact that we find no effects on borrowers with an established credit history, who would presumably be similarly emboldened. But we note the borrowers with established credit histories who are approved as a result of surpassing the screening threshold are those who could not receive a loan on the basis of their credit history alone. Thus, it is possible that they did redouble their efforts to find an alternative source of credit but that these efforts were fruitless. This interpretation is consistent with the results of Karlan and Zinman (2009), who find that borrowers randomly selected to receive a loan updated their perceptions about their own creditworthiness.

On the supply side, we can rule out mechanisms in which information spillovers occur through the credit registry, as loan approvals are not recorded in the Peruvian credit registry. And on the demand side, we can rule out that increased demand for loans from competitors was driven by increased loans from our partner bank and complementarities in demand for credit. Of all thin file borrowers to the right of the EFL threshold in the bandwidth of our primary specification, only a quarter of those who took loans from competing lenders also took a loan from our partner bank.

While we cannot pin down a single mechanism, we note that the consequences of each of these are largely the same. In either case, the additional loans resulting from our partner's efforts to expand financial inclusion were primarily issued by other financial institutions. Whether this was a direct response to the loan approvals of our partner bank, or merely an indirect consequence, the justification for subsidizing the expansion of financial inclusion remains.

References

Agarwal, Sumit, Thomas Kigabo, Camelia Minoiu, Andrea Presbitero, and André F. Silva (2018). Financial access under the microscope. Working Paper.

Azevedo, Viviane, Jeanne Lafortune, Liliana Olarte, and José Tessada (2019). Does Formal Credit Lead to More financial inclusion or distress? Results using a strict scoring rule amongst marginal clients in Paraguay. Working Paper.

Burke, Jeremy, Julian Jamison, Dean Karlan, Kata Mihaly, and Jonathan Zinman (2019). Credit Building or Credit Crumbling? A Credit Builder Loan's Effects on Consumer Behavior, Credit Scores and Their Predictive Power. Working Paper.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik (2014a). Robust nonparametric confidence intervals for regression-discontinuity designs. Econometrica 82 (6), 2295–2326.

Calonico, S., M.D. Cattaneo, and R. Titiunik. (2014b). Robust data-driven inference in the regression-discontinuity design." Stata Journal 14 (4), 909–46.

Cattaneo, Matias D., Michael Jansson, and Xinwei Ma (2018). Manipulation testing based on density discontinuity. The Stata Journal 18, (1), 234-261.

Cattaneo, Matias D., Rocio Titiunik, and Gonzalo Vazquez-Bare (2016). Inference in regression discontinuity designs under local randomization. The Stata Journal 16 (2), 331–367.

Djankov, Simeon, Caralee McLiesh, and Andrei Shleifer (2007). "Private credit in 129 countries." Journal of financial Economics 84, (2), 299-329.

Dobbie, Will, Paul Goldsmith-Pinkham, Neale Mahoney, and Jae Song (2016). Bad credit, no problem? Credit and labor market consequences of bad credit reports. No. w22711. National Bureau of Economic Research.

Hahn, Jinyong, Petra Todd, and Wilbert Van der Klaauw (2001). Identification and estimation of treatment effects with a regression discontinuity design. Econometrica 69 (1), 201–209.

Foley, Fritz, Agustín Hurtado-Andrés, Andres Liberman, and Alberto Sepúlveda (2019). The effects of information on credit market competition: Evidence from credit cards." Working Paper.

Hertzberg, Andrew, Jose Maria Liberti, and Daniel Paravisini (2011). "Public information and coordination: evidence from a credit registry expansion." The Journal of Finance 66, (2), 379-412.

Jappelli, Tullio, and Marco Pagano (2002). "Information sharing, lending and defaults: Cross-country evidence." Journal of Banking & Finance 26, (10), 2017-2045.

Karlan, Dean, and Jonathan Zinman (2009). "Expanding credit access: Using randomized supply decisions to estimate the impacts." The Review of Financial Studies 23, (1), 433-464.

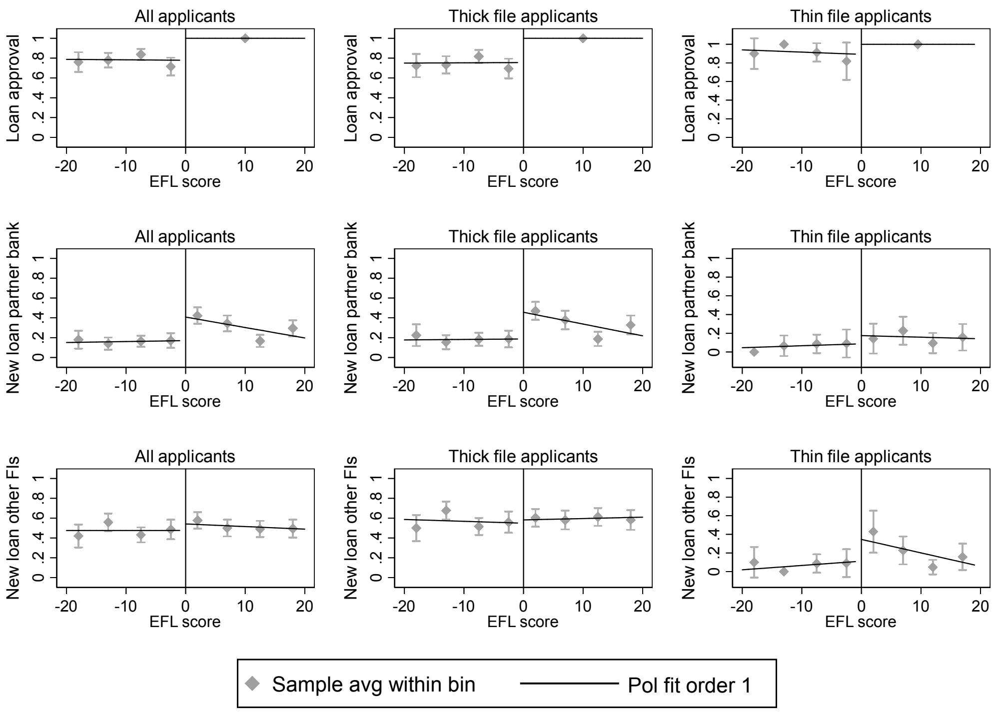
Karlan, Dean, and Jonathan Zinman (2018). "Long-run price elasticities of demand for credit: evidence from a countrywide field experiment in Mexico." The Review of Economic Studies 86, (4), 1704-1746.

Liberman, Andres (2016). "The value of a good credit reputation: Evidence from credit card renegotiations." Journal of Financial Economics 120, (3), 644-660.

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

Sutherland, Andrew (2018). "Does credit reporting lead to a decline in relationship lending? Evidence from information sharing technology." Journal of Accounting and Economics 66, (1), 123-141.

Figure 1: Probability of Loan Approval and Increased Borrowing from Partner and Competing Banks



Notes: The plots were generated using the "rdplot" Stata command developed by Calonico, Cattaneo, and Titiunik (2014b) for a bandwidth of 20 around the EFL score threshold with a global polynomial of order one and 95 percent confidence intervals for each bin. The probability of a new loan is measured by a dummy variable =1 if the amount outstanding of either the DNI or RUC associated with a loan applicant from a given financial institution increases by any amount number in the six months following their application to our partner bank.

Table 1. Background Characteristics

		(1) All sample	(2) Below cutoff		(3) Local pol. 0	(4) Local pol. 1	(5) Local pol. 2
Loan applicant's age	*Mean*	39.272	35.276	*Coeff*	1.023	0.541	-0.620
	Std Dev	(10.859)	(8.119)	*P-value*	(0.953)	(0.880)	(0.621)
	# Obs	1883	116	*# Obs*	243	702	1110
Applicant is female	*Mean*	0.499	0.539	*Coeff*	0.040	0.078	0.084
	Std Dev	(0.500)	(0.499)	*P-value*	(0.281)	(0.231)	(0.379)
	# Obs	1883	256	*# Obs*	538	933	1136
Log (business revenues)	*Mean*	9.983	9.565	*Coeff*	0.210	0.150	0.098
	Std Dev	(1.100)	(0.899)	*P-value*	(0.411)	(0.387)	(0.528)
	# Obs	1883	132	*# Obs*	279	676	996
EFX reject	*Mean*	0.193	0.201	*Coeff*	-0.019	-0.015	0.000
	Std Dev	(0.395)	(0.402)	*P-value*	(0.574)	(0.913)	(0.934)
	# Obs	1862	268	*# Obs*	559	743	839
Thin file at time of test	*Mean*	0.194	0.182	*Coeff*	0.008	-0.018	-0.015
	Std Dev	(0.396)	(0.386)	*P-value*	(0.935)	(0.739)	(0.782)
	# Obs	1883	314	*# Obs*	676	702	962
Equifax score at time of test	*Mean*	636.802	636.366	*Coeff*	4.074	-16.992	-15.254
	Std Dev	(216.553)	(223.292)	*P-value*	(0.748)	(0.526)	(0.829)
	# Obs	1883	287	*# Obs*	598	829	933

Notes: This table shows background characteristics of all loan applicants at the time of applying for a loan with our partner bank. Columns 1 and 2 show the mean, standard deviation and number of observations of each variable. The sample in Column 2 (below the cutoff) is within the optimal bandwidth calculated for the polynomial order 0. Columns 3 through 5 show the regression discontinuity impact estimates for the effect of the EFL tool on each variable using the Stata *rdrobust* command of local polynomial order 0 (Column 3), 1 (Column 4) and 2 (Column 5). The bandwidth, and therefore the number of observations, is optimally selected by the command. All regressions include as control variable the date when the applicant took the EFL tool. Robust bias-corrected p-values (in parentheses) and number of observations are reported below each coefficient.

Table 2. Loan offer and take-up six months after loan application

		(1)	(2)	(3)	(4)	(5)	(6)
		All applicants		Thick-file applicants		Thin-file applicants	
Additional controls		No	Yes	No	Yes	No	Yes
Prob loan approval	*Coeff*	0.220	0.231	0.247	0.252	0.089	0.102
	P-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.206)	(0.146)
	# Obs	1053	933	823	863	233	211
New loan partner bank	*Coeff*	0.289	0.276	0.335	0.330	0.139	0.128
	P-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.263)	(0.336)
	# Obs	648	676	511	524	157	162
New loan other FIs	*Coeff*	0.085	0.061	0.029	0.023	0.333	0.331
	P-value	(0.329)	(0.505)	(0.852)	(0.814)	(0.077)	(0.085)
	# Obs	783	848	672	632	111	111
New loan other banks	*Coeff*	0.008	-0.014	-0.009	-0.032	0.075	0.070
	P-value	(0.965)	(0.715)	(0.747)	(0.523)	(0.434)	(0.485)
	# Obs	752	829	650	708	119	119
New loan other NBFIs	*Coeff*	0.056	0.059	0.024	0.032	0.357	0.364
	P-value	(0.367)	(0.449)	(0.703)	(0.631)	(0.048)	(0.044)
	# Obs	917	906	586	567	111	107

Notes: This table shows regression discontinuity impact estimates for the effect of the EFL tool on the probability of loan approval and the probability of having a new loan in the first six months after the loan application. Each Row X Column represents the coefficient of interest from a seperate regression. All outcome variables are dummy variables. The estimates use the Stata *rdrobust* command of local polynomial order 1. The bandwidth, and therefore the number of observations, is optimally selected by the command. The first two columns show the results for all applicants. Columns 3 and 4 show the results for the sample of thick-file applicants. The last two columns restrict the sample to thin-file applicants. All regressions include as control variable the date when the applicant took the EFL tool. Additional controls in Columns 2, 4 and 6 include the loan applicant's age, gender, business revenues (in logs) and Equifax score at the time of the EFL tool. Heteroskedasticity robust p-values (in parentheses) and number of observations are reported below each coefficient.

Table 3. Pre-application loan take-up (placebo test)

		(1)	(2)	(3)	(4)	(5)	(6)
		All applicants		Thick-file applicants		Thin-file applicants	
Additional controls		No	Yes	No	Yes	No	Yes
New loan partner bank	*Coeff*	-0.013	-0.014	-0.012	-0.011	-0.005	-0.013
	P-value	(0.220)	(0.203)	(0.352)	(0.355)	(0.143)	(0.162)
	# Obs	648	630	606	567	104	111
New loan other FIs	*Coeff*	0.038	0.020	0.033	0.002	0.001	0.003
	P-value	(0.544)	(0.665)	(0.559)	(0.838)	(0.937)	(0.956)
	# Obs	676	702	586	650	169	157
New loan other banks	*Coeff*	0.053	0.038	0.055	0.036	0.063	0.062
	P-value	(0.499)	(0.574)	(0.462)	(0.559)	(0.358)	(0.361)
	# Obs	676	702	567	586	151	151
New loan other NBFIs	*Coeff*	-0.020	-0.035	-0.005	-0.006	-0.064	-0.059
	P-value	(0.956)	(0.707)	(0.885)	(0.916)	(0.109)	(0.143)
	# Obs	829	872	586	606	135	133

Notes: This table shows regression discontinuity impact estimates for the effect of the EFL tool on the probability of having a new loan six months before the loan application. Each Row X Column represents the coefficient of interest from a seperate regression. All outcome variables are dummy variables. The estimates use the Stata *rdrobust* command of local polynomial order 1. The bandwidth, and therefore the number of observations, is optimally selected by the command. The first two columns show the results for all applicants. Columns 3 and 4 show the results for the sample of thick-file applicants. The last two columns restrict the sample to thin-file applicants. All regressions include as control variable the date when the applicant took the EFL tool. Additional controls in Columns 2, 4 and 6 include the loan applicant's age, gender, business revenues (in logs) and Equifax score at the time of the EFL tool. Heteroskedasticity robust p-values (in parentheses) and number of observations are reported below each coefficient.

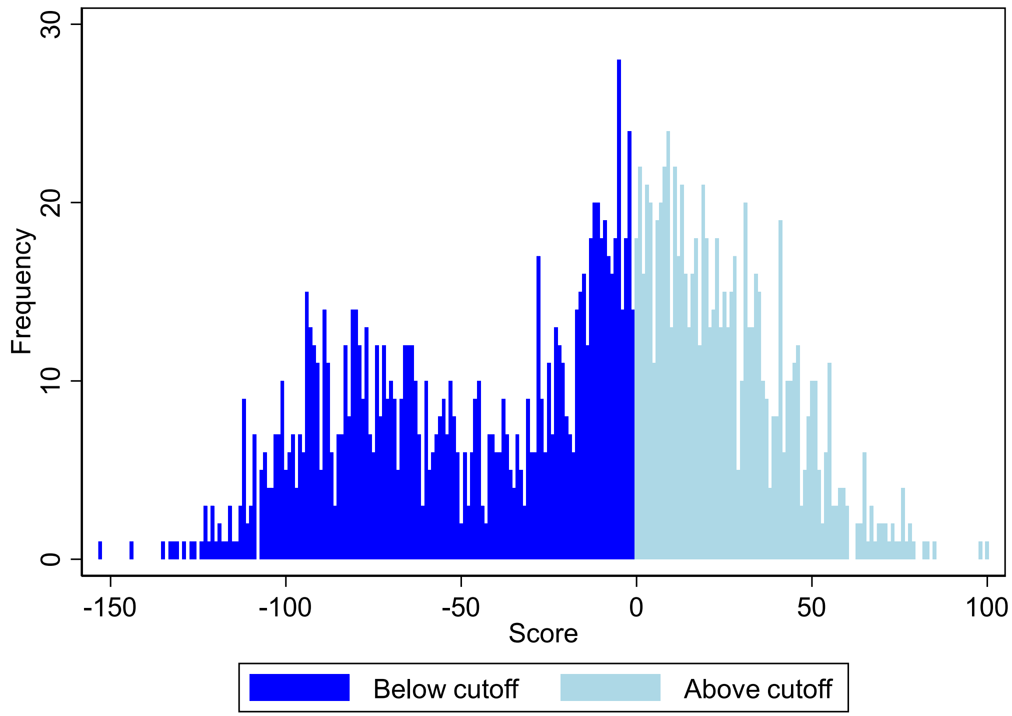
Table 4. Loan take-up six months after loan application, by level of competition

	(1)	(2)	(3)	(4)	(5)	(6)
	New SME loan with partner bank		New SME loan with other FIs		New SME loan with other NBFIs	
Additional controls	No	Yes	No	Yes	No	Yes
EFL pass	0.072	0.073	0.297*	0.306*	0.277*	0.287*
	(0.482)	(0.476)	(0.059)	(0.054)	(0.077)	(0.068)
Log(total NBFI credit per capita)*EFL pass	-0.088**	-0.074*	0.047	0.067	0.021	0.039
	(0.019)	(0.057)	(0.277)	(0.147)	(0.556)	(0.336)
Observations	157	157	111	111	111	111
R-squared	0.054	0.09	0.148	0.197	0.153	0.185

Notes: This table includes data on thin file loan applicants only. It shows OLS estimates of the effect of the EFL tool on the probability of getting a new loan in the first six months after the loan application. Each column is a different regression. Total NBFI credit per capita is measured at the district level. All regressions include as control variables the date when the applicant took the EFL tool, the EFL score, the EFL score interacted with being above the threshold, and log(total NBFI credit per capita). Additional controls in Columns 2, 4 and 6 include the loan applicant's age, gender, business revenues (in logs) and Equifax score at the time of the EFL tool. The constant is not reported. The bandwidth is chosen to be consistent with the regression discontinuity model in Table 2. Hetreoskedasticity obust p-values are reported in parentheses.

Appendix

Figure A1: Histogram of EFL Score (Normalized)



Notes: This figure shows the histogram of the EFL scores for the 1883 loan applicants in our sample. We normalized the EFL scores to zero at the threshold set by our partner institution. All applicants with EFL scores above zero were offered a loan.

Table A1. Background Characteristics (thick file applicants)

		(1)	(2)		(3)	(4)	(5)
		All sample	Below cutoff		Local pol. 0	Local pol. 1	Local pol. 2
Loan applicant's age	*Mean*	39.184	35.200	*Coeff*	1.691	0.819	0.481
	Std Dev	(10.575)	(7.873)	*P-value*	(0.633)	(0.732)	(0.767)
	# Obs	1517	110	*# Obs*	235	606	838
Applicant is female	*Mean*	0.504	0.543	*Coeff*	0.028	0.061	0.095
	Std Dev	(0.500)	(0.499)	*P-value*	(0.399)	(0.489)	(0.286)
	# Obs	1517	223	*# Obs*	460	632	934
Log (business revenues)	*Mean*	10.042	9.567	*Coeff*	0.258	0.185	0.167
	Std Dev	(1.098)	(0.870)	*P-value*	(0.321)	(0.326)	(0.350)
	# Obs	1517	123	*# Obs*	265	567	823
EFX reject	*Mean*	0.226	0.232	*Coeff*	-0.027	-0.030	-0.024
	Std Dev	(0.418)	(0.423)	*P-value*	(0.576)	(0.607)	(0.881)
	# Obs	1507	207	*# Obs*	430	681	751
Equifax score at time of test	*Mean*	628.448	631.843	*Coeff*	1.498	-18.036	-13.078
	Std Dev	(226.230)	(227.399)	*P-value*	(0.671)	(0.574)	(0.891)
	# Obs	1517	223	*# Obs*	460	708	809

Notes: This table shows background characteristics of the thick file loan applicants at the time of applying for an SME loan with our partner bank. Columns 1 and 2 show the mean, standard deviation and number of observations of each variable. The sample in Column 2 (below the cutoff) is within the optimal bandwidth calculated for the polynomial order 0. Columns 3 through 5 show the regression discontinuity impact estimates for the effect of the EFL tool on each variable using the Stata *rdrobust* command of local polynomial order 0 (Column 3), 1 (Column 4) and 2 (Column 5). The bandwidth, and therefore the number of observations, is optimally selected by the command. All regressions include as control variable the date when the applicant took the EFL tool. Heteroskedasticity robust p-values (in parentheses) and number of observations are reported below each coefficient.

Table A2. Background Characteristics (thin file applicants)

		(1)	(2)		(3)	(4)	(5)
		All sample	Below cutoff		Local pol. 0	Local pol. 1	Local pol. 2
Loan applicant's age	*Mean*	39.637	34.731	*Coeff*	0.460	-0.622	-1.813
	Std Dev	(11.973)	(9.392)	*P-value*	(0.624)	(0.686)	(0.623)
	# Obs	366	26	*# Obs*	53	153	206
Applicant is female	*Mean*	0.481	0.490	*Coeff*	0.074	0.085	0.056
	Std Dev	(0.500)	(0.505)	*P-value*	(0.557)	(0.792)	(0.874)
	# Obs	366	49	*# Obs*	107	133	172
Log (business revenues)	*Mean*	9.740	9.432	*Coeff*	0.266	0.106	0.282
	Std Dev	(1.078)	(1.020)	*P-value*	(0.656)	(0.781)	(0.280)
	# Obs	366	44	*# Obs*	92	153	169
EFX reject	*Mean*	0.056	0.045	*Coeff*	0.001	0.015	0.057
	Std Dev	(0.231)	(0.211)	*P-value*	(0.925)	(0.754)	(0.482)
	# Obs	355	44	*# Obs*	95	120	178
Equifax score at time of test	*Mean*	671.429	662.746	*Coeff*	19.812	-14.228	-60.839
	Std Dev	(166.537)	(179.376)	*P-value*	(0.923)	(0.696)	(0.440)
	# Obs	366	59	*# Obs*	135	169	162

Notes: This table shows background characteristics of the thin file loan applicants at the time of applying for an SME loan with our partner bank. Columns 1 and 2 show the mean, standard deviation and number of observations of each variable. The sample in Column 2 (below the cutoff) is within the optimal bandwidth calculated for the polynomial order 0. Columns 3 through 5 show the regression discontinuity impact estimates for the effect of the EFL tool on each variable using the Stata *rdrobust* command of local polynomial order 0 (Column 3), 1 (Column 4) and 2 (Column 5). The bandwidth, and therefore the number of observations, is optimally selected by the command. All regressions include as control variable the date when the applicant took the EFL tool. Heteroskedasticity robust p-values (in parentheses) and number of observations are reported below each coefficient.

Table A3. Loan offer and take-up six months after loan application (local polynomial order 0)

		(1)	(2)	(3)	(4)	(5)	(6)
		All applicants		Thick-file applicants		Thin-file applicants	
Additional controls		No	Yes	No	Yes	No	Yes
Prob loan approval	*Coeff*	0.218	0.228	0.249	0.254	0.082	0.105
	P-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.095)	(0.071)
	# Obs	702	726	650	708	175	140
New loan partner bank	*Coeff*	0.242	0.236	0.293	0.298	0.099	0.105
	P-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.209)	(0.231)
	# Obs	318	318	265	235	111	107
New loan other FIs	*Coeff*	0.061	0.059	0.028	0.011	0.254	0.295
	P-value	(0.266)	(0.300)	(0.594)	(0.729)	(0.039)	(0.026)
	# Obs	598	567	524	511	65	72
New loan other banks	*Coeff*	0.028	0.007	0.023	-0.010	0.061	0.067
	P-value	(0.898)	(0.951)	(0.995)	(0.849)	(0.293)	(0.288)
	# Obs	510	598	434	524	87	87
New loan other NBFIs	*Coeff*	0.033	0.054	-0.020	0.009	0.317	0.369
	P-value	(0.385)	(0.398)	(0.912)	(0.829)	(0.035)	(0.008)
	# Obs	471	567	511	586	44	44

Notes: This table shows regression discontinuity impact estimates for the effect of the EFL tool on the probability of loan approval and the probability of having a new loan in the first six months after the loan application. Each Row X Column represents the coefficient of interest from a seperate regression. All outcome variables are dummy variables. The estimates use the Stata *rdrobust* command of local polynomial order 0. The bandwidth, and therefore the number of observations, is optimally selected by the command. The first two columns show the results for all applicants. Columns 3 and 4 show the results for the sample of thick-file applicants. The last two columns restrict the sample to thin-file applicants. All regressions include as control variable the date when the applicant took the EFL tool. Additional controls in Columns 2, 4 and 6 include the loan applicant's age, gender, business revenues (in logs) and Equifax score at the time of the EFL tool. Heteroskedasticity robust p-values (in parentheses) and number of observations are reported below each coefficient.

Table A4. Loan offer and take-up six months after loan application (local polynomial order 2)

		(1)	(2)	(3)	(4)	(5)	(6)
		All applicants		Thick-file applicants		Thin-file applicants	
Additional controls		No	Yes	No	Yes	No	Yes
Prob loan approval	*Coeff*	0.213	0.225	0.242	0.242	0.136	0.136
	P-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.176)	(0.123)
	# Obs	1354	1288	1006	1085	211	187
New loan partner bank	*Coeff*	0.317	0.306	0.369	0.364	0.163	0.131
	P-value	(0.001)	(0.001)	(0.001)	(0.001)	(0.194)	(0.316)
	# Obs	848	848	672	672	169	169
New loan other FIs	*Coeff*	0.095	0.081	0.035	0.022	0.363	0.342
	P-value	(0.308)	(0.379)	(0.777)	(0.878)	(0.126)	(0.166)
	# Obs	933	917	838	823	157	157
New loan other banks	*Coeff*	-0.010	-0.015	0.000	-0.032	0.062	0.046
	P-value	(0.942)	(0.842)	(0.842)	(0.852)	(0.641)	(0.777)
	# Obs	1038	1053	745	758	153	153
New loan other NBFIs	*Coeff*	0.108	0.106	0.041	0.046	0.402	0.401
	P-value	(0.309)	(0.286)	(0.656)	(0.577)	(0.045)	(0.048)
	# Obs	829	829	758	758	164	162

Notes: This table shows regression discontinuity impact estimates for the effect of the EFL tool on the probability of loan approval and the probability of having a new loan in the first six months after the loan application. Each Row X Column represents the coefficient of interest from a seperate regression. All outcome variables are dummy variables. The estimates use the Stata *rdrobust* command of local polynomial order 2. The bandwidth, and therefore the number of observations, is optimally selected by the command. The first two columns show the results for all applicants. Columns 3 and 4 show the results for the sample of thick-file applicants. The last two columns restrict the sample to thin-file applicants. All regressions include as control variable the date when the applicant took the EFL tool. Additional controls in Columns 2, 4 and 6 include the loan applicant's age, gender, business revenues (in logs) and Equifax score at the time of the EFL tool. Heteroskedasticity robust p-values (in parentheses) and number of observations are reported below each coefficient.

Table A5. Pre-application loan take-up (placebo test, local polynomial 0)

		(1)	(2)	(3)	(4)	(5)	(6)
		All applicants		Thick-file applicants		Thin-file applicants	
Additional controls		No	Yes	No	Yes	No	Yes
New loan partner bank	*Coeff*	-0.009	-0.008	-0.005	-0.003	0.000	0.000
	P-value	(0.247)	(0.268)	(0.601)	(0.637)	(0.169)	(0.167)
	# Obs	361	361	379	379	65	72
New loan other FIs	*Coeff*	-0.003	-0.004	-0.001	0.008	0.027	0.040
	P-value	(0.846)	(0.683)	(0.772)	(0.732)	(0.738)	(0.677)
	# Obs	630	598	487	460	92	87
New loan other banks	*Coeff*	0.040	0.020	0.045	0.022	0.059	0.065
	P-value	(0.492)	(0.666)	(0.550)	(0.717)	(0.261)	(0.261)
	# Obs	434	471	411	411	87	92
New loan other NBFIs	*Coeff*	-0.039	-0.024	-0.043	-0.013	-0.030	-0.028
	P-value	(0.566)	(0.492)	(0.542)	(0.581)	(0.214)	(0.243)
	# Obs	598	538	524	487	87	72

Notes: This table shows regression discontinuity impact estimates for the effect of the EFL tool on the probability of having a new loan six months before the loan application. Each Row X Column represents the coefficient of interest from a seperate regression. All outcome variables are dummy variables. The estimates use the Stata *rdrobust* command of local polynomial order 0. The bandwidth, and therefore the number of observations, is optimally selected by the command. The first two columns show the results for all applicants. Columns 3 and 4 show the results for the sample of thick-file applicants. The last two columns restrict the sample to thin-file applicants. All regressions include as control variable the date when the applicant took the EFL tool. Additional controls in Columns 2, 4 and 6 include the loan applicant's age, gender, business revenues (in logs) and Equifax score at the time of the EFL tool. Heteroskedasticity robust p-values (in parentheses) and number of observations are reported below each coefficient.

Table A6. Pre-application loan take-up (placebo test, local polynomial 2)

		(1)	(2)	(3)	(4)	(5)	(6)
		All applicants		Thick-file applicants		Thin-file applicants	
Additional controls		No	Yes	No	Yes	No	Yes
New loan partner bank	*Coeff*	-0.021	-0.021	-0.018	-0.018	-0.021	-0.011
	P-value	(0.162)	(0.144)	(0.298)	(0.297)	(0.106)	(0.606)
	# Obs	962	933	795	782	196	151
New loan other FIs	*Coeff*	0.044	0.025	0.046	0.028	-0.021	-0.036
	P-value	(0.561)	(0.662)	(0.545)	(0.652)	(0.837)	(0.652)
	# Obs	962	962	838	838	183	172
New loan other banks	*Coeff*	0.049	0.031	0.071	0.050	0.046	0.036
	P-value	(0.543)	(0.714)	(0.512)	(0.648)	(0.828)	(0.954)
	# Obs	917	917	686	686	169	164
New loan other NBFIs	*Coeff*	0.014	0.004	0.016	0.020	-0.089	-0.086
	P-value	(0.865)	(0.872)	(0.865)	(0.766)	(0.058)	(0.105)
	# Obs	933	933	782	758	187	175

Notes: This table shows regression discontinuity impact estimates for the effect of the EFL tool on the probability of having a new loan six months before the loan application. Each Row X Column represents the coefficient of interest from a seperate regression. All outcome variables are dummy variables. The estimates use the Stata *rdrobust* command of local polynomial order 2. The bandwidth, and therefore the number of observations, is optimally selected by the command. The first two columns show the results for all applicants. Columns 3 and 4 show the results for the sample of thick-file applicants. The last two columns restrict the sample to thin-file applicants. All regressions include as control variable the date when the applicant took the EFL tool. Additional controls in Columns 2, 4 and 6 include the loan applicant's age, gender, business revenues (in logs) and Equifax score at the time of the EFL tool. Heteroskedasticity robust p-values (in parentheses) and number of observations are reported below each coefficient.

Table A7. Loan take-up six months after loan application, by level of competition (alternative measure of competition)

	(1)	(2)	(3)	(4)	(5)	(6)
	New SME loan with partner bank		New SME loan with other FIs		New SME loan with other NBFIs	
Additional controls	No	Yes	No	Yes	No	Yes
EFL pass	0.049	0.053	0.307*	0.321**	0.278*	0.292*
	(0.642)	(0.608)	(0.050)	(0.044)	(0.071)	(0.059)
Log(NBFI branches per capita)*EFL pass	-0.121**	-0.102*	0.055	0.079	0.019	0.04
	(0.024)	(0.066)	(0.358)	(0.214)	(0.705)	(0.468)
Observations	157	157	111	111	111	111
R-squared	0.052	0.088	0.146	0.194	0.154	0.186

Notes: This table includes data on thin file loan applicants only. It shows OLS estimates of the effect of the EFL tool on the probability of getting a new loan in the first six months after the loan application. Each column is a different regression. NBFI branches per capita is measured at the district level. All regressions include as control variable the date when the applicant took the EFL tool, the EFL score, the EFL score interacted with being above the threshold, and log(total NBFI branches per capita). Additional controls in Columns 2, 4 and 6 include the loan applicant's age, gender, business revenues (in logs) and Equifax score at the time of the EFL tool. The constant is not reported. The bandwidth is chosen to be consistent with the regression discontinuity model in Table 2. Hetreoskedasticity obust p-values are reported in parentheses.

Table A8. Profits from loans granted six months after loan application

		(1)	(2)	(3)	(4)	(5)	(6)
		All applicants		Thick-file applicants		Thin-file applicants	
Additional controls		No	Yes	No	Yes	No	Yes
Profits all banks	*Coeff*	1.539	1.410	1.105	0.959	4.240	4.036
	P-value	(0.070)	(0.075)	(0.218)	(0.243)	(0.005)	(0.008)
	# Obs	676	702	586	606	104	107
Profits partner bank	*Coeff*	2.152	2.035	2.520	2.417	0.961	0.879
	P-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.189)	(0.243)
	# Obs	648	676	511	524	146	151
Profits other FIs	*Coeff*	0.667	0.542	0.229	0.138	3.032	3.080
	P-value	(0.458)	(0.516)	(0.923)	(0.891)	(0.046)	(0.038)
	# Obs	726	752	672	809	107	107
Profits other banks	*Coeff*	-0.043	-0.193	-0.209	-0.341	0.949	0.922
	P-value	(0.890)	(0.745)	(0.717)	(0.483)	(0.231)	(0.234)
	# Obs	726	783	632	745	111	111
Profits other NBFIs	*Coeff*	0.671	0.578	0.329	0.333	2.951	3.006
	P-value	(0.273)	(0.337)	(0.717)	(0.719)	(0.047)	(0.040)
	# Obs	917	962	650	632	107	107

Notes: This table shows regression discontinuity impact estimates for the effect of the EFL tool on the profits of financial institutions from granting loans to applicants in the first six months after the loan application. Each Row X Column represents the coefficient of interest from a seperate regression. Profits are set to zero for applicants with no new loans with a financial institution, and are measured using the Inverse Hyperbolic Sine (IHS) function. The estimates use the Stata *rdrobust* command of local polynomial order 1. The bandwidth, and therefore the number of observations, is optimally selected by the command. The first two columns show the results for all applicants. Columns 3 and 4 show the results for the sample of thick-file applicants. The last two columns restrict the sample to thin-file applicants. All regressions include as control variable the date when the applicant took the EFL tool. Additional controls in Columns 2, 4 and 6 include the loan applicant's age, gender, business revenues (in logs) and Equifax score at the time of the EFL tool. Heteroskedasticity robust p-values (in parentheses) and number of observations are reported below each coefficient.

Table A9. Amount in default from loans granted six months after loan application

		(1)	(2)	(3)	(4)	(5)	(6)
		All applicants		Thick-file applicants		Thin-file applicants	
Additional controls		No	Yes	No	Yes	No	Yes
Amount in default from loans granted by partner bank	*Coeff*	-0.449	-0.422	-0.510	-0.472	0.111	0.075
	P-value	(0.280)	(0.308)	(0.392)	(0.426)	(0.765)	(0.973)
	# Obs	783	752	586	586	162	157
Amount in default from loans granted by partner bank	*Coeff*	-0.035	-0.019	-0.058	-0.059	-0.077	-0.003
	P-value	(0.797)	(0.903)	(0.659)	(0.608)	(0.086)	(0.446)
	# Obs	726	702	650	686	361	211
Amount in default from loans granted by other Fis	*Coeff*	-0.355	-0.343	-0.438	-0.411	-0.027	-0.090
	P-value	(0.394)	(0.413)	(0.438)	(0.466)	(0.614)	(0.374)
	# Obs	752	726	586	586	135	133
Amount in default from loans granted by other banks	*Coeff*	-0.189	-0.203	-0.248	-0.276	-0.043	-0.108
	P-value	(0.408)	(0.348)	(0.351)	(0.295)	(0.529)	(0.312)
	# Obs	726	803	632	686	135	133
Amount in default from loans granted by other NBFIs	*Coeff*	-0.124	-0.115	-0.159	-0.119	-0.015	0.000
	P-value	(0.901)	(0.902)	(0.897)	(0.979)	(0.548)	.
	# Obs	702	702	543	543	218	172

Notes: This table shows regression discontinuity impact estimates for the effect of the EFL tool on the defaulted amount of applicants 24 months after the loan application. Each Row X Column represents the coefficient of interest from a seperate regression. Defaulted amounts are set to zero for applicants with no new loans with a financial institution, and are measured using the Inverse Hyperbolic Sine (IHS) function. The estimates use the Stata *rdrobust* command of local polynomial order 1. The bandwidth, and therefore the number of observations, is optimally selected by the command. The first two columns show the results for all applicants. Columns 3 and 4 show the results for the sample of thick-file applicants. The last two columns restrict the sample to thin-file applicants. All regressions include as control variable the date when the applicant took the EFL tool. Additional controls in Columns 2, 4 and 6 include the loan applicant's age, gender, business revenues (in logs) and Equifax score at the time of the EFL tool. Heteroskedasticity robust p-values (in parentheses) and number of observations are reported below each coefficient.