Show Me Yours and I'll Show You Mine: Information Sharing in a Competitive Microcredit Market*

Ralph De Haas *(EBRD and Tilburg University: dehaasr@ebrd.com)*

Matteo Millone *(VU University Amsterdam: m.millone@vu.nl)*

Jaap Bos *(Maastricht University: j.bos@maastrichtuniversity.nl)*

This version: 13 November 2018

Abstract

We analyze contract-level data on approved and rejected microloans to assess the impact of a new credit registry in Bosnia and Herzegovina, a country with a competitive microcredit market. Our findings are threefold. First, lending tightens at the extensive margin as loan officers, using the new registry, reject more applications. Second, lending also tightens at the intensive margin: microloans become smaller, shorter and more expensive. This affects both new borrowers and lending relationships established before the registry. In contrast, repeat borrowers whose lending relationship started *after* the registry introduction begin to benefit from larger loans at lower interest rates. Third, information sharing reduces defaults, especially among new borrowers, and increases the return on lending. These findings illustrate how mandatory information sharing can reduce agency problems and safeguard loan quality in competitive credit markets in which borrowers are prone to overborrowing.

JEL codes: D04, D82, G21, G28

Keywords: Microcredit, information sharing, overborrowing, credit registry

* The authors thank Borislav Petric and Igor Duspara for their help with accessing EKI data and Thorsten Beck, Mehdi Beyhaghi, Martin Brown, Claudia Buch, Fabio Castiglionesi, Stijn Claessens, Hans Degryse, Sergei Guriev, Vasso Ioannidou, Artashes Karapetyan, José Liberti, Adrien Matray, Steven Ongena, Andrea Presbitero, Reinhard Schmidt, Jason Sturgess, Andrew Sutherland, Guillaume Vuillemey, Alberto Zazzaro and seminar participants at the EBRD, University of Zurich, Norges Bank, Leuven University, Sveriges Riksbank, VU University Amsterdam, Maastricht University, Tilburg University, University College London (UCL), Netherlands Bureau for Economic Policy Analysis, Bank of Lithuania, ASSA Annual Meetings, European Banking Center Conference (Tilburg University), International Conference on Credit Risk Evaluation (Venice), Development Economics Workshop (Tilburg University), Chicago Financial Institutions Conference, CEPR Annual Spring Symposium in Financial Economics, European Banking Center Network Workshop (Lancaster University), World Bank ABCDE Conference, MoFiR Workshop on Banking (Chicago), Conference of the Society for Institutional & Organizational Economics (Paris), China International Conference in Finance (Xiamen), International Banking Conference (Bocconi University), and the Frontiers in Empirical Finance workshop (University of Swansea) for useful comments. The views expressed are those of the authors and not necessarily of the EBRD. Financial support from the Bocconi Centre for Applied Research in Finance is gratefully acknowledged.

1. Introduction

Over the past three decades, microcredit—granting small loans to poor people—has experienced unprecedented growth in many emerging markets and developing countries. Due to this rapid expansion there are currently about 117 million microcredit clients worldwide.[1] The screening and monitoring of all these borrowers has proven challenging for many microfinance institutions. As microcredit markets have become increasingly competitive and saturated, numerous clients have started to engage in multiple loan-taking (McIntosh, de Janvry and Sadoulet, 2005). Such 'double-dipping' has eroded loan quality and contributed to microcredit repayment crises in countries as diverse as Bangladesh, Bolivia, Cambodia, India, Morocco, Nicaragua, Nigeria and Pakistan.

The spectre of widespread repayment problems among microcredit borrowers raises the question how to financially include poorer segments of the global population without eroding financial stability. Policy makers view public credit registries, which require lenders to share borrower information, as an important tool to manage this trade-off (Schicks and Rosenberg, 2011). Yet, while many countries have recently introduced such registries, these typically only involve commercial banks.[2] Only few countries have made credit reporting mandatory for microlenders (Lyman et al., 2011). Empirical evidence on whether and how credit registries can improve the functioning of saturated microcredit markets hence remains scarce.

To help fill this gap, we use data from an early adopter of a credit registry that includes microfinance institutions—Bosnia and Herzegovina—to trace the impact of mandatory information sharing on the quantity and quality of microcredit. Evaluating the impact of a new credit registry is challenging for two main reasons. First, borrower information is typically only publicly available *after* the registry is set up. Second, even if pre-registry data exist, it is difficult to identify the impact of information sharing if all borrowers are similarly affected. Our data have some unique features that help us surmount these challenges. In particular, we use contract-level information from a large microfinance institution about both accepted and rejected microcredit applications, the reason *why* applications were rejected, and the complete repayment history of each approved loan. Importantly, we have these data for the period before and after the credit registry introduction. This enables us to observe

[1] This is the estimated number of active borrowers in 2015 (source: http://www.themix.org/mixmarket).

[2] In 2014, 80 countries operated a public credit registry that covered at least five percent of the adult population (source: World Bank Doing Business database).

decisions by the same loan officers under different information-sharing regimes as well as to disentangle immediate and longer-term effects.

We combine the time variation in information sharing with cross-sectional borrower variation. For each applicant and approved borrower, we know whether they were new to the lender or a repeat client. Loan officers build up proprietary information through repeat lending (Rajan, 1992 and Boot, 2000) and can re-use this information when lending to the same borrower (Greenbaum and Thakor, 1995 and Agarwal and Hauswald, 2010). We therefore expect that the introduction of a credit registry has a larger impact on new as compared with repeat borrowers.

By way of preview, we find that mandatory information sharing can be an effective tool to improve the quality of microcredit. Information sharing tightens lending at the extensive margin: we observe loan officers rejecting more applications using the new registry information. Using loan-officer fixed effects, we show that lending standards also tighten at the intensive margin. Borrowers receive smaller, shorter and more expensive loans that require more collateral. This affects not only first-time borrowers—that is to say, clients that are new to the lender whose portfolio we study—but also existing lending relationships that had been established before the registry. In contrast, new relationships established *after* the registry introduction start to benefit from larger and longer loans at lower interest rates. This suggests that repeat borrowers can now signal their quality to competing lenders, thus forcing the incumbent lender to offer better terms. Lastly, information sharing results in higher loan quality, especially among first-time borrowers, and increases the return on microcredit. Various robustness and placebo tests confirm that our results reflect the introduction of the credit registry—and the associated improvement in available borrower information—rather than secular trends or model-specification choices.

This paper builds on an extensive theoretical literature, which we review in Section 2, and contributes to an expanding empirical literature on mandatory information sharing. Cross-country evidence suggests that information sharing is associated with less risk taking by banks (Houston et al., 2010; Büyükkarabacak and Valev, 2012) and more lending to the private sector, fewer defaults and lower interest rates (Jappelli and Pagano, 2002). These effects appear stronger in developing countries (Djankov, McLiesh and Shleifer, 2007) and for opaque firms (Brown, Jappelli and Pagano, 2009). Yet, cross-country studies only imperfectly control for confounding factors that may lead to a spurious correlation between information sharing and credit outcomes. They also remain silent about the mechanisms through which information sharing affects credit markets.

A small literature has therefore started to exploit contract-level information to identify the impact of information sharing. These papers study changes in the coverage (of borrowers) or participation (of lenders) of *existing* credit registries. Doblas-Madrid and Minetti (2013) focus on the staggered entry of lenders into a credit bureau for the US equipment financing industry. Entry improved repayment for opaque firms but reduced loan size. Hertzberg, Liberti and Paravisini (2011) find that lowering the reporting threshold of the Argentinian credit registry resulted in less lending to firms with multiple lending relationships. Banks that had negative (but private) information about borrowers reduced their exposure to these borrowers when it was announced that this information would become public. Lastly, Ioannidou and Ongena (2010) find that Bolivian firms switch banks once information about prior defaults is erased and the incumbent lender no longer holds them up.

We also exploit contract-level data and contribute to the literature in three important ways. First and foremost, we are among the first to assess the role of information sharing in a mature microcredit market, a part of the financial system characterized by stark information asymmetries and, increasingly, overindebtedness and repayment problems.[3] Existing work on this topic remains scarce. De Janvry, McIntosh and Sadoulet (2010) analyze the staggered use of a registry by the branches of a Guatemalan MFI. They document tighter screening of borrowers and an improvement in loan quality. Our empirical setting is quite different as, unlike the Guatemalan registry, participation in our setting is mandatory for all MFIs and banks. Moreover, we analyze the impact of the actual registry introduction rather than the (non-random) staggered increase in its use.

Second, by comparing the effect on existing versus new borrowers we can differentiate between the immediate impact of the new registry and its longer-term effects. We show that these effects are very different.

Third, our data are rich in that we observe both rejected loan applications and approved loans; the identity of the loan officer (so that we can observe one and the same loan officer under different information regimes); and *why* individual loan applications were rejected. That is, we see which type of information ('positive' or 'negative') loan officers use to reject

[3] In our setting, microcredit takes the form of individual-liability loans instead of joint-liability (group) loans as pioneered by the Bangladeshi Grameen bank in the 1970s. Microfinance institutions (MFIs) are increasingly moving from joint towards individual-liability credit as the latter is less time consuming and less onerous for borrowers (Cull, Demirgüç-Kunt and Morduch, 2009; de Quidt, Fetzer and Ghatak, 2018). The trend of liability individualization has, by its very nature, eroded the protective role of joint liability. This underlines the need for alternative mechanisms, such as information sharing, to contain agency problems in microcredit markets.

applications. This allows us to document directly *how* lenders use the registry once it becomes available. Lastly, we also observe interest rates, which in earlier studies were either unobserved (Doblas-Madrid and Minetti, 2013) or fixed in the short term (de Janvry, McIntosh and Sadoulet, 2010).

We proceed as follows. Section 2 surveys the related theoretical literature and develops testable hypotheses. Section 3, 4 and 5 then describe our empirical setting, data, and identification strategy, respectively. Section 6 presents our empirical results after which Section 7 concludes.

2. Hypotheses development

The literature on information sharing builds on theories that explore how asymmetric information causes lenders to provide either too little credit or too much credit. Stiglitz and Weiss (1981) show that lenders ration credit when they fear that a market-clearing interest rate will attract riskier borrowers. Some entrepreneurs with ex ante profitable projects are then denied credit. Making borrower information public can reduce such rationing. In contrast, de Meza and Webb (1987) and de Meza (2002) show that when information about entrepreneurial ability is private, too many individuals apply for a loan and some negative NPV projects receive credit. If entrepreneurial ability would instead be publicly observable, then lenders could better tailor interest rates. Marginal entrepreneurs consequently no longer apply for credit and overall lending declines.

Building on these seminal contributions, subsequent theoretical work has explored in detail how information sharing can reduce moral hazard, adverse selection, and over-borrowing. First, moral hazard may decline as borrowers no longer fear that their bank will extract rents by exploiting proprietary information (Padilla and Pagano, 1997). Hold-up problems due to informational lock-in (Sharpe, 1990; Rajan, 1992; von Thadden, 2004) diminish in particular for repeat borrowers. Moreover, with a registry in place, defaulting borrowers lose their reputation in the whole credit market and not just with their current lender. This further reduces moral hazard (Padilla and Pagano, 2000). Theory suggests that both mechanisms improve loan quality and lead to more lending at lower interest rates.

Second, the availability of centralized credit data can reduce adverse selection and bring safe borrowers back into the market (Pagano and Jappelli, 1993). While such improved screening boosts loan quality, the effect on the quantity of lending is ambiguous as more lending to safe borrowers may be offset by less lending to riskier clients.

Third, a credit registry can prevent borrowers from taking loans from multiple banks ('double dipping') instead of applying for one single loan (Hoff and Stiglitz, 1997; McIntosh and Wydick, 2005; and Bennardo, Pagano and Piccolo, 2015). When borrowers can hide total outstanding debt, each loan will be under-priced as new lenders ignore that their loan increases the default risk of existing debt. Sharing (positive) information about pre-existing loans reduces such negative externalities and makes lenders more careful.[4] This leads to fewer and smaller loans with a better repayment record. Because lenders now observe total outstanding debt of existing and new borrowers, they may *increase* their lending rates as under-pricing disappears in the new equilibrium.

To sum up, theory predicts an unambiguously positive effect of information sharing on loan quality. However, the impact on the quantity and price of credit depends on whether the main channel is a reduction in initial overborrowing or a reduction in agency problems. In case of the former channel, lenders reduce lending and increase interest rates while in the latter channel interest rates gradually go down while lending expands. Importantly, the applicability of both channels differs across borrower types. Figure 1 illustrates this in the form of three stylized lending relationships that each consist of a first loan (N) and a repeat loan (R) to the same borrower. At the time of the first loans (N1, N2, and N3) these borrowers are all new to the lender.

[Insert Figure 1 here]

To assess the registry impact on the screening of first-time clients, we can compare loans to new borrowers after the registry introduction (N3) with those to observationally identical new borrowers before the introduction (N1 and N2). We expect that for these new borrowers the immediate impact of information sharing is that the lender now observes all their outstanding debt across all lenders. In case of widespread overborrowing, 'double-dipping' theories predict that the lender will respond to this increased transparency by rejecting more loan applications, reducing loan size, and increasing interest rates. We thus expect:

[4] Degryse, Ioannidou and von Schedvin (2016) use data from a Swedish bank to show that when a previously exclusive firm obtains a loan from another bank, the initial bank decreases its internal limit, suggesting that information sharing allows lenders to condition their terms on loans from others.

Hypothesis 1: *For new loan applicants (N3), the credit registry results in a higher rejection probability as compared with new applicants before the registry (N1 and N2). For new borrowers (N3), the registry leads to smaller, more expensive, and more collateralized loans, of higher quality, as compared with new borrowers before the registry (N1 and N2).*

We expect the immediate tightening of lending at the extensive and intensive margin due to information sharing to be smaller for repeat loans. There are two reasons for this. First, the lender has already built up private information about existing borrowers and the newly available public information (for instance about outstanding debt elsewhere) therefore carries less weight. Second, the credit registry allows repeat borrowers to develop a publicly observable track record of good repayment behavior. This reduces the ability of the incumbent lender to hold-up repeat borrowers. This leads to our second hypothesis:

Hypothesis 2: *For repeat borrowers, the credit registry leads to a higher rejection probability as compared with repeat borrowers before the registry, but this increase in rejection probability is smaller than for new borrowers. Likewise, the credit registry leads to smaller, more expensive, and more collateralized repeat loans, with a higher repayment probability. This tightening at the intensive margin is less than for new borrowers.*

Moreover, we can distinguish between three types of repeat loans: those before the registry (R1), those granted with the registry in place while the previous loan to the same borrower had been granted before the registry (R2), and those granted with the registry in place while the previous loan had also been granted with the registry in place (R3). This allows for two additional predictions.

First, comparing R2 to R1 gives the effect of the registry on lending relationships that were already in place when information began to be shared. This shows whether loan officers update their view of *existing* clients, using the new public information that was not available when they first made a loan to these borrowers. If the new information sheds a negative light on existing clients, we expect that loan conditions for R2 repeat loans tighten as compared with similar pre-registry repeat loans (R1). Such a downward correction may nevertheless be partially offset by the increased ability of loan officers to monitor clients in the new regime.

Hypothesis 3: *For repeat loan applicants whose lending relationship started* before *the registry introduction (R2), the registry results in a higher probability of rejection compared* *with repeat applications before the registry (R1). For repeat borrowers (R2) whose lending relationship started* before *the registry introduction, the registry leads to smaller, more expensive, and more collateralized loans, and an increase in loan quality, compared with repeat borrowers before the registry (R1).*

Hypotheses 1 to 3 state that in a saturated microcredit market with overborrowing, the immediate impact of information sharing is to provide lenders with a complete picture of total outstanding debt of both new *and* existing clients. This 'reality check' reduces loan under-pricing and tightens lending on the extensive and intensive margins.

Second, when we compare repeat loans R3 to R1, the aforementioned 'surprise' element of the registry introduction is no longer present. In case of R3, loan officers have been able to use the registry information right from the start of the lending relationship. Comparing repeat loans R3 and R1 thus gauges the change in the lending equilibrium for repeat borrowers. We expect an improvement in lending outcomes during relationships established after the registry introduction (R3-N3) that exceeds the change during pre-registry relationships (R1-N1). Such a steeper improvement in lending terms (and reduction in the rejection of repeat applications) indicates an increased ability of loan officers to monitor clients as well as a reduction in the scope to extract rents from them. This leads to our fourth and final hypothesis:

<u>Hypothesis 4</u>: *For repeat loan applicants (R3) whose lending relationship started* after *the registry introduction, the registry results in a lower probability of rejection compared with repeat applications* before *the registry (R1). For repeat borrowers (R3) whose lending relationship started* after *the registry introduction, the registry leads to larger, cheaper, and less collateralized loans—and an increase in loan quality—compared with repeat loans* before *the registry (R1).*

3. Empirical setting

3.1. Bosnia and Herzegovina

Bosnia and Herzegovina is a middle-income country with a competitive financial sector that includes 12 microfinance institutions. Domestic credit expanded from 23.4 percent of GDP in 2001 to 67.7 percent of GDP in 2013.[5] Microcredit in Bosnia and Herzegovina almost

[5] Source: World Bank (http://data.worldbank.org/country/bosnia-and-herzegovina).

exclusively takes the form of individual-liability loans. While a private data-collection agency had been present since 2000, most lenders neither used it nor contributed information to it. Participation was voluntary and expensive, and coverage therefore incomplete. Lenders could not check whether loan applicants had already borrowed elsewhere. Loan officers of competing lenders sometimes even disseminated false information about their borrowers. Coordination failures thus prevented voluntary information sharing. This allowed many borrowers to take out multiple microloans at the same time (Maurer and Pytkowska, 2011).[6]

In response to this institutional gap and growing overindebtedness, the Bosnian central bank began to establish a public credit registry (Centralni Registrar Kredita, CRK) in 2006. Yet, it was only in July 2009 that participation became mandatory for all lenders, both banks and microfinance institutions. This is also the month in which EKI, the microfinance institution whose loan portfolio we analyze, started to provide information to the registry and began using it. Discussions with loan officers indicate that the July 2009 registry introduction marked a sudden improvement in the available information about loan applicants. No other financial regulation was introduced at this time.

The Bosnian credit registry requires lenders to submit a report each time a loan is disbursed, repaid, late, or written off. The registry contains both 'negative' information on past loan defaults and 'positive' information on pre-existing loans of the applicant. It also includes data on whether applicants have a guarantor or are a guarantor themselves. When loan officers contact the registry, they buy separate files that contain either negative or positive information. The registry keeps borrower information for five years. Each loan applicant also has a credit score that reflects current debt (if any) and their past repayment performance. This score is calculated using uniform regulatory guidelines for credit-risk assessment and ranges from A (best) to E (worst). For instance, after 15 days of late payment a borrower moves from category A ('Good') to B ('Late').

The central bank monitors whether reporting follows the appropriate formatting and undertakes random checks on data quality. Registry information is therefore regarded as

[6] The saturated and competitive nature of the Bosnian microcredit market was also revealed by a randomized controlled trial conducted in 2009 with a large local MFI (Augsburg et al., 2015). The goal of the experiment was to assess whether the MFI could profitably target somewhat poorer and riskier clients by incentivizing a random set of loan officers to take more risk. The resulting portfolio of marginal clients showed a default rate that was three times as high as the regular portfolio, suggesting that loan officers had already been pushing lending to its limits and that there were no viable borrowers left unserved in this competitive credit market.

comprehensive and reliable.[7] Lenders are required to include a clause in each loan contract in which the borrower agrees to a credit check at the registry. Borrowers are thus aware that their repayment performance is recorded and may be shared with other lenders.

3.2. The lender

We use data from EKI, a provider of individual-liability microcredit. Founded in 1996, EKI lends through a network of 15 branches across both parts of the country (the Republika Srpska and the Federation of Bosnia and Herzegovina). Most borrowers are sole proprietorships: formal or informal firms without a legal distinction between the owner and the business. Borrowers are therefore personally liable for their loans. Most of them are small entrepreneurs that are not covered by rating agencies or auditing firms.

EKI loan officers collect all loan-applicant information, including from the credit registry, to make an initial lending decision. They fill out an electronic site-visit form with information on the borrower, their credit history, and the available collateral. Initial lending decisions are discussed by the branch-level loan committee and applications are then approved or rejected. A branch employs on average 14 loan officers. Officers' pay is a function of both the quantity of new loans disbursed (flow) and the quality of their outstanding loan portfolio (stock). Like other lenders in the saturated Bosnian microcredit market, EKI loan officers could not observe pre-existing debt of either new loan applicants or existing clients before the registry introduction. As one loan officer put it: "*Before the introduction of the credit registry, we were basically blind.*"

EKI did not make any changes to its lending policies around the time of the introduction of the credit registry. Throughout the period 2007-2010, it had ample access to funding and funding costs did not change materially.

4. Data

We have access to all loan applications received and all loans granted by EKI. Figure A1 in the Appendix summarizes the loan applications (panel A) and approved loans (panel B) during the window June 2007-July 2011 around the introduction of the credit registry.[8] We

[7] Some information manipulation may occur. Yet, while submitting information to the registry is mandatory, using the data is voluntary and subject to a small fee. Our data show that the registry is actively used, suggesting that lenders attach value to it. The registry receives over 240,000 requests a month.

[8] Table A2 in the Appendix provides variable definitions and data sources.

also show the distribution of new versus repeat borrowers. Before the registry, 55 percent of all loan applications and 55 percent of all approved loans concerned new borrowers. After the introduction of the registry, these percentages drop to 43 and 41 percent, respectively.

Table 1 shows that the rejection rate almost doubles, from 8.8 to 16.4 percent, after the introduction of the credit registry (the remainder of the applications was approved or, in a few cases, withdrawn by the applicant). Appendix Table A1 indicates that the rejection rate increases for both new and repeat applicants. As expected, new applicants are still rejected more frequently. Conditional on acceptance, we observe a reduction—from 89.8 to 82.7 percent—in the portion of the requested amount that is granted (Table 1). This holds again for both new and repeat borrowers (Table A1).

An interesting feature of our data is that we know *why* loans were rejected, as loan officers are required to enter the reason for declining an application into the management information system. We split rejections into those using private versus public information. The former are based on data that EKI collected itself, either in the past or during the current screening. Rejections due to public information are based on either 'positive' information about outstanding debt elsewhere or 'negative' information about repayment problems. Both types of public information became easily available with the introduction of the registry while they were much more difficult to access before.

Panel A of Table 1 shows a clear shift in the rejection reasons with the credit registry in place. Loan officers now rely significantly more on public data about applicants, both 'positive' information about loans elsewhere and 'negative' information about repayment difficulties. The likelihood that a loan is rejected based on public information increases almost four times, suggesting that the registry led to an important change in loan officer behavior. Table A1 (panel A) shows that the new public information not only leads to more rejections of new but also of repeat borrowers. This indicates that the registry provides loan officers with information that complements the private information they already have about existing clients. Note that also before the registry some loans were rejected because of public information as such data were available for some larger applicants.

[Insert Table 1 here]

For the 116,517 loans approved between June 2007 and July 2011, we have information on their size, maturity, interest rate, collateral, and purpose. We know whether and when there was a late payment, whether the loan was written off and, if so, how much principal and interest was recovered. In all, we observe the complete borrowing history of 79,937 borrowers. We also know the identity of the 375 loan officers that granted the loans. The average officer approved 20 (17) loans per month before (after) the registry introduction.

For each borrower, we know their income, education, gender, and employment status (panel B). After the introduction of the registry, the composition of the borrower pool does not change much. The median borrower age is 41 years and around 60 percent of borrowers is male. The median loan amount is almost three times the median monthly household income of borrowers. The median maturity is two years and the annual nominal interest rate is 21 percent (20.1 before and 21.6 after the registry introduction).[9] Borrowers use the loans mainly for business purposes, with about half of all loans used to buy movable assets such as equipment and vehicles. Most loans are collateralized, typically by some form of personal collateral and/or a guarantor. In line with progressive lending, repeat loans tend to be larger, longer, and cheaper (Table A1, panel B).

Our measure of loan quality, *Problem loan*, is a dummy equal to one if a loan was written off. For each non-performing loan, we observe the date when the borrower was first in arrears (>30 days) and we use this as the default event in our hazard analysis (see Section 6.5). We do not take the write-off date as our default indicator because its timing depends more on the bank's discretion than on borrower behavior. It would therefore be a less clean signal of the start of repayment problems. Before the introduction of the registry, 10.1 percent of all microloans defaulted, and this number went down to 2.8 percent after the introduction (Table 1). At the same time, the number of days that the average loan was late declined from 4.2 to 4.1 days although there is wide variation. Due to this improved repayment performance, the return on microcredit went up from 18.1 percent before to 21.6 percent after the registry introduction (Section 6.7 analyzes lending profitability in more detail).

5. Empirical methodology

5.1. Impact on the extensive and intensive lending margins

This section introduces the empirical framework to test our four hypotheses. We use a difference-in-differences framework to systematically compare loan applications by new versus repeat borrowers. We first assess the extensive margin—the probability that an application is rejected—and then the intensive margin (loan amount, maturity, interest rate

[9] Consumer price inflation was 7 percent in 2008 (http://data.worldbank.org/country/bosnia-and-herzegovina).

and collateral). Our baseline specification focuses on applications and approved loans in a time window of one year before and one year after the introduction of the credit registry. This baseline OLS regression model is:

$$Y_{ilt} = \alpha_1 \cdot Credit\ registry_t + \alpha_2 \cdot New_{il} + \beta \cdot I_{ilt} + \gamma \cdot X_{ilt} + \varepsilon_{ilt} \tag{1}$$

where Y_{ilt} is one of our outcomes for loan or loan application i to loan officer l in month t; $Credit\ registry_t$ is a dummy variable that is one for observations after June 2009 (when the credit registry was in place); New_{il} is a dummy that is one for loans and loan applications by clients of loan officer l who had never borrowed before from EKI; I_{ilt} is an interaction term between $Credit\ registry_t$ and New_{il}; X_{ilt} is a matrix of covariates and ε_{ilt} is the error term. Our covariates X_{ilt} include loan-level controls, such as dummies for loan types, and key borrower characteristics, such as age and gender. We cluster standard errors at the month-loan officer level. Results are quantitatively and qualitatively similar when we cluster by month, branch, or month-branch.

A key parameter of interest is β: the additional impact of mandatory information sharing on loan outcomes for new borrowers. If mandatory information sharing has a larger impact on new borrowers as compared with repeat borrowers, as per Hypothesis 2, then the interaction between the registry dummy and *New* will be positive for the outcomes *Loan rejected*, *Interest rate* and *Collateral* and negative for *Loan amount* and *Loan maturity*.

To measure this interaction coefficient more precisely, we also estimate:

$$Y_{ilt} = A_l + B_t + \beta \cdot I_{ilt} + \gamma \cdot X_{ilt} + \varepsilon_{ilt} \tag{2}$$

where A_l and B_t are loan officer and month fixed effects to control for omitted lender characteristics and economy-wide shocks, respectively. If information sharing matters differentially for new borrowers, even after controlling for loan officer fixed effects, then this is strong evidence that our results are not driven by omitted local variables.

Finally, we estimate:

$$Y_{ilt} = C_{lt} + \beta \cdot I_{ilt} + \gamma \cdot X_{ilt} + \varepsilon_{ilt} \tag{3}$$

where C_{lt} identify loan officer-month fixed effects to absorb all factors, such as local business cycle effects, that affect all borrowers of the same loan officer in the same period.

Unbiased estimates should reflect the introduction of information sharing rather than differences between new and repeat borrowers. We therefore use propensity-score matching to ensure that new and repeat borrowers are comparable. Matching borrower and loan characteristics also circumvents the issue of jointness of loan terms (Brick and Palia, 2007). We match microloans on borrower and loan characteristics and calculate propensity scores with bias-corrected nearest-neighbor matching with replacement (Abadie, Drukker, Herr, and Imbens, 2004). This double-robust estimator yields unbiased estimates when either the propensity-score matching model or the linear regression model is correctly specified (Robins, 2000). There is ample common support with less than one percent of all observations not being supported.

Like most countries, Bosnia and Herzegovina was not immune to the global financial crisis. One may therefore wonder whether any effects we find should be partly attributed to the crisis rather than the introduction of the credit registry. We provide three pieces of evidence to show that this is unlikely. First, and most importantly, our data show that immediately after (but not before) the introduction of the credit registry, loan officers started to reject more loan applications based on registry information. This 'smoking gun' points directly to the registry causing the observed changes in lending behavior. Second, we provide an extensive set of placebo tests that show that our results quickly disappear if we let our registry 'treatment' start just one or two quarters earlier (that is, when we move the start date closer to the crisis but further from the actual registry introduction). Third, we find a strong *positive* effect of the new registry on loan quality. This is difficult to reconcile with the idea that our results would pick up a crisis effect, as the crisis would arguably have had a negative rather than a positive effect on borrower quality.

5.2. Impact on loan quality

The second part of our analysis deals with the impact of mandatory information sharing on repayment performance and loan quality. We start by estimating a simple linear probability model using the same specifications as equations (1) to (3), but using a binary variable indicating default as the dependent variable. In a next step, we then define the hazard rate as the probability that a borrower is late on their repayment at time t conditional on regular repayment up to that point. A hazard function allows us to model not only whether a borrower defaults but also how the default probability changes over time.[10] Our variable of interest is the time between disbursement and the first instance of significant late (>30 days) repayment. We do not use the write-off date as our default indicator because its timing depends more on the lender's discretion than on the borrower's default date. The hazard model allows us to compare the development of hazard rates before and after the introduction of mandatory information sharing and for new versus repeat borrowers.

An advantage of hazard models is their ability to deal with censoring, which occurs when a loan is repaid or when the life of a loan extends beyond the sample period. Such right censoring may yield biased and inconsistent estimates in static probability models (Ongena and Smith, 2001). A semi-parametric model (Cox and Lewis, 1966; Cox, 1972) can deal with right censoring as the log-likelihood function accounts for the ratio of completed versus non-completed loans.[11] Let T measure the amount of time before the first late repayment of the loan. The hazard function $h(t)$ is the probability of repayment being late at time t conditional on regular repayments until then:

$$h(t) = \lim_{\Delta t \to 0} \left\{ \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} \right\} \tag{4}$$

We can model the distribution of time until first late repayment as a survivor function:

$$S(t) = P(T \geq t) \tag{5}$$

The relationship between the survivor function and the hazard function is then:

$$h(t) = \frac{-d \log S(t)}{dt} \tag{6}$$

We can now estimate the effect of a set of time-varying covariates Xt on the distribution of time to default with the proportional hazard model:

[10] See Ongena and Smith (2001), Ioannidou, Ongena and Peydró (2015) and Jiménez, Ongena, Peydró and Saurina (2014) for recent applications of duration analysis in the empirical banking literature.

[11] Left censoring can bias estimates as well but is not an issue in our case as we only observe new loans.

$$h(t) = \lim_{\Delta t \to 0} \left\{ \frac{P(t \leq T < t + \Delta t | T \geq t, X_t, \beta)}{\Delta t} \right\} = h_0(t) exp(\beta' X_t) \quad (7)$$

where h_0 represents the baseline hazard when covariates are set to zero: $X=0$. Covariates shift the baseline hazard without affecting the shape of the hazard function. In the Cox (1972) semi-parametric approach the functional form of h_0 is not specified. The model uses the ranking of duration times to estimate the β parameters via maximum likelihood methods and relies on two assumptions. First, it assumes continuous time, as the presence of tied events in discrete time makes ranking impossible. Since late repayments are only observed at intervals, we deal with tied events with the approximation by Breslow (1974). Second, the Cox model assumes proportionality, which implies time fixed β coefficients. We relax this assumption by also estimating a model where the effect of covariates X_t can change over the life of the loan.

We will check the robustness of our results to the functional form of the hazard rate by estimating two parametric specifications using a Weibull and an exponential distribution. The Weibull distribution is expressed as:

$$h(t) = h\alpha t^{\alpha - 1} \quad (8)$$

where α measures duration dependence. If $\alpha>1$ the hazard rate increases with time. The exponential distribution is a Weibull distribution characterized by a constant hazard rate ($\alpha=1$): the probability of late repayment is constant over time (Kiefer, 1988).

6. Results

6.1. Information sharing and loan rejections

Table 2 provides estimation results to explain the probability that a microcredit application is rejected. In addition to *Credit registry*, *New borrower* and their interaction term, all specifications include our standard applicant and loan covariates. Moreover, each specification is based on a bias-corrected matched sample to ensure that new and repeat borrowers are comparable. Columns 1 to 3 show the baseline specification estimated with a linear probability model, so that we can use month and loan officer fixed effects (column 2) or their interaction (column 3); can interpret the coefficients as marginal effects; and prevent problems with interaction effects in nonlinear models (Ai and Norton, 2003). A possible disadvantage of linear probability models is that fitted values might fall outside the 0,1 bounds. However, in our case more than 92 percent of the linear predictions have a value between zero and one.

[Insert Table 2 here]

The introduction of the credit registry is associated with a large and statistically significant increase in the probability that a loan application gets rejected, all else equal. According to column 1, the marginal probability of rejection increases by 7.2 percentage points for repeat borrowers. This impact is clearly stronger for new borrowers: the interaction term of *New borrower* and *Credit registry* is positive and significant. In line with Hypothesis 1, the rejection probability increases by an additional 3.8 percentage points for new borrowers. In line with Hypothesis 2, the total impact of the registry is about 50 percent stronger for new borrowers than it is for repeat borrowers. This also holds when we add month and loan officer fixed effects (column 2) or month *x* loan officer fixed effects (column 3).

Columns 4 to 6 present Tobit regressions to assess whether the introduction of the registry also made loan officers more cautious, conditional on loan acceptance, in terms of the percentage of the requested loan amount that they granted. We find this to be the case. The registry led to a reduction in the percentage of the requested loan amount that was granted of 5.5 percentage points for repeat borrowers and 9 percentage points (that is, an additional 3.5 percentage points) for new borrowers.

The finding that information sharing increases the probability that an application is rejected, for new borrowers but also for repeat ones, suggests that the newly available information made loan officers more cautious. Table 3 looks at this more closely by estimating regressions for different types of repeat loans as per the classification of Figure 1. Column 1 confirms that the registry increased the probability that a repeat loan got rejected. Conditional on approval, loan officers also started to grant a smaller portion of the requested loan amount (column 4). Columns 2 and 5 show that this also holds for a more constrained sample in which we only consider second and third repeat loans (dropping the less common 4^{th}, 5^{th}, etc. loans) that have the same purpose as the previous loan to the same client (for example, both loans were intended to buy fixed assets). We continue to find a strong negative impact of information sharing and the estimated coefficients are only marginally smaller.

In columns 3 and 6, we differentiate between the effect of the registry on repeat loans where the previous loan was granted before the registry introduction (R2 in Figure 1) versus repeat loans where the previous loan was granted with the registry already in place (R3 in

Figure 1). The comparison group consists of pre-registry repeat loans (R1 in Figure 1). In line with Hypothesis 3, we find that the negative impact of the registry is driven by repeat loans to borrowers whose previous loan was disbursed *before* the registry. In contrast, the coefficient for repeat loans to clients that already received at least one loan after the registry is smaller and estimated imprecisely.

These results indicate that the credit registry made loan officers adjust their views about existing clients downwards. This effect is absent for repeat loan applicants about whom public information was available right from the beginning (R3). For them the rejection probability is no different than for otherwise similar repeat clients before the introduction of information sharing (R1). When we run a Wald test on the difference of coefficients we find that the coefficients are significantly different from each other at the 1 percent level.

To recap, the evidence so far indicates that the registry led to a decline in the likelihood of loan approval for new clients and, to a lesser extent, existing clients. The latter effect is driven by repeat loans to borrowers with whom the lender had already established a lending relationship before the registry introduction and about whom information sharing revealed new information that made loan officers revise their views downwards.

[Insert Table 3 here]

In Table 4, we assess what information causes the increased scrutiny of loan officers. We present multinomial logit regressions that link the probability of loan rejection to various types of borrower information. The dependent variable is categorical and indicates whether an application was rejected due to negative registry information (i.e., information about past defaults), positive registry information (i.e., information about outstanding debt elsewhere), or private information. The baseline option is that the loan application got accepted. As discussed before, private information are data that EKI collected itself, either in the past or during the current screening. This includes information about the character of the loan applicant or the quality of the business proposal. It also includes rejections due to negative feedback from neighbors or other clients as well as unsatisfactory financial ratios or a bad credit history with EKI itself.

We estimate the effect of the credit registry on rejections due to the newly available public information—negative (column 1) or positive (column 2)—or private information (column 3). We do this separately for new borrowers (panel A), all repeat borrowers (panel B), our narrow set of repeat borrowers (i.e., only second and third loans for the same purpose as the previous loan, panel C), and while splitting repeat borrowers into those whose previous loans were granted before or after the introduction of information sharing (panel D).

Panel A reveals that the increased scrutiny of new applicants is indeed driven by the registry information—both positive and negative. The average marginal probability that a new client is rejected due to unsatisfactory negative (positive) public information is 4.7 (3.8) percentage points higher after the introduction of the credit registry. The registry does not affect the likelihood of rejection due to private information. Panels B and C show that both types of registry information also reduce the probability that an application by a repeat borrower was accepted. Especially new information about previous defaults or late repayments with other lenders led loan officers to revise their views downwards.

Panel D shows interesting variation across repeat loans. Column 1 indicates that the newly available negative information (on repayment problems with other lenders) affects repeat loans irrespective of whether the lending relationship was started before or after the registry introduction. The probability of loan rejection based on negative information increases by about four times (an increase in the marginal rejection probability of 2.9 percentage points) for repeat loans to clients whose previous loan was granted before the registry. This captures the combined effect of negative registry information on both the screening and monitoring of clients. In contrast, rejection rates due to negative information increase 2.5 times (an increase in the marginal rejection probability of 1.9 percentage points) for repeat loans to borrowers whose previous loan had been granted with the registry already in place. This captures how negative registry information helped strengthen the monitoring of existing clients. The data show that one important role of the registry is to provide loan officers with up-to-date information on existing clients, allowing them to reject applications from clients that have defaulted elsewhere since they last took a loan from EKI.

In contrast, the new positive information (on outstanding debt elsewhere) *only* affects borrowers when the previous loan was granted before the registry. The average marginal rejection probability goes up by two percentage points as compared with an equivalent repeat request just before the introduction of information sharing. There is no impact on repeat loans to borrowers with whom a relationship was started after the registry introduction. This shows that 'positive' information (about outstanding debt elsewhere) mainly helps loan officers at the screening stage and not so much the monitoring stage of the lending process. The increase in rejections due to public information on pre-existing debt is in line with theories that stress that loans from other lenders act as strategic substitutes when firms' debt capacity is limited (Bizer and DeMarzo, 1992; Kahn and Mookherjee, 1998; Parlour and Rajan, 2001).

[Insert Table 4 here]

6.2. Information sharing and loan terms

We proceed by analyzing the change in lending conditions at the time of the credit registry introduction. We consider the *Loan amount*, *Loan maturity*, *Interest rate* and *Collateral* (the sum of personal, social and third-party collateral) and again compare the impact of the registry on repeat versus new borrowers. As before, new clients are new to EKI but may have borrowed from other lenders in the past.

Table 5 shows that information sharing is accompanied by a reduction in loan amounts (panel A) and maturities (panel B) and an increase in interest rates (panel C) and required collateral (panel D). These effects are statistically significant, stronger for new borrowers and hold when including the standard borrower and other covariates. The unreported covariate coefficients show that older, more educated, higher-income and urban borrowers receive larger loans at lower interest rates. All these results hold when adding loan officer fixed effects (column 2), loan officer and month fixed effects (column 3), or loan officer *x* month fixed effects (column 4).[12]

[Insert Table 5 here]

After the introduction of the registry, average loan size drops by 15.5 percent for repeat borrowers and, in line with Hypothesis 2, by just over 20 percent for new borrowers. The same pattern emerges when looking at maturity, with loans shortening by 12.4 percent (97 days) for existing borrowers and by 15.3 percent (120 days) for new borrowers. These smaller and shorter loans also become more expensive: interest rates increase by 0.68 and 0.36 percentage points for new and repeat borrowers, respectively. In a similar vein, collateral requirements go up after the introduction of the credit registry by 0.19 extra collateral items per loan. For new borrowers the number of required collateral items increases by 0.27. The increased reliance on collateral is in line with US evidence by Doblas-Madrid and Minetti (2013) and theoretical work by Karapetyan and Stacescu (2014) on the complementarity between information sharing and collateral. Borrowers with a (now

[12] The same holds when we match to correct for possible longitudinal changes in the borrower pool. New borrowers before and after the registry introduction are very similar along various observable characteristics. This suggests that the lender did not react to information sharing by shifting to different borrower types.

observable) blemished credit history become more likely to face collateral requirements.[13] In all, our results indicate that the introduction of the registry leads loan officers to significantly tighten lending conditions along several margins.

While the substantial reduction in average loan size after the registry introduction is in line with a tightening of credit standards, it could also reflect borrowers starting to take out smaller loans from multiple lenders. If successful, such borrowers may partially or fully offset the lower loan amounts we observe. However, such a multiple-borrowing strategy would be difficult to implement for loans that are collateralized by fixed assets that can only be pledged once (Bosnia and Herzegovina has a well-functioning collateral registry). We therefore re-ran the regressions in Table 5 for the subset of loans used to finance either buildings or other fixed assets and which are collateralized by these assets. In addition, we reran the regressions for the remainder of the sample, i.e., loans not used to finance fixed assets. We find significant coefficients with a magnitude comparable to those in Table 5 for both sub samples, suggesting that our results indeed reflect tighter supply-side conditions.

6.3. Information sharing and loan terms: Repeat borrowers

Table 5 showed that the registry tightened lending conditions for new as well as repeat borrowers. In Table 6, we focus on repeat borrowers and compare—as we did for the extensive margin in Table 3—three types of repeat loans: those granted before the introduction of the registry (the benchmark group shown as relationship 1 in Figure 1), those granted after the registry introduction but where the previous loan was granted before (relationship 2), and those that were granted after the registry introduction and where the previous loan had also been granted with the registry in place (relationship 3). The dependent variables now measure the percentage change in loan outcomes (amount, maturity, interest rate, collateral) since the previous loan to the same borrower.

[Insert Table 6 here]

The odd columns in Table 6 show the effect of the registry on the change in terms (compared to the previous loan to the same borrower) for all second and third loans. In line

[13] This fits with a broader empirical literature (Roszbach, 2004; Rice and Strahan, 2010; Berger, Scott Frame and Ioannidou, 2011) and theoretical work (Boot, Thakor, and Udell, 1991; Inderst and Mueller, 2007) highlighting that observably riskier borrowers are more likely to be required to pledge collateral.

with Table 5, there appears to be a decline in the progressiveness of lending terms. That is, after the credit registry, repeat loans increase less in size and become costlier relative to the previous loan to the same borrower. The even columns split these effects by type of repeat loan. It becomes clear that, in line with Hypothesis 3, they are driven by repeat loans after the introduction of information sharing where the previous loan occurred *before* the registry (relationship 2 in Figure 1). All else equal, the size of these repeat loans declines 39.5 percentage points faster and the interest rate increase is 8.7 percentage points higher (relative to the previous loan). The registry revealed new information (about outstanding debt or repayment problems elsewhere) that made loan officers tighten the lending terms for borrowers to whom they had already been lending *before* the registry.

To filter out such one-off 'surprise effects' due to the introduction of the registry during ongoing lending relationships, we compare the change in loan terms between first and repeat loans during relationships that were *started just after* the registry introduction with the change in loan terms during relationships that *ended just before* the registry introduction. This comparison, summarized in the last row of coefficients in Table 6, gives a cleaner estimate of the equilibrium change in lending conditions during lending relationships. Importantly, and in line with Hypothesis 4, we find that with the registry in place, repeat loans grow faster in size and length (compared to first-time loans) while interest rates decline more rapidly.

Figure 2 visualizes the results of Tables 5 and 6 for relationships consisting of three consecutive microloans. We compare relationships that ended just before information sharing (type 1) with those started soon after the registry introduction (type 3). In line with Table 5, the terms for first-time loans tighten significantly. Loan amounts drop by 15.3 percent and interest rates go up by 3.1 percent. Moreover, the maturity of first-time loans declines on average by 12.2 percent and collateralization increases by 6.6 percent (not shown).[14]

At the same time, however, the progressiveness of microcredit increases due to the registry. In line with Hypothesis 4, lending relationships that *start* in the new regime, so that public borrower information is available right from the beginning, see sharper increases in loan amount (and maturity) and stronger reductions in interest rates (and collateral) as compared with similar relationships that ended just before the registry. Figure 2 shows that at the time of the third loan, the initial tightening is overcome. Compared to the pre-registry equilibrium, third loans are 12.4 percent larger and 3.6 percent cheaper. In equilibrium, when

[14] These numbers differ slightly from those in Table 5. While that table is based on all first-time loans, Figure 2 focuses on first-time loans in relationships that consist of at least three loans with the same loan purpose.

the registry impacts both initial screening and subsequent monitoring, information sharing has a positive effect on the loan terms for borrowers who successfully repay their loans.

We can interpret Figure 2 through the lens of existing theories on relationship lending and information sharing. Note that the intertemporal pattern of loan amounts and interest rates is relatively flat before the registry. Successful borrowers are rewarded with only slightly larger and cheaper repeat loans. Qualitatively this appears in line with Boot and Thakor (1994), who show that interest rates can decrease with the length of the relationship as the build-up of private information reduces the riskiness of lending over time.

However, the intertemporal interest-rate curve steepens once the credit registry is introduced. First-time borrowers start to pay more while repeat borrowers start to pay less. This is difficult to reconcile with Boot and Thakor (1994) who would predict a reduction in interest rates especially for first-time borrowers. The steeper downward-sloping curve instead aligns with theories that stress that information sharing increases lender competition. Before information sharing, EKI charged first-time borrowers a lower-than-competitive interest rate. At the same time, successful repeat borrowers were charged a higher-than-competitive interest rate. This pattern is in line with Sharpe (1990), Petersen and Rajan (1995), Bouckaert and Degryse (2006) and Gehrig and Stenbacka (2007), who all predict that lenders in less competitive markets (such as in the absence of information sharing) smooth interest rates over time. That is, repeat borrowers get charged more because it is difficult for them to switch to an outside lender. They know that they may get pooled with low-quality borrowers and be offered an unattractive interest rate. The incumbent lender knows this as well and can therefore hold up these clients and extract rents. These rents can then be used to cross-subsidize first-time borrowers (for whom agency problems are most severe).

Implicit cross-subsidization becomes difficult in a more competitive market.[15] With a credit registry in place, lenders anticipate that good clients may eventually be poached by outside lenders who can now observe borrower performance (Boot, 2000 and Ioannidou and Ongena, 2010). The reduction in market power of the incumbent lender and the increased ability of (reputable) clients to switch to competitors, forces the incumbent to reward repeat clients with larger and cheaper loans. With information sharing, competition for repeat

[15] In fact, the bottom panel of Figure 2 closely resembles Figure II in Petersen and Rajan (1995). Note that their theoretical model only predicts this relationship for low-quality firms. Whether it also holds for the average firm depends on the extent to which high-quality firms are charged a mark-up in the low-competition equilibrium.

borrowers goes up (and interest rates down) while competition for first-time loans goes down (and interest rates up). This is exactly what we see in the data.

If information sharing enables outside lenders to identify high-quality borrowers, poaching can be targeted towards those borrowers. The ability to poach good borrowers should have the biggest impact on borrowers that are observationally more opaque. In unreported regressions, we therefore test whether the credit registry had a different effect on high versus low-income clients. We expect that access to borrower information is more salient for riskier clients as proxied by a lower monthly income. Without information sharing, it will be especially difficult for outside lenders to distinguish between good and bad clients in this pool of relatively risky clients. When we interact *Credit registry* with borrower income, we continue to find a negative base effect on loan amount and a positive effect on the interest rate and collateral for first-time borrowers. As expected, the effect of the credit registry is significantly stronger for low-income borrowers (cf. footnote 15).

[Insert Figure 2 here]

6.4. Information sharing and loan quality: Non-parametric results

Figure 3 provides a non-parametric view of microcredit quality in the form of a Kaplan and Meier (1958) survival analysis for the period June 2008 to July 2010. The graphs show how the probability that a borrower has not (yet) defaulted changes over time (horizontal axis, in quarters). At disbursement (*t=0*) the probability of survival is 1 but then gradually erodes over time. The graphs thus show the inverse of the cumulative default probability.

Panel A compares, for the whole sample period, the survival probability of new versus repeat borrowers. In this context, right censoring will affect disproportionately the more recent loans. The correct hazard rate is then calculated as the ratio of loans that have defaulted at time *t* over the remaining loans (Ongena and Smith, 2001). The key point to take away from this panel is the slightly lower survival probability of new borrowers (the small difference between both curves is nevertheless statistically insignificant as shown by a logrank test (p-value=0.00)).

[Insert Figure 3 here]

In panel B, we start to compare the survival behavior of loans granted before and after the introduction of the credit registry. On the one hand, we expect the impact of the credit registry to be concentrated among new borrowers as the information asymmetry between lender and loan applicant is largest. On the other hand, to the extent that the registry (also) had an impact on borrower behavior, we expect an improvement in repayment behavior among repeat borrowers as well as these now realize that a default will 'cost' them more in terms of foregone future borrowing opportunities.

Already after a few quarters a large gap opens between both curves: loans granted with the credit registry in place have a significantly higher survival probability compared with loans approved without mandatory information sharing. After a year, this difference in repayment behavior is a substantial 6.5 percentage points. This is the first piece of evidence that points to a positive impact of information sharing on loan quality. A striking aspect of panel B is that the difference between both loan types already emerges during the first quarters after loan disbursement. Indeed, the probability of a loan not being late in the first six quarters after disbursement increases from 89.8 percent before the credit registry introduction to 96.4 percent afterwards. This difference declines only very little over time.

Panels C and D split panel B into a panel for first-time loans (C) and one for repeat loans (D). As expected, this shows that most (but not all) of the increase in loan quality that became apparent in panel B, is driven by new borrowers. In line with Hypothesis 2, the impact of the credit registry is larger for new borrowers. The one-year survival probability of these loans in the year after the credit registry introduction (97 percent) is eight percentage points higher than in the preceding year (89 percent). The increase in survival probability is only four percentage points for repeat loans. Yet, even in panel D we see that the registry introduction is accompanied by a clear upward shift of the survival function: at each point in time repeat borrowers are less likely to default, suggesting that information sharing also increases borrower discipline.

A logical next step is to further split panel D of Figure 3 into the three types of lending relationships of Figure 1. We do this in Figure 4 which is based on the full population of repeat (second and third) loans. The solid line at the bottom shows the benchmark survival probability for repeat loans granted before the introduction of the credit registry (implying that the first loan was also granted before the registry). The striped and dotted lines indicate the survival probability of repeat loans granted after the introduction of information sharing. Both lines indicate an upward shift in survival probability. This shift is largest for repeat loans as part of lending relationships that were started after the registry introduction. Here, the survival probability is highest as both the initial screening at the start of the lending relationship and the subsequent monitoring during the lending relationship benefitted from the new registry information. In contrast, the striped line in the middle shows a smaller upward shift as these repeat loans are part of relationships that were started when public borrower information was not yet available.

[Insert Figure 4 here]

6.5. Information sharing and loan quality: (Semi-)parametric results

In Table 7, we first present simple linear probability models of loan default. Column (2) shows that once the credit registry is introduced, the probability of default is 2.4 percentage points lower for repeat borrowers and 3.3 percentage points lower for first time borrowers.

[Insert Table 7 here]

In Table 8 we then proceed by providing semi-parametric and parametric evidence on the impact of mandatory information sharing on loan quality. As discussed in Section 5.2, an important advantage of hazard models—where the hazard rate is the probability of a borrower defaulting at time *t* conditional on having repaid up to that point—is that they deal properly with right censoring. We stratify by branch so that the form of the underlying hazard function varies across branches (the coefficients of the remaining covariates are assumed constant across strata). Hence, we do not need to assume a particular form of interaction between the stratifying covariates and time.

[Insert Table 8 here]

In column 1, we present the results of a semi-parametric Cox proportional hazard model while columns 2 and 3 show equivalent specifications using a parametric exponential and Weibull model, respectively. In line with Figure 3, the results show that the registry introduction is associated with a statistically significant reduction in the hazard rate. Importantly, this effect is almost 50 percent higher for new EKI borrowers. The second line shows that this difference between new and repeat borrowers only emerges after the registry introduction. This is in line with the very similar solid lines in panels C and D of Figure 3. While before the registry, the survival probability of new borrowers is somewhat lower than that of repeat borrowers, the (semi-)parametric results in Tables 7 and 8 show that this difference is not significant when controlling for an extensive set of covariates and fixed effects. The estimated coefficients for these control variables are not shown but have the expected sign and, in most cases, display a statistically significant relationship with the hazard rate. For instance, we find that older and more educated borrowers pose less risk while longer and larger loans tend to have higher repayment risk, all else equal. The parametric exponential model in column 2 and the parametric Weibull model in column 3 produce very similar results. The latter shows an *Ln(alpha)* of -0.645, indicating that the hazard rate decreases over time as borrower risk is front loaded.

We also estimated (unreported) models that allow covariates to change during the life of the loan. To do so, we modify the structure of our dataset so that the number of observations on each loan equals the number of periods between disbursement and either repayment or default (Singer and Willett, 1993). The hazard rate now not only depends on the loan and borrower characteristics at the time of disbursement, but also on a set of other variables—including the registry introduction—that may change during the life of the loan. The results of these models with time-varying predictors are fully in line with Table 8: default risk is lower once the registry is introduced and this holds especially for first-time borrowers.

Table 9 provides semi-parametric evidence similar to the graphical evidence in Figure 4. We focus on the Cox proportional hazard model and the sample consists of repeat loans only. The first column uses the sample of all repeat loans while the second and third columns focus on second and third loans that have the same loan purpose as the previous loan. Columns 1 and 2 show that repeat loans granted after the introduction of information sharing are significantly less likely to default even when controlling for a battery of borrower and loan covariates. The size of the effect declines when we move from column 1 to 2, suggesting that about a third of the improvement in loan quality stems from changes in the observable characteristics of repeat versus first-time loans. Still, two thirds of the quality improvement results from a better screening and/or monitoring of observationally similar clients.

[Insert Table 9 here]

In column 3, we again split the post-registry repeat loans into those where the first loan was pre-registry and those where the first loan was post-registry as well. That is, as before, we compare the two types of post-registry repeat loans with a benchmark group of observationally similar repeat loans that were disbursed before the start of information sharing. In line with Figure 4, these estimates show that the improvement in loan quality (compared to the benchmark group of repeat loans before the registry introduction) is about

2.5 times as large for repeat loans that are part of lending relationships that were started when information sharing was already in place.

6.6. *Information sharing and loan quality: Late repayment*

Table 10 analyzes the impact of the credit registry on the number of days that microcredit is paid late. As most loans are paid on time (zero late payment) we estimate Tobit regressions. As before, we include *Credit registry* and *New borrower* as well as their interaction term. The introduction of the registry is accompanied by a significant reduction in the number of days that the typical first-time loan is repaid late. The coefficients in column 4 imply an average reduction of 0.6 days (from 4.4 days late repayment by first-time borrowers before the registry). We find no impact on late payment by repeat borrowers.

Columns 5 and 6 show multinomial logit regressions where the base outcome is that microcredit is always paid on time (at most one day late). We are interested whether the credit registry affected the likelihood that borrowers were at any time either between 2 and 15 days late (column 5) or between 16 and 30 days late (column 6). The 15-day threshold is important because this is when loans get reclassified from credit score A ('Good') to B ('Late'). While this classification was private information before the introduction of the credit registry, it became public afterwards. Borrowers may thus have started to try harder to avoid being downgraded from score A to B once the registry was in place. Indeed, our results indicate that the reduction in late payment among new borrowers is exclusively driven by a sharp decline in the likelihood that borrowers were between 16 and 30 days late. This is in line with borrowers exerting additional effort to stay below the 15-day threshold to avoid blemishing their (now public) track record.

[Insert Table 10 here]

6.7. *The impact of information sharing on lender profitability*

The introduction of mandatory information sharing affected microcredit along several margins: more applications were rejected while granted loans became smaller, shorter and more expensive. At the same time, loan quality increased as repayment went up. What has been the combined impact of these adjustments on the lender's profitability?

To answer this question, we first evaluate the profitability of EKI in the year before (June 2008–June 2009) and the year after (July 2009–July 2010) the introduction of the credit registry. We calculate the present value of all microcredit disbursed in each of those years.

For the first year all values are discounted to June 2008 and for the second one to July 2009. We use a weighted average of the interest rate on all debt funding to EKI as the discount rate.

For each year we then calculate the present value of total loan disbursements, the probability of loan default, the net present value of the loans, and the net present value per dollar lent. The light-grey bars in Figure 5 show for the year after the introduction of the credit registry a substantial decline in the present value of lending (measured as the total amount of new lending, net of fees, discounted back to the beginning of each period using the lender's average funding cost). The present value of total lending goes down by 49.7 percent due to the combined effect of more loan rejections and smaller loans.

At the same time, however, we know that the credit registry led to a substantial decline in the probability of default (loans that were at least 30 days late and were subsequently written off) from 10 to 4 percent (right axis). Because of this strong increase in repayment performance, the *net* present value of all loans (disbursements minus repayments) declined by only 31.2 percent. Indeed, the net present value per US$ lent increased from 11 to 14 cents (right axis) and the internal rate of return (IRR) on lending increased from 17.6 to 21.8 percent (an increase of 23.9 percent, not shown). Given that the cost of capital was roughly the same during both periods, and under the assumption that operational costs did not change substantially, these numbers indicate that mandatory information sharing significantly increased the profitability of the lender.

[Insert Figure 5 here]

In Table 11, we analyze the impact of information sharing on lender profitability at the loan level. We calculate the realized return (Haselmann, Schoenherr and Vig, 2018) on microcredit earned in the year before and the year after the introduction of the credit registry. For loans that were fully repaid, this return is simply the interest rate charged. For loans that were defaulted on, the realized return is the weighted average of the return before the moment of default and the return after default took place. Before default, the return is again simply the interest rate charged over the (gradually declining) outstanding amount. After the default, the return is negative and reflects the amount of the loan outstanding at the time of default as well as the portion of that amount that the lender managed to recover (if any).

Table 11 shows a significant increase in the return on microcredit of about 1 percentage point, reflecting the better repayment behavior due to the registry. This is an economically meaningful improvement equal to 6 percent of the pre-registry average return on loans. The interaction terms in columns 2 to 4 indicate that this positive effect is particularly prominent among new loans. These results can be interpreted in light of the model by Padilla and Pagano (1997) which highlights the ambiguous effect of information sharing on lender profitability. On the one hand, borrowers increase their efforts, and this boosts loan quality and profitability. On the other hand, information sharing increases lender competition as informational rents are reduced. This puts pressure on interest rates and (future) profit margins. While we find evidence for both mechanisms, the net effect is clearly positive for the lender whose portfolio we study. While information sharing reduces the interest rates that the lender can charge to well-performing repeat borrowers, this effect is more than offset by the substantial increase in borrower quality—especially among new borrowers.

[Insert Table 11 here]

6.8. Robustness and placebo tests

We subject our results to several robustness and placebo tests. Appendix Table A3 presents such tests for both the *New borrower* variable and the interaction between *Credit registry* and *New Borrower*. In the first three columns we vary the time window to estimate the effect of the registry introduction. Our regular window is one year before and one year after the introduction. In column 1, we use a narrower symmetric window (September 2008-April 2010). In column 2, we use a wider window (January 2008-December 2010) while in column 3 we use the widest window that our data permit (June 2007-July 2011). We show results for *Loan rejected* (*Proportion granted*) in the top (bottom) half of the table. In all cases the statistical and economic significance is very similar to our baseline estimates in Table 2.

Columns 4 to 6 provide placebo tests to check that hitherto undetected secular trends do not drive our results. This is also a test of the parallel-trends assumption: since no credit registry was introduced at the placebo dates, we should not detect any impact. In column 4, we move our two-year window a year forward. We thus take the true treatment period as the control period and let the treatment start in July 2010 (basically assuming that the credit registry was introduced a year later than in reality). In column 5, we move our two-year window a year backwards. We take the true control period as the treatment period and assume that the registry was already introduced in July 2008. This placebo test is useful because it checks whether we are not picking up any impact of the global financial crisis. As expected, in both columns the interaction term between *Credit registry* and *New borrower* is no longer statistically significant while the coefficient for *New borrower* continues to be.

Finally, in column 6, we randomly allocate borrowers to the category of new or repeat borrowers. We repeat this random allocation a thousand times and show the average result (here the treatment starts in July 2009, the actual date of the registry introduction). As expected, now both the coefficient for *New borrower* and for the interaction term are no longer statistically significant. Together these robustness and placebo tests give us confidence that the results in Table 2 indeed reflect a change in lending behavior due to the registry introduction in July 2009.

Appendix Tables A4 and A5 report similar placebo and robustness tests for the intensive margin. The same covariates as in Table 5 are included but not shown for reasons of brevity. Our results again disappear if we move the start of the treatment to a fictitious date one year earlier or later (Table A4). And as before, our coefficient of interest is robust to broadening or widening the window around the start date of the registry (Table A5).

Figures A2 and A3 visualize related placebo tests. Figure A2 shows the coefficient estimates and a 95 percent confidence interval for the interaction terms *Credit registry*New borrower* as in column 3 of Table 5. The value at time *t* shows the coefficients for the actual timing of the registry introduction. The values at *t-1*, *t-2*, and *t-3* indicate the estimates when introducing the registry one, two, or three quarters earlier than the real date. Figure A2 shows that when we artificially bring the registry introduction forward, the placebo impacts quickly dissipate and essentially become zero just one or two quarters before the actual introduction date. Figure A3 shows a similar placebo test for the impact of information sharing on loan quality (based on the interaction term in column 1 of Table 8). We conclude that our findings indeed capture the shift in information-sharing regime and not a secular longer-term trend.[16]

Lastly, we have assumed that outcomes would have developed in parallel for new versus repeat borrowers if no credit registry had been introduced. Any trend differences that appear once information sharing is introduced can then be attributed to the registry. We follow the strategy of Hertzberg, Liberman and Paravisini (2016) to shed further light on this assumption. To do so, we include a series of dummies in equation (2) that activate each

[16] In unreported tests we start the treatment period in October 2010 for *Loan amount*, September 2006 for *Loan maturity* and February 2009 for *Interest rate*. These placebo start times are based on a Clemente-Montañés-Reyes unit-root test, which indicates a possible break point in that month for each dependent variable. We also perform a test where the placebo treatment starts in September 2008—the collapse of Lehman Brothers—and ends with the introduction of the registry in July 2009. If we simply picked up a crisis effect, it should show up here. Our original results disappear in all these placebo tests as well.

bimonthly period for up to one year before and after the registry introduction. Each dummy is interacted with our *New borrower* variable:

$$Y_{ibt} = A_b + B_t + \sum_{\tau=-12}^{12} \beta_\tau \cdot I(\tau)_{ibt} + \gamma \cdot X_{ibt} + \varepsilon_{ibt} \qquad (9)$$

Appendix Figure A4 visualizes the estimated coefficients and 99 percent confidence intervals for the 12 interaction terms (6 before and 6 after the registry introduction) for our four dependent variables. If the set of interaction terms is insignificant before the introduction of the credit registry, we cannot reject the hypothesis of parallel trends. The graphs indeed show that the interaction terms are typically statistically insignificant before the registry introduction—indicating the absence of significant pre-trends—but significantly different from zero afterwards. F-tests confirm this visual impression.

7. Conclusion

Microfinance is rapidly coming of age. Lender competition is intensifying, microcredit markets are becoming increasingly saturated, and liability individualization is exposing lenders to more credit risk. Are credit registries a useful mechanism to broaden financial access without endangering financial stability? To help answer this question, we present direct evidence of what happens when lenders are required to start sharing borrower information. We use data from a microfinance institution in a middle-income country, a setting typical of the many markets where microfinance is currently expanding rapidly.

We document how information sharing makes loan officers initially more cautious at the extensive and intensive margins. The new public data lead to a significantly higher rejection probability for both new loan applicants and pre-existing lending relationships. Loan officers update their views of existing clients downwards on the basis of both the newly available negative information (on repayment problems with other lenders) and positive information (on outstanding debt elsewhere). The registry thus provides lenders with information that complements the private information they already had about their clients.

For new lending relationships that were established *after* the registry introduction, and about whom new public information was therefore available right from the start, we find a strong positive effect on the evolution of loan terms for borrowers who successfully repay their loans. While information sharing makes first-time microloans smaller and more expensive, repeat borrowers are now better off once they have successfully repaid two loans. Because of these impacts, we observe a clear improvement in repayment performance that has an economically meaningful impact on the average return on loans (equal to almost one percentage point).

Our results underline the importance of making both negative and positive borrower information available. Both types of information are valuable and are actively used once they become public. Negative information is used both at the start and during relationships. We observe that loan officers reject existing clients that ask for a loan renewal if they observe in the registry that the client has experienced repayment problems with another lender. In contrast, information on outstanding debt elsewhere is mainly used at the start of a new lending relationship (and for ongoing relationships at the time of the registry introduction). The strong increase in loan rejections due to debt elsewhere indicates that outside loans act as strategic substitutes. One role of information sharing in microfinance is thus to reduce coordination problems (as in Bolton and Scharfstein, 1996) and limit overall indebtedness. As such, our findings support a benign view of transparency in credit markets: more public information allows loan officers to make better lending decisions after taking outstanding debt elsewhere into account.

We do not find evidence that the availability of public borrower information reduces (or increases) the efforts of loan officers to collect private information. The data on loan rejections show only a very slight increase (decrease) in the probability of rejection due to private information for repeat (new) borrowers. We also do not find evidence of public information being inferior than private information (possibly leading to adverse selection problems). Indeed, the increased reliance on public information is accompanied by a strong and persistent improvement in loan quality.

In all, our findings illustrate how mandatory information sharing can help microfinance institutions to make better lending decisions. Yet, we also show that the introduction of a registry does not necessarily lead to an immediate increase in microcredit availability. Indeed, the short-term impact can be a reduction in lending as the newly available information leads to a rational reassessment of borrowers' total indebtedness and repayment performance. Our findings therefore help to explain why, when a new credit registry was recently introduced in the United Arab Emirates, one that was widely expected to increase banks' appetite to lend, the introduction was instead followed by a sharp increase in loan rejections.[17] Our results should help to manage similar expectations in countries that will soon introduce new credit registries (like Ukraine) or that are contemplating doing so (like India).

References

Abadie, A., Drukker, D., Herr, J.L. and Imbens, G.W., 2004. Implementing matching estimators for average treatment effects in Stata. *Stata Journal* 4, 290-311.

Abadie, A. and Imbens, G.W., 2011. Bias-corrected matching estimators for average treatment effects. *Journal of Business & Economic Statistics* 29(1), 1-11.

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

Ai, C., Norton, E.C., 2003. Interaction terms in logit and probit models. *Economics Letters* 80, 123-139.

Augsburg, B., De Haas, R., Harmgart, H. and Meghir, C., 2015. The impacts of microcredit: Evidence from Bosnia and Herzegovina. *American Economic Journal: Applied Economics* 7(1), 183-203.

Bennardo, A., Pagano, M., Piccolo, S., 2015. Multiple bank lending, creditor rights, and information sharing. *Review of Finance* 19, 519-570.

Berger, A.N., Scott Frame, W., Ioannidou, V., 2011. Tests of ex ante versus ex post theories of collateral using private and public information. *Journal of Financial Economics* 100(1), 85-97.

Bizer, D.S. and DeMarzo, M., 1992. Sequential banking. *Journal of Political Economy* 100(1), 41-61.

Bolton, P. and Scharfstein, D.S., 1996. Optimal debt structure and the optimal number of creditors. *Journal of Political Economy* 104(1), 1-15.

Boot, A.W.A., 2000., Relationship banking: What do we know? *Journal of Financial Intermediation* 9, 7-25.

Boot, A.W.A., Thakor, A.V., 1994., Moral hazard and secured lending in an infinitely repeated credit market game, *International Economic Review* 35, 899-920.

Boot, A.W.A., Thakor, A.V., Udell, G.F., 1991. Secured lending and default risk: equilibrium analysis, policy implications and empirical results. *Economic Journal* 101, 458-472.

[17] See http://www.thenational.ae/business/banking/adib-consumer-loan-rejections-soar-10-after-bank-adopts-credit-bureau-data.

Bouckaert, J., Degryse, H., 2006. Entry and strategic information display in credit markets. *Economic Journal* 116(513), 702-720.

Breslow, N., 1974. Covariance analysis of censored survival data. *Biometrics* 30(1), 89-99.

Brick, I.E, Palia, D., 2007. Evidence of jointness in the terms of relationship lending. *Journal of Financial Intermediation* 16, 452-476.

Brown M., Jappelli, T., Pagano, M., 2009. Information sharing and credit: Firm-level evidence from transition countries. *Journal of Financial Intermediation* 18(2), 151-172.

Büyükkarabacak, B., Valev, N., 2012. Credit information sharing and banking crises: An empirical investigation. *Journal of Macroeconomics* 34, 788-800.

Cox, D.R., 1972. Regression models and life-tables. *Journal of the Royal Statistical Society. Series B (Methodological)* 34(2), 187-220.

Cox D.R., Lewis, P.A. 1966. *The Statistical Analysis of Series of Events*. London Chapman and Hall.

Cull, R., Demirgüç-Kunt, A., Morduch, J., 2009. Microfinance meets the market. *Journal of Economic Perspectives* 23(1), 167-92.

de Janvry, A., McIntosh, C., Sadoulet E., 2010. The supply-and demand-side impacts of credit market information. *Journal of Development Economics* 93(2), 173-188.

de Meza, D., 2002. Overlending? *Economic Journal* 112(477), F17-F31.

de Quidt, J., Fetzer, T., Ghatak, M., 2018. Commercialization and the decline of joint liability microcredit. *Journal of Development Economics* 134, 209-225.

de Meza, D., Webb, D.C., 1987. Too much investment: a problem of asymmetric information. *Quarterly Journal of Economics* 102, 281-292.

Degryse, H., Ioannidou, V., von Schedvin, E., 2016. On the non-exclusivity of loan contracts: An empirical investigation. *Management Science* 62(12), 3510-3533.

Djankov, S., McLiesh, C., Shleifer, A., 2007. Private credit in 129 countries. *Journal of Financial Economics* 84(2), 299-329.

Doblas-Madrid, A., Minetti, R., 2013. Sharing information in the credit market: Contract-level evidence from U.S. firms. *Journal of Financial Economics* 109(1), 198-223.

Gehrig, T., Stenbacka, R., 2007. Information sharing and lending market competition with switching costs and poaching. *European Economic Review* 51, 77-99.

Greenbaum, S.I. and Thakor, A.V., 1995. *Contemporary Financial Intermediation*, Dryden Press, New York.

Haselmann, R., Schoenherr, D., Vig, V., 2018. Rent seeking in elite networks. *Journal of Political Economy*, 126(4), 1638-1690.

Hertzberg, A., Liberti, J.M., Paravisini, D., 2011. Public information and coordination: Evidence from a credit registry expansion. *Journal of Finance* 66(2), 379-412.

Hoff, K., Stiglitz, J.E., 1997. Moneylenders and bankers: Price-increasing subsidies in a monopolistically competitive market. *Journal of Development Economics* 55(2), 485-518.

Houston, J.F., Lin, C., Lin. P., Ma, Y., 2010. Creditor rights, information sharing, and bank risk taking. *Journal of Financial Economics* 96(3), 485-512.

Inderst, R., Mueller, H.M., 2007. A lender-based theory of collateral. *Journal of Financial Economics* 84, 826-859.

Ioannidou, V., Ongena, S., 2010. "Time for a change": Loan conditions and bank behavior when firms switch banks. *Journal of Finance* 65(5), 1847-1877.

Ioannidou V., Ongena, S., Peydró, J-L., 2015. Monetary policy, risk-taking and pricing: Evidence from a quasi-natural experiment. *Review of Finance* 19(1), 95-144.

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

Jiménez. G., Ongena, S., Peydró, J-L., Saurina, J., 2014. Hazardous times for monetary policy: What do twenty-three million bank loans say about the effects of monetary policy on credit risk? *Econometrica* 82(2), 463-505.

Kahn, C. and Mookherjee, D., 1998. Competition and incentives with nonexclusive contracts. *Rand Journal of Economics* 29, 443-465.

Kaplan, E.L., Meier, P., 1958. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association* 53(282), 457-481.

Karapetyan, A., Stacescu, B., 2014. Does information sharing reduce the role of collateral as a screening device? *Journal of Banking & Finance* 43, 48-57.

Kiefer, N.M., 1988. Economic duration data and hazard functions. *Journal of Economic Literature* 26(2), 646-679.

Lyman, T., Lythgoe, T., Miller, M., Reille, X., and Sankaranarayan, S., 2011. Credit reporting at the base of the pyramid: Key issues and success factors. FORUM 1, Washington, D.C.: CGAP and IFC.

Maurer, K., Pytkowska, T., 2011. Indebtedness of microfinance clients in Bosnia and Herzegovina. Results from a comprehensive field study. European Fund for Southeast Europe Development Facility, mimeo.

McIntosh, C, de Janvry, A., Saoulet, E., 2005. How rising competition among microfinance institutions affects incumbent lenders. *Economic Journal* 115(506), 987-1004.

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

Ongena, S., Smith, D.C., 2001. The duration of bank relationships. *Journal of Financial Economics* 61(3), 449-475.

Padilla, J., Pagano, M., 1997. Endogenous communication among lenders and entrepreneurial incentives. *Review of Financial Studies* 10(1), 205-36.

Padilla, J., Pagano, M., 2000. Sharing default information as a borrower discipline device. *European Economic Review* 44(10), 1951-1980.

Pagano, M., Jappelli, T., 1993. Information sharing in credit markets. *Journal of Finance*. 48(5), 1693-1718.

Parlour, C.A., Rajan, U., 2001. Competition in loan contracts. *American Economic Review* 91(5), 1311-1328.

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

Rajan, R.G., 1992. Insiders and outsiders: The choice between informed and arm's length debt. *Journal of Finance* 47, 1367-1400.

Rice, T., Strahan, P.E., 2010. Does credit competition affect small-firm finance? *Journal of Finance* 65(3), 861-889.

Robins, J.M., 2000. Robust estimation in sequentially ignorable missing data and causal inference models. *Proceedings of the American Statistical Association Section on Bayesian Statistical Science 1999*, 6-10.

Roszbach, K., 2004. Bank lending policy, credit scoring, and the survival of loans. *Review of Economics and Statistics* 86(4), 946-958.

Schicks, J., Rosenberg, R., 2011. Too much microcredit? A survey on the evidence on over-indebtedness, Occasional Paper 19, CGAP, Washington, D.C.

Sharpe, S.A., 1990. Asymmetric information, bank lending, and implicit contracts: A stylized model of customer relationships. *Journal of Finance* 45, 1069-1087.

Singer, J.D., Willett, J.B., 1993. It's about time: Using discrete-time survival analysis to study duration and the timing of events. *Journal of Educational and Behavioral Statistics* 18(2), 155-195.

Stiglitz, J.E., Weiss, A., 1981. Credit rationing in markets with imperfect information. *American Economic Review* 71(3), 393-410.

von Thadden, E-L., 2004. Asymmetric information, bank lending, and implicit contracts: The winner's curse. *Finance Research Letters* 1, 11-23.

Table 1. Summary statistics

	Mean pre Credit Registry	Mean post Credit Registry	Obs.	Median	St. dev.	Min	Max
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Panel A: Extensive margin							
Loan rejected	0.088	0.164***	136,557	0	0.322	0	1
Proportion granted	0.898	0.827***	136,557	1	0.327	0	1
Loan rejected: private information	0.056	0.044***	136,557	0	0.221	0	1
Loan rejected: credit registry positive	0.015	0.060***	136,557	0	0.176	0	1
Loan rejected: credit registry negative	0.017	0.059***	136,557	0	0.181	0	1
Panel B: Intensive margin							
Dependent variables:							
Loan amount (BAM)	3,907	3,101***	116,517	3,000	3,136	300	30,000
Loan maturity	26.951	23.391***	116,517	24	12.649	1	86
Interest rate	20.060	21.616***	116,517	21	2.351	12	30
Collateral	2.928	2.675***	116,517	2	1.513	1	10
Problem loan	0.101	0.028***	116,517	0	0.261	0	1
Days late	4.200	4.095**	97,968	2	6.522	0	182
Return on loan	18.122	21.590***	92,313	21	8.496	-81	30
Independent variables:							
New borrower	0.546	0.412***	116,517	0	0.500	0	1
Borrower age	40	42***	116,517	41	12.216	18	82
Borrower male	0.576	0.609***	116,517	1	0.492	0	1
Borrower education	2.933	2.943***	116,487	3	0.388	1	4
Borrower monthly income (BAM)	1,242	1,172***	116,517	1,078	713	50	36,500
Borrower rural	0.601	0.660***	116,517	1	0.485	0	1
Loan immovable	0.099	0.106***	116,517	0	0.302	0	1
Loan movable	0.445	0.505***	116,517	0	0.499	0	1
Loan stock	0.299	0.200***	116,517	0	0.440	0	1
Personal collateral	0.473	0.463***	116,517	0	0.684	0	8
Social collateral	2.386	2.115***	116,517	2	1.083	0	10
Third-party collateral	0.069	0.097**	116,517	0	0.389	0	6

Notes: Sample period is June 2007-July 2011. Asterisks refer to the p-value of a t-test of equality of means. *** and ** indicates significance at the 1% and 5% level, respectively. BAM is Bosnian Convertible Mark.

Table 2. Extensive margin: Information sharing and loan rejections

Dependent variable →	Loan rejected			Proportion granted		
	OLS			Tobit		
	[1]	[2]	[3]	[4]	[5]	[6]
Credit registry	0.072***			-0.055***		
	(0.005)			(0.005)		
New borrower	0.107***	0.105***	0.106***	-0.104***	-0.105***	-0.105***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Credit registry*New Borrower	0.038***	0.038***	0.037***	-0.035***	-0.034***	-0.037***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
No. of applications	64,009	64,009	64,009	64,009	64,009	64,009
Adjusted (Pseudo) R^2	0.054	0.097	0.137	0.081	0.086	0.137
Applicant and loan covariates	Yes	Yes	Yes	Yes	Yes	Yes
Matching	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	No	Yes	No	No	Yes	No
Loan officer FE	No	Yes	No	No	Yes	No
Loan officer x month FE	No	No	Yes	No	No	Yes

Notes: This table shows linear probability (columns 1-3) regression results to explain the probability that a loan application was rejected and Tobit regression results (columns 4-6) explaining the ratio of loan amount granted to loan amount requested. Robust standard errors are clustered by month-loan officer and appear in parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively. Before credit registry: June 2008-June 2009. During credit registry: July 2009-July 2010. All specifications include applicant covariates, loan amount requested (in columns 1-3), loan type, and a time-varying night-light measure of local economic activity (in columns 1-2 and 4-5). Constant not shown. New borrower: Dummy variable that is '1' if the loan applicant has never borrowed from EKI (the lender) before; '0' otherwise. Table A2 in the Appendix contains all variable definitions.

Table 3. Extensive margin: Information sharing and loan rejections for repeat borrowers

Dependent variable →	Loan rejected			Proportion granted		
	OLS			Tobit		
	[1]	[2]	[3]	[4]	[5]	[6]
Credit registry	0.066***	0.052***		-0.073***	-0.061***	
	(0.006)	(0.008)		(0.007)	(0.009)	
Credit registry			0.054***			-0.063***
No registry at time of previous loan			(0.008)			(0.009)
Credit registry			0.018			-0.023
Registry at time of previous loan			(0.013)			(0.015)
No. of applications	32,034	12,198	12,198	32,034	12,198	12,198
Adjusted (Pseudo) R^2	0.045	0.051	0.052	0.074	0.078	0.079
Applicant and loan covariates	Yes	Yes	Yes	Yes	Yes	Yes
Loan officer FE	Yes	Yes	Yes	Yes	Yes	Yes
Sample	All repeat	Narrow	Narrow	All repeat	Narrow	Narrow

Notes: This table shows linear probability (columns 1-3) regression results to explain the probability that a loan application was rejected and Tobit regression results (columns 4-6) explaining the ratio of loan amount granted to loan amount requested. The sample contains repeat loans only. Robust standard errors are clustered by month-loan officer and appear in parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively. Before credit registry: June 2008-June 2009. During credit registry: July 2009-July 2010. All specifications include applicant covariates, loan amount requested (in columns 1-3), loan type, and a time-varying night-light measure of local economic activity. Constant not shown. Credit registry (*No registry at time of previous loan*): Dummy variable that is '1' if the previous loan was disbursed before the introduction of the credit registry and the current loan after the introduction of the credit registry; '0' otherwise. Credit registry (*Registry at time of previous loan*): Dummy variable that is '1' if the previous loan and the current loan were both disbursed after the introduction of the credit registry; '0' otherwise. In columns 2, 3, 5 and 6 the reference group is 2nd and 3rd loans disbursed before the introduction of the credit registry. Narrow sample refers to 2nd and 3rd loans which have the same purpose (e.g. agricultural inputs, fixed assets, working capital, etc.) and product type (e.g. small business, housing, revolving, etc.) as the previous loan disbursed to that client. Table A2 in the Appendix contains all variable definitions.

Table 4. Extensive margin: Loan rejections due to positive and negative registry information

(a) New borrowers

	Negative registry information	Positive registry information	Private information
	[1]	[2]	[3]
Credit registry	0.822***	0.704***	0.053
	(0.121)	(0.108)	(0.156)
No. of loans		32,192	
Pseudo R^2		0.031	

(b) Repeat borrowers (all)

	Negative registry information	Positive registry information	Private information
	[1]	[2]	[3]
Credit registry	1.575***	0.688***	0.364**
	(0.162)	(0.121)	(0.171)
No. of loans		31,625	
Pseudo R^2		0.032	

(c) Repeat borrowers (narrow)

	Negative registry information	Positive registry information	Private information
	[1]	[2]	[3]
Credit registry	1.681***	0.545***	0.226
	(0.249)	(0.146)	(0.220)
No. of loans		11,969	
Pseudo R^2		0.039	

(d) Types of repeat loans

	Negative registry information	Positive registry information	Private information
	[1]	[2]	[3]
Credit registry	1.460***	0.596***	0.234
No registry at time of previous loan	(0.234)	(0.139)	(0.202)
Credit registry	0.943***	0.116	-0.512
Registry at time of previous loan	(0.352)	(0.234)	(0.363)
No. of loans		11,969	
Pseudo R^2		0.040	

Notes: This table presents multinomial logit regressions to explain the probability that a loan application was rejected due to negative information from the credit registry (column 1), positive information from the credit registry (column 2) or due to private information (column 3). The base probability is that the application was accepted. Robust standard errors are clustered by month-loan officer and appear in parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively. Before credit registry: June 2008-June 2009. During credit registry: July 2009-July 2010. All specifications include applicant covariates, loan size and type, and a time-varying night-light measure of local economic activity. Constant not shown. Credit registry (*No registry at time of previous loan*): Dummy variable that is '1' if the previous loan was disbursed before the introduction of the credit registry and the current loan after the introduction of the registry; '0' otherwise. Credit registry (*Registry at time of previous loan*): Dummy variable that is '1' if the previous loan and the current loan were both disbursed after the introduction of the credit registry; '0' otherwise. In panels (C-D) the reference group is 2nd and 3rd loans disbursed before the introduction of the credit registry. Narrow sample refers to 2nd and 3rd loans with the same purpose (e.g. agricultural inputs, fixed assets, working capital, etc.) and product type (e.g. small business, housing, revolving, etc.) as the previous loan disbursed to that client. Table A2 in the Appendix contains all variable definitions.

Table 5. Intensive margin: Information sharing and loan terms

(a) Loan amount

	[1]	[2]	[3]	[4]
Credit registry	-0.177***	-0.155***		
	(0.007)	(0.008)		
New borrower		-0.052***	-0.053***	-0.061***
		(0.006)	(0.006)	(0.006)
Credit registry*New Borrower		-0.047***	-0.051***	-0.051***
		(0.010)	(0.010)	(0.010)
No. of loans	57,417	57,417	57,417	57,417
Adj. R^2	0.554	0.556	0.561	0.541

(b) Loan maturity

	[1]	[2]	[3]	[4]
Credit registry	-0.137***	-0.124***		
	(0.005)	(0.006)		
New borrower		-0.038***	-0.039***	-0.037***
		(0.005)	(0.005)	(0.005)
Credit registry*New Borrower		-0.029***	-0.033***	-0.035***
		(0.008)	(0.008)	(0.008)
No. of loans	57,417	57,417	57,417	57,417
Adj. R^2	0.450	0.452	0.457	0.464

(c) Interest rates

	[1]	[2]	[3]	[4]
Credit registry	0.511***	0.361***		
	(0.019)	(0.024)		
New borrower		0.048***	0.042***	0.042***
		(0.016)	(0.015)	(0.015)
Credit registry*New Borrower		0.315***	0.334***	0.327***
		(0.026)	(0.025)	(0.025)
No. of loans	57,417	57,417	57,417	57,417
Adj. R^2	0.643	0.645	0.657	0.663

(d) Collateral

	[1]	[2]	[3]	[4]
Credit registry	0.223***	0.191***		
	(0.015)	(0.017)		
New borrower		0.101***	0.106***	0.103***
		(0.011)	(0.011)	(0.011)
Credit registry*New Borrower		0.071***	0.068***	0.077***
		(0.019)	(0.018)	(0.018)
No. of loans	57,417	57,417	57,417	57,417
Adj. R^2	0.459	0.461	0.476	0.490
Borrower and loan covariates	Yes	Yes	Yes	Yes
Matching	Yes	Yes	Yes	Yes
Month FE	No	No	Yes	No
Loan officer FE	No	Yes	Yes	No
Loan officer x month FE	No	No	No	Yes

Notes: This table shows OLS regressions to estimate the impact of the introduction of the credit registry on loan amount (Panel A); loan maturity (Panel B); interest rate (Panel C) and number of pledged collateral items (Panel D). Before credit registry: June 2008-June 2009. During credit registry: July 2009-July 2010. Specifications in columns 1-3 include a time-varying night-light measure of local economic activity. Constant not shown. New borrower: Dummy variable that is '1' if the loan applicant has never borrowed from EKI (the lender) before; '0' otherwise. Table A2 in the Appendix contains all variable definitions. Robust standard errors are clustered by month-loan officer and appear in parentheses. ***, **, * correspond to the 1%, 5%, and 10% level of significance, respectively.

Table 6. Intensive margin: Information sharing and repeat borrowers

Dependent variable →	Δ% Loan amount		Δ% Loan maturity		Δ% Interest rate		Δ% Collateral	
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Credit registry	-0.341***		-0.194***		0.075***		0.160***	
	(0.029)		(0.027)		(0.005)		(0.026)	
Credit registry		-0.395***		-0.233***		0.087***		0.205***
No registry at time of previous loan		(0.030)		(0.028)		(0.005)		(0.027)
Credit registry		0.161**		0.176***		-0.033***		-0.259***
Registry at time of previous loan		(0.065)		(0.048)		(0.006)		(0.047)
No. of loans	8,414	8,414	8,414	8,414	8,414	8,414	8,414	8,414
Adjusted R^2	0.098	0.105	0.088	0.092	0.073	0.090	0.118	0.127
Loan and branch covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Loan officer FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table shows OLS regressions to estimate the impact of the introduction of the credit registry on the rate of change of loan amount [1-2]; loan maturity [3-4]; interest rate [5-6] and total number of collateral contracts [7-8]. Robust standard errors are clustered by month-loan officer and appear in parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively. Before credit registry: January 2008-June 2009. During credit registry: July 2009-December 2010. All specifications include a time-varying night-light measure of local economic activity and control dummies for product type. Constant not shown. Credit registry (*No registry at time of previous loan*): Dummy variable that is '1' if the previous loan was disbursed before the introduction of the credit registry and the current loan after the introduction of the credit registry; '0' otherwise. Credit registry (*Registry at time of previous loan*): Dummy variable that is '1' if the previous loan and the current loan were both disbursed after the introduction of the credit registry; '0' otherwise. The sample consists of all 2nd and 3rd loans that have the same purpose (e.g. agricultural inputs, fixed assets, working capital, etc.) and product type (e.g. small business, housing, revolving, etc.) as the previous loan disbursed to that client. The reference group is 2nd and 3rd loans disbursed before the introduction of the credit registry. Table A2 in the Appendix contains all variable definitions.

Table 7. Information sharing and loan quality: Regression analysis

	[1]	[2]	[3]	[4]
Credit registry	-0.048***	-0.024***		
	(0.005)	(0.003)		
New borrower	-0.000	-0.001	-0.001	-0.001
	(0.003)	(0.003)	(0.003)	(0.003)
Credit registry*New borrower	-0.018***	-0.009**	-0.009**	-0.009**
	(0.004)	(0.004)	(0.004)	(0.004)
No. of loans	57,445	57,445	57,445	57,445
Adjusted R^2	0.049	0.196	0.196	0.199
Borrower and loan covariates	Yes	Yes	Yes	Yes
Matching	Yes	Yes	Yes	Yes
Month FE	No	No	Yes	No
Loan officer FE	No	Yes	Yes	No
Loan officer x month FE	No	No	No	Yes

Notes: This table shows loan-level linear probability regressions to estimate the impact of the introduction of the credit registry on the probability of a loan defaulting. Before credit registry: June 2008-June 2009. During credit registry: July 2009-July 2010. Specifications in columns 1-3 include a time-varying night-light measure of local economic activity. Constant not shown. New borrower: Dummy variable that is '1' if the loan applicant has never borrowed from EKI (the lender) before; '0' otherwise. Table A2 in the Appendix contains all variable definitions. Robust standard errors are clustered by month-loan officer and appear in parentheses. ***, **, * correspond to the 1%, 5%, and 10% level of significance, respectively.

Table 8. Information sharing and loan quality: Hazard analysis

Dependent variable →	Hazard ratio		
Functional form →	Cox	Exponential	Weibull
	[1]	[2]	[3]
Credit registry	-0.674***	-0.610***	-0.642***
	(0.067)	(0.071)	(0.068)
New borrower	0.031	0.004	0.017
	(0.037)	(0.041)	(0.039)
Credit registry*New borrower	-0.330***	-0.326***	-0.326***
	(0.107)	(0.113)	(0.110)
Ln(Alpha)			-0.645***
			(0.023)
No. of loans	57,581	57,581	57,581
Log-likelihood ratio	-36,176	-22,904	-21,628
Borrower and loan covariates	Yes	Yes	Yes
Branch stratification	Yes	Yes	Yes

Notes: This table shows the results of a Cox proportional hazard model [1], a parametric exponential hazard model [2] and a parametric Weibull hazard model [3]. The dependent variable is the hazard rate, the probability that a loan i is defaulted on in a given month t given that default did not occur earlier. A default event occurs when a borrower is at least 30 days late in making a payment and the loan was eventually written off. Sample period: June 2008-July 2010. Credit registry is a dummy variable that is '1' if the credit registry was in place in a given month, zero otherwise. New borrower: Dummy variable that is '1' if the loan applicant has never borrowed from EKI (the lender) before; '0' otherwise. All specifications include a time-varying night-light measure of local economic activity and are stratified at the branch level. Robust standard errors are clustered by month-loan officer and appear in parentheses. ***, **, * correspond to the 1%, 5%, and 10% significance level, respectively. Table A2 in the Appendix contains all variable definitions.

Table 9. Information sharing and loan quality: Hazard analysis for repeat borrowers

Dependent variable →	Hazard ratio (Cox)		
	[1]	[2]	[3]
Credit registry	-0.605***	-0.379***	
	(0.070)	(0.110)	
Credit registry			-0.343***
No registry at time of previous loan			(0.111)
Credit registry			-0.864**
Registry at time of previous loan			(0.346)
No. of loans	29,472	8,434	8,434
Log-likelihood ratio	-22,165	-6,405	-6,404
Borrower and loan covariates	Yes	Yes	Yes
Sample	All repeat	Narrow	Narrow

Notes: This table shows the results of a Cox proportional hazard model where the dependent variable is the hazard rate, the probability that a loan i is defaulted on in a given month t given that default did not occur earlier. A default event occurs when a borrower is at least 30 days late in making a payment and the loan was eventually written off. Sample period: June 2008-July 2010. All specifications include a time-varying night-light measure of local economic activity and are stratified at the branch level. Robust standard errors are clustered by month-loan officer and appear in parentheses. ***, **, * correspond to the 1%, 5%, and 10% significance level, respectively. Credit registry is a dummy variable that is '1' if the credit registry was in place in a given month, zero otherwise. Credit registry (*No registry at time of previous loan*): Dummy variable that is '1' if the previous loan was disbursed before the introduction of the credit registry and the current loan after the introduction of the credit registry; '0' otherwise. Credit registry (*Registry at time of previous loan*): Dummy variable that is '1' if the previous loan and the current loan were both disbursed after the introduction of the credit registry; '0' otherwise. In columns 2 and 3 the reference group is 2nd and 3rd loans disbursed before the introduction of the credit registry. Narrow sample refers to 2nd and 3rd loans which have the same purpose (e.g. agricultural inputs, fixed assets, working capital, etc.) and product type (e.g. small business, housing, revolving, etc.) as the previous loan disbursed to that client. Table A2 in the Appendix contains all variable definitions.

Table 10. Information sharing and late repayment

	Days late				2-15 days late	16-30 days late
	[1]	[2]	[3]	[4]	[5]	[6]
Credit registry	0.085	0.084				
	(0.058)	(0.113)				
New borrower	0.026	0.037	0.032	0.034	-0.033	0.131***
	(0.042)	(0.042)	(0.041)	(0.042)	(0.025)	(0.035)
Credit registry*New borrower	-0.660***	-0.603***	-0.583***	-0.627***	-0.042	-0.105**
	(0.073)	(0.071)	(0.071)	(0.073)	(0.042)	(0.053)
No. of loans	48,217	48,217	48,217	48,217	47,942	
Log likelihood	-163,082	-162,347	-162,330	-162,838	-56,289	
Adjusted R^2	0.002	0.003	0.007	0.003	0.040	
Borrower and loan covariates	Yes	Yes	Yes	Yes	Yes	
Matching	Yes	Yes	Yes	Yes	Yes	
Month FE	No	No	Yes	Yes	Yes	
Loan officer FE	No	Yes	Yes	No	Yes	
Loan officer x month FE	No	No	No	Yes	No	

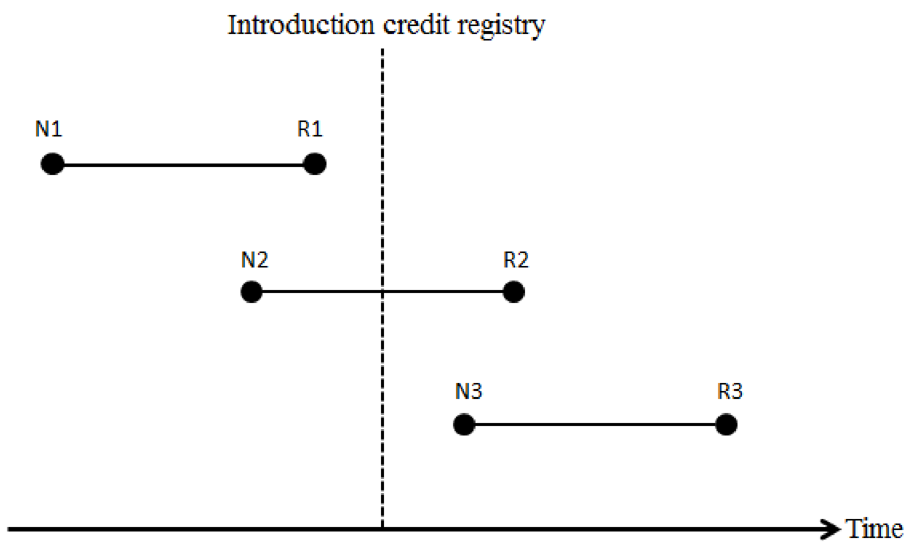
Notes: This table shows loan-level Tobit regressions (columns [1] to [4]) and multinomial logit regressions (columns [5] and [6]) to estimate the impact of the introduction of the credit registry on the number of days a loan is being paid late. Being more than 15 days late leads to a downgrade from class A to B in the loan-quality classification of the credit registry. The base category in columns [6] and [7] consists of loans that are either on time or only one day late. Before credit registry: June 2008-June 2009. During credit registry: July 2009-July 2010. Specifications in columns 1-3 include a time-varying night-light measure of local economic activity. Constant not shown. New borrower: Dummy variable that is '1' if the loan applicant has never borrowed from EKI (the lender) before; '0' otherwise. Table A2 in the Appendix contains all variable definitions. Robust standard errors are clustered by month-loan officer and appear in parentheses. ***, **, * correspond to the 1%, 5%, and 10% level of significance, respectively.

Table 11. Information sharing and return on loans

	[1]	[2]	[3]	[4]
Credit registry	0.011***	0.007***		
	(0.001)	(0.002)		
New borrower		-0.001	-0.001	-0.001
		(0.002)	(0.002)	(0.001)
Credit registry*New Borrower		0.010***	0.010***	0.010***
		(0.002)	(0.002)	(0.002)
No. of loans	57,392	57,392	57,392	57,392
Adjusted R^2	0.182	0.182	0.184	0.188
Borrower and loan covariates	Yes	Yes	Yes	Yes
Matching	Yes	Yes	Yes	Yes
Month FE	No	No	Yes	No
Loan officer FE	No	Yes	Yes	No
Loan officer x month FE	No	No	No	Yes

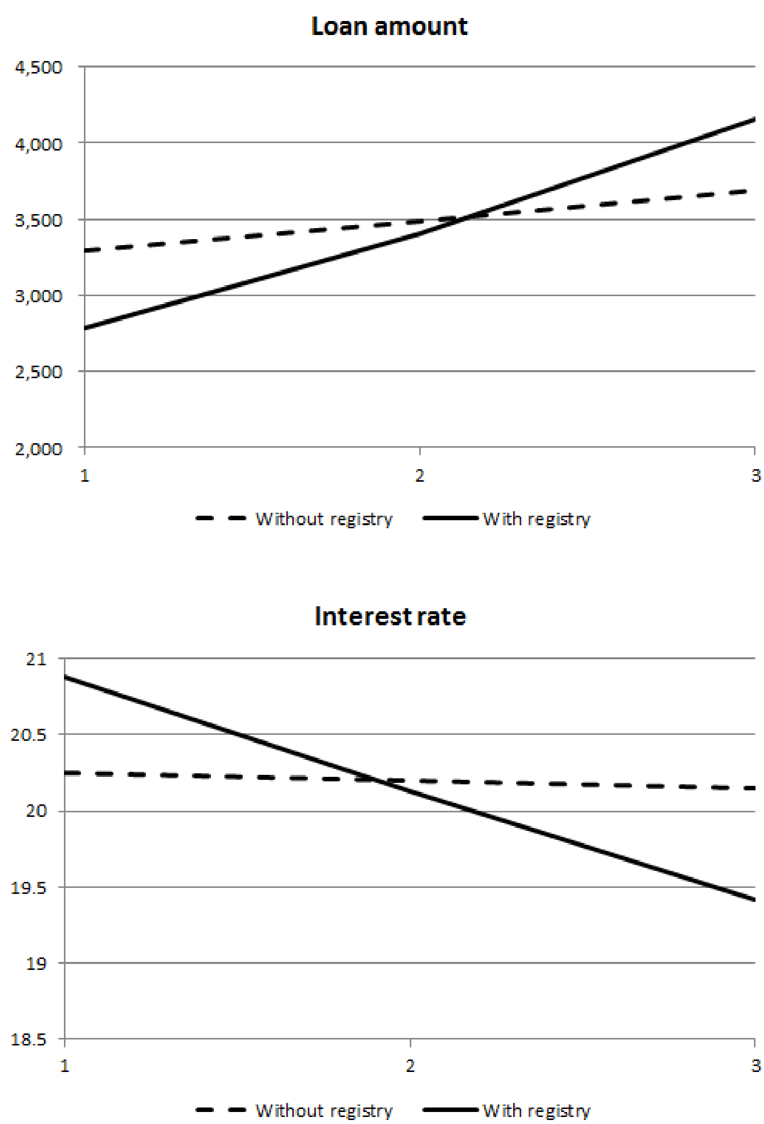
Notes: This table shows loan-level OLS regressions to estimate the impact of the introduction of the credit registry on the return on loans. Before credit registry: June 2008-June 2009. During credit registry: July 2009-July 2010. Specifications in columns 1-3 include a time-varying night-light measure of local economic activity. Constant not shown. New borrower: Dummy variable that is '1' if the loan applicant has never borrowed from EKI (the lender) before; '0' otherwise. Table A2 in the Appendix contains all variable definitions. Robust standard errors are clustered by month-loan officer and appear in parentheses. ***, **, * correspond to the 1%, 5%, and 10% level of significance, respectively.

Figure 1. Information sharing: New versus repeat borrowers



Notes: This figure provides a schematic overview of the three types of repeat borrowers in our empirical analysis. Dots indicate loans and horizontal lines indicate lending relationships. Each lending relationship consists of two subsequent loans: a first-time loan (N) to a new borrower and then a repeat loan (R) to that same borrower. Lending relationship 1 (top) consists of two loans that were both granted before the introduction of the credit registry. Lending relationship 2 (middle) consists of two loans, the first of which was granted before the introduction of the credit registry and the second one afterwards. Lending relationship 3 (bottom) consists of two loans that were both granted after the introduction of the credit registry.

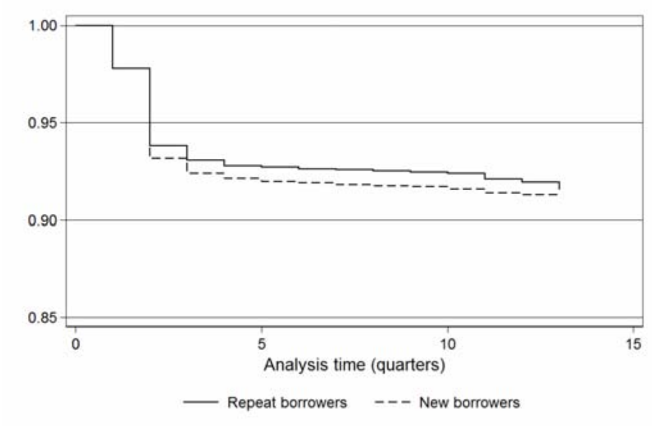
Figure 2. Intensive margin: Information sharing and repeat borrowers



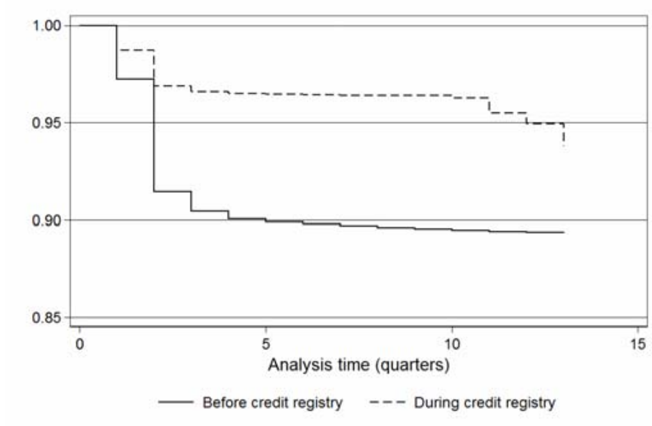
Notes: These charts summarize the impact of the credit registry on first, second, and third loans in terms of loan amount (in Bosnian Convertible Mark (BAM), top) and interest rate (in percent, bottom). The horizontal axis indicates the first, second and third loan in a lending relationship. The solid lines show for these three consecutive loans the amount (top) and interest rate (bottom) for relationships established after the introduction of the credit registry, so that all three loans were granted with information sharing in place. The striped lines show the loan amount (top) and interest rate (bottom) for relationships before the introduction of the credit registry, so that all three loans were granted without information sharing in place. The amounts and interest rates are conditional on our standard set of borrower covariates.

Figure 3. Information sharing and loan quality: Kaplan-Meier survival analysis

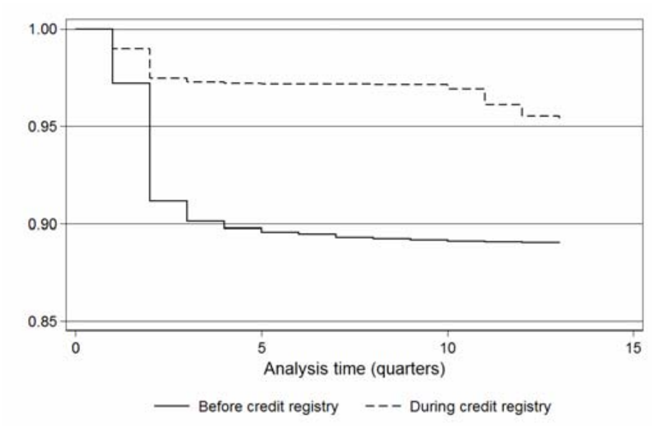
(a) New versus repeat borrowers



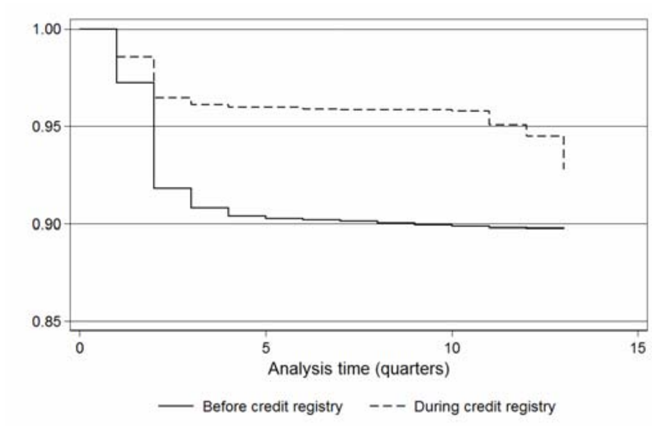
(b) Before versus during credit registry



(c) Before versus during credit registry: new

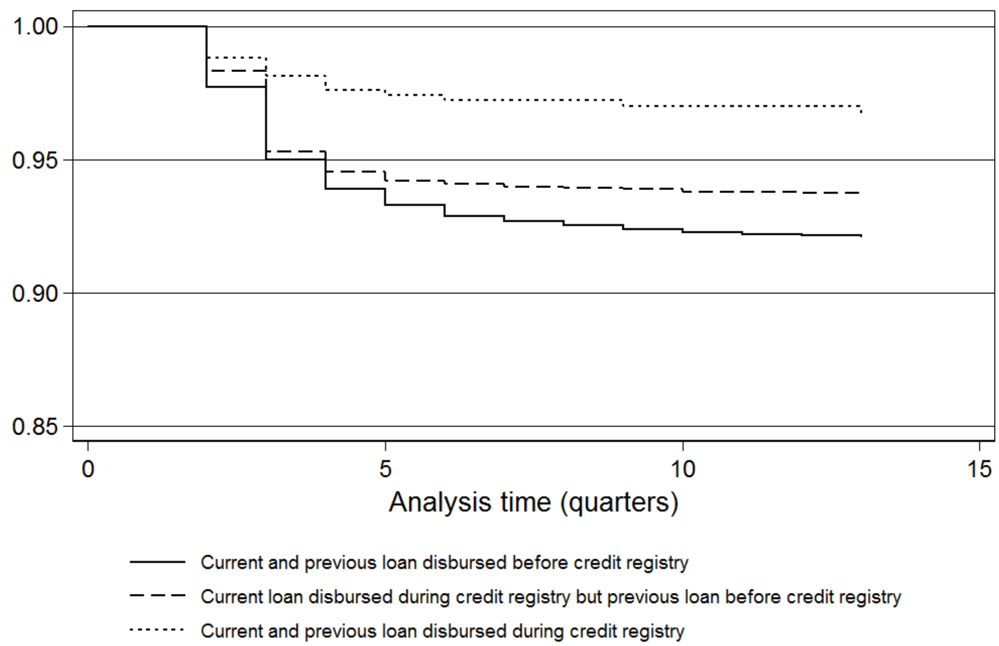


(d) Before versus during credit registry: repeat



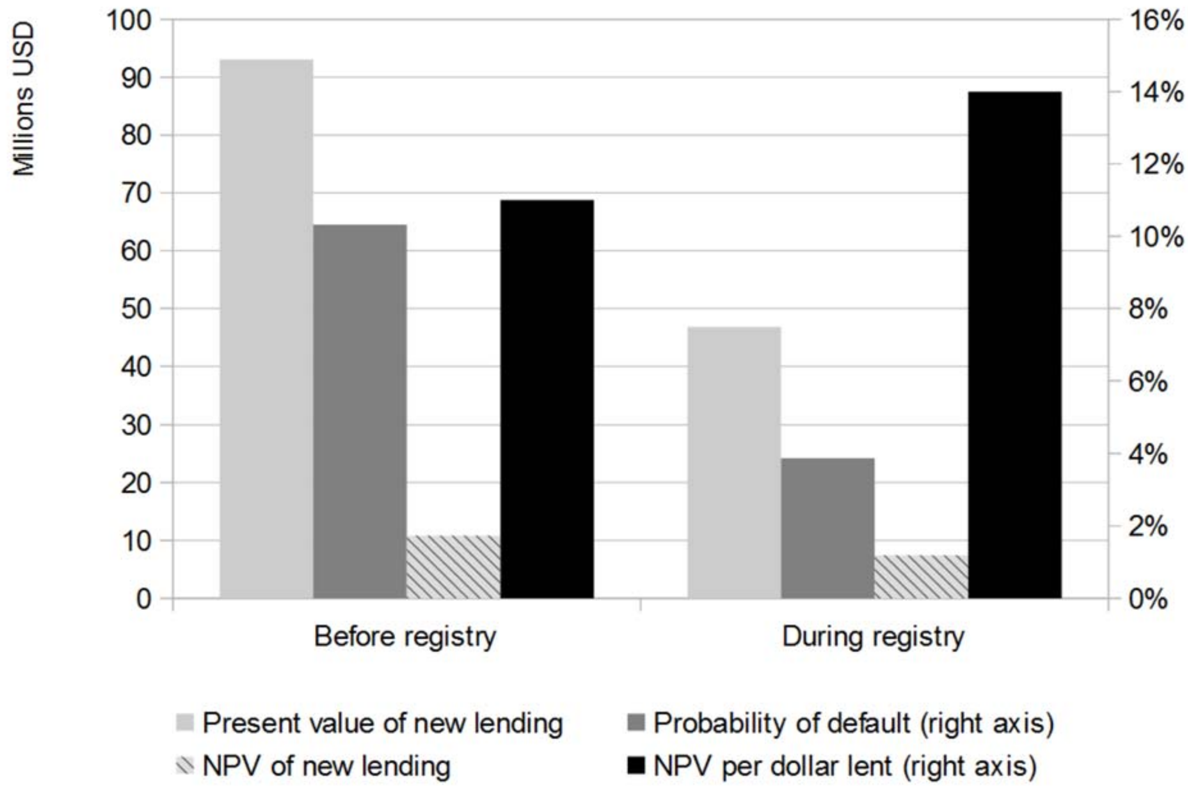
Notes: These four graphs show Kaplan-Meier survival estimates over the sample period June 2008-July 2010. Logrank test statistics for differences between the curves:
Panel A: $\chi^2(1) = 8.19$ (*p-value*= 0.00). Panel B: $\chi^2(1) = 706.30$ (*p-value*= 0.00).
Panel C: $\chi^2(1) = 431.52$ (*p-value*= 0.00); Panel D: $\chi^2(1) = 278.62$ (*p-value*= 0.00).

Figure 4. Information sharing and loan quality: Effect on different types of repeat borrowers



Notes: These graphs show Kaplan-Meier survival estimates over the sample period June 2008-July 2010. These estimates are based on a sample of 2nd or 3rd loans that have the same purpose and product type as the previous loan disbursed to the same client.

Figure 5. Information sharing and aggregate lending profitability



Notes: This figure compares the portfolio of all loans disbursed in the year before (June 2008-June 2009, left) and after (July 2009-July 2010, right) the introduction of the credit registry. The present value of new lending (left axis) is the total amount of new lending, net of fees, disbursed in the year before (after) the credit-registry introduction. The present value of these amounts is calculated by discounting back to the beginning of each period, using the average funding cost of the lender, and is shown as a positive number. The probability of default is the probability that repayment on a loan is at least 30 days late and that eventually part of or all the loan is written off (right axis). The NPV is the net present value of new lending in the year before (after) the credit-registry introduction (left axis). The net present value is calculated as the present value of all loan repayments and interest payments minus the initial loan disbursements (net of fees). This amount is discounted back to the beginning of each period using the average funding cost of the lender. The NPV per dollar lent is the net present value divided by the present value of total lending in each of the two periods (right axis).

Appendix

Table A1. Summary statistics: new vs. repeat borrowers

	Mean pre Credit registry	Mean post Credit registry	Obs.	Median	St. dev.	Min	Max
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Panel A: Extensive margin							
Loan rejected							
New borrowers	0.131	0.247***	69,381	0	0.375	0	1
Repeat borrowers	0.033	0.101***	67,176	0	0.244	0	1
Proportion granted							
New borrowers	0.855	0.745***	69,381	1	0.378	0	1
Repeat borrowers	0.952	0.889***	67,176	1	0.252	0	1
Loan rejected: private information							
New borrowers	0.088	0.077***	69,381	0	0.278	0	1
Repeat borrowers	0.015	0.019***	67,176	0	0.129	0	1
Loan rejected: credit registry positive							
New borrowers	0.022	0.093***	69,381	0	0.209	0	1
Repeat borrowers	0.005	0.035***	67,176	0	0.134	0	1
Loan rejected: credit registry negative							
New borrowers	0.021	0.077***	69,381	0	0.194	0	1
Repeat borrowers	0.013	0.047***	67,176	0	0.165	0	1
Panel B: Intensive margin							
Dependent variables:							
Loan amount (BAM)							
New borrowers	3,589	2,845***	57,823	3,000	2,875	300	30,000
Repeat borrowers	4,293	3,281***	58,694	3,000	3,357	300	30,000
Loan maturity							
New borrowers	26.331	22.354***	57,823	24	12.003	1	86
Repeat borrowers	27.700	24.117***	58,694	24	13.232	1	86
Interest rate							
New borrowers	20.183	21.955***	57,823	21	2.328	12	30
Repeat borrowers	19.912	21.379***	58,694	21	2.37	12	30
Collateral							
New borrowers	2.907	2.709***	57,823	3	1.409	1	7
Repeat borrowers	2.884	2.640***	58,694	2	1.454	1	7
Problem loan							
New borrowers	0.1	0.027***	57,823	0	0.268	0	1
Repeat borrowers	0.102	0.028***	58,694	0	0.255	0	1
Days late							
New borrowers	4.259	3.966***	48,188	2	6.428	0	182
Repeat borrowers	4.131	4.184***	49,339	2	6.611	0	182
Return on loan							
New borrowers	18.4	21.9***	45,455	21	8.6	-81	26
Repeat borrowers	17.8	21.4***	46,858	21	8.4	-79	30

Notes: Sample period is June 2007-July 2011. Asterisks refer to the p-value of a t-test of equality of means. *** and ** indicates significance at the 1% and 5% level, respectively. BAM is Bosnian Convertible Mark.

Table A2. Variable definitions and data sources

	Definition	Source	Unit
Panel A: Extensive margin			
Dependent variables:	Definition	Source	Unit
Loan rejected	Dummy=1 if loan application is rejected.	EKI	Dummy
Proportion granted	Ratio of loan amount granted to loan amount requested.	EKI	%
Loan rejected: credit registry (negative)	Dummy=1 if loan application is rejected because of a low credit score or repayment history in the registry.	EKI	Dummy
Loan rejected: credit registry (positive)	Dummy=1 if loan application is rejected because of too many outstanding loans with competing lenders.	EKI	Dummy
Panel B: Intensive margin			
Dependent variables:	Definition	Source	Unit
Loan amount	Loan amount at time of disbursement.	EKI	BAM
Loan maturity	Maturity of the loan at time of disbursement.	EKI	Months
Interest rate	Annual nominal interest rate.	EKI	%
Collateral	Total number of collateral items pledged.	EKI	Discrete
Problem loan	Dummy=1 if borrower was at any time at least 30 days late in making a payment and the loan was subsequently written off.	EKI	Dummy
Days late	Number of days loan is late on first late repayment.	EKI	Discrete
Return on loan	Measure of loan profitability taking into account loss given default.	EKI	%
Independent variables:			
Credit registry	Dummy=1 for all months after and including July 2009 (time of CRK introduction); 0 otherwise.	Central Bank of Bosnia	Dummy
New borrower	Dummy =1 if the loan applicant has never borrowed from EKI (the lender) before; 0 otherwise.	EKI	Dummy
Borrower age	Borrower age.	EKI	Years
Borrower male	Dummy= 1 if borrower is male; 0 otherwise.	EKI	Dummy
Borrower education	1 = None, 2 = Primary, 3 = Secondary, 4 = Tertiary (College/University/Post Graduate).	EKI	1 to 4
Borrower monthly income	Total monthly borrower income (primary plus secondary income source).	EKI	BAM
Borrower rural	0 = Urban; 1 = Rural.	EKI	Dummy
Loan immovable	Loan purpose = Purchase immovable assets (land and/or buildings).	EKI	Dummy
Loan movable	Loan purpose = Purchase movable assets (equipment, fixed assets, vehicles).	EKI	Dummy
Loan stock	Loan purpose = Purchase of stock (merchandise, raw material, working capital, agricultural inputs, livestock for reproduction, seedlings for orchards).	EKI	Dummy
Personal collateral	Number of personal collateral pledges for each loan (includes mortgages, administrative bans on the borrower's salary, and pledges of movable assets).	EKI	Discrete
Social collateral	Number of social collateral pledges for each loan (includes total and partial guarantees provided by family and friends of the borrower).	EKI	Discrete
Third-party collateral	Number of third party collateral pledges for each loan (includes checks or bills of exchange issued by a guarantor company).	EKI	Discrete

Notes: BAM is Bosnian Convertible Marka. BEPS is the EBRD Banking Environment and Performance Survey. MIX: www.mixmarket.org/.

TABLE A3. Extensive margin: Robustness and placebo tests

Dependent variable →	**Loan rejected**					
	Robustness tests			Placebo tests		
	Narrow window	Broad window	Broadest window	Post is pre	Pre is post	Random assignment
	[1]	[2]	[3]	[4]	[5]	[6]
New borrower	0.107***	0.104***	0.107***	0.108***	0.143***	0.001
	(0.006)	(0.004)	(0.004)	(0.006)	(0.006)	(0.006)
Credit registry*New borrower	0.029***	0.039***	0.042***	0.008	0.005	0.000
	(0.009)	(0.007)	(0.006)	(0.007)	(0.010)	(0.000)
Applicant covariates	Yes	Yes	Yes	Yes	Yes	Yes
Loan covariates	Yes	Yes	Yes	Yes	Yes	Yes
Loan officer x month FE	Yes	Yes	Yes	Yes	Yes	Yes
No. of applications	46,484	88,929	108,162	76,860	37,185	64,148
Adjusted R^2	0.133	0.140	0.133	0.133	0.138	0.120

Dependent variable →	**Proportion granted**					
	Robustness tests			Placebo tests		
	Narrow window	Broad window	Broadest window	Post is pre	Pre is post	Random assignment
	[1]	[2]	[3]	[4]	[5]	[6]
New borrower	-0.103***	-0.101***	-0.103***	-0.107***	-0.139***	-0.001
	(0.006)	(0.004)	(0.004)	(0.005)	(0.006)	(0.006)
Credit registry*New borrower	-0.030***	-0.038***	-0.036***	-0.009	-0.001	0.000
	(0.007)	(0.007)	(0.006)	(0.007)	(0.010)	(0.001)
Applicant covariates	Yes	Yes	Yes	Yes	Yes	Yes
Loan covariates	Yes	Yes	Yes	Yes	Yes	Yes
Loan officer x month FE	Yes	Yes	Yes	Yes	Yes	Yes
No. of applications	46,484	88,929	108,162	76,860	37,185	64,148
Pseudo R^2	0.124	0.130	0.133	0.124	0.129	0.120

Notes: Columns [1], [2] and [3] show robustness tests of our main results as reported in Table 3. In columns [1] we use a shorter time window September 2008-April 2010. In column [2] the window is January 2008-December 2010. In column [3] we use an even larger window June 2007-July 2011. Columns [4], [5] and [6] show placebo tests of our main results as reported in Table 3. In columns [4] and [5] we move the two-year window one year forward and backward, respectively. In column [6], we randomly allocate loans to either the new or repeat borrower group. We repeat this random allocation a thousand times and show the average result. The treatment period starts in July 2009. Credit registry is a dummy variable that is '1' if the credit registry was in place in a given month, zero otherwise. New borrower: Dummy variable that is '1' if the loan applicant has never borrowed from EKI (the lender) before; '0' otherwise. A dummy for new borrowers is included but not shown. Robust standard errors are clustered by month-loan officer and appear in parentheses. ***, **, * correspond to the 1%, 5%, and 10% level of significance, respectively. Table A2 in the Appendix contains all variable definitions. The same borrower and loan covariates as in Table 3 are included but not shown.

Figure A1. Data structure

Panel A

<table>
<tr><td colspan="4">Loan applications
(N=136,557)</td></tr>
<tr><td colspan="2">Loan applications before credit registry
(N=83,726)</td><td colspan="2">Loan applications during credit registry
(N=52,831)</td></tr>
<tr><td>Repeat borrowers
(N=37,291)</td><td>New borrowers
(N=46,435)</td><td>Repeat borrowers
(N=29,885)</td><td>New borrowers
(N=22,946)</td></tr>
</table>

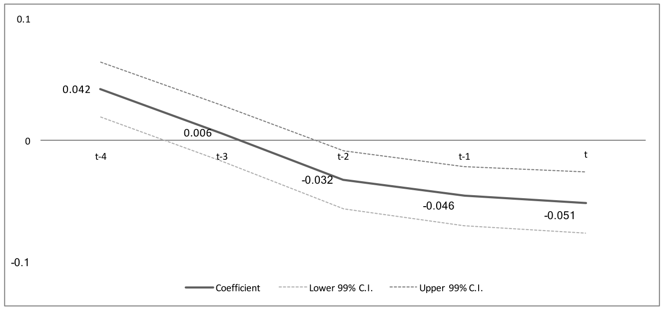
Panel B

<table>
<tr><td colspan="4">Approved loans
(N=116,517)</td></tr>
<tr><td colspan="2">Approved before credit registry
(N=73,191)</td><td colspan="2">Approved during credit registry
(N=43,326)</td></tr>
<tr><td>Repeat borrowers
(N=33,206)</td><td>New borrowers
(N=39,985)</td><td>Repeat borrowers
(N=25,484)</td><td>New borrowers
(N=17,842)</td></tr>
</table>

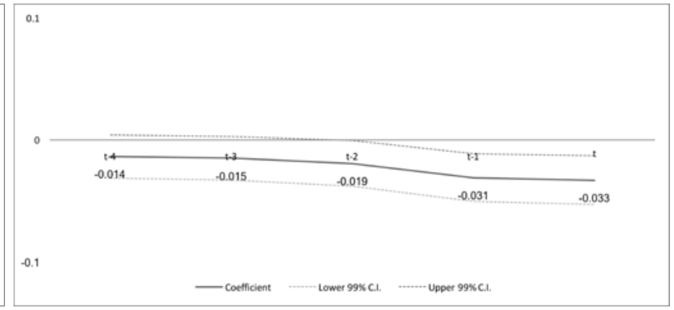
Notes: This figure summarizes the sample of loan applications and approved loans for the period June 2007-July 2011. During credit registry: July 2009-July 2010. Of all applications, 12,017 were rejected by the lender and 8,023 were withdrawn by the borrower before a lending decision was taken or the loan was disbursed.

Figure A2. Intensive margin: Placebo tests

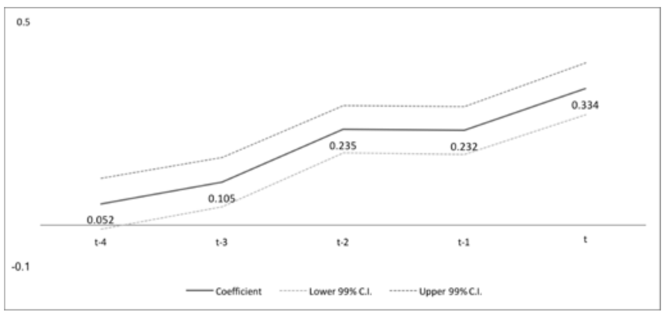
(A) Loan amount



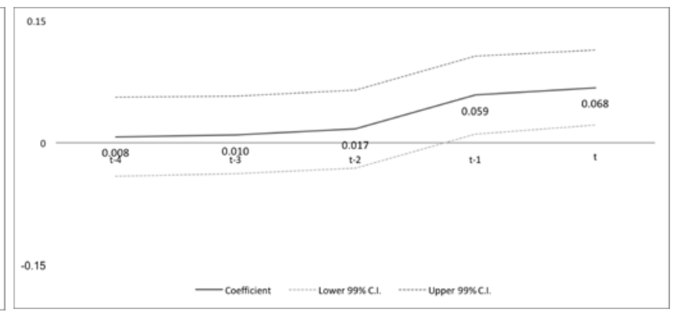
(B) Loan maturity



(C) Interest rate

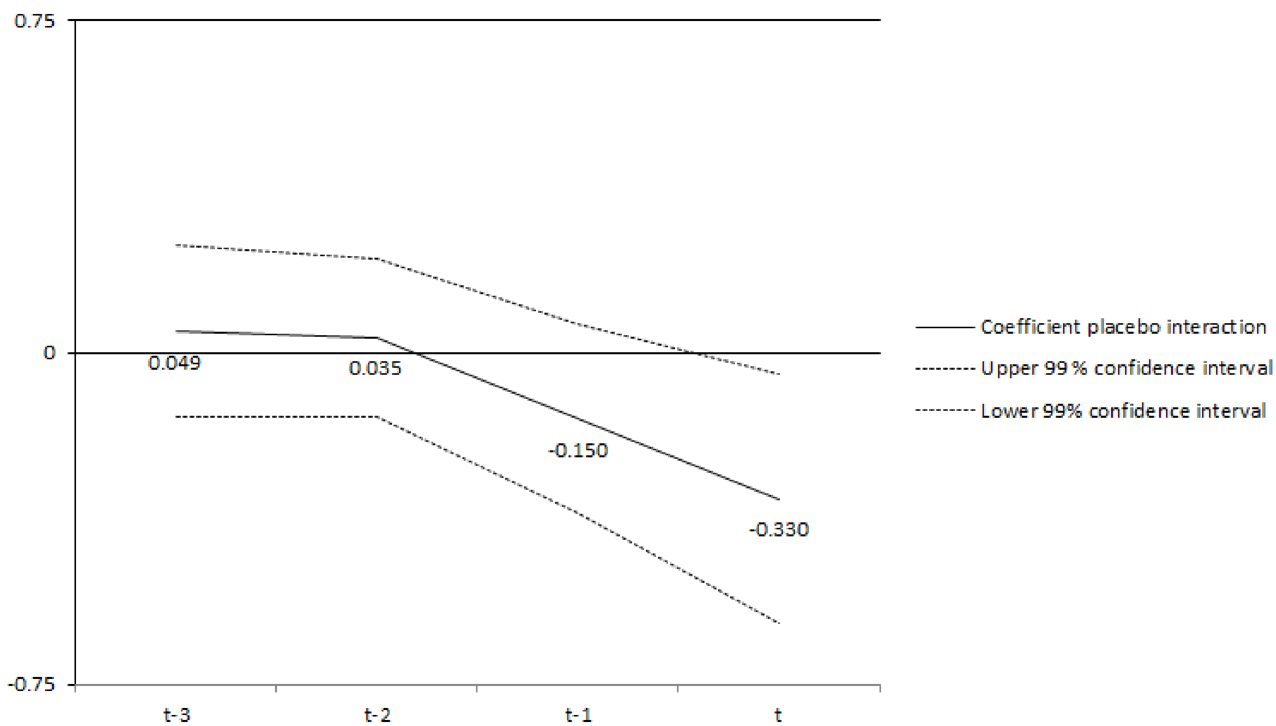


(D) Collateral



Notes: This graph shows the coefficient estimates (and a 99% confidence interval) for the interaction term *Creditregistry***Newborrower* from column 3 of Table 5. The value at t shows the coefficient when using the actual timing of the credit registry introduction. The values at $t-1$, $t-2$, etc. show the coefficient estimates when introducing the credit registry 1 quarter, 2 quarters, etc. earlier than the real introduction date.

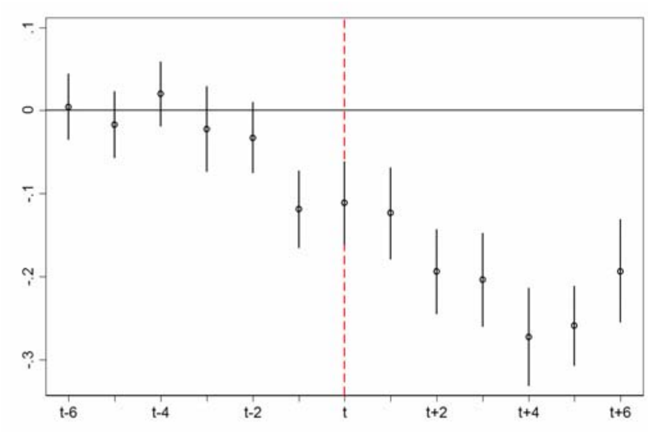
Figure A3. Cox proportional hazard model: Placebo test



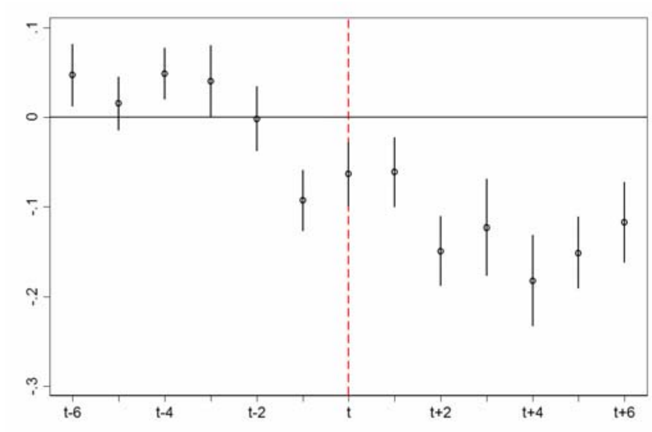
Notes: This graph shows coefficient estimates (and a 95% confidence interval) for the interaction term $Creditregistry*Newborrower$ from column 1 of Table 8. The value at t shows the coefficient when using the actual timing of the credit registry introduction. The values at $t-1$, $t-2$, etc. show the coefficient estimates when introducing the registry 1 quarter, 2 quarters, etc. earlier than the real introduction date.

Figure A4. Loan terms: Parallel trends for new and repeat borrowers

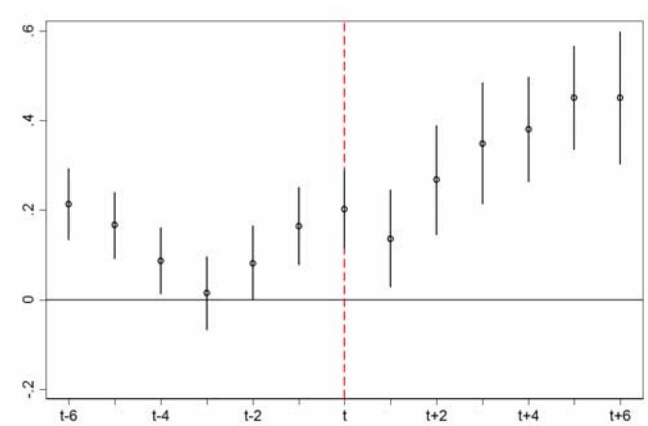
(a) Loan amount



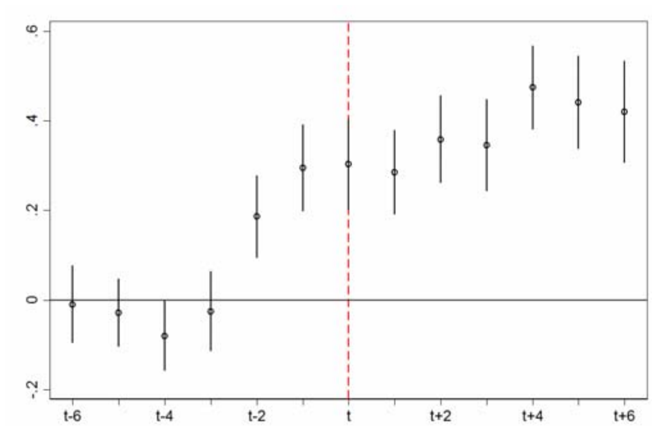
(b) Loan maturity



(c) Interest rate



(d) Collateral



Notes: Parallel trend test over the sample period June 2008-July 2010. We add to our specification (2) twelve interaction terms between our treatment variable and dummy variables that are one in a single month of the period consisting of the year before and the year after the introduction of the credit registry. The graphs report the estimated coefficients and 99% confidence intervals for these interaction terms.